Human–Centered System Design for Electronic Governance

Saqib Saeed
Bahria University Islamabad, Pakistan

Christopher G. Reddick
University of Texas at San Antonio, USA

Information Science
REFERENCE

Managing Director:	Lindsay Johnston
Editorial Director:	Joel Gamon
Book Production Manager:	Jennifer Yoder
Publishing Systems Analyst:	Adrienne Freeland
Development Editor:	Austin DeMarco
Assistant Acquisitions Editor:	Kayla Wolfe
Typesetter:	Christy Fic
Cover Design:	Jason Mull

Published in the United States of America by
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Library of Congress Cataloging-in-Publication Data

Human-centered system design for electronic governance / Saqib Saeed, Christopher G. Reddick, Editors.
pages cm
Includes bibliographical references and index.
Summary: “This book provides special attention to the most successful practices for implementing e-government technologies, highlighting the benefits of well designed systems in this field, while investigating the implications of poor practices and designs”-- Provided by publisher.
ISBN 978-1-4666-3640-8 (hardcover) -- ISBN (invalid) 978-1-4666-3641-5 (ebook) -- ISBN 978-1-4666-3642-2 (print & perpetual access) 1. Internet in public administration--Technological innovations. I. Saeed, Saqib, 1970- editor of compilation. II. Reddick, Christopher G., editor of compilation.
JF1525.A8H86 2013
352.3’802854678--dc23
2012045322

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

Chapter 11

E-Voting System Usability:
Lessons for Interface Design, User Studies, and Usability Criteria

M. Maina Olembo
Technische Universität Darmstadt, Germany

Melanie Volkamer
Technische Universität Darmstadt, Germany

ABSTRACT

The authors present a literature review, carried out by searching through conference proceedings, journal articles, and other secondary sources for papers focusing on the usability of electronic voting (e-voting) systems and related aspects such as ballot design and verifiability. They include both user studies and usability reviews carried out by HCI experts and/or researchers, and analyze the literature specifically for lessons on designing e-voting system interfaces, carrying out user studies in e-voting and applying usability criteria. From these lessons learned, the authors deduce recommendations addressing the same three aspects. In addition, they identify for future research open questions that are not answered in the literature. The recommendations hold for e-voting systems in general, but this chapter especially focuses on remote e-voting systems providing cryptographic verifiability, as the authors consider these forms as most promising for the future.

INTRODUCTION

Electronic Voting (e-voting) systems continue to be used in different countries and contexts around the globe, enabling governments to obtain information on citizens' preferences more quickly and efficiently. Systems in use are both e-voting machines and remote e-voting systems. Four reasons can be stated why usability is important in e-voting systems. First, due to the election principle of universal suffrage, anyone who meets the voting age requirement[1] should be able to use these systems to cast his vote. This includes first time voters, elderly persons, and even those who

DOI: 10.4018/978-1-4666-3640-8.ch011

do not frequently interact with technology. Second, a voter should be able to easily express his wishes. The interface design should neither cause him to make mistakes nor influence his decision. Poor interface design can easily cause a voter not to cast a vote for his desired candidate. Third, voters remain novices due to a lack of training and irregular interaction with these systems, since elections are held infrequently in many countries and contexts. As such, the learning that occurs from continuous interaction with systems over a period of time is less likely to occur. Finally, if the usability of e-voting systems is not considered, frustration is likely to occur, reducing acceptance among voters, and thus decreasing voter turnout. Usability issues are especially important in e-voting systems that provide verifiability and in particular when using cryptographic verifiability.

Verifiable voting systems have been discussed since their proposal by Cohen and Fischer (1985). We mainly distinguish e-voting systems which implement voter verifiable paper audit trails (VVPAT) and those using cryptographic means for verifiability. The premise is that individual voters are able to verify that their vote is cast as they intended, and stored as they cast it. In addition, voters and any interested parties are able to verify that all votes are tallied as stored. Voters will then have to carry out certain steps to verify votes, and may encounter unfamiliar terminology, such as encryption. However, verifiable systems are only beneficial if any voter who wants to verify his vote can do so without being a specialist. The German Federal Court Decision (2009) backs this stance, requiring that the correctness of all essential steps in the election are publicly examinable without having specialist knowledge. Verifiable e-voting systems are thus even more challenging from a usability point of view; however, the future of e-voting lies in verifiable e-voting as black box systems (where one has to trust that technical and organizational processes are correct, yet no way of testing this is provided) continue to face criticism (Alvarez & Hall, 2008, p. 31).

There is a lot of literature available on usability studies of e-voting systems. Researchers and developers of future e-voting systems should take this into account, especially if they address the most complex systems, namely verifiable and cryptographically-verifiable remote e-voting systems.

In this book chapter, we review existing literature on usability of e-voting systems. Note that the literature is primarily from a western perspective, specifically from American and European contexts. The literature is available either as user studies or usability reviews carried out by HCI experts and/or researchers. We *summarize lessons learned* for e-voting system interface design, user studies, and usability criteria, from which we *extract relevant recommendations* for the same. We argue for extensions in these three areas to take into account verifiable and cryptographically-verifiable e-voting systems.

The content of this book chapter is especially relevant to designers of e-voting systems for whom the information provided will form basic input for designing future e-voting systems, in particular, verifiable e-voting systems, where user interaction for verifiability is required, and therefore understanding is critical. Researchers who are interested in replicating user studies or carrying out their own studies will obtain invaluable information, as will researchers and practitioners who are interested in the usability of e-voting systems, and the criteria used to determine usability. Further we identify future research beneficial to researchers seeking open questions in this field.

The reader is advised that accessibility issues in e-voting systems will not be discussed, and are left for future work. However the Voluntary Voting Systems Guidelines (VVSG), which we review, addresses the evaluation of accessibility aspects of e-voting systems.

In the next section we give background information for the reader to better understand the content presented. Following that, in the methodology section, we describe the approach used to identify literature included in this book chapter, and the approach by which the lessons learned and recommendations were obtained. We then present lessons learned from the literature regarding interface design, user studies, and usability criteria for e-voting systems. In each of the lessons learned sections, we derive recommendations from the literature as a take-away for the reader. In addition we present future research directions and conclude with a brief discussion on the future of usable e-voting systems.

BACKGROUND INFORMATION

In order for the reader to understand the content presented in this book chapter, we briefly introduce and categorize the *E-Voting Systems* which have been studied in the literature. We introduce terms for, and briefly describe the working of, *Verifiability*, and present the *Voluntary Voting Systems Guidelines (VVSG)* which have been recommended for use in determining the usability of e-voting systems. These guidelines have also been used in the literature surveyed. We give an overview of the *ISO 9241-11 standard (ISO, 1998)* which, in the literature surveyed, has primarily been used as criteria to evaluate the usability of e-voting systems. We will review the *Human Computer Interaction (HCI) Usability Evaluation Techniques* applied in the literature surveyed, summarize the approach used to design many of the user studies, and present terms we will use to refer to some of the experiments carried out. Finally we will touch on *Mental Models*, which determine how people perceive voting and voting aspects. This is a recent yet relevant consideration in the design of e-voting systems, especially those systems providing verifiability.

E-Voting Systems

E-voting systems can be categorized based on their use, whether in person *at the polling station* or *remotely*. We only discuss those e-voting systems that are later presented in the book chapter:

- **Mechanical Voting Machines:** Though the lever voting machine is a type of *mechanical voting machine* we include it in our discussion as its usability has been evaluated in the literature surveyed. Lever voting machines are used to tabulate votes at polling stations. They maintain no record of individual votes, only the tally of votes, which is kept by a mechanical register (Jones, 2001). The lever voting machine was commonly used in polling station elections in the USA.
- **E-Voting Machines:** Nowadays, e-voting machines, also referred to as *direct recording electronic* (DRE) voting machines, are more commonly used in polling stations. These devices are designed to electronically record votes cast by voters, and can be classified as:
    - **Classical or early DRE voting machines:** These mostly use touch buttons. An example is the NEDAP voting machine which was used in the Netherlands until 2007 (Loeber, 2008).
    - **Advanced DRE voting machines:** Provide voters with a touchscreen interface. An example is the Diebold AccuVote TS.
    - **DREs with paper audit trail (PAT):** Also referred to as a voter verifiable paper audit trail (VVPAT) (Mercuri, 2001). These provide a printout of the voter's selection. The VVPAT is a paper record generated during vote casting that can be checked at the point of casting a vote, or audited at a later stage, to verify correctness of the digital tally. With the implementation of VVPATs, it is common to have DRE machines fitted with printers for this

purpose[2]. An example is the Avante Vote-Trakker, which prints a paper record of the vote, behind a plastic screen.
  - **DRE with cryptography:** These provide verifiability using cryptographic means. In many of these systems, a voter can leave the polling station with a paper receipt which he can later use to verify that his vote was stored correctly. An example is Bingo Voting (Bohli, Müller-Quade & Röhrich, 2007).
- **Paper-based or scan-based e-voting systems:** These are used in polling stations. Punch card voting systems use paper ballots onto which the voter punches holes at designated points to indicate the candidate he has selected. Optical scan systems use paper ballots that are then scanned with scanning devices, either at the polling station (precinct-count), or at a central tallying location (central-count) (Jones, 2003). In Demirel, Frankland and Volkamer (2011), optical scan systems are further classified into three. First, a scanning device can be employed by voters during the vote casting process, for example the Digital Voting Pen (Arzt-Mergemeier, Beiss & Steffens, 2007). In the second category, voters have to scan their paper ballots with a Precinct Optical Scan (PCOS) system after marking their candidates on the ballot. In the third category, the paper ballots are scanned by the election authority during the tallying process, using an optical scanner or a barcode scanner.
- **Remote E-Voting Systems (REV):** These include SMS voting systems (Storer, Little & Duncan, 2006), and Internet or online voting systems. Some Internet-based e-voting systems have been deployed, for example in Estonia (Madise & Martens, 2006) and in Switzerland (Gerlach & Gasser, 2009). These systems are considered black box systems as they provide no verifiability of the internal processes (non-verifiable). We can distinguish between fully-verifiable Internet voting systems, for example Helios (Adida, 2008), and partially-verifiable Internet-based e-voting systems, for example the Norwegian e-voting system (Stenerud & Bull, 2012). We briefly discuss verifiability in the next subsection.

Verifiability

Verifiability has been introduced in e-voting systems to enable voters to check that the systems do not manipulate their votes. There are three aspects of verifiability that we consider. The first is referred to as *cast-as-intended*. Here the assurance is that the vote has been sent out from the system as the voter intended it, without any unauthorized modification. The second aspect is *stored-as-cast* which provides assurance that the vote is stored in the ballot box as it was sent, and has not been manipulated. The availability of stored-as-cast verifiability is dependent on cast-as-intended verifiability being provided. These two aspects constitute *individual verifiability*, which is specifically available to an individual voter. *Tallied-as-stored*, which refers to *universal verifiability*, can be made available to any interested party to ensure that all votes counted in a given election have not been modified since the time that voters cast them. A system that provides both individual and universal verifiability is *end-to-end* or *fully-verifiable*.

