RICE UNIVERSITY

The Usability of Electronic Voting Machines and How Votes Can Be Changed Without Detection

by

Sarah P. Everett

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

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

_____
Michael Byrne, Associate Professor
Psychology

_____
Phil Kortum, Professor-in-the-Practice
Psychology

_____
Margaret Beier, Assistant Professor
Psychology

_____
Dan Wallach, Associate Professor
Computer Science

HOUSTON, TEXAS

MAY 2007

ABSTRACT

The Usability of Electronic Voting Machines and

How Votes Can Be Changed Without Detection

by

Sarah P. Everett

The problems in the 2000 election in Florida focused national attention on the need for usable voting systems. As a result, the Help America Vote Act (HAVA) of 2002 provided funding for updating voting equipment and many states purchased direct recording electronic (DRE) systems. Although these electronic systems have been widely adopted, they have not been empirically proven to be more usable than their predecessors in terms of ballot completion times, error rates, or satisfaction levels for the average voter. The series of studies reported here provides usability data on DREs to compare with that of previous voting technologies (paper ballots, punch cards, and lever machines). Results indicate that there are not differences between DREs and older methods in efficiency or effectiveness. However, in terms of user satisfaction, the DREs are significantly better than the older methods. Paper ballots also perform well, but participants are much more satisfied with their experiences voting on the DREs.

These studies also go beyond usability comparisons and test whether voters notice if their final ballots on the DRE reflect choices other than what the voters selected. Results indicate that over 60% of voters do not notice if their votes as shown on the review screen are different than how they were selected. Entire races can be added or removed from ballots and voter's candidate selections can be flipped and the majority of

users do not notice. Beyond discovering that most voters do not detect the changes, these studies also identify several characteristics of the voter and the voting situation that are important in whether participants will or will not notice the changes. This means that attacks could be targeted to only those people who will most likely not notice the changes. The result is that malicious software installed on a DRE could steal votes right in front of voters with a low probability of being detected.

ACKNOWLEDGEMENTS

First, thanks to my advisor, Dr. Mike Byrne, for the guidance and support he has given me during the last five years. I have enjoyed being part of TeamVoting and am grateful for all the experiences this project has provided.

I would also like to thank the rest of my committee members: Dr. Phil Kortum, Dr. Margaret Beier, and Dr. Dan Wallach. Their valuable feedback improved the quality of this dissertation in many ways.

Thanks to Kristen Greene for all her help and ideas, and to the entire VoteBox team (Dan Sandler, Kyle Derr, Ted Torous, Michael O'Conner, and Amanda Flato) for including my ideas in the design of their system and for altering their system to accommodate my research needs. This work would not have been possible without them.

Finally, and most importantly, I would like to thank my husband, parents, and sister for their unending love, support, and encouragement. This dissertation is dedicated to them.

This work was supported by National Science Foundation's grant #CNS-0524211 to the group entitled A Center for Correct, Usable, Reliable, Auditable, and Transparent Elections (ACCURATE).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

ADA – Americans with Disabilities Act

BDF - Ballot Definition File

COTS - Commercial Off-The-Shelf

DRE - Direct Recording Electronic

FEC - Federal Election Commission

HAVA - Help America Vote Act

HCI - Human-Computer Interaction

IDV - Independent Dual Verification

IPIP - International Personality Item Pool

ISO - International Standards and Organization

ITA - Independent Testing Authorities

NIST - National Institute of Standards and Technology

SES - Socioeconomic Status

SUS - System Usability Scale

VGA - Video Graphics Array

VVPAT - Voter-Verified Paper Audit Trails

VVSG – Voluntary Voting System Guidelines

CHAPTER 1

INTRODUCTION

The problems in the 2000 election in Florida focused national attention on the need for usable voting systems. As the country became familiar with terms such as "butterfly ballot" and "hanging chads," many states realized the importance of replacing these systems to avoid such problems in future elections. The Help America Vote Act (HAVA) of 2002 provided funding for updating voting equipment and intended for states to replace their outdated voting methods with newer, more reliable systems. As a result of this legislation, many states have spent millions to purchase direct recording electronic (DRE) systems to replace older technologies such as punch cards and lever machines.

In the 2006 election, it was estimated that over 66 million people would be voting on DREs in 34% of the nation's counties. While this is less than the percent of counties using paper ballots and optical scan technology (54%), DREs were much more widely used than punch cards at 4% of counties and also lever machines at 4%. The replacement of older voting systems with newer, electronic ones has been steadily increasing since 2000 when they were used in a mere 10% of counties, to 18% in 2002, to 21% in 2004, and finally to the previously mentioned 34% in 2006 (Election Data Services, 2006).

Although these electronic systems have been widely adopted, they have not been empirically proven to be more usable than their predecessors in terms of ballot completion times, error rates, or satisfaction levels for the average voter. There have been many reports of problems in the recent 2002, 2004, and 2006 elections (e.g. Kimball, 2004; Kimball & Kropf, 2006; and VotersUnite, 2006), but not much empirical research. Several studies that have done good investigative work in this area (Bederson, Lee,

Sherman, Herrnson, & Niemi, 2003; Conrad et al., 2006; and Herrnson et al., 2006), but there is still much work to be done. What these studies do not tell us is how DREs compare directly to previous voting methods, or whether voters will even use the additional features that DREs provide, such as the opportunity to check their choices in the form of a review screen before voters commit their selections.

The series of studies reported here provides usability data on DREs to compare against that of the previous voting technologies. Data were collected on three aspects of usability: efficiency, effectiveness, and satisfaction. These studies also went beyond performance comparisons and focused on a more specific aspect of the electronic voting experience: whether voters would notice changes made to their ballots. The first study examined whether there were usability differences between DREs with and without an opportunity for voters to review their choices before casting their ballot. The second study investigated whether voters noticed if races were added to or missing from their review screen. The third study in the series explored whether voters noticed any changes to their selections that were inserted into their review screen. These last two studies have security as well as usability implications; if voters do not notice when changes are made to their votes, hackers could steal elections by changing votes right in front of users. If the voters who do not notice changes can be identified, attacks could be targeted to those individuals, lowering the risk of detection.

CHAPTER 2

VOTING METHODS

Traditional Methods and History

Before discussing current and future directions in voting, it is worth surveying what voting methods have previously been used. One of the oldest and most traditional forms of voting, dating back to 1629 in the United States alone, is to cast a vote using a paper ballot. While paper ballots may seem simple, there are numerous forms of these ballots and they have undergone tremendous changes. Former versions of voting by paper ballot have required the voter to write in the candidate's name and have allowed political parties to distribute pre-printed ballots. For this latter form of paper ballot, the only actions required on the part of the voter were to bring this pre-printed ballot to the polling place and deposit it in a ballot box (Jones, 2003).

In 1888, Australian ballots were introduced for the first time in the United States. These were early versions of the current paper ballots. The Australian ballot lists all races and all the candidates running in these contests. Various forms of the paper ballot have required voters to scratch out the names of candidates they do not want, or more commonly, to put an X or other mark besides the name of their selected candidate (Jones, 2003). Three forms of paper ballots currently used in the United States include the open-response ballot, arrow ballot, and bubble ballot. Examples of these ballots can be seen in Figures 1, 2, and 3. A special form of paper ballot, the optical scan ballot, can be read by a machine and does not require hand counting.

FOR
UNITED STATES
SENATE
(VOTE FOR ONE)

(Republican Party)
Cecile Cadieux……..( )
(Democratic Party)
Fern Brzezinski…….( )
(Independent)
Corey Dery……( )

Figure 1. The open-response ballot.

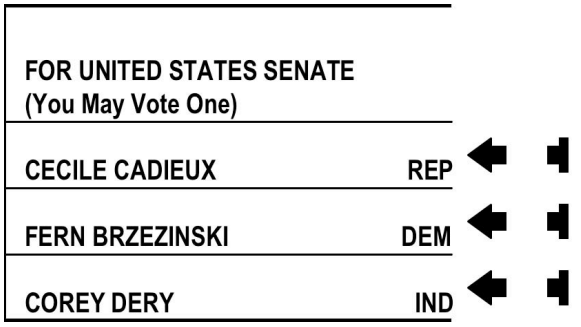


Figure 2. The arrow ballot.

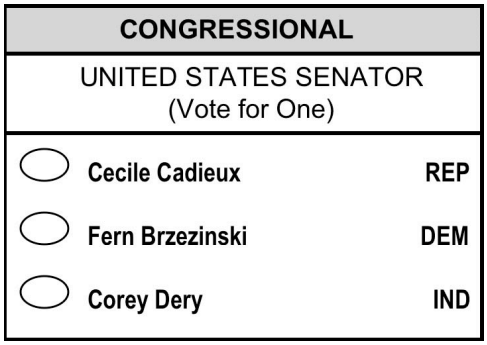


Figure 3. The bubble ballot.

Another form of voting introduced in the United States at the end of the nineteenth century was the lever machine (Jones, 2003). Although these machines are no longer being produced, they are still in use throughout the country. The lever machine has a full-face design; all the races and choices on the ballot can be seen at one time. See Figure 4 for a picture of a lever machine booth and Figure 5 for a close-up picture of the lever machine. Voters indicate their selections by moving a lever beneath the candidate's name. When the voter leaves the voting booth, the counters at the back of the machine increment to record his choices. One criticism of lever machines is that they only maintain total counts of votes; each individual vote is not recorded, only the total number of them. This means that recounts can only be performed on a machine-by-machine basis instead of by counting individual ballots.

Figure 4. The lever machine booth.

Figure 5. A close-up of the lever machine.

Although not originally intended for use in elections, punch card ballots were first seen in polling places in the 1960's. Punch card ballots themselves contain no election-specific information; they simply contain an array of numbered positions with perforated outlines or "chads" (Jones, 2003). Each voting station (Figure 6) contains a booklet with the races and candidates on the ballot (Figure 7). The voter slips the punch card behind the booklet and uses a stylus to punch out the chads corresponding to the candidates of their choice. Mainly as a result of the problems caused by the butterfly ballot in the 2000 Florida election, punch card systems are rapidly being replaced across the country.

Figure 6. The punch card station.

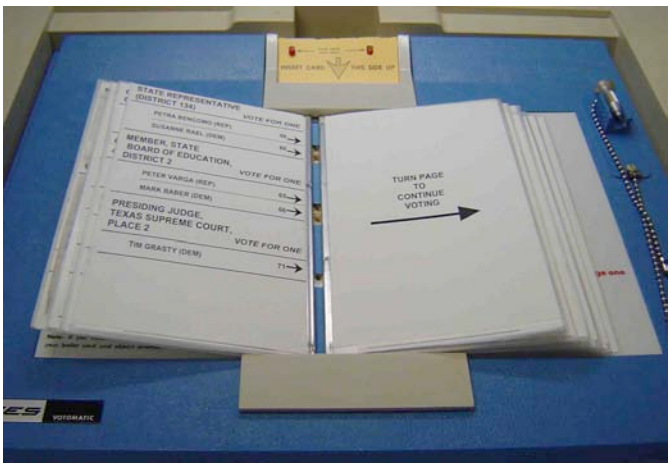


Figure 7. A page of the punch card ballot.

The newest forms of voting are DRE systems. Although they are, in many respects, a novel form of voting, the idea for DREs is not as recent as might be believed. The first patents for electrical vote recording machines began appearing in the middle of the 19th century. However, the idea was not truly realized until 1974 when the first commercial electronic vote recording machine was available (Jones, 2003). Since that time, there have been many improvements to, and has been much controversy about these machines, as will be further discussed here.

Types of DREs

There are currently many types of DREs on the market and each is slightly different. The Election Assistance Commission (EAC) has produced the Voluntary

Voting System Guidelines (VVSG) that provide basic design guidelines, but the design space for these machines is large and there are still many decisions to be made. These range from simple interface decisions such as the choice of colors and fonts to more complicated choices including how many contests to display on a page, whether to use multiple columns on a page, whether to allow scrolling, and how to display lengthy lists of candidate names. Still other design decisions must be made regarding input devices and whether to provide a review screen (Bremer, n.d.).

There are many ways to design a review screen for a DRE. Among the issues that must be addressed when designing these screens are display decisions such as whether to simply list all the contests and the voter's choice for each, and whether to highlight undervotes (Alvarez, 2002). The review screen also must provide a method for voters to change their choices or correct errors. This is an area where much research is needed on how voters use review screens so that they can be optimally designed.

Another DRE design decision, whether to include a printed display of the voter's selections, is more of a hardware issue. Many DREs do not provide a voter-verified paper audit trail (VVPAT) of the votes they electronically record. This is similar to lever machines that do not produce a paper record of each individual's vote. Lever machines simply provide summary counts for each contest on each machine and a recount only could be done by counting the paper records at the level of the machine. With paperless DREs, not only is there no paper record of every individual's vote, but there are also increased concerns about security and error. Many people feel that a printer attached to each machine that would provide a paper record of every vote would help alleviate some of the security concerns. Supporters of this paper trail also believe that in addition to

alleviating concerns about erroneous or fraudulent results, the paper record would increase voter confidence in DREs (Tokaji, 2005).

To give an idea for the types of DREs currently on the market and the dimensions on which they differ, Table 1 displays six commercially available electronic voting machines. The table below was taken mostly from Herrnson, Niemi, Horowitz, and Richman (2004). The Avante VoteTrakker, Diebold AccuVote TSx, ES&S iVotronic, Hart InterCivic eSlate, Nedap ES LibertyVote, and the Sequoia AVC Edge all have been used in recent elections. The table also includes the VoteBox system that was developed by Wallach and colleagues at Rice University and was used in the studies reported in this paper. The information about VoteBox included below is for the system as configured for use in these studies; it is a tool under development and may soon provide more capabilities.

	Avante Vote Trakker	Diebold AccuVote TSx	ES&S iVotronic	Hart InterCivic eSlate	Nedap ES Liberty Vote	Sequoia AVC Edge	VoteBox
Touchscreen	Yes	Yes	Yes	No	No	Yes	No
Screen size	14" or 15"	15"	12.1" or 15"	12.1"	24" x 34.5"	15"	17"
Type of buttons	Software	Software	Hardware & software	Hardware	Hardware	Software	Software using mouse
Separate navigation and selection buttons	Yes	Yes	N/A	Yes	Yes	Yes	Yes
Adjustable font size	Yes	Yes	Yes	Yes	Yes	Yes	No
Audio	Yes	Yes	Yes	Yes	Yes	Yes	No
Braille	Yes	No	Yes	Yes	Yes	Yes	No
Activation	Voter inserts card	Poll worker controls	Poll worker controls	Voter types ID number	Optional	Voter inserts card	Poll worker controls
Number of races on a screen	One at a time	Program-mable	More than one	Program-mable	All at once	Program-mable	One at a time
Page number displayed	Yes	Yes	Yes	Yes	All contests shown at once	Yes	No
Jump to a specific page	No	Yes	Yes	No	Yes	Yes	No
Selection method	Name or near name	Name or near name	Name or near name	Hardware button	Membrane button	Name or near name	Name or near name
Changing a vote	Push directly or after unselect	Unselect current vote first	Push directly or after unselect	Push new choice directly	Unselect current vote first	Push directly or after unselect	Push directly or after unselect
Write-in mechanism	Software - QWERTY keyboard	Software - QWERTY keyboard	Software – ABC keyboard	Software – ABC keyboard	Software - QWERTY keyboard	Software - QWERTY keyboard	No
Overvotes Prevented	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Review screen	Yes	Yes	Yes	Yes	All contests shown at once	Yes	Yes

Table 1. Six commercially available DREs and the VoteBox system, and where each stands on a variety of design and interface dimensions. Taken mostly from Herrnson et al. (2004).

CHAPTER 3

USABILITY AND HUMAN FACTORS

Why Is Usability Important?

When discussing voting systems, it is immediately obvious that issues such as accuracy and security are important. However, there is another facet of voting systems that can be overlooked: usability. Why is usability so critical in voting systems? To ensure the integrity of elections, voters must actually be able to cast their votes as intended. Unintentional undervotes (i.e., not casting a vote in a race), overvotes (i.e., voting for more candidates in a race than is allowed), or votes for the wrong candidates can substantially impact elections, as evidenced in the 2000 election upset in Florida. Wand et al. (2001) showed that the butterfly ballot used in Palm Beach County, Florida caused over 2,000 voters to vote for Pat Buchanan instead of Al Gore, tipping the election to George W. Bush. In another analysis of this election, Mebane (2004) used ballot-level data to show that 50,000 votes intended for Bush or Gore were lost to overvotes. He claims that had technology been available to warn voters of overvotes, Gore would have won by more than 30,000 votes. Similarly, Herron and Sekhon (2003) focused on presidential overvotes in the 2000 election and found that many more of the overvotes came from voters with a tendency to vote Democratically further down the ballot. In a user test performed soon after the 2000 election, Sinclair, Mark, Moore, Lavis, and Soldat (2000) found experimental evidence that the butterfly ballot was indeed more confusing than a single column ballot and that its use led to systematic errors in a mock election.

Although the design of the butterfly ballot in the 2000 presidential election has proven to be exceptionally confusing, poor ballot design is not limited to punch card ballots. Niemi and Herrnson (2003) studied paper ballots and found many designs to be unnecessarily complex and inconsistent between states. Some ballots include a symbol of the political party to help illiterate voters. However, the emblems used to represent parties vary widely between states, from shapes and objects to animals. Instructions can be misleading (e.g., "Vote for one" when electing a president and vice-president together) and options can be confusing (e.g., in some states straight-party voting requires multiple votes or allows voters to vote a straight-party ticket with exception). In addition, candidates may be listed with nicknames or without their political party. All of these ballot design inconsistencies can serve to confuse users and create additional potential for errors.

Confusion is especially problematic in elections because it is the voters themselves who must consider their voting experience to be a success. To have confidence in the outcome of an election, voters must believe that their votes were cast as intended and recorded. They must believe that they successfully used the voting system. Without this belief, the outcome of the election may be questioned. Voters may wonder if their vote was recorded as intended or even counted at all.

Usability problems can also cause long lines at the polls. It is important that voting not take an inordinate amount of time. Because of the vast number of voters casting a ballot at each election, it is essential that this process be done efficiently. If users are taking longer to cast their vote than anticipated, lines can quickly stack up to unexpected lengths, resulting in frustration and discontent.

Finally, usability can affect future voter turnout. If voters have had a bad experience, believe their vote will not count, have to wait an inordinate about of time to vote, or worry about figuring out how to use a voting system, they may choose to abstain from voting in the future. In fact, Roseman and Stephenson (2005) found just such an effect when they studied Georgia's switch to touchscreen electronic voting systems for the 2002 gubernatorial election. There was a statistically significant decrease in voter turnout among older people, who might be expected to be less comfortable with computers than younger people.

Assessing Usability

To assess the usability of voting systems, the National Institute of Standards and Technology (NIST) recommends using the International Standards and Organization's (ISO 9241-11, 1998) metrics of efficiency, effectiveness, and satisfaction (Laskowski, Autry, Cugini, Killam, & Yen, 2004). As NIST is responsible for setting voting system testing standards, it is important for research on the usability of such systems to use these metrics.

The efficiency of a system can be determined by studying the relationship between the level of effectiveness achieved and the amount of resources used (Industry Usability Reporting Project, 2001). The level of efficiency of a system can be measured objectively, usually by recording time on task. In the voting domain, efficient systems will take an acceptable amount of time and effort by the voters to cast votes for intended candidates.

Effectiveness is defined as the relationship between the goal of using the system and the accuracy and completeness with which this goal is achieved (Industry Usability

Reporting Project, 2001). In an effective voting system, voters would complete and cast their ballot, with their choices marked correctly and without outside help (e.g., an assist from a pollworker). Effectiveness can be measured by recording completion rates and number of assists, but is usually measured by collecting error rates.

The third measure of usability recommended by NIST is satisfaction, defined as the user's subjective response to working with the system (Industry Usability Reporting Project, 2001). This is the only subjective measure of the three, and it answers questions about how satisfied voters are with their voting experience, and how confident they are that their votes were recorded. User satisfaction with a system can be assessed by administering the System Usability Scale (SUS; Brooke, 1996). This is a standardized instrument that consists of 10 questions addressing different aspects of satisfaction. It has been used in a variety of domains for a range of tasks.

It is worth noting that the three usability metrics recommended by NIST match closely with several performance variables suggested by Card, Moran, and Newell (1983). Among other measures, these Human-Computer Interaction (HCI) researchers advise studying the time taken to accomplish tasks and the quality of performance, including how well tasks are done and the frequency and severity of errors. Card, Moran, and Newell also recommend including the user's subjective feeling about the system as a performance variable.

Frøkjær, Hertzum, and Hornbæk (2000) studied the correlations between efficiency, effectiveness, and satisfaction to determine whether all three are necessary when measuring usability. They found that the correlation between efficiency and effectiveness was negligible and report that user satisfaction was not determined by the

efficiency with which people can use a system. (This performance versus preference distinction is further discussed below.) Overall, Frøkjær et al. caution against discussing the usability of a system without including all three measures.

Individual Characteristics Affecting Voting

One reason that creating a usable voting system is challenging because there is such a diverse population of voters. Voters differ in many factors such as education, socioeconomic status (SES), age, race, gender, and previous voting experience, to name a few. At each election there are many first-time voters, and training cannot be depended upon to teach these new voters. In addition, voting must be accessible to non-English speaking voters and voters with disabilities.

Voters who do not read or write fluently in English may wish to vote in their primary language. In some counties, the printing of bilingual ballots is required by law due to large language minority populations (Saltman, 2006). However, it is not feasible for precincts to have copies of their ballots printed in all the languages in which voters may wish to vote. The use of DREs greatly simplifies this problem as they can support many languages without much additional cost.

Many advocacy groups have supported the adoption of DREs because of their enormous potential to be more accessible to voters with disabilities. One of the main guidelines for designing for physical usability is to allow for device independence such that users can choose their own input devices and are not required to use a standard keyboard or mouse (Trewin, 2006). Although some options like speech input may not be practical for voting because of privacy concerns, others such as specialized controllers, tongue switches, or sip-and-puff devices may be useful. Some, but not all, current DREs

allow voters to use their own input devices to the polling place. Even when users can bring their own devices, there may be additional issues that affect their ability to cast their own secret ballot, such as needing additional time to vote or accuracy with text input tasks like writing in a candidate. Systems need to allow for easy error correction, regardless of input device used, and systems must not have time constraints for completing ballots.

While electronic voting systems may improve accessibility for many groups of voters, computerized voting could lead to new problems for other groups of voters. Many users may not be experienced using computers and may not be comfortable performing a task as important and official as voting on such machines. It is possible that the "digital divide" will be seen in voting and may discourage older voters, who are typically the more experienced voters but less experienced computer users, from voting (Traugott et al., 2005).

Although there is a common assumption that older people are less willing to use computers and other new technologies, studies such as Czaja and Lee (2003) have found that older users are receptive to using computers. Differences may arise, however, in their ability to complete tasks efficiently and independently, especially on new interfaces that are used as infrequently as voting machines. As Czaja and Lee report, many studies in which older adults learn to use a computer for a variety of applications show that the older adults are slower to acquire computer skills and need more help to complete tasks.

Interface design decisions can help mitigate the effects of the visual impairment and reduced control of movements that sometimes accompany aging. Because older adults are more sensitive to glare problems, designers should be careful to minimize the

glare that users will experience (Welford, 1985). Using high-contrast displays and larger font sizes, or at least providing an option for users to increase the font size, can help those with decreased vision. Allowing users to adjust the gain ratio on their computer mouse can help users with reduced movement control (Fisk, Rogers, Charness, Czaja, & Sharit, 2004), although this may not be a viable option for voting machines. However, using an input device such as a touchscreen instead of a mouse has been shown to dramatically reduce the time taken to complete a computerized questionnaire for older users (Ellis, Joo, & Gross, 1991), and many DREs take advantage of this technology.

Even with these improvements to interfaces, older users may perform tasks more slowly than younger users, a phenomenon that may be seen on DREs as well as other methods. Studies have shown than with age, almost all aspects of performance become slower, especially for difficult tasks (Welford, 1985). Fisk et al. give a general rule of thumb that movement times in older adults will be about one and a half to two times what they would be in younger adults. In the voting arena, this could lead increased wait times if users are taking longer to vote than expected. For all these reasons, it is important that user tests involve participants with a range of ages.

In addition to differences due to demographics variables such as age or education, personality traits may play a role in voting behavior as well. There has been much debate as to how many personality factors exist and various researchers have proposed models with 3, 7, 12, or even 16 individual factors (McCrae & Costa, 2003; Pervin, 1996). However, the emerging consensus is that there are five basic dimensions of personality and this is known as the five-factor model (FFM). Although some controversy surrounds the names of these factors, they are commonly referred to as Neuroticism, Extraversion,

Agreeableness, Conscientiousness, and Openness to Experience (or Intellect) (Costa & McCrae, 1992).

Individuals scoring high in neuroticism tend to worry, feel insecure, and experience negative affect such as sadness, guilt, anger, or depression. It is a measure of maladjustment versus emotional stability. People scoring high in extraversion are sociable, talkative, assertive, and warm. They tend to seek excitement and have a need for stimulation and interpersonal interaction. High scorers in agreeableness are often soft-hearted, forgiving, and good-natured (Pervin, 1996). They are usually eager to help and expect others to be equally helpful (Costa & McCrae, 1992)

People scoring high in conscientiousness are hard-working, dependable, organized, effective, self-disciplined, ambitious, and persevering (McCrae & Costa, 1992). Although this is not determined by age, Goldberg, Sweeney, Merenda, and Hughes (1998) have found a positive correlation between age and conscientiousness, indicating that older adults tend to view themselves higher in this dimension than younger ones. Conscientiousness has been found to be a valid predictor of work performance in different job types, as reported by Barrick and Mount (1991). This suggests that people with the persistence, obligation, and sense of purpose associated with high scores in conscientiousness perform better in general than those scoring low on this dimension.

High scorers on openness to experience tend to have broad interests and are intellectually curious, untraditional, creative, and imaginative (McCrae & Costa, 1992). Scores on this dimension are associated with education and intelligence measures, and Goldberg et al. (1998) found a positive correlation between education and openness to

experience. However, openness to experience has been shown to be separate factor from intelligence and the two are considered separate dimensions of individual differences (McCrae & Costa, 1987).

The first three factors described here may not be predictive of voting behavior. Neuroticism is a measure of emotional stability and is much more likely predict experiencing inward negative emotions such as sadness, embarrassment, anger, and disgust than outward behaviors such as voting performance. Extraversion and agreeableness are really dimensions of interpersonal tendencies (Costa & McCrae, 1992) and so are not likely to influence a specific act of voting performance (i.e., checking one's vote carefully on the review screen).

It is the two factors of the Big 5, conscientiousness and openness to experience, that may affect voting behavior. As noted above, conscientiousness has been shown to be related to work performance and people high in conscientiousness are reliable and hard-working. In addition, these individuals tend to be cautious and to avoid making mistakes (Gully, Payne, Koles, & Whiteman, 2002). Thus, voters who are high in conscientiousness may vote more carefully and check their ballots closely. Individuals who score high in openness to experience are willing to question authority and think about new political ideas (Costa & McCrae, 1992). When voting, people who score high on the openness to experience factor may be curious about how the system is working and whether or not it is operating properly. Thus, these voters may check their ballots more carefully.

Human Factors in Voting

The previous section discussed individual differences that may impact voting behavior. The focus now turns to discussions of people in general, without focusing on the differences between them. The human factors literature on human performance can contribute much to the discussion and understanding of issues and potential problems in voting. One of these areas is the errors that people make while voting. When studying errors, it is useful to distinguish between the different types of errors that are possible. Reason (1990) divides errors into planning failures and execution failures. Planning failures, or mistakes, occur when a prior intention is not adequate to achieve a desired result. Execution errors, which include slips and lapses, occur when the prior intention was correct, but the action was not as intended. The voting world is more concerned with slips and lapses, that is, the errors that occur when the voter's intent is correct (the voter knows he wants to vote for Candidate A), but the execution in erroneous (the voter accidentally casts a vote for Candidate B or no candidate at all).

Another manner in which to divide errors is into errors of commission and errors of omission (Wickens & Hollands, 2000). Errors of commission occur when the user does the wrong thing, similar to execution errors in which a slip has occurred. In a commission error, the voter mistakenly votes for an unintended candidate. Errors of commission can also occur when a person votes for more candidates in a race than is allowed, for example, voting for two candidates in a vote-for-one race such as for governor. These types of errors are known as overvotes. Errors of omission, on the other hand, when the user does not do anything when he should have performed some action, similar to execution errors in which a lapse has occurred. An omission error would occur

if the voter meant to choose a candidate but did not choose anyone at all. These are known as undervotes in the voting literature.

Postcompletion errors are a specific type of omission error that can occur there is an additional step in a procedure after the main goal is completed (Byrne & Bovair, 1997). Examples include leaving the original of the photocopy machine after making copies or forgetting to put the gas cap back on after refueling. These situations are similar to when voters leave the voting booth upon the appearance of the review screen instead of after actually casting the ballot. The main goal of selecting the intended candidates is complete and voters consider the task finished. However, there is one more crucial step – the voter must actually press the button to record the vote, the terminal subgoal. In several respects, forgetting to complete the voting process by pressing the last button to submit the ballot is similar to other tasks in which postcompletion errors have been seen. Although it seems reasonable that this behavior is a type of postcompletion error, voting is an infrequently performed task and it is not always clear that the voters know they must perform the last step, as is the case in the commonly-cited examples.

Voting on Election Day can be a stressful experience for many voters, possibly leading to emotional reactions like anxiety and frustration. Polling places can be loud and distracting, and voters may feel time pressure to complete their ballots quickly. Studies of stress on human performance have examined a variety of stressful situations ranging from life-threatening situations such as believing one's plane was crashing (Berkun, 1964) to everyday stressors such as distraction and time pressure (Wright, 1974). The effect of stress on human performance often follows an inverted U-shaped pattern defined by the Yerkes Dodson Law; high stress can lead to errors, but moderate levels of stress may

actually improve performance over situations without any stress (Wickens & Hollands, 2000). In many studies, intense noise has been shown to slow performance on manual tasks and to cause increases in errors, and high levels of time pressure may speed performance, but at the expense of accuracy (Salas, Driskell, & Hughes, 1996).

On most tasks, people tend to make more mistakes as they increase their speed. People often can consciously increase the accuracy of their performance, but this comes at a loss of speed. This sacrificing of accuracy for speed or vice versa is known as the speed/accuracy trade-off (Wickens & Hollands, 2000). Voting at a polling place on Election Day likely follows this trade-off as it can be considered a speeded task; voters want to complete and cast their ballots as quickly as possible. Studies have shown that participants can impose some control over the direction of their speed-accuracy trade-off when instructed which aspect of performance to optimize (Howell & Kreidler, 1963). However, when left on their own, people tend to work at levels that achieve maximum performance efficiency (Fitts, 1966). Taken together, these findings mean that voters should be able to optimize their voting performance under normal conditions. However, when there is increased pressure to vote quickly, as in a situation when there is a very long line to vote or voters are in a hurry to get to work, accuracy may be sacrificed to finish the task quickly. On the other hand, because voting is such an important activity, voters could consciously increase their accuracy to ensure that their vote is expressed correctly, which would lead to increased vote times.

On a DRE, voters are given a unique opportunity to check the accuracy of their votes via the review screen. They are presented with a list of all the contests and their choices in each, and voters may confirm that each selection in correct. They are

attempting to distinguish a possible signal (a mistake or error) from noise (all the correctly displayed choices). This is an example of a signal-detection problem (Green & Swets, 1966). A hit would indicate that a voter correctly identified a mistake on the review screen, and a false alarm would occur if the voter mistakenly reported a problem with the review screen when no problem was actually present. In this situation, confirming the review screen without noticing an error would constitute a miss, and voters would be making a correct rejection if they approve the review screen when there are not any problems. People can adjust their response criterions to reflect a conservative or risky strategy (Wickens & Holland, 2000). If a voter has been following the news about potential accuracy or security problems with DREs, they may adopt a lower response criterion such that misses are less likely, which would also result in more false alarms.