Verifiability is available in e-voting systems in polling stations, for example through voter verifiable paper audit trails (VVPAT) (Mercuri, 2001), and in remote electronic voting systems through cryptographic means, for example, the Benaloh Challenge (Benaloh, 2006). According to de Jong, van Hoof and Gosselt (2008), paper audit trails may be incorporated in the voting process in various ways. Voters may be required to verify the printout (voter verified paper audit trail), or they may be enabled to do so (voter verifiable paper audit trail). The paper printouts may also have different functions in the elections: a) They are considered the final election result, therefore all printouts are tallied; b) They are used to verify the results of voting machines and results are compared from a random sample; and c) They are recounted when disputes arise.

The Benaloh Challenge provides the first step of individual verifiability. The voter makes his candidate selection, which the system encrypts and then commits to. This commitment is displayed to the voter, for example in Helios (Adida, 2008), as a hash value. A voter can now choose whether to cast a vote after making his candidate selections, or verify it (that is, challenge the system). If he chooses to verify it, the system has to provide the information used for encrypting the vote, which can be verified independently to prove that it correctly encrypted the vote. Since the system does not know which vote a voter will verify and which one he will cast, it is forced to behave correctly. The voter can then check that the same commitment is posted on a bulletin board. This provides the second step of verifiability. In the third step of verifying, interested parties can download election data and cryptographic material to then tally the votes and verify correct processing and tallying. We do not go into the detail of these steps in this book chapter, but further reading on Electronic Voting is available for the interested reader in the Additional Reading section.

Voluntary Voting Systems Guidelines (VVSG)

The Voluntary Voting Systems Guidelines (VVSG) have been adopted by the United States Election Assistance Commission (EAC) (2005) for the certification of voting systems. A revision (VVSG 1.1.) was proposed in 2009. The VVSG provide requirements against which to test voting systems for compliance regarding functionality, accessibility and security. Volume 1 of the VVSG focuses on voting system performance. Section three of this same volume focuses specifically on usability and accessibility and section nine focuses on VVPAT usability. The recommendations for consideration regarding usability are summarized below:

- The vendor should carry out usability tests on the voting system using participants that are representative of the general population.
- The voting process should provide functional capabilities, such as alert the voter of overvotes[3] or undervotes[4] on the ballot.
- The voting system should present information in any language required by law.
- The voting process should minimize cognitive difficulties for the voter, for example, by providing clear instructions.
- The voting process should minimize perceptual difficulties for the voter, for example, taking into account color-blind voters.
- The voting system should minimize interaction difficulties for the voter, for example, minimizing accidental activation.
- The voting process should maintain privacy of the voter's ballot.

ISO 9241-11 Standard

In the literature surveyed, the ISO 9241-11 standard (ISO, 1998) from the International Organization for Standardization (ISO) has commonly been used to evaluate the usability of e-voting systems. According to the standard, usability is defined as *'the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use'*.

Effectiveness focuses on accuracy and completeness of user tasks. It is commonly measured using error rate. Other measures prescribed by the standard are percentage of goals achieved, percentage of users successfully completing a task and average accuracy of completed tasks. Efficiency considers resources used to achieve effectiveness in user tasks. It is commonly measured using time taken to perform a task. Other measures include mental or physical effort, materials or financial cost of performing the task.

Satisfaction focuses on user's attitudes towards the system, and is measured subjectively by asking users to report on their opinions of a given system. The standard also gives frequency of discretionary use and frequency of complaints as measures.

One approach in subjectively measuring user satisfaction is to use a standardized instrument, usually in the form of a Likert scale (Laskowski, Autry, Cugini, Killam & Yen, 2004). The instrument most commonly used in the literature reviewed has been the System Usability Scale (SUS), (Brooke, 1996). SUS is a ten-item scale, which is presented to a user immediately after they have interacted with the system under evaluation. A subjective usability (SU) score is then calculated based on the responses given. A high SUS score indicates high voter satisfaction with the e-voting system.

According to Rubin and Chisnell, (2008), the attributes of usability, in addition to efficiency, effectiveness and satisfaction, are usefulness, learnability and accessibility. The authors explain that learnability is a part of effectiveness and relates to a user's ability to operate the system to some defined level of competence after some predetermined amount and period of training, and also considers how a user performs with an interface after long periods of disuse.

Note that these three metrics for usability, namely, effectiveness, efficiency, and satisfaction, have also been recommended by the National Institute of Standards and Technology (NIST) (Laskowski et al., 2004) to assess the usability of voting systems.

Human Computer Interaction (HCI) Usability Evaluation Techniques

We discuss in this section the variety of techniques from the human computer interaction (HCI) field that have been used to evaluate the usability of e-voting systems in the literature surveyed. Readers are directed to Sharp, Rogers and Preece (2007) for more detail on HCI usability evaluation techniques. Further literature on HCI Research Techniques, Usability, and Design is also provided in the Additional Reading section.

The techniques used in the literature surveyed are:

- **Interviews:** There are *structured, unstructured, and semi-structured interviews*, which involve interviewing individuals, as well as *group interviews*, for example, focus groups (Sharp et al. 2007). In interviews, participants answer questions asked by an interviewer and the situation is similar to the interviewer and participant having a discussion about a topic of interest. The interviewer can ask for clarification where responses are unclear, or further explore comments that participants make:
    - **Focus groups:** These fall under group interviews, where a group of three to ten people are interviewed, and the discussion is moderated by a trained facilitator. *Scenarios* or *prototypes* can be used to help guide the discussion in interviews.
        - **Scenarios:** These are narrative descriptions of tasks, for example one could write a narrative describing how a voter would cast his vote using an e-voting system. Participants would then read the scenario and discuss the same with the facilitator and other participants.

        - **Prototypes:** These are mock-ups, ranging from paper-based to software-based, that are designed early in the development process, to give stakeholders an idea of a design, for example, of an interface. Participants would then interact with these prototypes during the focus group.
- **Analytical evaluation:** These are carried out by experts who have a background in HCI. The methods in this category include heuristic evaluation and cognitive walkthroughs:
    - **Heuristic evaluation:** This is an inspection technique where experts use usability principles to evaluate the e-voting system. These heuristics can be obtained from literature, for example, Nielsen's Heuristics (Nielsen, 1993).
    - **Cognitive walkthrough:** Experts mimic a user's interaction with the e-voting system for a specific scenario of use. Their thoughts and observations are captured for further analysis and give insight to potential problem areas in the design of the e-voting system.
- **Personas:** These are descriptions (using fictional characters) of typical users that would use a system. These are not real users, but are designed to represent profiles of real users of systems. Designers then focus on these users and design the e-voting system to fit their needs.
- **Lab Studies:** Participants interact with the e-voting system to cast a vote, and the time taken to do so, and any errors made are recorded. The lab may be set up to represent the environment in a real polling station.
- **Field study:** Here participants interact with the e-voting system in a 'normal' uncontrolled environment, as opposed to the controlled environment of a lab setting.

Techniques for data collection in user studies, as seen in the literature surveyed, are:

- **Pre- and post-voting questionnaires:** The pre-voting questionnaire frequently collects demographic data or data that can be used to include or exclude participants from the study. Post-voting questionnaires collect participants' opinions after they have interacted with the system.
- **Exit poll:** Participants are asked a series of questions at the end of a voting session. In many cases, this is after they have cast a vote in a real election.
- **Direct Observation:** A participant may be observed as they interact with the e-voting system. In this type of observation, the participant is aware of the observation, for example, since the researcher is present in the lab, or recording tools are utilized:
    - **Video recording:** A camera can be set up to record the interaction.
    - **Audio recording:** Audio capture tools can be used to capture the participants' thoughts if they are asked to think aloud.
    - **Collecting eye-tracking data:** This data captures the points on the screen that the participant focuses on, which can then inform placement of interface elements for greatest impact.
- **Indirect Observation:** In this type of observation, data can be collected unobtrusively, without the participant being aware of it, unless they are informed:
    - **Log file analysis:** Researchers use software to collect information about participants' interaction with the e-voting system. This data is stored in a log file that is later examined.

The literature surveyed shows that a common approach in the design of user studies for e-voting has been to split participants into several groups. One group of participants receives a slate instructing them which candidates to vote for. This is the '*directed*' group. A second group of participants receives a voter guide with relevant information on the candidates to help guide their selection. Participants in this '*voter guide*' group are then free to vote for candidates of their choice. A third group is '*undirected*', receiving neither a slate nor a guide, and participants are free to vote as they choose.

In some voting experiments, a group of participants are asked to skip some races in order to simulate *roll-off*[5]. In other cases, participants in the undirected group are given an exit interview for researchers to determine voting intent. This is then compared to votes cast to obtain the error rate. Finally, the order of participants' interaction with the systems is usually randomized (using a Latin Square) in order not to introduce bias.

Mental Models

Mental models describe people's perceptions of objects, determining how people then interact with the objects (Staggers & Norgio, 1993). Mental models are an important consideration in design, as designers are then able to use knowledge of the user's mental model in designing an object, thus lessening frustration, and greatly improving satisfaction. It is only recently (Campbell & Byrne, 2009b) that researchers have considered voters' mental models in e-voting systems.

METHODOLOGY

In this section, we discuss and justify how we selected the papers reviewed, how the lessons learned are extracted, and how the recommendations and future work are identified.

How the literature was identified: The literature reviewed in this book chapter was identified by searching scientific literature repositories, such as www.scholar.google.com, and the digital libraries of the Institute of Electrical and Electronics Engineers (IEEE), and the Association for Computing Machinery (ACM). Search criteria included 'usability and e-voting', 'usability and remote e-voting' and 'ballot design'. The search was not limited to a specific time period. From the results of the electronic repository search, we reviewed conference papers and journal articles. Further papers and reports were obtained from the references of the identified papers. The papers reviewed were published between the years 1998 and 2012.

The conference papers were proceedings of HCI, usability, e-voting, security, democracy, and governance conferences and workshops. We reviewed these to identify papers that focused on user studies and usability of e-voting systems. These conferences are the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI), the International Conference on Human Computer Interaction (HCII), the IADIS International Conference Interfaces and Human Computer Interaction, the Symposium on Usable Privacy and Security (SOUPS), the USENIX Security Symposium, the International Conference on E-Voting and Identity (VOTE-ID), the Electronic Voting Technology Workshop/Workshop on Trustworthy Elections (EVT/WOTE), the International Workshop on Requirements Engineering for Electronic Voting Systems (REVOTE), the International Conference on Electronic Governance (EGOV), the International Conference for E-Democracy and Open Government (CEDEM), the International Conference on Theory and Practice of Electronic Governance (ICE-GOV), the Workshop on Socio-Technical Aspects in Security and Trust (STAST), and the International Conference on Information Integration and Web-based Applications and Services (iiWAS). We included a paper (Everett, Byrne, & Greene, 2006) that was published in the proceedings of the Annual Meeting of the Human Factors and Ergonomics Society (HFES). References from the identified papers were reviewed and further relevant papers identified and selected for inclusion.