In addition to measuring how long voters take to cast their ballots and the accuracy with which they can complete this task, it is important to assess how satisfied voters are with their experience. It has been well established in the human factors literature that a user's subjective preference does not always match their performance. Although a meta-analysis by Nielsen and Levy (1994) found that on average there is a positive association between preference and performance, there were still many cases in which there was a performance/preference mismatch. Numerous studies have shown that users often will favor a system on which they do not perform optimally (e.g., Andre & Wickens, 1995). Merwin and Wickens (1993) found that performance judging continuous 2D data was best with grayscale displays, but that grayscale displays were not preferred over various color displays, as might be expected. This dissociation between performance

and preference applies to interface designers as well as system users. In a study by Bailey (1993) using designers as participants, 95% of the participants preferred an interface that did not produce the fastest performance. For reason such as this, Andre and Wickens strongly recommend collecting user performance ratings in addition to preference ratings.

Previous Research on the Usability of Voting Systems

It is only recently that controlled usability tests of voting systems have been conducted. Previous to this research, human performance on voting systems was gauged by studying voter performance in actual elections. However, these methods are limited as very few, if any, polling places have any record of how long it took voters to complete their ballots or how confident they were with their performance (Brennan, 2006). Errors have been estimated by examining residual vote rates in actual elections. The residual vote rate for a particular contest is the difference between how many valid votes were cast in that contest and how many total ballots were cast. This is an imperfect measure of voting error rates, however. It is fairly safe to assume that voters do not intentionally spoil their ballots by overvoting. Conversely, with undervotes it is impossible to know whether voters intentionally abstained from voting in a particular contest or made a mistake. Exit polls and previous voting records of the precinct can provide insight into typical patterns of voter preferences, and this can inform, but not confirm, error estimates.

By studying nationwide residual vote rates in the 1988 to 2000 elections, Ansolabehere and Stewart (2005) determined that the type of voting system used in an election significantly impacts how many votes will be counted. They found that the difference between the best and worst voting systems was up to 2% of the ballots cast.

Paper ballots counted either by hand or by optical scanner performed the best and were consistently better than punch cards and lever machines. DREs were found to have worse residual rates than paper ballots and optically scanned ballots.

Kimball and Kropf (2005) focused solely on paper ballots and design effects on unrecorded votes. They collected ballots from 250 counties in 5 states. Based on the belief that unrecorded votes are the equivalent of item nonresponse in questionnaires, Kimball and Kropf coded ballots on conformity to good design practices suggested in the literature on self-administered questions. Combining ballot features into one index of ballot design and comparing this to residual vote rates in the 2002 gubernatorial election showed that ballot design played a significant role in unrecorded votes. The authors suggest that trying to solve the residual vote problem by buying new technology instead of improving ballot design may not have the desired effect.

Indeed, other studies have found differences in residual vote rates even between versions of the same voting technology. In an early study, Mather (1964) compared election data from counties using older lever machines and those using newer versions, and found a significant increase in the number of valid ballots cast with the new machines. The difference was likely due to a locking mechanism that prevented voters from opening the booth's curtain without having any levers pulled down, effectively preventing unintentional undervotes for complete ballots. Kimball (2005) studied optical scan ballots used in the 2004 election and found that connect-the-arrow ballots produced residual vote rates of 0.9%, but fill-in-the-bubble rates of were only 0.6%. Similarly, in a large-scale study, Herrnson et al. (2006) found error rates for commercially available DREs ranging from 2.7% on the Avante VoteTrakker to 3.7% on the Hart InterCivic

eSlate and Nedap LibertyVote.

Previous work collected baseline usability data for several traditional voting systems. Everett, Byrne, and Greene (2006) studied three types of paper ballots: bubble ballots, arrow ballots, and open-response ballots. Greene, Byrne, and Everett (2006) looked at lever machines in addition to paper ballots, and Byrne, Greene, and Everett (2007) added punch cards. Overall, these results indicated that paper ballots, especially bubble ballots, seem to be the voting method that is the most usable for the greatest range of users. One reason that paper ballots are the most usable is due to their error rate of about 1.5%. While this number is certainly higher than we would like to accept in a task as important as voting, it is lower than the error rates for both the punch cards and lever machines. In addition, although there are not yet directly comparable empirical findings from DREs, preliminary results from work by Herrnson and colleagues show higher error rates than were found with old-fashioned paper ballots.

Although they have not directly compared their findings on DREs to traditional forms of voting, Herrnson and colleagues have performed several studies of voters using DREs (Bederson et al., 2003; Conrad et al., 2006; Herrnson et al., 2006). In recent work that has not yet been published, they found evidence of usability problems with DREs. In a large-scale field study, error rates (including residual votes and votes for the wrong candidate) on the presidential race were as high as 4.2% on one electronic voting system, as noted above. Of all voters in the study, 2.5% voted for the wrong candidate for president. Herrnson and colleagues point out that this is an especially serious type of error because not only does the preferred candidate lose a vote, but an opponent also gains a vote. This study also showed evidence that error rates were affected by voter

demographics such as age, education, computer experience, race, and English as primary language. These studies are important as they compare different types of DREs, and their results begin to shed light on the usability of electronic voting systems. Further work will be needed to directly compare the usability of DREs with that of older forms of voting.

CHAPTER 4

SECURITY

Security Background

To ensure the integrity of our elections, the voting process and its outcomes must be accurate and secure. Although DREs have been touted as the best solution for accurate elections, many computer security experts have expressed major concerns about their susceptibility to fraud and errors (e.g., Dill 2005). A study of the source code for one type of DRE, the Diebold AccuVote-TS revealed many security flaws that could be exploited to affect an election outcome (Kohno, Stubblefield, Rubin, & Wallach, 2004). These included bugs that allowed voters to cast multiple votes without being later detected and allowing voters administrator access to the systems, as discussed below. Another in-depth study of the AccuVote-TS revealed that malicious code could be easily installed on voting machines on Election Day, and could spread like a computer virus, infecting other machines through normal voting activities (Feldman, Halderman, & Felten, 2006).

The Hack-a-Vote project from a computer security class at Rice University (Bannet, Price, Rudys, Singer, & Wallach, 2004) demonstrated how easily even students could introduce hacks into a voting system. Students incorporated several types of hacks into voting machine code that could affect the final result. Hackers then switched to the role of auditors and had difficulty finding the bugs in others' software. This project demonstrated what many security experts have been concerned about: how easily attackers could insert malicious code into systems and steal elections. In fact, a report on voting system vulnerabilities by the Brennan Center for Justice (Brennan, 2006) has

identified this sort of threat, using corrupt software programs, as the least difficult type of attack.

Although independent testing authorities (ITAs) conduct the testing and certification of voting equipment, voting software cannot be assumed to be trustworthy even at the time of certification. While this testing process should help safeguard against buggy or malicious software, its tests are incomplete and the results are secret. Commercial off-the-shelf (COTS) software and the ballot definition files (BDFs) containing election-specific information are not included in the testing process. ITAs do not specifically check for security problems or potential attacks, but focus on adherence to Federal Election Commission (FEC) standards (Simons, 2004). The result is that the software certified for use in elections cannot be assumed to be secure.

Types of Attacks

Numerous potential threats to voting have been identified and many attacks apply only to DREs. However, some threats concern a range of voting methods. Ballot box stuffing, for example, traditionally could be accomplished with paper ballots or punch cards. However, the problem of a single person casting of multiple votes applies to any voting method. In lever machines, mechanical devices prevent this problem by only allowing a voter to cast one ballot when inside the machine. Poll workers easily could notice if a voter was repeatedly entering and exiting the booth, as would be required to cast multiple votes. With newer systems such as DREs, however, this could be accomplished much more subtly. As noted above, malicious software that allowed the casting of many votes could be hidden in the voting machine software or installed later. Also, many of these machines are activated through the use of an insecure smartcard or

personal identification number (PIN) (i.e., a numeric password) to authenticate the user to the system. As Kohno et al. (2004) have shown with Diebold AccuVote-TS machines, a computer savvy-voter could bring a stack of smartcards to vote multiple times, or program a smartcard such that it would not be deactivated by the voting system after its use.

One of the requirements of elections is that a voter's choices must be kept secret so that voters can make decisions without fear or intimidation. Keeping ballot information private prevents voters from being forced into voting in a certain manner. It also prevents vote-buying attacks. As long as a voter's choices are secret, he cannot prove to a potential buyer for whom he actually voted. Schemes for vote buying, such as chain voting as described by Jones (2005), have long been in place for paper-based voting methods, and the threat is still very real for computerized systems. With systems that record the cast ballot information in a sequential manner, there is the possibility that someone, such as a poll worker with access to voting records, could match up voters with the ballots they cast. Kohno et al. (2004) found that the Diebold AccuVote-TS system stored vote information in just such a manner.

Also, a recent study on the Nedap/Groenendaal ES3B voting computer using in the Netherlands has shown that emissions from voting machines are detectable for up to 25 meters away. Researchers found that enough information can be gleaned from these emissions to compromise ballot secrecy. By analyzing the pattern and strength of emissions, researchers could determine for which candidates and political party a voter chose (Gonggrijp et al., 2006).

Although it is possible that a sort of denial-of-service attack could affect non-computer based voting systems, they are of much more concern with DREs. In this type of attack, DREs would not allow voters to cast any ballots at all. If an attacker gained administrative privileges to a voting terminal, he could prematurely end the election and prevent any more votes from being cast. Denial-of-service attacks could happen if the computers at a polling place all crashed simultaneously or if computers depending on internet connections to download ballots at the start of an election were denied service. These attacks can launched from a distance and on a very large scale (Kohno et al., 2004). Denial-of-service attacks could also be targeted so that machines crashed only in certain precincts or at certain times during Election Day, distorting the election results (Feldman et al., 2006). Researchers such as Simons (2004) have suggested that polling place should be prepared with paper ballots in case of DRE failures, such as could occur with denial-of-service attacks.

Corrupt vote-counting practices have long been of concern to elections, and with electronic voting systems, there is an increased threat of vote flipping (i.e., changing a person's vote to a candidate other than the one originally chosen). Kiayias, Michel, Russell, and Shvartsman (2006) found that Diebold's AccuVote Optical Scan voting terminal (AV-OS) is vulnerable to attacks that can be carried out with commercial off-the-shelf equipment and only take a few minutes. The vote tabulations can be manipulated so that votes for a certain candidate are not counted or are switched with another candidate, thus affecting the resulting totals. This could be done so that evidence vote stealing would be impossible to find; Feldman et al. (2006) showed that on the

AccuVote-TS, it would be possible to change all the voting records and counters such that the fraud would be undetectable.

A study by Di Franco, Petro, Shear, and Vladimirov (2004) has shown that effective vote flipping can be even subtler than might be assumed. They showed that in an election as close as the 2000 presidential election, even changing one vote per voting machine could have changed the outcome. They note that such small manipulations of votes may not even be detected at all, or may be ignored as simple noise in the results.

Previous Security Concerns

Although the recent widespread adoption of DREs has brought with it a wide array of security concerns, including those described above, security concerns have long plagued elections. As Jones (2006) notes, issues concerning vote manipulation and fraud were of much concern to 19th century inventors. In the 1870s, precinct-level fraud was a major security issue. Mechanical or electronic voting machines aid in the prevention of this type of fraud, as well as mitigating ballot box stuffing and vote buying, but their use can lead to other types of attacks. Insider attacks are of concern to voting machines, as they are in the hands of the technicians or developers. A sealed counter mechanism was suggested in the late 19th century to protect against insider attacks. Even the idea of voter-verified paper trails have been around for over 100 years, as a patent from 1899 included one.

More specific to current accuracy and security concerns are the computerized vote-tallying issues discussed by Saltman in 1988. He reports on two instances of potential inaccuracies and fraud, and writes of steps that should be taken to control for these. The first example concerns the November 1984 Carroll county school board

election. When a mandatory rerun of punch card ballots did not return the same results as the original outcome, concerns about fraud or vote-tallying inaccuracy were raised. It was later discovered that the administrator of the Data Processing Center of the county had accidentally installed a test version of the vote-tallying software that did not allow for the current ballot configuration. The second example describes the May 1986 county commissioner election in Stark County, Ohio. The race was so close that a recount by computer was ordered. The recount flipped the election outcome, but a manual recount showed evidence in favor with the original outcome. A program error in the recount software that prevented the computer from accurately distinguishing between different types of ballots was soon discovered.

While neither of the two computerized vote-tallying software problems described above was determined to be fraud, they did raise public awareness as to the potential for fraud. To prevent future fraud and error, Saltman (1988) brings to awareness certain vulnerabilities of vote-tallying software. These include logical errors and hidden code in software, undocumented changes to software, and unauthorized access to election computers. Although these issues are almost 20 years old, they correspond closely with vulnerabilities of DREs currently in use.

Proposed Solutions

Although groups such as VotersUnite (2007) have been advocating a return to paper ballots and complete ban on DRE use, others have suggested a wide variety of strategies to help ameliorate the security risks of electronic voting. One such strategy (Chaum, 2004) involves letting voters take home a receipt of their vote. This receipt would let them check (via a webpage) that their vote was included in the final vote tally.

However, because the voter's choices are encrypted, the receipt would not let voters prove to others how they voted. This idea is similar to a proposal for a receipt-dispensing paper-based system (Rivest, 2006), but allows voters to cast their ballot electronically. In both methods, it would be difficult for hackers to add, change, or delete ballots without detection, as voters will be able to check the website's information against their own receipt.

Because of increasing concerns about DRE security, independent dual verification (IDV) systems such as DREs with VVPATs have been suggested by researchers as well as advocates and lawmakers. The Verified Voting Foundation has been calling for voter-verifiable audit trails for electronic voting systems for several years (Verified Voting Foundation, 2004). They believe that it is imperative that a permanent record of every ballot is kept that can be verified by the voter and cannot be changed after this verification. Similarly, a recent report from NIST (2006) has called for software independence in voting systems. In a software-independent system, malicious code or errors could not change an election outcome without being detected. An outside record such as a paper ballot or paper trail could be verified independently of the voting system and used in an audit. This NIST report recommends that the VVSG 2007 should ban the use of software-dependent voting systems, which would outlaw DREs without VVPATs.

Cohen (2005) studied paper and auditory audit trails to assess participant preference and performance. Results indicated that participants caught more errors with the auditory version of the voter-verifiable audit trail. Also, participants reported that they would prefer the paper audit trail for their own county elections. However, the two tasks were fairly dissimilar because in one condition participants were forced to listen to

auditory confirmation after each selection, but those in the other condition were not forced to read through the paper record that was presented only at the end of voting. This study provides a start, but more research is needed on VVPATs and whether people actually check their selections for accuracy before committing their vote.

Another type of IDV system is witness systems (Gerck, 2001). One type of witness system has a device that taps into the video graphics array (VGA) line and the line back from buttons (Kelsey, 2006). This device records all screen images and button presses, enabling auditors to see the ballot as it was seen and cast by the voter, not just as the computer recorded it. Another proposed witness system has a camera mounted above the DRE screen that takes pictures of the voter's interactions with the machine, as described in Appendix D of the VVSG 2005 (Election Assistance Commission, 2005). These types of IDV system would allow vote flipping not only to be detected during an audit, but also corrected.

A different type of solution to security problems that has been advocated is to publicly publish the code used on DREs. Currently, the companies producing DREs are not releasing the source code used on their machines and are depending on this secrecy to ensure the security of their systems. However, this secrecy usually does not last, as was demonstrated in 2004 when the Diebold source code was published on the Internet. Instead of depending on "security through obscurity," open source software allows computer experts to examine the code for errors or security vulnerabilities (Hoepman & Jacobs, 2007).

While open source software allows for detailed study of the code used to run voting machines, the code can be extremely lengthy and difficult to examine. Yee,

Wagner, Hearst, and Bellovin (2006) have suggested using prerendered user interfaces as a partial solution to this problem. In this scheme, all of the user interface components are laid out ahead of time and are separate from the machine that will be used on Election Day. These prerendered user interfaces serve as electronic sample ballots and are made public before the election. This allows public verification of the ballot information and layout. Also, because all the user interface elements are removed from the actual software used to run the election, the amount of code that needs to be verified and trusted is significantly reduced.

A voting machine prototype using this type of prerendered interface was built by Wallach and colleagues at Rice University. The group has created a system, VoteBox, that demonstrates the usefulness of security properties such creating a system where only a small part of the code has to be trusted. VoteBox is currently a paperless DRE with a mouse as the input device, but this prototype will be developed further to include features such as VVPAT capability, touchscreen input, and multiple language support. A picture of the VoteBox system can be seen in Figure 8 and a screenshot of a page allowing users to vote on a proposition is shown in Figure 9. (A comparison of many properties of this system and commercially available systems was shown in Chapter 1.) The interface of the system was developed in consultation with a human factors group at the same university, and the usability of this system was assessed in the studies reported here.

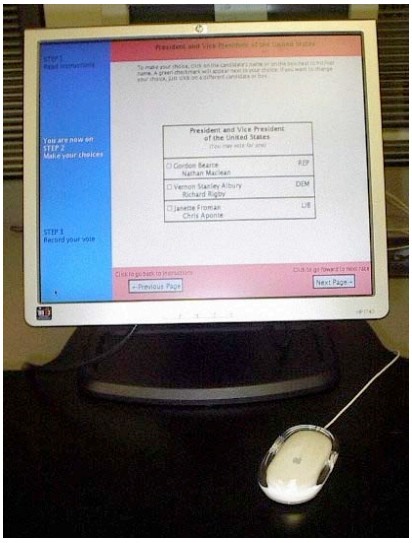


Figure 8. The VoteBox system setup.

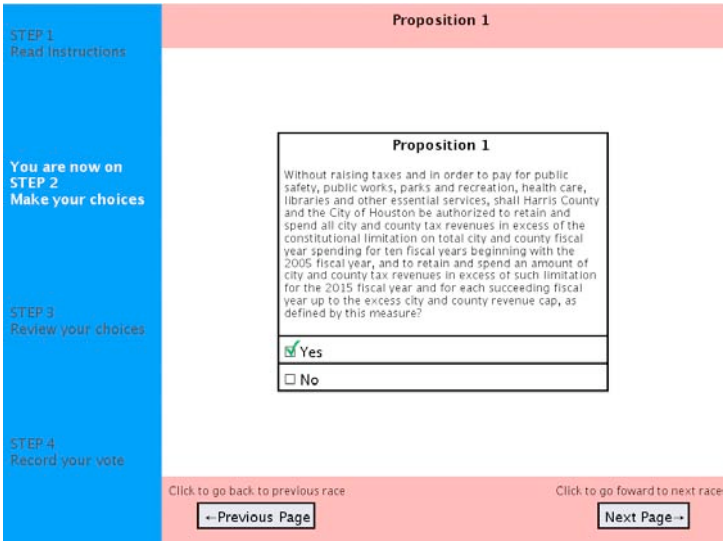


Figure 9. A screenshot from VoteBox allowing users to vote on a proposition.

CHAPTER 5

STUDY 1

Overview

The main purpose of the first study was to collect baseline performance data on participants' voting with the VoteBox electronic voting machine that was described above (from here simply referred to as the DRE). Participants voted on the DRE and one additional voting method (lever machine, punch card ballot, or paper ballot). The usability of each was assessed and compared to previous work on existing voting methods. Results from this study begin to shed light on how performance on DREs lines up with performance on older, more traditional forms of voting.

The main manipulation of this study was whether participants saw a review screen after selecting their choices. A review screen is a screen that appears on the DRE after the voter has made all of their selections and displays each contest on the ballot and the voter's selection for each. (Figure 10 shows a screenshot of an example DRE review screen taken from the VoteBox system described above.) In this way the review screen serves as a summary of the voter's ballot and is the last opportunity participants have to check their ballot before its submission. The review screen is a feature that DREs can provide that older voting methods cannot. It has the potential to be useful in drawing voter attention to undervotes and allowing them to check their selections before casting their ballot. This study assessed whether the usability of DREs as measured in terms of efficiency, effectiveness, and satisfaction is affected by the presence of such a review screen.

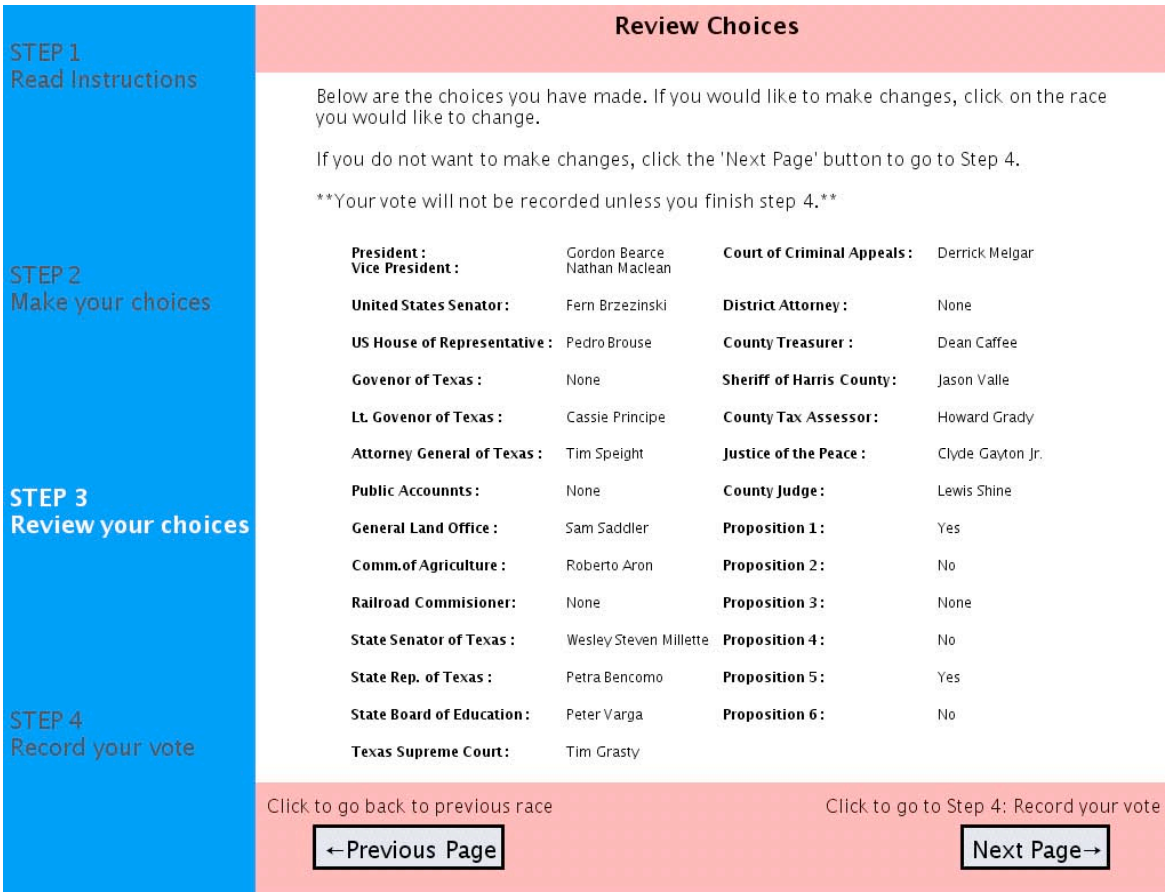


Figure 10. A screenshot of VoteBox's review screen.

Methods

Participants

Participants for this study were 48 Rice University undergraduates (21 male, 27 female). The average age of participants was 19.8 years (*SD* = 4.6) with median age of 19 years. All participants had normal or corrected-to-normal vision. The participants were comfortable with computers; the average self-rating of computer expertise was 6.7 on a 10-point scale (*SD* = 1.5).

Participants represented a variety of political affiliations, with Democrats being the most strongly represented (see Table 2). For the racial distribution of the sample, see

Table 3. Thirty-four participants (71%) had previous voting experience. The methods with which the most participants had had experience were the bubble ballot and electronic voting methods (see Table 4 for how many participants had had previous experience with each voting method).

Political affiliation	Frequency
Republican	7
Democrat	22
Libertarian	6
Independent	10
Other	3
Total	48

Table 2. Frequency of each political affiliation.

Race	Frequency
African American	3
American Indian	0
Asian American	7
Caucasian	28
Mexican American or Chicano	1
Other Hispanic or Latino	1
Multiracial	6
Other/No report	2
Total	48

Table 3. Frequency of each race.

Previous Voting Method	Frequency
Bubble ballot	17
Arrow ballot	3
Open response ballot	5
Lever machine	3
Punch card	2
Electronic - touchscreen	11
Electronic - other	7

Table 4. Frequency of having previous experience with each voting method.

All participants received credit towards a course requirement for their participation.

Design

The primary manipulation in this study was whether participants were presented with a review screen displaying their selections at the end of voting. This was a between-subjects manipulation with half the participants seeing a review screen after making their selections, and the rest continuing straight from selections to casting their ballots. The remainder of the design was very similar to that of previous studies (Everett, et al., 2006; Greene, et al. 2006; and Byrne, et al., 2007). There were two information conditions: a directed condition in which participants were given a paper (a "slate") instructing them for whom to vote and an undirected condition in which participants were offered a voter's guide and made their own selections. The type of additional voting method the participant used was also a between-subjects variable. Participants voted on one of paper, punch card, or lever machine in addition to the electronic voting machine. The order of presentation was counterbalanced.

Materials

The electronic voting machine used in the study was the VoteBox system described in Chapter 4. (Picture of the VoteBox system were shown in Chapter 4. See Figure 8 for a picture of VoteBox and Figures 9 and 10 for screenshots of a proposition contest and VoteBox's review screen. Appendix A contains more screenshots from the system). The VoteBox team has included features in their system that support running typical user testing studies (e.g. recording response times and user actions), as well as adding support for making malicious changes to a user's vote. Study 1 only took advantage of the user testing support the system provides, while the second and third studies also used the malicious code part. The instructions on the DRE were standard voting instructions. On the actual review screen itself, there was an instruction to click on a candidate's name to change your selection for that race. Consistent with other voting methods used in the study, the DRE did not require users to enter a pin or insert a smartcard to begin, or to choose a language.

There were three voting methods used in addition to the DRE described above. The first was a paper ballot, specifically a bubble ballot on which voters indicated their choices by filling in an oval next to the candidate name of their choice. The bubble ballot used in the study was based on a real ballot used in a previous national election. The second was a punch card ballot on which participants indicated their choices by using a stylus to punch a perforated hole next to the name of the candidate of their choice. The third method was a lever machine, on which participants indicated their choices by pulling a lever beneath the name of the candidate of their choice. The punch card systems and lever machines used in the study were purchased at auction from local counties and

had previously been used in national elections. (The punch card ballot did not split candidates for one contest between two pages, as was done in the butterfly ballots in 2000.) Pictures of each voting method can be seen in Chapter 2. The bubble ballot can be seen in its entirety in Appendix B.

There were 27 contests on each ballot. Twenty-one were office contests including federal, state, and local races, and six contests were proposition decisions. The races and propositions used in the study are included in Appendix C. As a previous study (Everett et al., 2006) had found no differences between using fictitious names and the names of actual politicians, only fictitious names were used in this study. When using fictitious names, the names on ballots do not have to be changed between experiments as politicians change offices. Also, there are no effects of current events or campaigns on the results. The use of fictitious names allowed for greater consistency between the materials used in this study, past studies, and any future studies. The candidate names were created by a random-name generator found at http://www.kleimo.com/random/name.cfm. The propositions were ones that had not been voted on in the local area, but which had relevance to the area.

The slates contained the names of the races and the candidate name or proposition position for which the voter should vote. Within the directed condition, three versions of the slate were used to ensure a fair mix of candidates and positions for which participants are instructed to vote. There was a slate in which 85% of the candidates were Republicans, one in which 85% were Democrats, and one that contained a more even mixture of candidate affiliations. All slates instructed participants to vote "Yes" for four propositions and "No" for two.

The voter guides were based on the guides distributed by the Houston League of Women Voters. For each contest on the ballot, the guide listed the candidates and their responses to a few questions on key issues (see Appendix D for the full voter guide used in the study). For the propositions, the guide listed reasons for and against each proposition. Disclaimers on all guides, instructions, and debriefing forms informed participants that all the material in the guides was solely for research purposes and may not represent the views of any real person.

None of the voting methods offered a straight-party ticket option. Because each state has its own rules about how to handle straight-party voting and many states forbid this option, it was not offered in this study. Of course, voters still could vote a straight-party ballot, but it would require them to mark a choice in every race.

The survey that participants completed at the end of the study contained questions about general demographics, previous voting experience, previous computer experience, and a survey of subjective user experience, the SUS (Brooke, 1996). A measure of the Big 5 personality characteristics was also included to see whether personality characteristics of voters played a role in their performance. This measure was taken from the International Personality Item Pool (IPIP), a website of public-domain personality measures. The Big 5 survey contained 50 items that resulted in scores on each of the Big 5 factors. Results from the use of this fairly short 50-item inventory have been shown to be highly correlated with the original NEO-PI-R by Costa and McCrae in 1992 (Goldberg et al., 2006).

Procedure

After signing an informed consent form and reading the study instructions, participants voted on both voting technologies to which they were assigned (the DRE and one of the additional methods). Before voting each time, they were reminded to vote consistently if they were in the undirected condition, or to vote as their slate instructed in the directed condition. If the participants were in the undirected condition, a brief exit interview was conducted after voting to obtain the voter's intent. After they completed this process, voters filled out a survey packet. Finally, participants were debriefed as to the purpose and goals of the study and departed.

Because lever machines only record the total counts for each candidate or position and do not provide records of each individual ballot cast, ballots on these machines were scored immediately. Punch card ballots, paper ballots, and those cast on the DRE were scored and errors were tallied later. Although punch card ballots and optical-scan ballots like our paper ballot are typically scored by machine, they were hand-counted in this study. Scorers judged the intent of the voter, counting marks that may not have been counted if the ballots were run through a machine. Because of this, the error rates reported here may represent a best-case scenario; even if participants did not mark ballots as instructed, their marks were counted if their intent could be inferred. This system is similar to manual recounts that are occasionally mandated for auditing purposes or in extremely close races.

Results

DRE Review Screen Performance

Besides comparing performance on DREs to that on older voting methods, the purpose of Study 1 was to compare performance on a DRE with a review screen and one without. The two version of the DRE will be evaluated on the three usability measures of efficiency, effectiveness, and satisfaction. Outliers were identified as data points which were outside 3.5 standard deviations of the mean for the voting method. When outliers were detected, they were replaced by the mean for the voting method. In the timing data, one outlier was replaced in the DRE condition. In the data on error rates, one outlier in the bubble ballot rate and one in the DRE rate were found and replaced. There were no outliers identified in the SUS data.

ANOVAs revealed no effect of review screen presence on time ($F(1, 46) = 1.119$, $p = .30$), error rate ($F(1, 46) = 0.003$, $p = .95$), or satisfaction ($F(1, 46) = 1.13$, $p = .29$). For the remaining Study 1 results, the no review screen/review screen conditions were collapsed, as there was no difference between them.

DREs Versus Older Voting Methods

In this study, timing, error, and satisfaction data were collected as in previous studies (Everett et al., 2006; Greene et al., 2006; and Byrne et al., 2007). Because studies have already been conducted looking at older forms of voting, the results here focus on how performance on the DRE compares to these previous methods.

As mentioned above, each participant did not vote on each voting method, but only on two: the DRE and one of the bubble ballot, lever machine, and punch card. Because of this, comparisons between the methods were made using paired *t*-tests. Next,

ANOVAs were run for each measure of voting performance using the difference scores between the DRE and "other" method. This "other" method included the bubble ballot, lever machine, and punch card. These were done to examine the effect of other independent variables of information condition and specific voting method on time, error rates, and satisfaction.

Time

For all four of the voting methods, ballot completion time was measured beginning when the participant entered the voting room and ending when they exited the room. This was recorded in seconds using a stopwatch. For the DRE, the computer itself also recorded the time taken to vote. The correlation between stopwatch-recorded times and computer-recorded times was very high ($r = .998$, $p < .001$). Although this correlation reveals that the two times were significantly highly related, the correlation is not perfect. There are two reasons for this discrepancy: 1) the computer started recording time once the person was already in front of the machine (not walking into the room) and stopped as soon as the person clicked the cast vote button (instead of when they exited the booth), and 2) occasionally participants would walk away from the computer before they had actually finished voting (i.e. while the DRE was showing the review screen instead of after casting their ballot). Because of this, all efficiency measures reported here will use the time recorded by the stopwatch instead of by the computer to ensure that the times are as accurate as possible and to keep consistency between the four voting methods.