Relevant journals were from the fields of usability, user studies, computing, electronic governance, information design, and political science. Here we reviewed articles from the Journal of Usability Studies, International Journal of Human-Computer Studies, IEEE Transactions on Information Forensics and Security, International Journal Universal Access in the Information Society, the European Journal of Information Systems, the Social Science Computer Review, the International Journal of Electronic Governance, Information Design Journal, American Politics Research, the Journal for the American Political Science Association, Perspectives on Politics, and Public Opinion Quarterly, to identify relevant papers. Once again, the references from the selected papers were reviewed to identify other relevant articles.

How the lessons learned were obtained: The literature was reviewed for findings that could guide researchers on (a) designing e-voting interfaces, (b) conducting user studies, and (c) determining usability criteria used in e-voting usability studies.

How the recommendations were made: Recommendations are given as a take-away pointing to what researchers should do, in light of the research findings. For example, where studies found that first time and elderly voters need help while voting, we recommend integrating help features in e-voting system interfaces. Recommendations will be numbered R-ID-(subsection)-(number) in the Interface Design section, R-US-(subsection)-(number) in the user studies section, and R-UC-(subsection)-(number) in the Usability Criteria section. For example, recommendation R-ID-BD-1 is the first recommendation focusing on Ballot Design, contained in the Interface Design section.

How the future work is deduced: Where we identified open questions in the literature surveyed, we recommend these as future work. Future work recommendations are identified as FW-ID-(subsection)-(number), FW-US-(subsection)-(number), and FW-UC-(subsection)-(number) for the three sections. For example, FW-ID-BD-1 is first recommendation for future work on ballot design, in the Interface Design section.

LESSONS LEARNED AND RECOMMENDATIONS FOR INTERFACE DESIGN

In this section, we summarize lessons learned from the literature regarding design of e-voting interfaces. Additionally we provide recommendations as a take-away for e-voting interface designers, and researchers looking for future research and open questions in this field. We first address issues related to the ballot and vote casting, namely *Standardized Ballot/Interface Design*, *Simple and Clear Ballot Instructions*, *and Providing of Review/Confirmation Screens*. *Considering Time, Speed, Voting Tasks and Effort*, *Providing Help Features*, and *Educating Voters and Poll Workers* are the next subsections, and relate to supporting voters and poll workers. *Identifying Mental Models* is the next subsection and here the lessons learned are applicable for providing voters with help during voting and educating voters and poll workers, in addition to systems providing Voter Verifiable Paper Audit Trails (VVPAT) and cryptographic verifiability. Further individual lessons learned and recommendations for interface design of these two types of e-voting systems are discussed in the next two subsections, that is, *Use of the Voter Verifiable Paper Audit Trail (VVPAT),* and *Understanding in Cryptographically-Verifiable Voting*. We concentrate on these e-voting systems separately as they are especially challenging from a usability point of view. Note that in this section we refer to voters, as opposed to participants.

Standardized Ballot/Interface Design

In the literature surveyed, we find attention is given to screen guides for voters. For example, the e-voting machine in Roth (1998) used flashing lights to indicate to voters races that were not yet completed. The importance of appropriate ballot format is seen in Kimball and Kropf (2005) where they indicate that poor ballot format can result in unrecorded votes. In addition, ballot interfaces that are not standardized can be unfamiliar to voters and cause confusion (Niemi & Herrnson, 2003). Participants in the studies by Greene, Byrne and Everett (2006), Everett, Byrne and Greene (2006), and Byrne, Greene and Everett (2007) were most satisfied with the bubble ballot. The bubble ballot design is similar to a multiple-choice exam question where voters are presented with candidate information in a listing. Next to each candidate name, is an oval or square box on which the voter can mark their selection with an X or by shading the box. Considering large candidate listings, MacNamara, Gibson and Oakley (2012) adapted the Prêt à Voter backend to accompany the Dual Vote frontend, and acknowledge that the paper ballot used was not suitable for a large candidate list. Intuitively, one might think that a full-face ballot, that is, one that displays all offices and candidates on a single screen is a solution. However, Norden et al. (2006) found a high rate of residual votes[6] with full-face DREs. Selker, Hockenberry, Goler and Sullivan (2005) describe a Low Error Voting Interface (LEVI) designed after several iterations, with the aim of reducing the number of errors by voters. This was achieved by providing voters with information on which races they had voted in, which races were pending, and if they had made a selection on a screen.

Recommendations

- **R-ID-BD-1:** Ballot design should be standardized such that the ballots are familiar to voters, for example, imitating the paper ballot design on the e-voting system interface.
- **R-ID-BD-2:** The ballot should indicate to voters when their vote has been successfully cast and should clearly indicate if the vote casting process has been completed.
- **R-ID-BD-3:** The interface should inform voters if their vote is invalid based on the selections they have made.
- **R-ID-BD-4:** Researchers should use the bubble ballot as a standard design where the ballots and candidate listing supports it.

Future Work

- **FW-ID-BD-1:** Future research should investigate the design of large ballots on e-voting systems and their usability.

Simple and Clear Ballot Instructions

Ballot instructions guide voters on how they can successfully mark their desired candidate(s) on the ballot presented. These can be placed at the top of the ballot, or the instructions can precede each race that voters have to fill out, if the ballot contains more than one race. Roth (1998) calls for the use of simple and clear instructions in the ballot to avoid voter confusion. The literature surveyed points to the use of the Flesch-Kincaid readability test (Kincaid, Fishburne, Rogers & Chissom, 1975), to test the comprehension level of instructions in the ballot. The location of the instructions on the ballot has been considered as in Western culture, for example, voters usually look at the upper left-hand corner of the ballot (Kimball & Kropf, 2005). Instructions should appear before the task as voters may not read the instructions. Laskwoski and Redish (2006) recommend four best practices for ballot instruction design. First the consequences of an action, for example, `*You will not be able to make further changes'* should be mentioned before the action to be carried out, for example, `*Click to Confirm'*. This is because voters are likely to carry out the action if it appears first, and then read of the consequences later. Second, at the sentence level, voters should be given the context of use before the action, for example: `*To vote for your desired candidate, mark an x in the* *box next to the candidate name*'. In this example, the sentence is in two parts, the first is the context, the second is the action to be carried out. Third, is to use words that are familiar to voters, and the fourth best practice is to place instructions in logical order, following the structure of the ballot the voter will encounter, for example, `*To vote for your desired candidate, mark an x in the box next to the candidate name. To proceed to the next race, click the Next button*'.

Recommendations

- **R-ID-BI-1:** Simple and clear instructions should be designed for ballots.
- **R-ID-BI-2:** The consequences of an action in an instruction should precede the call to act.
- **R-ID-BI-3:** Each instruction sentence on the ballot should give the result of the action before stating the action to be carried out.
- **R-ID-BI-4:** The instructions should use words that are familiar to voters, and additionally should match the order of tasks on the ballot, in a logical sequence.
- **R-ID-BI-5:** Instructions should be placed at the upper left-hand corner of the ballot, and should be given before the task to be carried out, for instance, having instructions how to mark a candidate correctly, just above where this action will be carried out.

Providing Review/ Confirmation Screens

In the literature surveyed there are calls for the implementation of review screens in e-voting systems. A review screen presents the voter with all the candidate selections they have made on the ballot, and supports their re-checking the correctness of the selections, before casting the ballot. Herrnson et al. (2006) recommend that voters should be given an option to review their vote(s) and correct the ballot when errors are noted. Norden, Creelan, Kimball and Quesenbery (2006) found that lower residual vote rates were observed when the voter was given the chance to correct errors before casting his ballot. Voters are also seen to expect this, as MacNamara et al. (2010) report that a system that did not implement a confirmation graphical user interface received a low voter satisfaction rating.

The question arises whether the review screen would help voters notice changes made to their votes. Everett (2007) found only 32% of participants noticed if races were added to or missing from the review screen, and only 37% of participants noticed vote flipping on their review screen. Campbell and Byrne (2009a) did a follow up on these findings, since participants in Everett's study were not explicitly instructed to check the review screen, and the interface may not have simplified the process of checking. They investigated if the rate of anomaly detection could be improved by instructing participants to check the review screen, and re-designing the review screen such that undervotes were highlighted. Participants were later directly asked if they noted any anomalies on the review screen, and 50% reported noticing anomalies. As undervotes were highlighted on the review screen, more participants noticed them (61%) compared to 39% who noticed their votes were changed to another candidate.

Recommendations

- **R-ID-RS-1:** Review screens should be implemented in e-voting systems.
- **R-ID-RS-2:** Voters should be instructed to pay attention to the review screen, in order to ensure that they notice any changes made to their votes.
- **R-ID-RS-3:** Techniques such as additional coloring or highlighting should be used to draw voters' attention to races where they have not yet voted.

Future Work

- **FW-ID-RS-1:** Future research should investigate how to instruct voters to pay more attention to the review screen especially for them to notice when their candidate selections are modified.

Considering Time, Speed, Voting Tasks, and Effort

Note that we consider voting tasks to be the activities that voters would carry out while casting their vote, for example, changing their vote and writing in a candidate's name for a write-in vote. Herrnson et al. (2008) found that more actions meant more time to complete the voting process and lowered voter satisfaction. Kiosk-voting participants in Oostveen and Van den Besselaar (2009) commented that they found it faster to mark a cross (X) next to their candidate's name, fold the ballot or put it in an envelope and then in to a ballot box, compared to inserting a smart card, typing in a PIN code, scrolling through several computer screens before making a selection and confirming their vote.

Herrnson et al. (2006) and Herrnson et al. (2007) found that participants' accuracy declined when they were asked to carry out less straight forward tasks, such as voting for two candidates in one office. Everett et al. (2008) showed that increasing voting speed for voters may decrease accuracy. They also noted a 'fleeing voter' error, where some participants left the DRE before completing a final step. Greene (2008) compared direct access and sequential access interfaces for DREs and found that voters were dissatisfied with the direct access voting interface, which was faster, but also had more undervote errors and premature ballot casting. Conrad et al. (2009) found that the more effort users needed to put into casting their vote, the less satisfied they were. Lower satisfaction also resulted from voting incorrectly and an increased effort to vote.

Recommendations

- **R-ID-TS-1:** Researchers should aim to reduce the amount of time and effort that voters need to take in order to cast their vote, that is, the number of voting tasks should be reduced, in particular regarding poll workers and voters enabling e-voting machines or in checking the voter's right to vote in Internet voting.
- **R-ID-TS-2:** Care should be taken in facilitating voters to vote quickly, as in direct access DREs, since faster voting may lead to more voter errors.

Future Work

- **FW-ID-TS-1:** Further research should investigate how much time voters would be willing to spend in carrying out voting tasks, and how much time they would consider to be too much.

Providing Help Features

Both Herrnson et al. (2006) and Herrnson et al. (2008) found that although voters can cast their votes unassisted, not all of them are able to do so, and help facilities need to be provided. Prosser, Schiessl and Fleischhacker (2007) found in their study that voters hardly took notice of the help information provided, and recommend providing help 'just-in-time' when it is needed.