Participants completed their ballots fairly quickly, with an average ballot completion time of 237 seconds ($SD = 114$). Completion times for the four voting

methods can be seen in Figure 11. Paired *t*-tests revealed no differences between ballot completion times on the DRE and any of the three other voting methods, although the DRE-punch card difference approached significance $t(15) = -2.12$, $p = .051$.

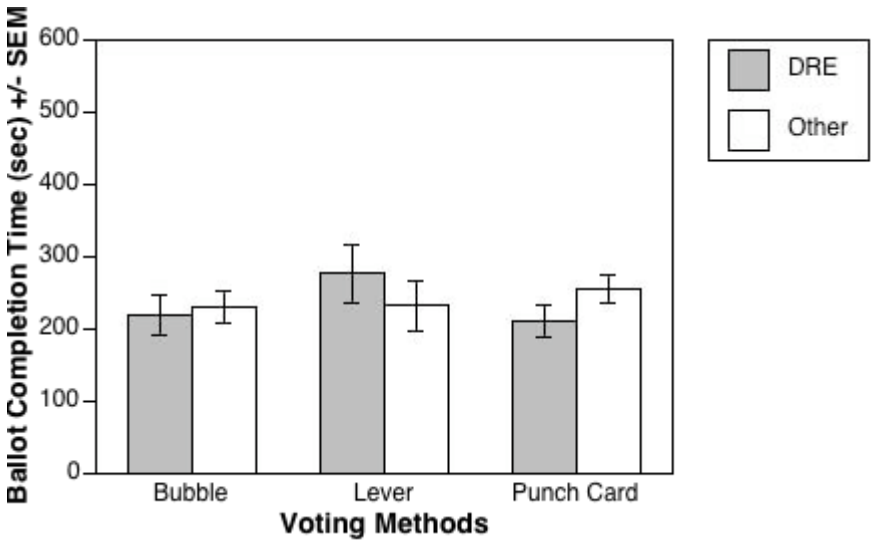


Figure 11. Pairs of mean ballot completion times (in sec) by voting method. Gray = DRE, white = other method (bubble, lever, or punchcard).

An ANOVA was run on the difference between ballot completion times on the DRE and the other method on which participants voted by information condition and other method. This revealed no effect of which other method was used or information condition. There was also no interaction of the other method and information condition on the ballot completion time differences.

Errors

Errors were counted when the voter made an incorrect selection in the directed condition, or, in the undirected condition, when their selection was inconsistent with their other vote cast and their intent as given in the exit interview. See Table 5 for an example of an error in the undirected condition.

Voting Method	Voter's Candidate Choice	Error Determination
DRE	Gordon Bearce	Correct – Intended candidate
Lever	Janette Froman	*Error* – Inconsistent choice
Exit Interview	Gordon Bearce	Correct – Intended candidate

Table 5. Example of an error in the undirected condition.

There are two ways to consider errors: by race and by ballot. Errors by race will be discussed first. On every ballot, there were 27 races and each of these represented a potential for error. Error rates were computed by simply summing the errors committed, and dividing by the total possibilities for errors.

Average error rates in this study were very low, with an average rate of 0.0008 (*SD* = 0.0053). Error rates for each voting method can be seen in Figure 12. (Note that the scale for this graph is different for this study than for Studies 2 and 3 due to especially low error rates.) Paired *t*-tests performed on the error rates of the DRE versus each other method showed no differences between any of the pairs.

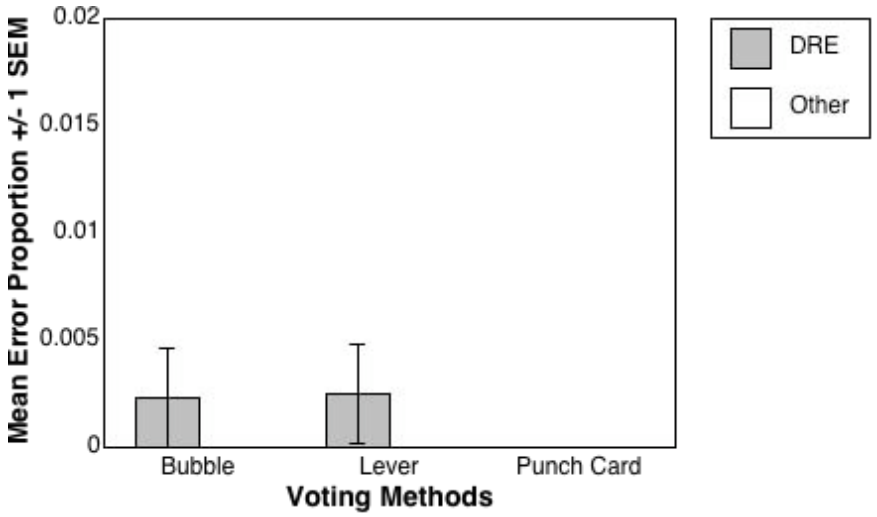


Figure 12. Pairs of mean error rates by voting method. Gray = DRE, white = other method (bubble, lever, or punchcard). Note that several error rates are so low that they do not appear on the figure.

An ANOVA performed on the difference scores between the error rates of the DRE and the other method by information condition and other method revealed no effect of which other voting method was used or in which information condition participants were. The interaction of other method and information condition was not significant.

The other way to consider errors is by ballot: each ballot contains zero errors or at least one error. In this first study, 9% of ballots contained at least one error. A chi-squared test reveals no difference between the by-ballot errors for the four voting methods. Table 6 shows the number of ballots containing zero or at least one error for each voting method.

	Errors		
	Zero	At least 1	Total
DRE	45	3	48
Bubble ballot	13	3	16
Lever machine	14	2	16
Punch card	15	1	16
Total	87	9	96

Table 6. Number of ballots containing zero or at least 1 error by voting method.

Satisfaction

Differences between the DRE and the other 3 methods of voting are seen in participants' levels of satisfaction. Satisfaction ratings for the four voting methods can be seen in Figure 13. The DRE consistently produced the highest levels of satisfaction with scores averaging 90.8 (SD = 8.2) on a scale from 0 to 100. The bubble ballot with an average of 82.2 (SD = 13.5) produced respectable satisfaction scores as well, as was expected from findings in earlier work (Everett et al., 2006; Greene et al., 2006; and Byrne et al., 2007).

Paired t-tests revealed significant differences between the DRE and each of the other three voting methods on this measure ($t(15) = 3.49$, $p = .003$ for the DRE-bubble comparison; $t(15) = 4.81$, $p < .001$ for the DRE-lever comparison; and $t(15) = 5.41$, $p < .001$ for the DRE-punch card comparison).

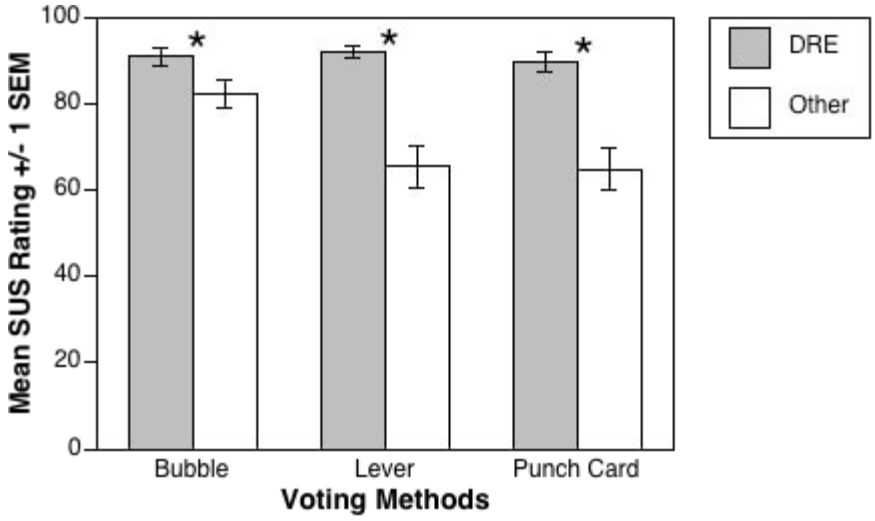


Figure 13. Pairs of mean SUS ratings by voting method. Gray = DRE, white = other method (bubble, lever, or punchcard). * $p < .05$.

An ANOVA was performed on the SUS score differences between the DRE and other method the participants used, including the variables of information condition and which other method was used. The ANOVA revealed a reliable effect of which other method the participant used ($F(2, 42) = 5.00$, $p = .01$). A post-hoc test was performed to examine the exact nature of this effect. This test showed that the difference between the DRE and the bubble was different from both the DRE and lever and the DRE and punch, which were not reliably different from each other. This meant that while voters were more satisfied with their experience on the DRE than on the bubble ballot, the difference between these two methods was smaller than the difference between the DRE and the lever machine or punch card.

Discussion

Study 1 did not find differences between the DRE with the review screen and the one without a review screen in any of the three usability measures used. It is possible that ceiling effects played a role in effectiveness. Participants were performing at such high

levels and producing hardly any errors, that there really was no room for effectiveness increases due to the review screen presence. While there were not any improvements in errors rates due to the review screen's presence, there also were not any changes in satisfaction or efficiency. This study shows that a potentially helpful tool, a chance to review choices before finally casting a ballot, can be added to voting machines without additional time cost. However, this could be because voters are not using this opportunity to review their choices. The data from this study cannot distinguish between these two possibilities.

Comparing performance on the DRE versus the three other voting methods did not reveal differences in time between the voting methods, similar to previous studies on older technologies (Everett et al., 2006; Greene et al., 2006; and Byrne et al., 2007). Also, there were no reliable differences between error rates. However, SUS scores revealed that the DRE was strongly preferred, as indicated by significant differences in satisfaction between the DRE and each other method. The result that the computerized method of voting is preferred is not surprising as the participants were all very familiar with technology. This finding needs to be replicated with a broader sample of users, but does begin to point towards the willingness of voters to use new, electronic forms of voting.

It is important to note that the participants in Study 1 were all bright, motivated college students who are very familiar with technology. While they may not be representative of the larger voting population, they probably serve a best-case scenario for what voting performance could be. Voters in general may take longer to cast ballots and will probably make more errors. Thus, the subject pool was broadened in Studies 2 and 3.

CHAPTER 6

STUDY 2

Overview

Study 2 was similar to the first study except that the main manipulation was the addition or removal of contests to/from the review screen. This study examined whether participants noticed if contests (with candidate/proposition position selections) were displayed on the review screen that the participant had not previously seen, or if contests were not displayed on the review screen for which participant had had the opportunity to vote. This is one method by which a voter's ballot potentially could be changed after he/she has made choices. This study examined whether voters even notice the problem.

This study also began to identify the individual characteristics of voters who notice the changes to their review screen. If it is possible to identify which groups of people do not notice changes made to their choices on the review screen, hackers could target vote-flipping schemes to impact these people only. Characteristics examined were age, education, previous voting experience on DREs, computer expertise, and whether participants have been following the news, as well as the personality characteristics of conscientiousness and openness to experience. This study also sought to identify in what situations people noticed the changes and if performance differences were present in these people. Again, if these variables predict who will notice changes, hackers could target attacks such that they only will be activated in certain situations or when voters show certain performance characteristics.

Methods

The methods for the second study will be the same as Study 1 except where noted below.

Participants

Seventy participants were recruited from the general Houston population to increase the diversity of the sample. Four were removed from analysis due to self-reported vision problems. All remaining participants had normal or correct-to-normal vision.

Participants ranged in age from 18-76, with an average age of 43.1 years (*SD* = 15.5) and a median age of 45. Participants had an even distribution of educational levels ranging from having a high school degree or less to postgraduate degrees (such as MA, PhD, JD, etc.). See Table 7 for a breakdown of education levels. The participant population was also fairly diverse in terms of race and income (see Tables 8 and 9). Democrats were the most strongly represented in the sample, as can be seen in Table 10. Fifty-eight participants (88%) had previous voting experience and this was with a variety of voting methods, as can be seen in Table 11.

Education level	Frequency
High school degree or less	16
Some college	16
Bachelor's degree or equivalent	16
Postgraduate degree	18
Total	66

Table 7. Frequency of each education level in Study 2.

Race	Frequency
African American	24
American Indian	1
Asian American	5
Caucasian	27
Mexican American or Chicano	4
Other Hispanic or Latino	0
Multiracial	1
Other/No report	4
Total	66

Table 8. Frequency of each race in Study 2.

Income	Frequency
< $20,000	27
$20,000 - $40,000	15
$40,000 - $60,000	11
$60,000 – $80,000	4
> $80,000	9
Total	66

Table 9. Frequency of each income level in Study 2.

Political affiliation	Frequency
Republican	19
Democrat	28
Independent	13
Libertarian	0
Other/No report	6
Total	66

Table 10. Frequency of each political affiliation in Study 2.

Voting method	Frequency
Bubble ballot	33
Arrow ballot	3
Open response ballot	1
Lever machine	20
Punch card	28
Electronic - touchscreen	22
Electronic - other	15

Table 11. Frequency of having previous experience with each voting method in Study 2.

Participants were fairly comfortable with computers, rating themselves on average at a 5.7 for computer expertise on a 10-point scale (*SD* = 2.5). All participants were paid $25 for their participation in the one-hour study.

Design

As in Study 1, participants voted on the DRE and one additional voting method and the order of presentation was counterbalanced. This study also had the between-

subjects manipulation of information condition (participants were given either a slate in the directed condition or offered a voter guide in the undirected condition).

Unlike Study 1, all participants saw a review screen after making their selections. In this study, however, there were additions to or removals from the review screen of entire contests, a between-subjects manipulation. The change variables were the type of change (either removals or additions) and the number of changes (2 or 8 contests). Thus for each participant, the review screen did one of the following: accurately reflected their choice (the control condition), removed 2 races, removed 8 races, added 2 races, or added 8 races. The locations of these additions/removals were randomized to control for ballot position or serial position effects.

To look at who noticed the additions/removals and under what conditions, several types of variables were considered. Seven *individual characteristic variables* were examined to look at who did and did not notice the changes: age, highest education level achieved (high school or less, some college, college degree, or postgraduate degree), self-rated computer expertise, previous electronic voting experience, whether the participant had been following the news on potential DRE security problems, and conscientiousness and openness to experience (as measured by the Big 5 assessment). To look at under what conditions participants are noticing the changes, three *situational variables* were examined: how many contests were changed on the review screen, whether these changes were additions to or removals from the review screen, and whether participants were in the undirected or directed conditions. Finally, to look at what types of performance led participants to notice the changes, four *performance variables* were examined: total time

taken to vote, total time spent on the review screen, time spent on the review screen the first time it was viewed, and number of times the review screen was viewed.

Materials

The same materials were used as in Study 1 with a few exceptions on the survey. Each SUS scale measuring satisfaction was given to participants immediately after they voted on the relevant method, instead of participants completing both of the SUS scales at the end as part of the survey packet. The survey also included more questions about whether participants noticed the additional/missing races, etc. An example of a survey used in Study 2 (not including the SUS or Big 5 assessment) can be seen in Appendix E.

As in the Study 1, the instructions on the DRE were standard voting instructions. They did not give instructions on what to do if the review screen was inaccurate or in any way refer to the fact that there may be changes made into the review screen. On the actual review screen itself, there was the typical instruction to click on a candidate's name to select a different candidate for that race, but it did not mention the potential for experimenter-inserted changes.

The DRE did not allow voters to correct the experimenter-inserted changes. This meant that if the review screen was missing 2 contests, the participant could look for them or report a problem, but the contests were not recoverable. It was not felt necessary to include the ability to correct these changes, as an actual hacker probably would not provide such an opportunity. Eight additional offices with fictitious candidate names and three new propositions were created for use in the review screen changes. The new offices and propositions are included in Appendix F.

Procedure

The procedure was the same as in the previous study, with the exception of the SUS survey mentioned above. Also, if participants reported the change to the experimenter, they were told to go ahead and cast their vote as it was displayed, even if there was a mistake on it. They were told that there was a place on the survey packet where they could report anything "funny" that happened while voting.

Results

DRE Review Screen Changes

The main manipulation of Study 2 was the addition and removal of entire contests from the review screen. If users are not noticing these changes to their ballots, hackers could steal votes in this manner with less of risk of detection. If some voters are noticing these manipulations, it is important to figure out who is noticing and under what circumstances.

Responses to questions in the survey packet indicated participants liked the review screen and 97% felt having a review screen was useful. Eighty-five percent of participants reported that having a review screen made them feel confident that their vote would be counted correctly. A further question was included in the survey packet for all participants except those in the control condition. This question informed participants that changes had been made to their selections and asked if they had noticed these changes. Responses to this questions were used as the dependent variable, notice-change. Because the survey packets for the participants in the control condition did not include this questions, conclusive false alarm rates cannot be established, although no control-

condition participants verbally reported any problems with their review screens to the experimenters.

After removing the participants who were in the control condition (no change to their review screens), the data from 53 participants remained. Of these 53 voters, only 17 people (32%) noticed the manipulation to their review screen. For the other 68% of voters, the computer program had changed their votes and the voters did not notice. This meant that voted contests had been added or removed from participants' ballots without detection.

As this has major security implications, it is important to discover who is and who is not noticing these changes. To look at this, the individual characteristics, situational, and performance variables described above were considered. The correlation matrix of all the individual characteristic and performance variables can be seen in Table 12. Age, time spent on the review screen during it first view, and total time spent on the review screen all correlate with at least three other variables.

	Age	Education	Conscientiousness	Openness to experience	Computer expertise	Prev. elec. voting experience	Following news	Total time	First review screen time	Total review screen time	# review screen visits
Age	-	.22	.24	-.12	-.43**	.31*	.25	.20	.47**	.45**	-.15
Education	.22	-	.19	.24	.22	.41**	.10	-.11	.25	.29*	.04
Conscientiousness	.24	.19	-	.15	.01	.26	.03	-.07	.22	.26	.12
Openness to experience	-.12	.24	.15	-	.30*	.21	.14	-.01	.19	.22	-.11
Computer expertise	-.43**	.22	.01	.30*	-	.22	.20	-.38	-.17	-.13	.11
Prev. elec. voting experience	.31*	.41**	.26	.21	.22	-	.24	-.08	.26	.27	-.06
Following news	.25	.10	.03	.14	.20	.24	-	.05	-.03	-.02	.02
Total time	.20	-.11	-.07	-.01	-.38	-.08	.05	-	.51**	.49**	-.06
First review screen time	.47**	.25	.22	.19	-.17	.26	-.03	.51**	-	.98**	-.02
Total review screen time	.45**	.29*	.26	.22	-.13	.27	-.02	.49**	.98**	-	.14
# review screen visits	-.15	.04	.12	-.11	.11	-.06	.02	-.06	-.02	.14	-

Table 12. The correlation matrix of individual characteristic and performance variables. N = 50. * $p < .05$. ** $p < .01$.

As the dependent variable, notice-change, is dichotomous, logistic regression models were run to test the predictive power of the variables of interest. Regressions were run separately for the three types of variables. The measure of goodness of fit used

here was the Nagelkerke R^2, an index for which the maximum attainable value is 1.00 (Cohen, Cohen, West, & Aiken, 2003).

The number of participants who noticed the changes can be seen in Table 13 by the individual characteristics. When the notice-change variable was regressed on individual characteristics variables, only the personality trait openness to experience significantly predicted whether participants would notice the change ($b = 0.16$, $p = .04$). The individual characteristics variables accounted for almost 40% of the variance, with an R^2 or uncertainty coefficient of .39 in the significant model ($X^2(7, N = 49) = 16.28$, $p = .02$). Table 14 displays the results of this regression.

	% noticed	# noticed	# did not notice
Age – 18 to 27+	23%	3	10
Age – 28 to 44	22%	2	7
Age – 45 to 62	35%	9	17
Age – 63 and older	75%	3	1
Education – High school or less	9%	1	10
Education – Some college	21%	3	11
Education – Bachelor's degree	36%	5	9
Education – Postgraduate degree	57%	8	6
Openness to experience - Low	0%	0	5
Openness to experience - Medium	25%	5	15
Openness to experience - High	46%	12	14
Following the news – Not at all	33%	4	8
Following the news – Somewhat	34%	11	21
Following the news - Carefully	25%	2	6
Previous electronic voting experience? – No	24%	7	22
Previous electronic voting experience? - Yes	42%	10	14
Computer expertise – Low	18%	2	9
Computer expertise – Medium	38%	11	18
Computer expertise - High	31%	4	9

Table 13. Frequency of noticing the changes by individual characteristics.

Variable	*b*	*SE*	Wald	*df*	*p*
Age	0.05	0.03	2.07	1	.150
Education	0.54	0.44	1.52	1	.217
Conscientiousness	0.06	0.05	1.20	1	.273
Openness to experience	0.16	0.08	4.34	1	.037*
Following news	-0.33	0.86	0.15	1	.702
Prev. elec. voting experience	-0.81	0.69	1.38	1	.240
Computer expertise	0.07	0.19	0.14	1	.709
Constant	-8.58	2.71	10.02	1	.002**

	R^2	X^2	*df*	*p*
Model	.390	16.28	7	.023*

Table 14. Logistic regression analysis of individual characteristic variables on whether participants noticed the change. N = 49. *$p < .05$. **$p < .01$.

Next, the effect of the situational variables was evaluated. The number of participants in each condition who noticed the changes can be seen in Table 15. The notice-change variable was regressed on the situational variables (see Table 16 for results). The directed information condition significantly predicted whether participants would notice the change ($b = -2.14$, $p = .003$), but the number and type of changes did not. This model with the situational models predicted less of the variance ($R^2 = .26$) than the individual characteristics did, but the model was still significant ($X^2(3, N = 53) = 10.73$, $p = .01$).

	% noticed	# noticed	# did not notice
Information condition - Undirected	12%	3	23
Information condition - Directed	52%	14	13
Type of changes - Addition	31%	8	18
Type of changes - Removal	33%	9	18
Number of changes - Two	32%	8	17
Number of changes - Eight	32%	9	19

Table 15. Frequency of noticing the changes by situational variable conditions. N = 53.

Variable	*b*	*SE*	Wald	*df*	*p*
Information cond. (1)	2.14	0.73	8.58	1	.003**
# of changes	0.11	0.66	0.03	1	.863
Type of changes (1)	-0.27	0.66	0.17	1	.677
Constant	0.28	0.63	0.20	1	.657
	R^2		X^2	*df*	*p*
Model	.256		10.73	3	.013*

Table 16. Logistic regression analysis of situational variables on whether participants noticed the change. For information condition, undirected served as the reference variable in the dummy coding, with information cond. (1) indicating directed condition. For the type of changes, removing race(s) served as the reference variable and type of changes (1) indicated the insert condition. N = 53. *$p < .05$. **$p < .01$.

Finally, the notice-change variable was regressed on the performance variables. Results of this regression indicated that first review screen time was acting as a suppressor variable. It appeared that this variable and total time spent on the review

screen were so highly related that the regression results were being affected. Thus, the regression was run again without this variable and results can be seen in Table 17. The total time participants spent on the review screen had a significant effect on whether participants noticed the changes (b = .083, p = .005). It accounted for 79% of the variance in this highly significant model ($X^2(3, N = 51) = 42.235$, $p < .001$). Histograms of the total time spent on the review screen can be seen in Figure 14 for participants who noticed the changes and those that did not. Most of the voters who did not notice the changes to their review screens spent very little time on the review screen, while voters who did notice display more of a normal distribution of review screen times.

Variable	*b*	*SE*	Wald	*df*	*p*
Total time	-.01	.01	2.58	1	.108
Total review screen time	.08	.03	7.88	1	.005**
# review screen visits	-1.59	1.32	1.46	1	.228
Constant					
	R^2		X^2	*df*	*p*
Model	.791		42.24	3	<.001**

Table 17. Logistic regression analysis of performance variables on whether participants noticed the change. N = 51. **$p < .01$.

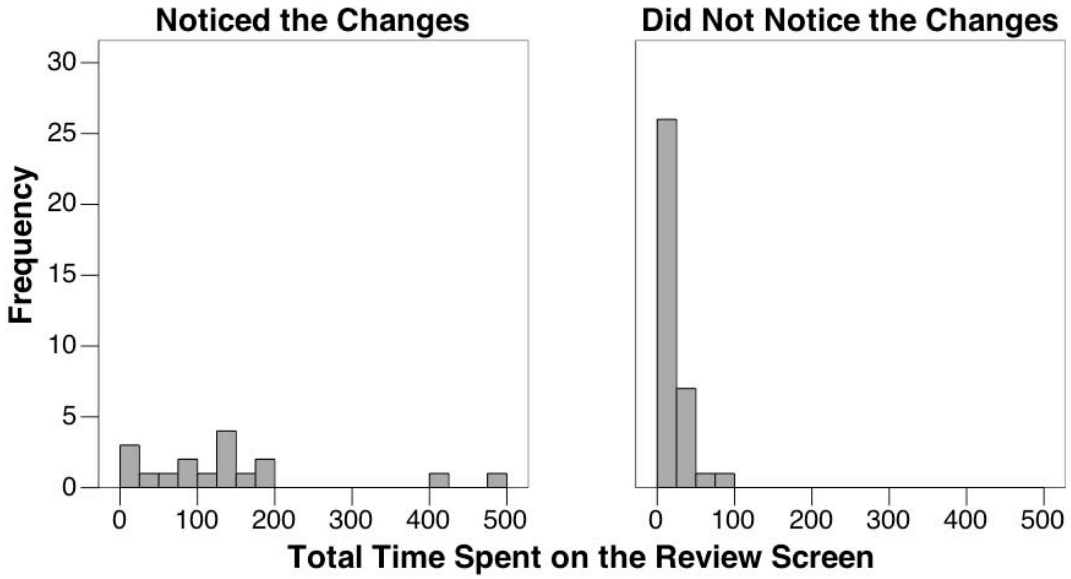


Figure 14. Histograms of the total time spent on the review screen for participants who noticed and participants who did not notice the changes.

As education, conscientiousness, computer expertise, previous electronic voting experience, following the news, and number of review screen visits were not significantly correlated with any other variables and did not contribute any predictive power to the regression, they were removed from further analysis. The time participants spent on the review screen the first time it was seen was also not included in further analysis as it had been removed from the regression analyses.

Using the results from the correlation matrix and regression models described above, four variables were determined to be the most important in predicting who would notice the changes. These were age, openness to experience, total time on the review screen, and information condition. To determine the relative sizes of the direct and indirect effects of these variables on the notice-change variable, a structural equation model was constructed that contained causal paths from each of these to the notice-

change variable. Total time on review screen had a direct path to the notice-change variable. Information condition, age, and openness to experience were hypothesized to have indirect effects on the notice-change variable (through total time on review screen) as well as direct effects. The model was analyzed using AMOS 7.0 and standardized path coefficients were computed, as can be seen in Figure 15. The model's fit to the data was evaluated by examining the results from three goodness-of-fit measures: chi-square, comparative fit index (CFI), and root-mean-square error of approximation (RMSEA). From these three measures ($X^2(2, N = 49) = 0.92$, $p = .63$; CFI = 1.00; and RMSEA = .00), the model was determined to be a good fit.

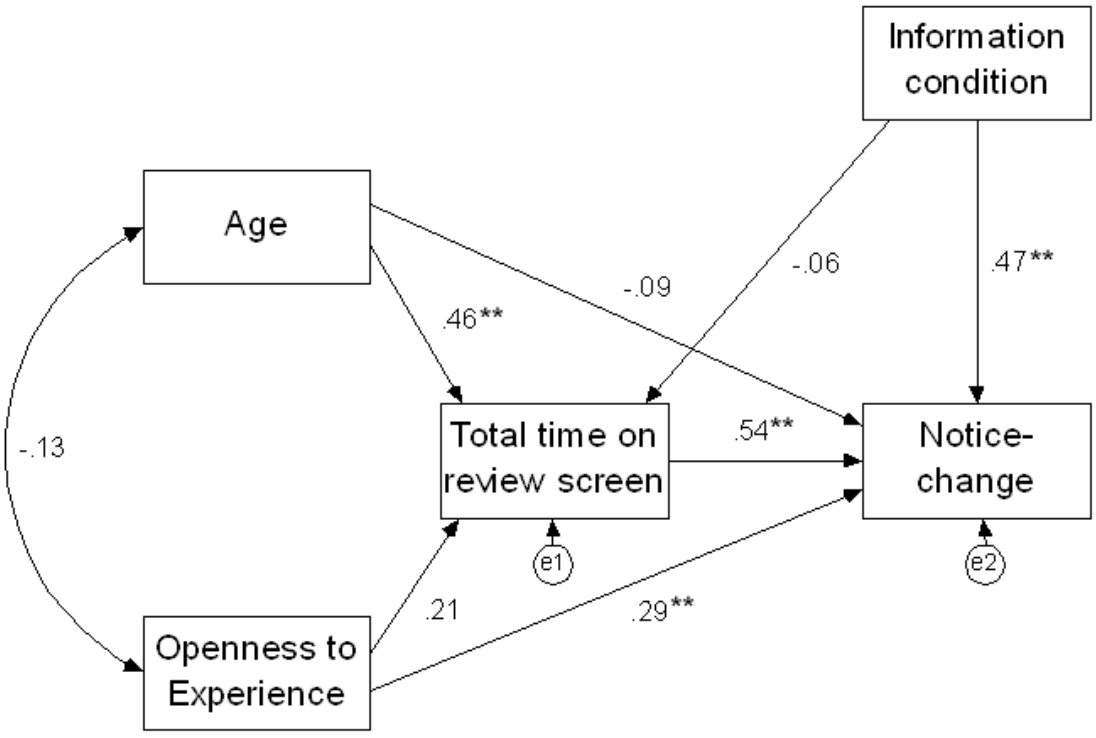


Figure 15. Path diagram and standardized parameter estimates for the Study 2 model. **$p < .01$.

As expected, the path from total time on review screen to the notice-change variable was highly significant (ß = .54, $p < .001$), indicating that people who spent more

time on the review screen were more likely to notice the changes. Age had an indirect effect on the notice-change variable through total time on review screen (ß = .46, $p <$ .001). Older voters spent more time on the review screen and people who spent longer on this screen were more likely to notice the changes. Openness to experience had a significant direct effect on the notice-change variable (ß = .29, $p = .002$); people scoring higher in openness to experience were more likely to notice the changes. Finally, information condition had a large direct effect on who noticed the changes (ß = .47, $p <$ .001). People in the directed condition who were given a slate were much more likely to notice the changes made on their review screens.

DREs Versus Older Voting Methods

For all results that report on differences between the DRE and other methods, results are only used from participants who were either in the control condition where no changes were made to their review screen or who did not notice the changes to their review screen. This was done to ensure cleaner comparisons with the older voting methods. Outliers were identified and replaced as described in Study 1. This resulted in one replacement of a ballot completion time for the lever machine, two replacements of error rates on the DRE, and one replacement of an error rate on the punch card. There were no outliers identified in the SUS scores. Two participants were removed from error analysis because their performance resulted in outliers on both of the methods on which they voted. An additional three participants were removed from the error analysis because their candidate and propositions selections could not be read from the data files.

Paired *t*-tests were performed to look at participant's performance on the DRE and the other methods to which they were assigned. The DRE was paired with each of the

lever machine, bubble ballot, and punch card ballot individually. Also, difference scores were computed for each usability measure between the DRE and the other methods. ANOVAs were conducted on these difference scores to look for effects of information condition or type of other method on difference scores.

MANOVAs were also performed to look at the effects of age, education, which other voting method was used, and information condition on voting behavior. Both age and education were treated as continuous variables in these analyses. The repeated measure was the participant's performance on the DRE and on the other method they used.

Time

The overall average ballot completion time for Study 2 was 377 seconds (*SD* = 257). Completion times by paired voting methods can be seen in Figure 16. Each pair represents scores from one group of participants. For example, in the two columns labeled "Bubble," the white column represents the ballot completion times on the bubble ballot, and the gray column represents DRE times of those same people. Paired *t*-tests revealed a significant difference between ballot completion times on the DRE and the lever machine ($t(17) = 2.87$, $p = .01$), but not between times on the DRE and bubble ballot, or the DRE and the punch card.