Recommendations

- **R-ID-HF-1:** It is recommended that designers integrate help facilities to give voters information when they need it and to guide them on what next steps are required. In Internet voting, for example, the help facilities should be placed on every page the voter will access on the e-voting system, as well as next to tasks that are likely to be confusing for voters.

Future Work

- **FW-ID-HF-1:** Future research should investigate appropriate help for voters given the different voting contexts, for example hotlines and email in Internet voting.

Educating Voters and Poll Workers

Participants in the study by Herrnson et al. (2005b) responded more favorably to the voting machines, and the authors credit this to an education campaign carried out to familiarize voters with the machines. Kalchgruber and Weippl (2009) report on a short case study where two groups of students, the first group with an IT security background, and the second group with basic IT security knowledge, were given explanations of how the Scratch and Vote system (Adida & Rivest, 2006) worked. They found that a lot of education is necessary for voters to understand the difference between non-verifiable, partially-verifiable and end-to-end verifiable systems. Three studies in the literature surveyed focused on the poll worker (Chisnell, Becker, Laskowski & Lowry, 2009a; Chisnell, Bachmann, Laskowski & Lowry, 2009b; Goggin, 2008), but none to our knowledge concentrates on poll worker education.

Recommendations

- **R-ID-ED-1:** Voters should be educated before introducing new e-voting technology.
- **R-ID-ED-2:** The techniques used for educating voters, for example, videos or handouts, should take into account the diversity of voters, in terms of age, experience with voting, and education.

Future Work

- **FW-ID-ED-1:** Future research should investigate effective means of voter education, particularly to introduce voters to, and enable them to use, new e-voting systems and new approaches to voting, such as verifiable e-voting systems.
- **FW-ID-ED-2:** As poll workers also require education for new e-voting systems and approaches, and as this has not yet been addressed comprehensively in the literature, future research should investigate poll worker education.

Identifying Mental Models

Campbell and Byrne (2009b) carried out an online web-based survey in order to identify voters' mental model of straight-party voting (SPV)[7]. The findings show that voters are confused by SPV, which can be due to a gap between how voters think SPV should work and how it actually does work. Karayumak, Kauer, Olembo, Volk, and Volkamer (2011b) found that voters were confused why they needed to verify their vote, and were concerned about compromising secrecy of the vote.

Recommendations

- **R-ID-MM-1:** Voters' mental model should be investigated as voter confusion may be due to differences between how voters think an e-voting system should work, and how it has been designed and implemented.
- **R-ID-MM-2:** Voters should be educated about verifying their vote to deal with the problem of confusion and secrecy concerns.

Future Work

- **FW-ID-MM-1:** Future research should investigate specific approaches to extending voters' mental model to take into account new e-voting systems and voting approaches, as the literature surveyed does not give information on how to close the gap identified.

Use of the Voter Verifiable Paper Audit Trail (VVPAT)

In the study by MacNamara et al. (2011a), 89% of participants thought a paper audit trail was important in e-voting systems. The following literature, however, shows that users have challenges with the VVPAT. Participants in the study by Selker et al. (2005) reported no errors after voting, although errors actually existed. Selker, Rosenzweig and Pandolfo (2006) also report on a study where it was observed that voters, when asked to pay close attention to the VVPAT, did not notice any errors, yet errors did exist. In de Jong et al. (2008), the paper audit trail was observed not to affect the voters' experiences with the voting machine. Herrnson et al. (2006) observed that some voters ignored the VVPAT, and those who did not, seemed to get confused. They also found that voters had less confidence that their vote was accurately recorded on the systems with a VVPAT. In another field study, although participants were instructed to pay attention to the verification systems, it was noted that they did not spend as much time as needed to verify every selection (Center for American Politics and Citizenship [CAPC], 2006).

Recommendations

- **R-ID-VV-1:** Voters should be provided with clear instructions that are easy to understand and follow, guiding them on how to check their vote selection when VVPATs are provided.

Future Work

- **FW-ID-VV-1:** Further research is necessary to identify the best approach for instructing voters, as well as to further investigate voter challenges with VVPATS.
- **FW-ID-VV-2:** Voters' mental model regarding VVPATS should be investigated in future research.

Understanding in Cryptographically-Verifiable Voting

Bär, Henrich, Müller-Quade, Röhrich, and Stüber (2008) discuss the use of Bingo voting (Bohli et al., 2007) in a student election at Karlsruhe University. They report that some voters did not trust the random number generator as they were not familiar with its internal operation, and that a majority of voters did not bother to check any random number in order to verify that their vote was properly recorded. Using a focus group approach, Schneider et al. (2011) assessed the usability of Prêt à Voter version 1.0 (Bismark et al., 2009). Although participants were able to cast their votes, audit them, and check for inclusion on the bulletin board, comments showed that they did not understand some of the design decisions, and in verifying their ballot, expected to see the name of the candidate they had voted for. Karayumak et al. (2011b) identified that although voters were able to verify their vote when instructed to do so using modified interfaces of the Helios voting system (Adida, 2008), they lacked an understanding of the need for verifiability. Carback et al. (2010) observed from comments by voters who had voted with Scantegrity II in a municipal election that voters did not understand about verifying their votes online.

Recommendations

- **R-ID-CV-1:** Voters should be provided with clear instructions that are easy to understand and follow to verify whether their vote is properly encrypted (cast as intended) and stored (stored as cast).
- **R-ID-CV-2:** Appropriate help features should be integrated for cryptographically-verifiable e-voting, taking into account the different types of voters, ranging from first-time voters, to frequent voters.
- **R-ID-CV-3:** Voter education is recommended in cryptographically-verifiable

e-voting as voters may not understand the need for verifiability, and instead expect a secure system that does not require one to carry out verifiability steps.

Future Work

- **FW-ID-CV-1:** It is recommended that voters' mental model and voter education be investigated in future research, with the aim of eliminating confusion and the reluctance to carry out the necessary verifiability steps.
- **FW-ID-CV-2:** Further research should investigate voters' view of the amount of time it takes to verify votes in the case of cryptographically-verifiable e-voting systems, whether they find it too long or acceptable.

LESSONS LEARNED AND RECOMMENDATIONS FOR USER STUDIES

In this section, we summarize lessons learned from the literature surveyed regarding conducting user studies to evaluate the usability of e-voting systems. As such, we use the term participants, rather than voters. The discussion here focuses on *Relevant Methodology*, *Considering Ecological Validity*, and *Maintaining Vote Secrecy*. We notice interaction between these three areas as they relate to the approach used in e-voting studies. *Incentives for Participants*, *Number of Participants in E-Voting User Studies*, *Considering Voter Information Processing Capabilities when Designing User Tasks for Studies,* and *Considering Ethical Issues* focus specifically on participants. We wrap up with *Errors due to Technology used in User Studies*. We make recommendations that are relevant for researchers interested in carrying out user studies, and researchers interested in open questions for further research. Contradictions and open research questions will be recommended for further research.

Relevant Methodology

Some researchers in the literature surveyed apply several HCI methodologies on the same system, beginning with an evaluation by experts, then moving on to user studies with actual users. Bederson, Lee, Sherman, Herrnson and Niemi (2003) used an expert review, close-up observation and a field study. Herrnson et al. (2006) used an expert review by computer-human interaction experts, a laboratory experiment, a large field study and natural experiments in two US states. Traugott et al. (2005) describe natural experiments as observing changes in the real world and investigating their consequences. Karayumak, Kauer, Olembo and Volkamer (2011a) started with a cognitive walkthrough on Helios version 3.0 and the improved interfaces were tested in a user study by Karayumak et al. (2011b).

A high number of interface and interaction issues are likely to be identified by experts and these can be dealt with before users are involved. System reviews with experts can take a shorter amount of time, and identify a high number of critical errors, compared to testing with users, a process which may take longer and cost more. Furthermore, an all-rounded view of the systems is obtained using different research methodologies. Pilot studies are critical in identifying issues before users interact with the systems (Sharp et al., 2007).

Recommendations

- **R-US-RM-1:** We recommend the following methodology: A usability study should begin with an evaluation involving experts, after which changes can be made to the e-voting aspect under study based on the feedback received. As a second step, a user study should be carried out. Note that before the user study is carried out, pilot studies are necessary to identify any difficult or unclear issues in the study design. If a lack of understanding of the e-voting system aspect under study is observed, the voters' mental model should be investigated. Based on feedback from participants after the user study, the e-voting should be re-designed. The re-design should be tested in subsequent user studies, and several iterations at this stage may be necessary, switching between re-design and small user studies for user feedback. Finally, field studies should be carried out, where the re-designed e-voting system can be tested in a real election with real voters. Exit polls should accompany the field studies, to obtain voters' feedback on the e-voting system, and related aspects being studied.

Considering Ecological Validity

Ecological validity considers the extent to which the results of an experiment can be applied to real-life conditions (Patrick, 2009). As a result, researchers want their studies to be as realistic as possible. Schneider et al. (2011) designed a ballot similar to that of UK elections in order to give participants a familiar voting experience. The study in Byrne et al. (2007) found that the frequency of errors was not affected by the use of fictional candidates.

Cross II et al. (2007) set up two voting booths in the lobby of a student union building, in order to provide a noisy environment similar to that available during actual elections. In Fuglerud and Røssvoll (2011) participants could decide in which location to conduct the test and were encouraged to use their own PC and equipment. Weber and Hengartner (2009) had each participant used their own laptop and email address.

Field studies, or exit polls, can also be carried out to obtain feedback from voters after they have voted in a real election. Carback et al. (2010), report on an exit poll carried out after voters voted in a real election using Scantegrity II at Takoma Park. Herrnson et al. (2005) carried out an exit poll out after participants voted in an actual election. Van Hoof, Gosselt & de Jong (2007), and de Jong, van Hoof and Gosselt (2008) also recruited participants immediately after they cast a vote in a real, national election. If this is done, however, there may be legal requirements to satisfy, for example, van Hoof et al. (2007) had to have election questions that were unrelated to the real election questions. A second challenge is that participants may cast a vote for the same candidate they voted for in the real election, or try to avoid revealing their true vote, in the case where real ballots are used, by randomly selecting candidates. This then makes error detection difficult as participants cannot recall their vote selection (Selker et al., 2006).

The tasks that participants carry out can be structured as real tasks similar to those in real elections. Herrnson et al. (2008) gave participants tasks to carry out, for example change a vote, cast a write-in vote or omit a vote for one office, but also asked them to make selections on their own in response to information about specific candidates to keep them engaged in the process. How participants interact with the voting processes can also resemble real scenarios. In Chisnell et al. (2009b), participants worked in pairs, to mimic a realistic situation for poll workers.

Rather than setting up a mock election, we find attempts made to give a real election whose results participants can be interested in. De Jong et al. (2008) had participants vote for a charity organization to receive a 1,000 Euro donation. This gave the participants a clear purpose and immediate implications to their voting. Winckler et al. (2009) also asked participants to vote for a charity organization. In MacNamara et al. (2012) participants voted for their favorite country from four available options. Participants in the study by Cross II et al. (2007) voted on the best burger and fries from options given.