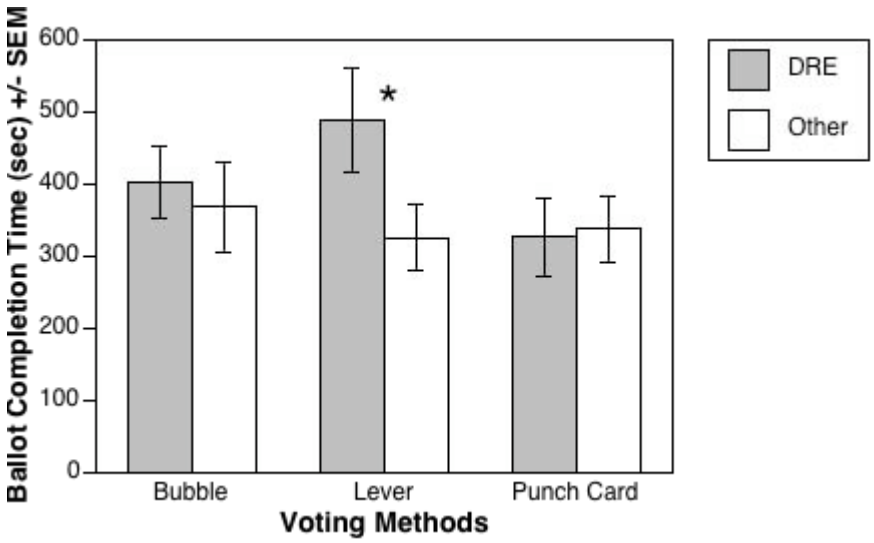


Figure 16. Pairs of mean ballot completion times (in sec) by voting method. Gray = DRE, white = other method (bubble, lever, or punchcard). * p < .05.

An ANOVA was performed on the difference between the ballot completion times of the DRE and other method that the participants used by information condition and other method. This did not reveal main effects of information condition or which other method was used. However, the interaction of these two variables on the ballot completion difference scores was significant ($F(2, 43) = 4.76$, $p = .01$). In the undirected condition, voters were slower on the DRE than the lever machine, but about equal on the DRE and the bubble ballot and punch card. However, in the directed condition, voters took longer on the bubble ballot and punch card than the DRE, but about the same on the DRE and lever machine.

A MANOVA performed on ballot completion time on the DRE and ballot completion time on the other method on which participants voted by age, education, information condition, and other method revealed a main effect of education on ballot completion time $F(1, 42) = 4.54$, $p = .04$, indicating that participants with more education

voted more quickly (see Figure 17 for ballot completion times by education level). There were no other significant main effects or interactions of information condition, which other method was used, age, or education.

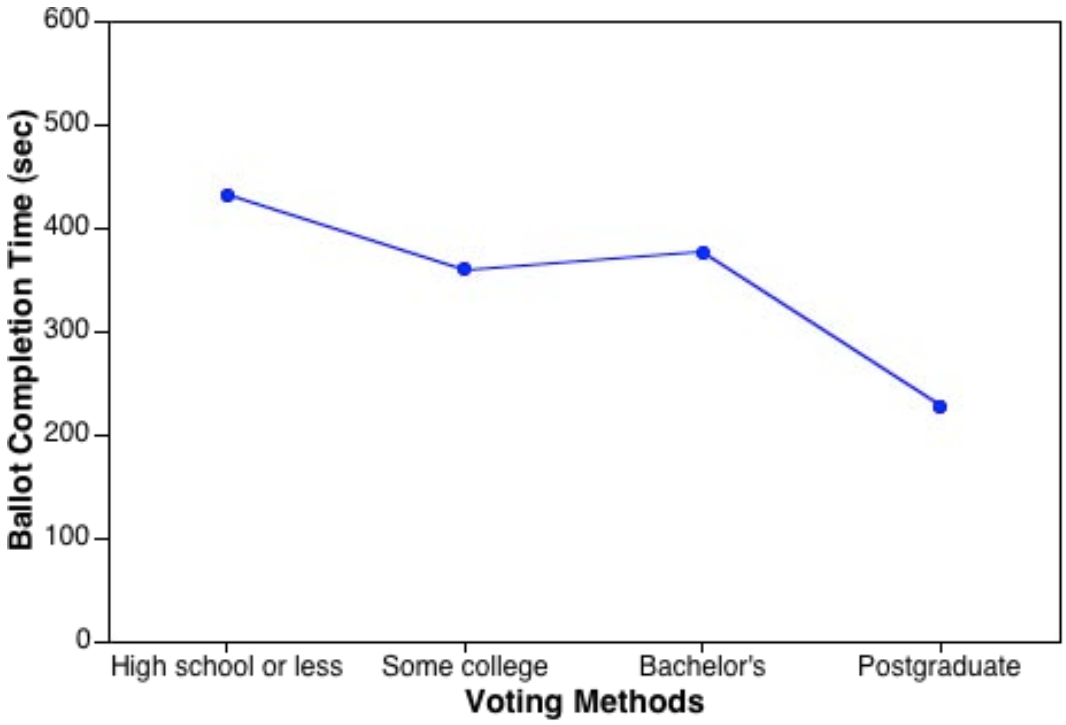


Figure 17. Ballot completion times by education levels.

In general, ballot completion times for each voting method were unrelated to their previous voting experience with the method. However, there was a negative relationship between ballot completion time on the DRE and self-rated computer expertise (r = -0.29, p = .04); times decreased as expertise increased. This relationship can be seen in Figure 18.

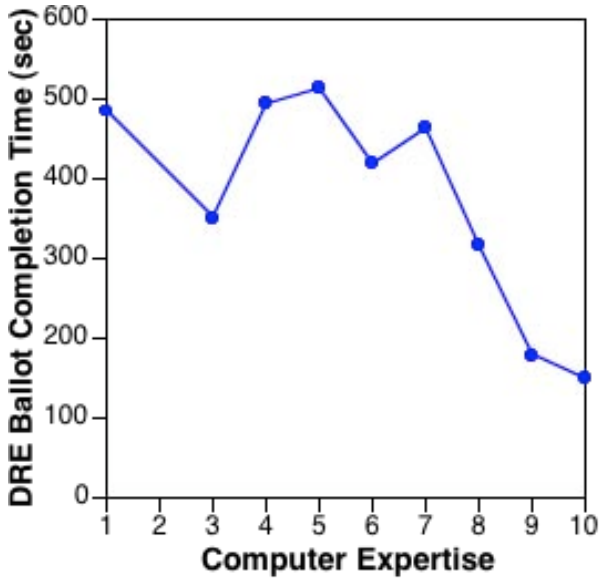


Figure 18. Average ballot completion times on the DRE by computer expertise. 1 = novice, 10 = expert.

Errors

Per-ballot error rates between the DRE and each other method were compared, as can be seen in Figure 19. Participants made more errors in this study than in the first one, with an average error rate of 0.016 (*SD* = 0.040) in the current study. Paired *t*-tests did not show significant differences in error rates between the DRE and any of the other three methods.

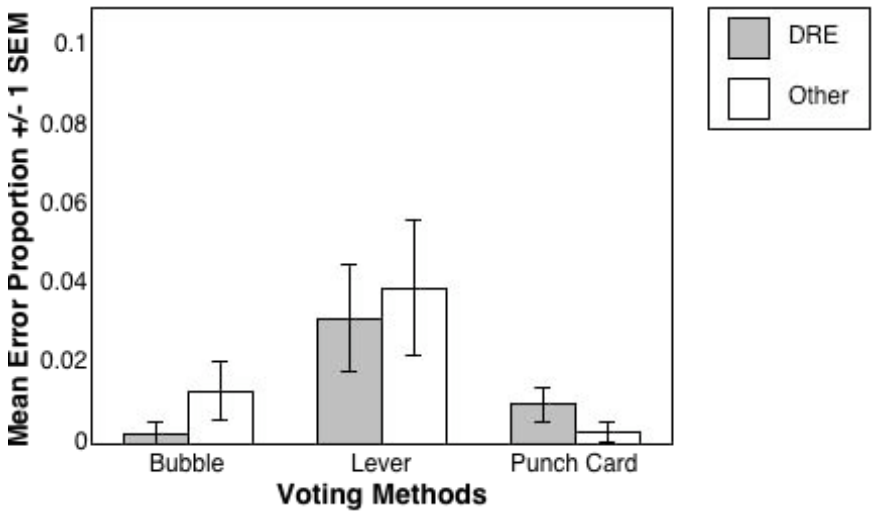


Figure 19. Pairs of mean error rates by voting method. Gray = DRE, white = other method (bubble, lever, or punch card).

An ANOVA was performed on the difference scores between the error rate on the DRE and other method on which the participants voted by information condition and other method. This revealed no effect of information condition, which other method was used, or their interaction.

A MANOVA including age, education, information condition, and other method was performed on the error rates of the DRE and other methods. There were no significant main effects or interactions of information condition, which other method the participant used, age, or education.

Error rates for each voting method were unrelated to their previous voting experience with the method. Error rates on the DRE were also unrelated to computer expertise.

Ballots were also examined to see whether different voting methods were more likely to produce ballots with at least one error. Errors by ballot can be seen in Table 18.

Over 27% percent of ballots contained at least one error. A chi-square test showed no differences between voting methods as to whether each ballot contained at least one error.

	Errors		
	Zero	At least 1	Total
DRE	34	13	47
Bubble ballot	11	3	14
Lever machine	11	7	18
Punch card	13	2	15
Total	69	25	94

Table 18. Number of ballots containing zero or at least 1 error by voting method.

Satisfaction

Satisfaction levels of the participants for the four voting methods can be seen in Figure 20. The average SUS score for the DRE was 86.1 (SD = 16.6), the average SUS for the bubble ballot was 81.3 (SD = 22.2), the average for the lever machine was 71.5 (SD = 14.8), and for the punch card ballot was 69.0 (SD = 22.2). Paired t-tests revealed reliable differences between the DRE and the bubble ($t(15) = 2.24$, $p = .04$), and the DRE and the lever machine ($t(17) = 2.37$, $p = .03$). The DRE and punch card pair approached significance ($t(14) = 1.96$, $p = .07$).

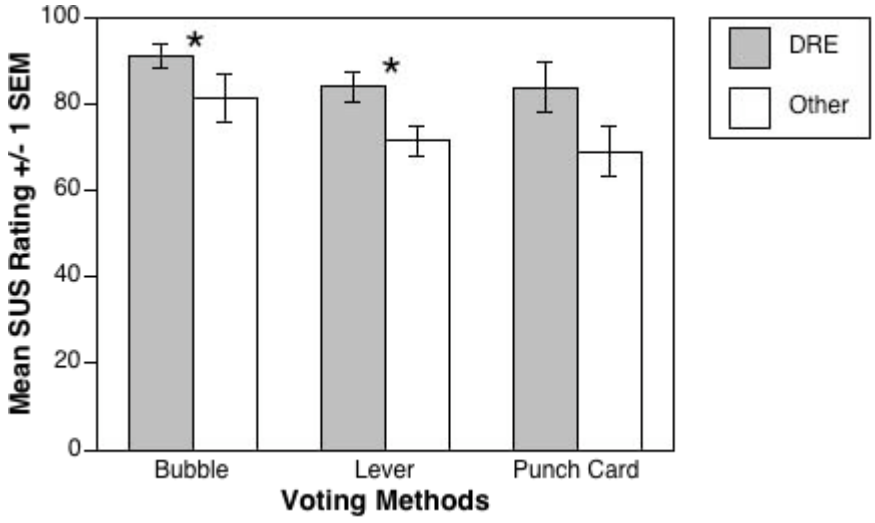


Figure 20. Pairs of mean SUS ratings by voting method. Gray = DRE, white = other method (bubble, lever, or punchcard). * $p < .05$.

An ANOVA was performed on the satisfaction difference scores between the DRE and the other method that the participant used by information condition and other method. According to this ANOVA, there was no effect on the difference scores between the DRE and other method of information condition, which other method the participants used, or their interaction.

A MANOVA performed on SUS scores on the DRE and other method by age, education, information condition, and which other method was used showed no significant main effects or interactions. The difference between the SUS scores on the DRE and other method approached significance ($F(1, 42) = 3.65$, $p = .06$), suggesting that voters might be more satisfied with experiences on the DRE than previous methods.

Satisfaction levels for each voting method were unrelated to their previous voting experience with the method. Satisfaction levels on the DRE were also unrelated to computer expertise.

Discussion

The most interesting result from this study is that only 32% of voters noticed that their malicious changes had been made to their review screens. After the voters made their selections, the computer changed their votes and this mostly was not detected. Voters simply breezed past the review screen and submitted the corrupted ballots. This was a robust effect and it did not matter how many contests were changed or whether the computer deleted races or added them. Eight of 27 races, or 30% of the ballot, could be changed mostly without detection and it is likely this effect would hold with even higher percentages of the ballot. This means that once a hacker installed malicious software on a voting machine, they could significantly impact the results of a large number of contests.

Beyond showing that the majority of voters did not detect the changes made to their ballots, this study identified several characteristics of the voter and the voting situation that affect the likelihood that the voter will notice the changes. Participants who spent longer on the review screen were more likely to notice the changes. It is important to note that the direction of causality cannot be determined from the current experimental design. People who notice the changes to their review screen spent much longer on the review screen than those who did not notice. However, this could be because they were carefully checking the information shown on the review screen which took awhile, or it could be the case that once voters noticed the changes, they spent longer on the review screen trying to figure out the problem. While the relationship between total time on review screen and the notice-change variable may not be surprising, it is certainly important for security reasons. Many DREs have multiple pages of review screens and a voting machine running malicious software could be programmed to record how much

time a voter spent on the first review screen page. For the additional review screen pages, the voting machine could then change the choices of only those voters who spent very little time on the first page.

In terms of individual characteristics, voters who scored high in openness to experience were more likely to notice the changes. While personality measures are obviously not given before allowing people to vote, the finding that a personality trait is a significant predictor of a specific behavior is important to personality research. Although previous research in this domain has found that certain traits such as conscientiousness can predict broad performance outcomes such as job performance (Barrick & Mount, 1991), the current study shows that openness to experience has significant predictive power for one very specific act, whether voters will carefully check their review screens and detect errors.

Through its effect on review screen times, the age of voters was also a significant predictor of whether they would notice the changes. Older people tended to spend more time on the review screen, which meant they were more likely to notice the changes. However, the effect of age on noticing the changes was only indirect, mediated through time spent on the review screen. More work with a larger sample is needed to explore whether there really is a direct effect of age on noticing.

The final variable that significantly predicted whether voters would notice the changes was information condition. Voters in the directed condition who had been given a paper with the names of candidates to choose were more likely to notice the changes. This effect makes sense as participants in this directed condition could easily make direct comparisons between the paper they held and the review screen in front of them. This

finding is important because when DREs undergo certification and testing, participants will most likely be directed who to vote for. This study shows that those testing processes will significantly overestimate the number of voters who will use the review screen to check the accuracy of their ballots and notice any changes.

When the DRE was compared against older voting methods, the biggest differences were seen in the satisfaction participants felt with each method. The DRE was much preferred to the bubble ballot, lever machine, and punch card ballot, regardless of any individual characteristics such as age or education. Because of these high satisfaction ratings of the DRE, it is likely that voters will support further adoption of these electronic voting machines and will not be intimidated from voting by the change in technology.

As the results indicated, voters with less computer expertise took more time voting on the DRE, but were not any less satisfied with the experience. Voters with more education also took less time voting, an effect has been previously reported by Byrne et al. (2007). Results such as these can help inform machine allocation decisions. Precincts in which voters are more highly educated and or which have more industry requiring computer skills need fewer voting machines since their voters will be able to cast their ballots more quickly. As shown in Ohio in 2004, improper allocation of voting equipment can lead to long lines at the polls, which can even affect the outcome of elections (Mebane, 2006).

Although the error results do not point to any overall differences between voting methods, error rates were still much higher than would be desirable in elections. While the average per race error rate in the current study is lower than has been found in previous studies, over 27% of ballots contained at least one error. This is a huge amount

of error in almost all races, especially some of the extremely close races that have been seen in recent elections.