Recommendations

- **R-US-EV-1:** Ballots used in user studies should be similar to those used in real elections to minimize confusion and the number of errors from participants.
- **R-US-EV-2:** Researchers should select one of the following approaches to provide ecological validity: a) The ballot used in the study should be similar to a real ballot, for example based either on the candidates listed, the design of the ballot, or the number of races provided. b) The voting environment should be similar to that of a real election, either by the voting machines used, or holding the election in a location where real elections are held, for example in a town hall. c) Giving voters tasks similar to tasks in a real election, for example, picking a ballot paper, marking their ballot in the voting booth, and dropping their ballot in a ballot box. Though this example is for polling station and paper-based elections, the same can be used for Internet voting where voters are made aware that they are in the voting booth (marking candidates) and dropping their ballot in the ballot box (submitting the vote). d) Running an election for which participants are more likely to be interested in the results, for example a charities' election.
- **R-US-EV-3:** This recommendation includes suggestions that are optional for the researcher. Where mock elections are set up for user studies, they should give a realistic feel to the participants, for example, in the design of the ballot, or in the location of the study. Fictitious candidates can be included in ballots for user studies. As argued in the literature surveyed, this increases the life span of the research instruments used, and they do not become obsolete in a short period of time. User studies can also be set up in the participants' natural environment, or in the case of Internet voting, should use the participant's equipment where possible, in order to be realistic.

Maintaining Vote Secrecy

A trade-off between vote secrecy and identifying participants' voting errors in user studies is seen, for example in van Hoof et al. (2007), where participant numbers were marked on the paper ballot to check if participants voted correctly. A camera recorded the voter interaction with the voting machine, and researchers could see the actual voter input. Some participants seemed to vote for the same candidate they voted for in the real election as participants, though instructed which candidates to vote for, instead cast votes for other candidates. In Conrad et al. (2009), the participant first indicated to an experimenter which candidates he intended to vote for. Some researchers seem to be aware of this challenge, for example, MacNamara et al. (2010) report that they did not identify voter intention in their study in order to preserve secrecy of the ballot.

Recommendations

- **R-US-VS-1:** Where possible, researchers should preserve vote secrecy, or inform participants when it will not be preserved. As an example, researchers can direct voters which candidates they should vote for, but should not link the vote to the participants (for example, by randomly giving voter slates in the study). Researchers can then compare expected votes with actual votes to identify error the rate.

Incentives for Participants

In user studies, participants are often given financial or in-kind incentives to motivate them and improve their engagement in the study. However, this might not apply in e-voting studies. Goggin (2008) attempted to mimic voter motivation as in a real election, and offered participants a $5 USD bonus for voting accurately, that is, voting for the same candidates in all ballots in the study. The bonus was found not to increase voter motivation for accuracy.

Recommendations

- **R-US-IP-1:** Researchers should offer financial or in-kind incentives to participants in user studies.

Future Work

- **FW-US-IP-1:** Future research should investigate if incentives will improve participants' engagement in e-voting studies.
- **FW-US-IP-2:** A second question should explore the role of intrinsic motivation in engaging participants in e-voting user studies, for example, by motivating them to participate in studies in order to be part of improving the usability of e-voting systems.

Number of Participants in E-Voting User Studies

The user studies reviewed in the literature surveyed have recruited varying numbers of participants; the smallest we identified was 7, and the largest 1,540. There are guidelines available on the number of users necessary for results of a study to have statistical significance (Cook and Campbell, 1979). With usability studies, the number of participants is determined by the resources available for the study, including time and finances, as well as the study design. Lazar, Feng and Hochheiser (2010) recommend 15 – 20 participants pointing out that smaller studies may miss out on some useful results.

Recommendations

- **R-US-NP-1:** It is recommended that researchers determine the number of participants for their e-voting studies based on the resources available, the study design, and whether statistically significant results are required. For statistically significant results, statistical techniques can be used, based on the desired degree of confidence, to identify the number of participants from the sample population. If statistically significant results are not required, studies should have, as a minimum, 15 – 20 participants, depending on the goals of the study.
- **R-US-NP-2:** Field studies should have a large number of participants, as the results will be representative of a larger population.

Considering Voter Information Processing Capabilities when Designing User Tasks for Studies

In the study by van Hoof et al. (2007), participants were instructed which candidate to vote for, and were expected to memorize this and vote for the same candidate twice, using a voting machine and a paper ballot. The paper reports that participants voted for the wrong candidate as they could not remember the voting task. Research in psychology points to human limits in processing information (Miller, 1956), with proposals being made for grouping information that is presented to aid recall.

Recommendations

- **R-US-VIP-1:** Researchers should not require participants to remember voting tasks, and instead should provide both written and verbal instructions on what tasks participants are expected to carry out.

Considering Ethical Issues

As user studies involve human participants, ethical issues have to be taken into account (Lazar et al., 2010). In the literature surveyed, we see that researchers may be required to obtain approval from a university institutional review board (IRB) (Carback et al., 2010). In some studies, we see participants being informed about the goals and objectives of the study and signing consent forms before taking part in the study (Everett et al. (2006); Greene et al. (2006); Fuglerud & Røssvoll (2011)). Informing participants fully of the objectives of a study however presents a challenge since participants may change their behavior once they are made aware of the goal of the study, or try to act in a manner they think appropriate, for example in a study on verifiable e-voting systems, they may verify their vote in the study, but they may not do so in a real election.

Recommendations

- **R-US-EI-1:** Usability study design as well as tasks for participants should be reviewed to ensure that they do not violate ethical requirements.
- **R-US-EI-2:** Participants should be informed about the goals of the study either before or after the study.
- **R-US-EI-3:** Participants should sign consent forms before participating in e-voting usability studies.
- **R-US-EI-4:** Where a university ethics board or institutional review board is available, these should check that the study design and materials consider ethical issues. Where these bodies are not available, or cannot offer necessary guidance, researchers should take the responsibility of separately reporting how they have met standard ethical requirements (see Burmeister, 2000).

Future Work

- **FW-US-EI-1:** Future research in usability of e-voting systems should investigate how to handle ethical issues, besides requiring full disclosure to participants as this may not match with e-voting study objectives.

Errors Due to Technology Used in User Studies

In MacNamara et al. (2010) the e-voting system being tested was found to have misclassified 38 out of 332 votes (an error rate of 11.4%), majority of which were caused by a faulty hybrid pen. When these misclassifications were removed, the error rate was much lower, standing at 3.9%.

Recommendations

- **R-US-ET-1:** Researchers should use equipment whose development has been completed and tested, in order to avoid errors arising in actual user studies.

LESSONS LEARNED AND RECOMMENDATIONS FOR USABILITY CRITERIA

In this section we discuss lessons learned from the literature regarding M*etrics for Usability Evaluation*, which describes the measures used in the literature surveyed to measure usability. In the next subsection, *Usability and Design Guidelines*, we focus on guidelines identified in the literature surveyed. We close by discussing *The Need for Baseline Data* in evaluating usability. Researchers seeking to identify usability criteria for their research will find useful information. We also indicate open research questions for future research.

Metrics for Usability Evaluation

In the literature surveyed, we observe that a large number of studies used the three metrics for usability from the International Organization for Standardization (ISO) namely, effectiveness, efficiency and satisfaction. (Goggin, 2007; Goggin, 2008; Campbell & Byrne, 2009a; Everett et al., 2006; Greene et al., 2006; Byrne et al., 2007; Everett, 2007). These three metrics have also been recommended by the National Institute of Standards and Technology (NIST) as appropriate to evaluate the usability of e-voting systems (Laskowski et al., 2004). In some studies, the criteria used were not identified as those recommended by ISO, yet the data collected concerned voter errors, ballot completion time, and voter satisfaction (MacNamara et al., 2011b; Conrad et al., 2009; MacNamara et al., 2010).

Everett et al. (2008) used errors and the System Usability Scale (SUS) (Brooke, 1996) while van Hoof et al. (2007) only used the error rate to evaluate usability of the e-voting system. The SUS has also been used by MacNamara et al. (2011a), while Winckler et al. (2009) used SUS in addition to the Unified Theory of Acceptance and Use of Technology (UTAUT) model (Venkatesh, Morris, Davis and Davis, 2003).

Finally, a number of other approaches have been used, for example, Fuglerud and Røssvoll (2011) used accessibility guidelines, for example, the Web Content Accessibility Guidelines (WCAG) 2.0 from the World Wide Web Consortium (W3C) (W3C, 2008), expert evaluation using personas, and subjective ranking of prototypes by participants and Cross II et al. (2007) used Likert-scale type questions to evaluate usability. Bär et al. (2008) compare the usability of Bingo Voting to Prêt à Voter (Bismark et al., 2009), Punchscan (Popoveniuc & Hosp, 2010), and a scheme by Moran and Naor (Moran & Naor, 2006), all verifiable voting schemes. The authors use the effort needed for voting and additional steps for the voter to ensure correctness, as measures to compare the usability of these schemes. CAPC (2006), Oostveen and Van den Besselaar (2009), Karayumak et al., (2011b) and Herrnson et al., (2008) gave voters questionnaires developed by the authors to obtain voters' subjective assessments, and used these to evaluate usability. CAPC (2006) also used heuristics developed by the researchers for the expert review. Chisnell et al. (2009a) and Chisnell et al. (2009b), Herrnson et al. (2006), Laskowski and Redish (2006), Kimball and Kropf (2005) and Roth (1998) used guidelines developed by the researchers themselves.

The challenge with using different approaches to evaluate usability of e-voting systems is that it makes it difficult for researchers to effectively compare the usability of the systems, particularly if these approaches are not replicable by other researchers.

Recommendations

- **R-UC-MUE-1:** It is recommended that a standardized approach to evaluate usability is adopted, for example, using the three ISO measures of effectiveness, efficiency and satisfaction.

Future Work

- **FW-UC-MUE-1:** Where existing usability metrics are insufficient to evaluate usability aspects, for example in verifiable e-voting systems, future research should explore new metrics, for example, number of actions for voters, and learnability of interfaces.

Need for Baseline Data

Baseline data provides a benchmark against which researchers can compare the usability of e-voting systems. Three studies were identified as having been carried out with the specific objective of providing baseline data. Everett et al. (2006) measured the usability of the three traditional paper ballots, namely bubble, arrow and open ballots. Greene et al. (2006) focused on lever machines, arrow and bubble ballots, and Byrne et al. (2007) investigated paper ballots, punch cards and lever machines.

Recommendations

- **R-UC-BD-1:** Usability evaluations of traditional e-voting systems should be carried out with the specific objective of providing baseline data to allow for comparison between traditional and new e-voting systems.

FUTURE RESEARCH DIRECTIONS

There is a need for further research to determine how to design usable e-voting systems, how to evaluate the usability and which usability criteria to apply. This is especially the case for verifiable and cryptographically-verifiable remote electronic voting systems, where users can carry out extra steps to verify their vote before casting it. Mental models have only recently been investigated in e-voting but they are an important aspect to understand so as to improve the usability of e-voting systems. Consequently, more research should be done to identify voters' mental model, particularly applied to cryptographically-verifiable e-voting systems, where voters are seen to lack understanding (see for example, Carback et al., 2010; Schneider et al., 2011; and Karayumak et al., 2011b).

Research has been carried out that compares the usability of DREs to traditional voting methods in order to provide baseline data for comparison with new e-voting systems. This should be extended for different types of elections, ballots, and races.

In many of the e-voting studies surveyed, the criteria that have commonly been used to evaluate usability have been efficiency, effectiveness and satisfaction. There are no studies where learnability of e-voting systems has been tested. We consider learnability to be relevant measures, especially for cryptographically-verifiable e-voting systems, because it can show how easy it is for a voter to re-learn how to use an interface after a period of time where they have not used it (Rubin & Chisnell, 2008). We therefore recommend that this be investigated in future research.

Furthermore, although there is a standard available for criteria with which to assess the usability of e-voting systems (Laskowski et al., 2004), our survey of the literature shows that this is currently not in use across studies of usability in e-voting. Researchers have applied different criteria to determine usability of e-voting systems making it difficult to compare the usability. Some research however has applied the ISO 9241-11 standard (ISO, 1998), which gives metrics for usability. More research is needed to identify criteria that can be applied uniformly for usable e-voting systems. If the existing criteria are insufficient, there is need for further research to expand them to accommodate newer e-voting systems or develop new criteria.

Bederson et al. (2003) refer to research on ballot position having an effect on the candidates that are selected. As such, they point out that ballots designed in user studies can incorporate techniques to randomize candidate order. They also indicate that this can create difficulty for voters who have pre-planned their voting. Given that Prêt à Voter (Bismark et al., 2009) randomizes the candidate order on the ballot, further research should investigate how this affects voters in elections where Prêt à Voter is used.

Most of the research found in the literature surveyed focused on the voter perspective, and in this book chapter, we also focus on the voter perspective. However, Chisnell et al. (2009a, 2009b), Goggin (2008), and Claasen, Magleby, Monson and Patterson (2008) specifically focus on poll workers. Further research should investigate usability issues for this group of stakeholders.

CONCLUSION

In this book chapter we have focused on the usability of e-voting systems, and have reviewed and summarized lessons learned regarding interface design principles, user studies, and usability criteria. We have made recommendations ideally for three groups of researchers: those interested in designing e-voting system interfaces, those interested in carrying out user studies in e-voting, and those seeking for further research in this field. We have, in addition, indicated open research questions in the field of usability and e-voting that researchers can carry out to extend knowledge in the field.

This work shows that a lot of research has been done on DREs and traditional voting methods, but that in comparison, research on the usability of cryptographically-verifiable e-voting systems, and Internet-voting systems is wanting. However, both are more challenging; cryptographically-verifiable e-voting systems use terms that many voters may not be familiar with, and in Internet voting there is no poll worker available who the voter can ask for help or instructions. Research efforts need to be geared in this direction.

With the adoption of verifiable and partially-verifiable e-voting systems, for example in Norway (Stenerud & Bull, 2012), usability is an important consideration to avoid disenfranchising voters. So far, we see research carried out to reduce the technical complexity of verifiability, for example, in Eperio (Essex, Clark, Hengartner & Adams, 2010). This is one step in improving the usability of e-voting systems, making it possible for all who choose to cast their vote by electronic means to do so without undue difficulty. More and more e-voting systems, however, need to focus on and improve usability.

REFERENCES

Adida, B. (2008). Helios: Web-based open-audit voting. In *Proceedings of the 17th Conference on Security Symposium*, San Jose, CA (pp. 335–348). Berkeley, CA: USENIX Association.

Adida, B., & Rivest, R. L. (2006). Scratch & vote: Self-contained paper-based cryptographic voting. In *Proceedings of the 5th ACM Workshop on Privacy in the Electronic Society, WPES '06* (pp. 29–40). New York, NY: ACM.

Alvarez, R. M., & Hall, T. E. (2008). *Electronic elections: The perils and promises of digital democracy*. Princeton, NJ: Princeton University Press.

Arzt-Mergemeier, J., Beiss, W., & Steffens, T. (2007). The digital voting pen at the Hamburg elections 2008: Electronic voting closest to conventional voting. In Alkassar, A., & Volkamer, M. (Eds.), *VOTE-ID 2007* (*Vol. 4896*, pp. 88–98). Lecture Notes in Computer Science Berlin, Germany: Springer-Verlag.

Bär, M., Henrich, C., Müller-Quade, J., Röhrich, S., & Stüber, C. (2008). *Real world experiences with bingo voting and a comparison of usability*. Paper presented at the Workshop on Trustworthy Elections (WOTE 2008), Leuven, Belgium.

Bederson, B. B., Lee, B., Sherman, R. M., Herrnson, P. S., & Niemi, R. G. (2003). Electronic voting system usability issues. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI 2003* (pp. 145 - 152). New York, NY: ACM.

Benaloh, J. (2006). Simple verifiable elections. In *Proceedings of the USENIX/Accurate Electronic Voting Technology Workshop 2006 EVT'06*. Berkeley, CA: USENIX Association.

Bismark, D., Heather, J., Peel, R. M. A., Ryan, P. Y. A., Schneider, S., & Xia, Z. (2009). Experiences gained from the first prêt à voter implementation. *First International Workshop on Requirements Engineering for E-voting Systems RE-VOTE '09* (pp. 19-28) Washington, DC: IEEE Computer Society.

Bohli, J., Müller-Quade, J., & Röhrich, S. (2007). Bingo voting: Secure and coercion-free voting using a trusted random number generator. In A. Alkassar & M. Volkamer (Eds.), *Proceedings of the 1st International Conference on E-voting and Identity* (pp. 111-124). Berlin, Germany: Springer-Verlag.

Brooke, J. (1996). SUS: A 'quick and dirty' usability scale. In Jordan, P. W., Thomas, B., Weerdmeester, B. A., & McClelland, A. L. (Eds.), *Usability evaluation in industry*. London, UK: Taylor and Francis.

Burmeister, O. K. (2000). Usability testing: Revisiting informed consent procedures for testing internet sites. In J. Weckert (Ed.) *Selected Papers from the Second Australian Institute Conference on Computer Ethics CRPIT '00* (pp. 3-9). Darlinghurst, Australia: Australian Computer Society, Inc.

Byrne, M. D., Greene, K. K., & Everett, S. P. (2007). 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 CHI 2007* (pp. 171–180). New York, NY: ACM.

Campbell, B. A., & Byrne, M. D. (2009a). Now do voters notice review screen anomalies? A look at voting system usability. In *Proceedings of the 2009 Conference on Electronic Voting Technology/Workshop on Trustworthy Elections EVT/WOTE'09*. Berkeley, CA: USENIX Association.

Campbell, B. A., & Byrne, M. D. (2009b). Straight-party voting: What do voters think? *IEEE Transactions on Information Forensics and Security*, *4*(4), 718–728. doi:10.1109/TIFS.2009.2031947

Carback, R., Chaum, D., Clark, J., Conway, J., Essex, A., & Herrnson, P. S. … Vora, P.L. (2010). Scantegrity II municipal election at Takoma park: The first E2E binding governmental election with ballot privacy. In *Proceedings of the 19th USENIX Conference on Security USENIX Security'10.* Berkeley, CA: USENIX Association.

Center for American Politics and Citizenship (CAPC). (2006). *A study of vote verification technology conducted for the Maryland state board of elections. Part II: usability study.* Retrieved August 3, 2012, from http://www.cs.umd.edu/~bederson/voting/verification-study-jan-2006.pdf

Chin, J. P., Diehl, V. A., & Norman, K. L. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI '88* (pp. 213-218). New York, NY: ACM.

Chisnell, D., Bachmann, K., Laskwoski, S., & Lowry, S. (2009b). Usability for poll workers: A voting system usability test protocol. In Jacko, J. A. (Ed.), *Human Computer Interaction, Part IV, HCII 2009* (*Vol. 5613*, pp. 458–467). Lecture Notes in Computer Science Berlin, Germany: Springer-Verlag. doi:10.1007/978-3-642-02583-9_50

Chisnell, D., Becker, S., Laskowski, S., & Lowry, S. (2009a). Style guide for voting system documentation: Why user-centered documentation matters to voting security. In *Proceedings of the 2009 Conference on Electronic Voting Technology/Workshop on Trustworthy Elections (EVT/WOTE 2009).* Berkeley, CA: USENIX Association.

Claassen, R. L., Magleby, D. B., Monson, J. Q., & Patterson, K. D. (2008). At your service: Voter evaluations of poll worker performance. *American Politics Research*, *36*, 612. doi:10.1177/1532673X08319006

Cohen, J. D., & Fischer, M. J. (1985). A robust and verifiable cryptographically secure election scheme. In *Proceedings of the 26th Annual Symposium on Foundations of Computer Science SFCS '85* (pp. 372-382). Washington, DC: IEEE Computer Society.

Conrad, F. G., Bederson, B. B., Lewis, B., Peytcheva, E., Traugott, M. W., & Hanmer, M. J. (2009). Electronic voting eliminates hanging chads but introduces new usability challenges. *International Journal of Human-Computer Studies*, *67*(1), 111–124. doi:10.1016/j.ijhcs.2008.09.010

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston, MA: Houghton Mifflin Company.

Cross, E. V. II, McMillian, Y., Gupta, P., Williams, P., Nobles, K., & Gilbert, J. E. (2007). Prime III: A user centered voting system. In *CHI '07 Extended Abstracts on Human Factors in Computing Systems CHI EA '07*. New York, NY: ACM. doi:10.1145/1240866.1241006

De Jong, M., van Hoof, J., & Gosselt, J. (2008). Voters' perceptions of voting technology: Paper ballots versus voting machine with and without paper audit trail. *Social Science Computer Review*, *26*(4), 399–410. doi:10.1177/0894439307312482

Demirel, D., Frankland, R., & Volkamer, M. (2011). *Readiness of various eVoting systems for complex elections*. Technical Report, Technische Universität Darmstadt.

Election Assistance Commission (EAC). (2005). *Voluntary voting systems guidelines volumes 1 and 2*. Retrieved April 16, 2011 from, http://www.eac.gov/testing_and_certification/voluntary_voting_system_guidelines.aspx#VVSG%20Version%201.1

Essex, A., Clark, J., Hengartner, U., & Adams, C. (2010). Eperio: Mitigating technical complexity in cryptographic election verification. In *Proceedings of the 2010 Conference on Electronic Voting Technology/Workshop on Trustworthy Elections EVT/WOTE 2010*. Retrieved August 3, 2012, from http://static.usenix.org/events/evtwote10/tech/full_papers/Essex.pdf

Everett, S. P. (2007). *The usability of electronic voting machines and how votes can be changed without detection*. Doctoral dissertation, Rice University, Houston, Texas.

Everett, S. P., Byrne, M. D., & Greene, K. K. (2006). Measuring the usability of paper ballots: Efficiency, effectiveness and satisfaction. In *Proceedings of the Human Factors and Ergonomic Society 50th Annual Meeting* (pp. 2547 - 2551).

Everett, S. P., Greene, K. K., Byrne, M. D., Wallach, D. S., Derr, K., Sandler, D., & Torous, T. (2008). Electronic voting machines versus traditional methods: Improved preference, similar performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI '08* (pp. 883 - 892). New York, NY: ACM.

Gerlach, J., & Gasser, U. (2009). *Three case studies from Switzerland: E-voting*. Technical Report. Berkman Center Research Publications, Berkman Center for Internet & Society.

German Federal Court Decision. (2009). *Bundesverfassungsgricht. Urteil des Zweiten Senats. German Federal Court Decisions 2 BvC 3/07*, (pp. 1-163).

Goggin, S. N. (2008). Usability of election technologies: Effects of political motivation and instruction use. *The Rice Cultivator*, *1*, 30–45.

Goggin, S. N., & Byrne, M. D. (2007). An examination of the auditability of voter verified paper audit trail (VVPAT) ballots. In *Proceedings of the USENIX Workshop on Accurate Electronic Voting Technology EVT '07*. Berkeley, CA: USENIX Association.

Greene, K. K. (2008). *Usability of new electronic voting systems and traditional methods: Comparisons between sequential and direct access electronic voting interfaces, paper ballots, punch cards, and lever machines*. Unpublished Master's Thesis, Rice University, Houston, Texas.

Greene, K. K., Byrne, M. D., & Everett, S. P. (2006). A comparison of usability between voting methods. In *Proceedings of the 2006 USENIX/ACCURATE Electronic Voting Technology Workshop*. Vancouver, BC: USENIX Association.

Herrnson, P. S., Bederson, B. B., Lee, B., Francia, P. L., Sherman, R. M., & Conrad, F. G. (2005). Early appraisals of electronic voting. *Social Science Computer Review*, *23*(3), 274–292. doi:10.1177/0894439305275850

Herrnson, P. S., Bederson, B. B., Niemi, R. G., Conrad, F. G., Hanmer, M. J., & Traugott, M. (2007). *The not so simple act of voting: An examination of voter errors with electronic voting*. Retrieved August 3, 2012, from http://www.bsos.umd.edu/gvpt/apworkshop/herrnson2007.pdf

Herrnson, P. S., Niemi, R. G., Hanmer, M. J., Bederson, B. B., Conrad, F. G., & Traugott, M. (2006). The importance of usability testing of voting systems. In *Proceedings of the USENIX/Accurate Electronic Voting Technology Workshop 2006 EVT '06*. Berkeley, CA: USENIX Association.

Herrnson, P. S., Niemi, R. G., Hanmer, M. J., Francia, P. L., Bederson, B. B., Conrad, F., & Traugott, M. (2005b). *The promise and pitfalls of electronic voting: results from a usability field test.* Paper presented at the Annual meeting of the Midwest Political Science Association. Chicago, IL.

Herrnson, P. S., Niemi, R. G., Hanmer, M. J., Francia, P. L., Bederson, B. B., Conrad, F. G., & Traugott, M. W. (2008). Voters' evaluations of electronic voting systems: Results from a usability field study. *American Politics Research*, *36*(4), 580–611. doi:10.1177/1532673X08316667

International Organization for Standardization. ISO 9241 – 11. (1998). *Ergonomic requirements for office work with visual display terminals (VDT) – Part 11: Guidelines on usability*. Geneva, Switzerland: ISO.

Jones, D. W. (2001). *A brief illustrated history of voting*. University of Iowa, Department of Computer Science. Retrieved August 1, 2012, from http://homepage.cs.uiowa.edu/~jones/voting/pictures/

Jones, D. W. (2003). The evaluation of voting technology. In Gritzalis, D. A. (Ed.), *Secure electronic voting* (pp. 3–16). Kluwer Academic Publishers. doi:10.1007/978-1-4615-0239-5_1

Karayumak, F., Kauer, M., Olembo, M. M., Volk, T., & Volkamer, M. (2011b). User study of the improved Helios voting system interface. In *1st Workshop on Socio-Technical Aspects in Security and Trust* (STAST) (pp. 37 – 44). doi: 10.1109/STAST.2011.6059254

Karayumak, F., Kauer, M., Olembo, M. M., & Volkamer, M. (2011a). Usability analysis of Helios - An open source verifiable remote electronic voting system. In *Proceedings of the 2011 USENIX Electronic Voting Technology Workshop/Workshop on Trustworthy Elections (EVT/WOTE 2011).* Berkeley, CA: USENIX Association.

Kimball, D. C., & Kropf, M. (2005). Ballot design and unrecorded votes on paper-based ballots. *Public Opinion Quarterly*, *69*(4), 508–529. doi:10.1093/poq/nfi054

Kincaid, J. P., Fishburne, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (Automated readability index, fog count, and Flesch reading ease formula) for navy enlisted personnel*. Research Branch Report 8-75. Chief of Naval Technical Training: Naval Air Station Memphis.

Kirakowski, J., & Corbett, M. (1993). SUMI: The software usability measurement inventory. *British Journal of Educational Technology*, *24*, 210–212. doi:10.1111/j.1467-8535.1993.tb00076.x

Laskowski, S. J., Autry, M., Cugini, J., Killam, W., & Yen, J. (2004). *Improving the usability and accessibility of voting systems and products*. (NIST Special Publication 500 – 256).

Laskowski, S. J., & Redish, J. (2006). Making ballot language understandable to voters. In *Proceedings of the USENIX/Accurate Electronic Voting Technology Workshop EVT'06*. Berkeley, CA: USENIX Association.

Lazar, J., Feng, J. H., & Hochheiser, H. (2010). *Research methods in human computer interaction*. Wiley Publishing.

Loeber, L. (2008). E-voting in the Netherlands: From general acceptance to general doubt in two years. In R. Krimmer & R. Grimm (Eds.), *Proceedings of the 3rd International Conference on Electronic Voting*, Bregenz, Austria (pp. 21 – 30).

MacNamara, D., Carmody, F., Scully, T., Oakley, K., Quane, E., & Gibson, J. P. (2010). Dual vote: A novel user interface for e-voting systems. *IADIS International Conference Interfaces and Human Computer Interaction 2010, IHCI10,* Freiburg, Germany, (pp. 129-138).

MacNamara, D., Gibson, J. P., & Oakley, K. (2012). *A preliminary study on a DualVote and Pret a Voter hybrid system*. In 2012 Conference for E-Democracy and Open Government (CeDEM12). Danube University, Krems.

MacNamara, D., Scully, T., Carmody, F., Oakley, K., Quane, E., & Gibson, J. P. (2011b). (in press). DualVote: A non-intrusive e-voting interface. [IJCISIM]. *International Journal of Computer Information Systems and Industrial Management Applications*.

MacNamara, D., Scully, T., Gibson, J. P., Carmody, F., Oakley, K., & Quane, E. (2011a). *DualVote: Addressing usability and verifiability issues in electronic voting systems*. In 2011 Conference for E-Democracy and Open Government (CeDEM11). Danube University, Krems.

Madise, U., & Martens, T. (2006). E-voting in Estonia 2005. The first practice of country-wide binding internet voting in the world. In A. Prosser & R. Krimmer (Eds.), *2nd International Workshop on Electronic Voting 2006, Lecture Notes in Informatics vol. 47*, (pp. 83-90).

Mercuri, R. T. (2001). *Electronic vote tabulation checks and balances.* Doctoral Dissertation. University of Pennsylvania, Philadelphia.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97. doi:10.1037/h0043158

Moran, T., & Naor, M. (2006). Receipt-free universally-verifiable voting with everlasting privacy. In C. Dwork (Ed.), *Proceedings of the 26th Annual International Conference on Advances in Cryptology CRYPTO '06*, (pp. 373-392). Berlin, Germany: Springer-Verlag.

Nielsen, J. (1993). *Usability engineering*. San Francisco, CA: Morgan Kaufmann Publishers.

Niemi, R. G., & Herrnson, P. S. (2003). Beyond the butterfly: The complexity of U.S. ballots. *Perspectives on Politics*, *1*(2), 317–326. doi:10.1017/S1537592703000239

Norden, L., Creelan, J. M., Kimball, D., & Quesenbery, W. (2006). *The machinery of democracy: Usability of voting systems*. Brennan Center for Justice, New York University School of Law.

Norden, L., Kimball, D., Quesenbery, W., & Chen, M. (2008). *Better ballots*. Brennan Center for Justice, New York University School of Law.

Oostveen, A., & Van den Besselaar, P. (2009). Users' experiences with e-voting: A comparative case study. *International Journal of Electronic Governance*, *2*(4). doi:10.1504/IJEG.2009.030527

Patrick, A. (2009). *Ecological validity in studies of security and human behaviour*. Key note talk: ISSNet Workshop. Retrieved April 16, 2012, from http://www.andrewpatrick.ca/cv/Andrew-Patrick-ecological-validity.pdf

Popoveniuc, S., & Hosp, B. (2010). An introduction to punchscan. In Chaum, D., Jakobsson, M., Rivest, R. L., Ryan, P. Y. A., & Benaloh, J. (Eds.), *Towards trustworthy elections* (pp. 242–259). Berlin, Germany: Springer-Verlag. doi:10.1007/978-3-642-12980-3_15

Prosser, A., Schiessl, K., & Fleischhacker, M. (2007). E-voting: Usability and accpetance of two-stage voting procedures. In Wimmer, M. A., Scholl, H. J., & Grönlund, A. (Eds.), *EGOV 2007* (*Vol. 4656*, pp. 378–387). Lecture Notes in Computer Science Berlin, Germany: Springer-Verlag.

Roth, S. K. (1998). Disenfranchised by design: Voting systems and the election process. *Information Design Journal*, *9*(1), 29–38. doi:10.1075/idj.9.1.08kin

Rubin, J., & Chisnell, D. (2008). *Handbook of usability testing: How to plan, design, and conduct effective tests* (2nd ed.). Indianapolis, IN: Wiley Publishing.

Selker, T., Hockenberry, M., Goler, J., & Sullivan, S. (2005). *Orienting graphical user interfaces reduce errors: The low error voting interface*. VTP Working Paper #23. Retrieved on August 3, 2012 from www.vote.caltech.edu.

Selker, T., Rosenzweig, E., & Pandolfo, A. (2006). A methodology for testing voting systems. *Journal of Usability Studies*, *2*(1), 7–21.

Sharp, H., Rogers, Y., & Preece, J. (2007). *Interaction design: Beyond human-computer interaction*. John Wiley and Sons.

Slaughter, L., Norman, K. L., & Shneiderman, B. (1995). Assessing users' subjective satisfaction with the information system for youth services (ISYS). In *Proceedings of the Third Annual Mid-Atlantic Human Factors Conference* (pp. 164-170).

Staggers, N., & Norgio, A. F. (1993). Mental models: Concepts for human computer interaction research. *International Journal of Man-Machine Studies*, *38*(4), 587–605. doi:10.1006/imms.1993.1028

Stenerud, I. S. G., & Bull, C. (2012). When reality comes knocking: Norwegian experiences with verifiable electronic voting. In M. J. Kripp, M. Volkamer, & R. Grimm. (Eds.) *5th International Conference on Electronic Voting EVOTE 2012* (pp. 21-33). Bregenz, Austria.

Storer, T., Little, L., & Duncan, I. (2006). An exploratory study of voter attitudes towards a pollsterless remote voting system. In D. Chaum, R. Rivest & P. Y. A. Ryan (Eds.), *IaVoSS Workshop on Trustworthy Elections (WOTE 06) Pre-Proceedings* (pp. 77–86).

Traugott, M. W., Hanmer, M. J., Park, W., Herrnson, P. S., Niemi, R. G., Bederson, B. B., & Conrad, F. G. (2005). *The impact of voting systems on residual votes, incomplete ballots, and other measures of voting behavior.* Paper presented at the Annual Meeting of the Midwest Political Science Association, Chicago, IL.

Van Hoof, J. J., Gosselt, J. F., & De Jong, M. D. T. (2007). T*he reliability and usability of the Nedap voting machine*. University of Twente, Faculty of Behavioural Sciences. Retrieved August 3, 2012, from http://wijvertrouwenstemcomputersniet.nl/images/c/ca/UT_rapportje_over_nedap.pdf

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *Management Information Systems Quarterly*, *27*(3), 425–478.

Weber, J., & Hengartner, U. (2009). *Usability study of the open audit voting system Helios*. Retrieved August 3, 2012, from http://www.jannaweber.com/wp-content/uploads/2009/09/858Helios.pdf

Winckler, M., Bernhaupt, R., Palanque, P., Lundin, D., Leach, K., & Ryan, P. … Strigini, L. (2009). Assessing the usability of open verifiable e-voting systems: A trial with the system Prêt à Voter. In *Proceedings of ICE-GOV,* (pp. 281 - 296).

World Wide Web Consortium. (W3C). (2008). *Web content accessibility guidelines (WCAG).* Retrieved August 10, 2012, from http://www.w3.org/TR/WCAG20/

ADDITIONAL READING

Adida, B., De Marneffe, O., Pereira, O., & Quisquater, J. (2009). Electing a university president using open-audit voting: Analysis of real-world use of Helios. In *Proceedings of the 2009 Conference on Electronic Voting Technology/Workshop on Trustworthy Elections EVT/WOTE'09*. Berkeley, CA: USENIX Association.

Benaloh, J. (2006). Simple verifiable elections. In *Proceedings of the USENIX/Accurate Electronic Voting Technology Workshop 2006* EVT'06. Berkeley, CA: USENIX Association.

Brewer, M. B. (2000). Research design and issues of validity. In Reis, H. T., & Judd, C. M. (Eds.), *Handbook of research methods in social and personality psychology* (pp. 3–16). Cambridge University Press.

Chaum, D. (2004). Secret-ballot receipts: True voter-verifiable elections. *Security & Privacy*, *2*(1), 38–47. doi:10.1109/MSECP.2004.1264852

Chaum, D., Carback, R., Clark, J., Essex, A., Popoveniuc, S., & Rivest, R. L. … Sherman, A. T. (2008). Scantegrity II: End-to-end verifiability for optical scan election systems using invisible ink confirmation codes. In *Proceedings of the USENIX/Accurate Electronic Voting Technology Workshop on Electronic Voting Technology EVT '08.* Berkeley, CA: USENIX Association

Clarkson, M. R., Chong, S., & Myers, A. C. (2008). Civitas: Toward a secure voting system. *IEEE Symposium on Security and Privacy*, (pp. 354-368).

Dillman, D. A. (2000). *Mail and internet surveys: The tailored design method*. New York, NY: John Wiley and Sons.

Guzdial, M., Santos, P., Badre, A., Hudson, S., & Gray, M. (1994). *Analyzing and visualizing log files: A computational science of usability.* Presented at HCI Consortium Workshop. Retrieved August 3, 2012, from http://smartech.gatech.edu/jspui/bitstream/1853/3586/1/94-08.pdf

Heather, J., Ryan, P. Y. A., & Teague, V. (2010). Pretty good democracy for more expressive voting schemes. In D. Gritzalis, B. Preneel, & M. Theoharidou. (Eds.), *Proceedings of the 15th European Conference on Research in Computer Security ESORICS '10* (pp. 405-423). Berlin, Germany: Springer-Verlag.

Juels, A., Catalano, D., & Jakobsson, M. (2005). Coercion-resistant electronic elections. In *Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society WPES '05* (pp. 61-70) New York, NY: ACM.

Kimball, D., & Kropf, M. (2007). *Do's and don'ts of ballot design*. AEI Brookings: Election Reform Project. Retrieved August 3, 2012, from http://www.electionreformproject.org/Resources/314eaad7-6dc9-4dad-b6e4-5daf8f51b79e/r1/Detail.aspx

Kish, L. (1965). *Survey sampling*. New York, NY: Wiley.

Krug, S. (2005). *Don't make me think: A common sense approach to the web* (2nd ed.). Thousand Oaks, CA: New Riders Publishing.

Lohr, S. L. (2010). *Sampling: Design and analysis* (2nd ed.). Pacific Grove, CA: Duxbury Press.

National Center for the Study of Elections of the Maryland Institute for Policy Analysis & Research. (2006). *A study of vote verification technologies, part 1: Technical study*. Retrieved August 3, 2012, from http://www.elections.state.md.us/pdf/Vote_Verification_Study_Report.pdf

Neff, C. A. (2001). A verifiable secret shuffle and its application to e-voting. In *Proceedings of the 8th ACM Conference on Computer and Communications Security CCS '01* (pp. 116-125). New York, NY: ACM.

Norman, D. A. (2002). *The design of everyday things*. New York, NY: Basic Books.

Pieters, W. (2006). *What proof do we prefer? Variants of verifiability in voting*. In Workshop on Electronic Voting and e-Government in the UK, Edinburgh, UK.

Puiggali, J., & Morales-Rocha, V. (2007). Remote voting schemes: A comparative analysis. In A. Alkassar & M. Volkamer (Eds.) *Proceedings of the 1st International Conference on E-voting and Identity VOTE-ID'07* (pp. 16-28). Berlin, Germany: Springer-Verlag.

Quesenbery, W. (2004). *Defining a summative usability test for voting systems*. A report from the UPA 2004 workshop on voting and usability.

Quesenbery, W. (2004). *Oops! They forgot the usability: Elections as a case study*. Second Annual Usability and Accessibility Conference.

Reinhard, K., & Jung, W. (2007). Compliance of POLYAS with the BSI protection profile - Basic requirements for remote electronic voting systems. In A. Alkassar & M. Volkamer (Eds.), *Proceedings of the 1st International Conference on E-voting and Identity VOTE-ID'07* (pp. 62-75). Berlin, Germany: Springer-Verlag.

Ryan, P. Y. A. (2004). *A variant of the chaum voter-verifiable scheme*. Technical Report CS-TR 864, University of Newcastle.

Sampigethaya, K., & Poovendran, R. (2006). A framework and taxonomy for comparison of electronic voting schemes. *Computers & Security*, *25*(2), 137–153. doi:10.1016/j.cose.2005.11.003

Sampigethaya, K., & Poovendran, R. (2006). A survey on mix networks and their secure applications. In *Proceedings of the IEEE, 94*(12), 2142-2181. doi = 10.1109/JPROC.2006.889687

Sherman, A. T., Gangopadhyay, A., Holden, S. H., Karabatis, G., Koru, A. G., & Law, C. M. … Zhang, D. (2006). An examination of vote verification technologies: Findings and experiences from the Maryland study. In *Proceedings of the USENIX/Accurate Electronic Voting Technology Workshop 2006 EVT'06*. Berkeley, CA: USENIX Association.

Smith, W. D. (2005). *Cryptography meets voting*. D.O.I = 10.1.1.58.9599

U.S. Election Assistance Commission (EAC). (2011). *A survey of internet voting: Testing and certification technical paper # 2*. Retrieved August 10, 2012, from http://www.eac.gov/assets/1/Documents/SIV-FINAL.pdf

Wallschlaeger, C., & Busic-Synder, C. (1992). *Basic visual concepts and principles for artists, architects and designers*. Dubuque, IA: William C. Brown Publishers.

Wolchok, S., Wustrow, E., Halderman, A. J., Prasad, H. K., Kankipati, A., & Sakhamuri, S. K. … Gonggrijp, R. (2010). Security analysis of India's electronic voting machines. In *Proceedings of the 17th ACM Conference on Computer and Communications Security* (pp. 1–14). New York, NY: ACM.

KEY TERMS AND DEFINITIONS

DRE Voting Machines: A voting machine with mechanical or electrical components for recording votes. DRE voting machines are used at the polling station, where the voter can cast an electronic vote, but in a supervised location, given the presence of election officials.

E-Voting: A means of voting where either the vote casting processes or the tallying of the results are carried out using electronic means. As an example, this can involve using electronic voting machines, or Internet voting systems for vote casting, and optical scanning techniques for vote capture and tallying.

Mental Models: This is considered a person's cognitive internal representation of how something works in the real world. Users create these conceptual models to explain how they understand an object to work, and their interaction with it. These models are not necessarily an accurate reflection of how the object actually works.

Remote Electronic Voting: This also refers to Internet voting. In this case, the voter is able to cast his vote in an unsupervised location, whether from his home or office. Given the unsupervised nature of voting, there are concerns of vote buying or selling.

Usability Criteria: Measures used to assess whether an e-voting system meets usability goals. These criteria include efficiency (speed of performance), effectiveness (rate of errors) and subjective satisfaction. Other criteria are time to learn and retention over time.

Usability Evaluation: Studying the extent to which an e-voting system is fit for use and meets the goals of the voters. This is a process by which designs are assessed and systems tested to ensure they behave as expected, and meet voter expectations. Usability evaluation can be carried out by expert analysis or by user participation.

User Studies: Unlike what the name might suggest, these are tests carried out to determine how users perform with an e-voting system. The goal of user studies is to test if the e-voting system can be used by voters to achieve their desired goals.

Verifiable E-Voting Systems: These are e-voting systems (whether polling station-based or remote) that allow the voter to check that his vote is received and recorded correctly by the voting system, and that it is included in the final tally.

ENDNOTES

1. Typically the voting age is set at 18 years, however some countries set it at 16 years, and others at 21 years. There is no maximum voting age requirement.
2. Note that Voter Verifiable Audio Audit Transcript Trails (VVAATT) have been proposed. We do not consider them here as they aim to provide verification to visually-impaired voters.
3. The voter has not made more than the allowable number of selections for any race.
4. The voter has not made less than the allowable number of selections for any race.
5. Roll-off is the failure to cast votes for some offices on a ballot (Bederson et al., 2003).
6. Residual vote rate is the difference between the number of ballots cast and the number of valid votes cast in a particular contest.
7. Straight party voting (SPV) allows a voter, by a single choice, to select all the candidates of a party on a given ballot