CHAPTER 7

STUDY 3

Overview

The purpose of the third study was again to examine what percentage of voters noticed changes that were inserted into the review screen. In Study 3, vote flipping occurred on the review screen, instead of the addition/removal of entire contests as in the second study. The experimenter-induced vote flipping caused candidates to be listed on the review screen that were different from the choices the participant actually made. For example, such a change would be if the review screen displayed "Gordon Bearce" as the participant's choice for President, but the participant actually chose "Janette Froman." This is another potential method by which a voter's ballot could be changed after he/she has made choices.

As in Study 2, this study examined whether voters notice this problem, and if any individual characteristics could be identified that would predict who noticed the changes. The impact of situational and performance variables on whether voters noticed the changes was also evaluated. Study 3 used a much larger sample than the previous study to gain more information on who is and who is not noticing the review screen changes.

Methods

The methods were exactly the same as in Study 2 except where otherwise noted below.

Participants

There were 108 participants in this study. Participants were recruited through advertisements in local newspapers. Seven participants were removed from analysis due

to self-reported vision problems. All the remaining participants had normal or corrected-to-normal vision.

There was an even gender split in the study with 51 male and 50 female participants. The age range of participants was 20-75 with a mean of 40.8 (*SD* = 14.8) and a median of 40. Participants in this study were diverse in terms of the highest level of education achieved, with most having had some college or holding a college degree. (See Table 19 for a breakdown of education levels.) Participants represented a variety of races and income levels (see Tables 20 and 21), although a greater proportion of participants fell into the lower income categories.

Education Level	Frequency
High school degree or less	18
Some college	41
Bachelor's degree or equivalent	31
Postgraduate degree	11
Total	101

Table 19. Frequency of each education level in Study 3.

Race	Frequency
African American	36
American Indian	0
Asian American	5
Caucasian	46
Mexican American or Chicano	5
Other Hispanic or Latino	3
Multiracial	5
Other	1
Total	101

Table 20. Frequency of each race in Study 3.

Income	Frequency
< $20,000	35
$20,000 - $40,000	35
$40,000 - $60,000	16
$60,000 – $80,000	4
> $80,000	11
Total	101

Table 21. Frequency of each income level in Study 3.

Participants were diverse in terms of political party affiliation, with Democrats being the most strongly represented as in Study 2 (see Table 22). Ninety-three participants (92%) had previous voting experience. As is shown in Table 23, the voting methods with which the most participants had had previous voting experience were punch card ballots and DREs. This was expected as these two methods had been used in recent elections in the local area.

Political affiliation	Frequency
Republican	23
Democrat	51
Libertarian	0
Independent	20
Other	7
Total	101

Table 22. Frequency of each political affiliation in Study 3.

Previous voting method	Frequency
Bubble ballot	39
Arrow ballot	3
Open response ballot	3
Lever machine	27
Punch card	54
DRE	59

Table 23. Frequency of having previous experience with each voting method in Study 3.

As in previous studies, participants were fairly comfortable with computers, with an average self-rating of expertise at 5.8 of 10 (SD = 2.4). All participants were paid $25 for their participation in the study.

Design

As in the previous two studies, participants voted on the DRE and one additional voting method, and the order was counterbalanced. This study also had the between-subjects manipulation of information condition (directed vs. undirected). However, in this study a third level of this variable was added: 1/3 of participants were given a paper

instructing for whom they should vote, but this slate directed them to skip some races and not select a candidate in these races. This was done to more accurately represent voting situations in which participants often do not vote for every office. The locations of skipped races on the slate were selected such that they represented real roll-off rates (i.e., the races further down the ballot, the more local races, were more likely to be skipped) (Nichols & Strizek, 1995). Thus, in Study 3, the information variable had 3 levels: directed, directed with roll-off, and undirected.

Unlike Study 2, this study did not manipulate the number of races displayed on the review screen. The manipulation of interest was the vote flipping that occurred on the review screen, and this was decided by 3 between-subjects variables: the number, location, and type of changes. The number of changes or flips was 1, 2, or 8. The location of these changes was either at the top or the bottom of the ballot, to allow for studying ballot position effects. Finally, the type of these changes was either that the program changed a voter's choice to a different candidate, or that the program changed a voter's choice to "None." A control condition was not run in this third study because both of the previous studies already ran this condition.

The design of the study was a 3 (other method) x 3 (information condition) x 3 (number of changes) x 3 (type of changes) x 2 (location of changes) resulting in 162 conditions. Because only 108 participants were run, the design was not fully crossed and higher order interactions were not examined. However, it is important to note that the first between-subjects variable, which other method was used in addition to the DRE, does not directly affect the DRE. This results in only 54 conditions for the notice changes analysis. Similarly, the three between-subjects variables pertaining to the review screen

changes do not affect the usability data gathered on the bubble ballot, lever machine, and punch card, resulting in only nine conditions for these analyses.

Materials

The same materials were used as in Study 2, with the addition of the reduced slates in the directed with roll-off condition.

Procedure

The procedure was the same as in the Study 2.

Results

DRE Review Screen Changes

The main manipulation of the third study was the vote flipping that occurred on the review screen. As in Study 2, participants were asked a series of questions in the survey packet about the review screen and its performance. They were finally informed that changes had in fact been made to their selections as shown on the review screen and were asked if they had noticed. One participant did not fully complete the survey packet and did not provide a response as to whether or not she noticed the change. Thus, one hundred participants were included in the analysis of noticing the change. Of these voters, only 37 people (37%) noticed the manipulation to their review screen, similar to the 32% found in Study 2. This means that over 60% of participants in the study did not detect that their votes had been changed.

All participants reported feeling that a review screen was useful and 95% reported that they had checked their ballot either somewhat or very carefully. Seventy percent of participants reported that having a review screen made them feel confident that their vote

would be counted correctly. Even with these positive reports of the review screen, most voters did not use it to verify that their selections were accurate.

To look at who noticed these changes, the same seven individual characteristic variables and four performance variables are examined as in Study 2. The situational variables examined here are similar to those in Study 2, but are updated to reflect the experimental design of Study 3. In this study, the situational variables examined were how many contests were changed on the review screen, whether these changes occurred at the top or bottom of the ballot, whether the change was to a different candidate or to "None", and whether participants were in the undirected, directed, or directed with roll-off conditions.

The correlation matrix of all the individual characteristic and performance variables can be seen in Table 24. Five variables are correlated with many other variables. Age, openness to experience, computer expertise, total time, and review screen time during its first view all correlate with at least four other variables.

	Age	Education	Conscientiousness	Openness to experience	Computer expertise	Prev. elec. voting experience	Following news	Total time	First review screen time	Total review screen time	# review screen visits
Age	-	.02	.17	-.31**	-.42	.27**	.29**	.44**	.20*	.05	-.05
Education	.02	-	.04	.29**	.35**	.13	.15	-.04	-.01	.11	.23*
Conscientiousness	.17	.04	-	.21*	.11	-.02	-.11	.03	-.07	.03	.00
Openness to experience	-.31**	.29**	.21*	-	.32**	-.08	.09	-.17	-.04	.15	.11
Computer expertise	-.42	.35**	.11	.32**	-	-.00	-.07	-.46**	-.33**	-.13	.03
Prev. elec. voting experience	.27**	.13	-.02	-.08	-.00	-	.11	.13	.17	.10	.10
Following news	.29**	.15	-.11	.09	-.07	.11	-	.16	.07	.15	.13
Total time	.44**	-.04	.03	-.17	-.46**	.13	.16	-	.59**	.48**	.16
First review screen time	.20*	-.01	-.07	-.04	-.33**	.17	.07	.59**	-	.65**	.18
Total review screen time	.05	.11	.03	.15	-.13	.10	.15	.48**	.65**	-	.66**
# review screen visits	-.05	.23*	.00	.11	.03	.10	.13	.16	.18	.66**	-

Table 24. The correlation matrix of individual characteristic and performance variables. N = 100. * p < .05, ** p < .01.

The number of participants who noticed the changes can be seen in Table 25 by the individual characteristics. As in Study 2, logistic regressions were run separately for the three types of variables. The notice-change variable was regressed first on individual characteristics variables, yielding a model ($X^2(7, N = 100) = 12.73, p = .08$) that

accounted for 17% of the variance. How carefully the participant had been following the news about potential security problems with DREs had a significant effect on noticing the change (b = 1.15, p = .012), and the effect from age approached significance (b = -0.04, p = .052). Results from this regression can be seen in Table 26.

	% noticed	# noticed	# did not notice
Age – 18 to 34	40%	16	24
Age – 35 to 49	37%	11	19
Age – 50 to 64	30%	7	16
Age – 65 and older	43%	3	4
Education – High school or less	33%	6	12
Education – Some college	29%	12	29
Education – Bachelor's degree	47%	14	16
Education – Postgraduate degree	45%	5	6
Openness to experience - Low	26%	5	14
Openness to experience - Medium	37%	22	38
Openness to experience - High	45%	9	11
Following the news – Not at all	15%	3	17
Following the news – Somewhat	39%	26	40
Following the news - Carefully	54%	7	6
Previous electronic voting experience? – No	31%	13	29
Previous electronic voting experience? - Yes	41%	24	34
Computer expertise – Low	39%	7	11
Computer expertise – Medium	30%	8	19
Computer expertise - High	40%	22	33

Table 25. Frequency of noticing the changes by individual characteristics.

Variable	*b*	*SE*	Wald	*df*	*p*
Age	-0.04	0.02	3.78	1	.052
Education	0.36	0.28	1.62	1	.203
Conscientiousness	0.04	0.04	1.57	1	.211
Openness to experience	0.004	0.04	0.01	1	.918
Following news	1.15	0.46	6.29	1	.012*
Prev. elec. voting experience	0.53	0.48	1.23	1	.267
Computer expertise	-0.15	0.12	1.72	1	.190
Constant	-1.09	1.48	0.54	1	.463
	R^2		X^2	*df*	*p*
Model	.165		12.73	7	.079

Table 26. Logistic regression analysis of individual characteristic variables on whether participants noticed the change. N = 100. *$p < .05$.

The number of participants who noticed the changes can be seen in Table 27 by levels of the situational variables. The notice-change variable was next regressed on these situational variables, as shown in Table 28. The resulting model ($X^2(5, N = 100) = 16.84$, $p = .005$) accounted for 21% of the variance in noticing the changes. Information condition, specifically the directed condition, had a significant effect on noticing the change ($b = 1.964$, $p = .001$). None of the other situational variables had a significant effect. It did not matter how many races were changed, what type of changes were made, or where on the ballot they were located; participants either checked the review screen or they did not.

	% noticed	# notice	# did not notice
Information condition – Undirected	18%	6	27
Information condition – Directed	60%	21	14
Information condition – Directed with roll-off	31%	10	22
Location of changes - Top	31%	16	36
Location of changes - Bottom	44%	21	27
Type of changes – To other candidate	35%	17	31
Type of changes – To "None"	38%	20	32
Number of changes - One	29%	10	25
Number of changes - Two	39%	13	20
Number of changes - Eight	44%	14	18

Table 27. Frequency of noticing the changes by levels of the situational variables. N = 100.

Variable	*b*	*SE*	Wald	*df*	*p*
Information cond.			12.40	2	.002**
Information cond. (1)	1.96	0.58	11.40	1	.001**
Information cond. (2)	0.75	0.60	1.57	1	.210
# of changes	0.08	0.07	1.12	1	.290
Location of changes (1)	-0.65	0.46	1.98	1	.160
Type of changes (1)	-0.09	0.46	0.04	1	.850
Constant	-0.70	0.55	1.65	1	.199
	R^2		X^2	*df*	*p*
Model	.212		16.84	5	.005**

Table 28. Logistic regression analysis of situational variables on whether participants noticed the change. For information condition, undirected served as the reference variable, with directed as Information Cond. (1) and directed-reduced as Information Cond. (2). For location of changes, change at the bottom of the ballot was the reference variable and change at the top of the ballot was Location of Changes (1). For type of changes, changes to "none" was the reference variable and changes to a different candidate was Type of Changes (1). N = 100. **$p < .01$.

Finally, the notice-change variable was regressed on the performance variables. The time participants spent on the review screen the first time they saw it was again removed from analysis because it was acting as a suppressor variable. The resulting regression table can be seen in Table 29. The total time the participant spent on the review screen had significant effects on whether participants noticed the change (b = 0.048, $p < .001$). This variable accounted for 70% of the variance in predicting whether participants noticed the review screen in the highly significant model ($X^2(3, N = 100) = 68.243$, $p < .001$). Histograms of the time spent on the review screen for participants who noticed the changes and those who did not notice the changes can be seen in Figure

21. As in Study 2, most participants who did not notice the review screen changes spent very little time on the review screen, while more of a normal distribution of review screen times is seen for those who did notice the changes.

Variable	*b*	*SE*	Wald	*df*	*p*
Total time	-0.002	0.002	0.77	1	.382
Total review screen time	0.05	0.01	12.98	1	<.001**
# review screen visits	0.42	0.30	1.96	1	.161
Constant	-3.29	0.85	14.95	1	<.001**
	R^2		X^2	*df*	*p*
Model	.700		68.24	3	<.001**

Table 29. Logistic regression analysis of performance variables on whether participants noticed the change. N = 100. **$p < .01$.

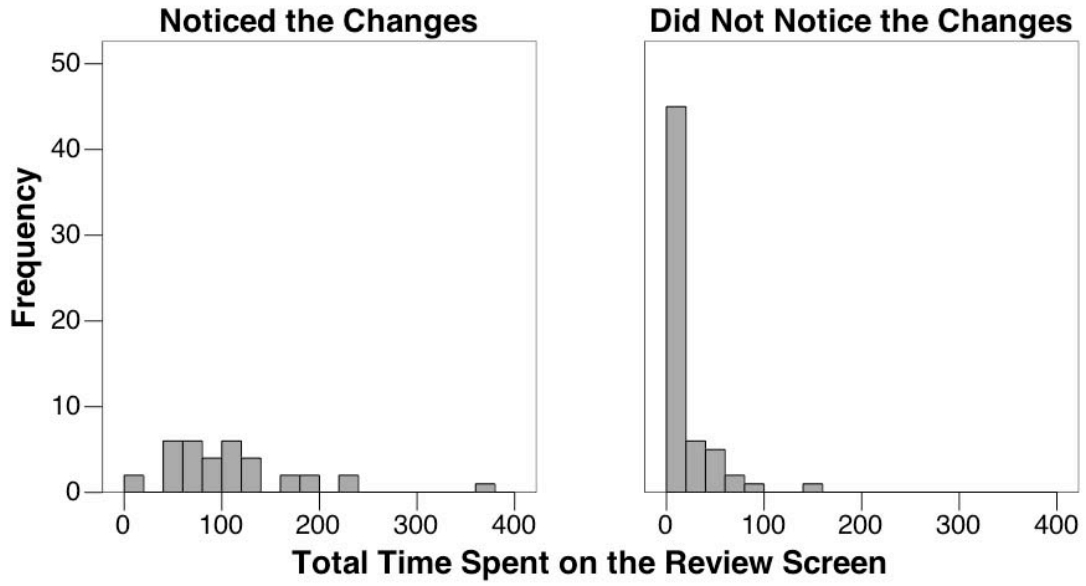


Figure 21. Histograms of the total time spent on the review screen for participants who noticed and participants who did not notice the changes.

The variables education, conscientiousness, previous electronic voting experience, and number of review screen visits were not significantly correlated with any other variables and did not contribute any predictive power to the regression. Thus, they were removed from further analysis. Time spent on the review screen the first time it was seen was also not included in further analysis as it was removed from regression analyses. As the total time spent on the review screen was the only performance variable that reliably predicted who noticed the changes, it was retained for further analysis.

To look at the relative weights of the variables that contributed significantly to whether participants noticed the changes, a structural equation model was constructed using the results from the correlation matrix and regression models, as in Study 2. Information condition was again included to control for the effects of that variable on the model. Total time on review screen had a direct path to the notice-change variable. Age and following the news were hypothesized to have indirect effects on the notice-change variable (through total time on review screen) as well as direct effects. Openness to experience and computer expertise also had indirect paths to the notice-change variable through total time on the review screen. The model was standardized path coefficients were computed, as can be seen in Figure 22. The model had a good fit with the data, as evidenced from three measures of fit ($X^2(6, N = 100) = 1.76$, $p = .94$; CFI = 1.00; and RMSEA = .00).

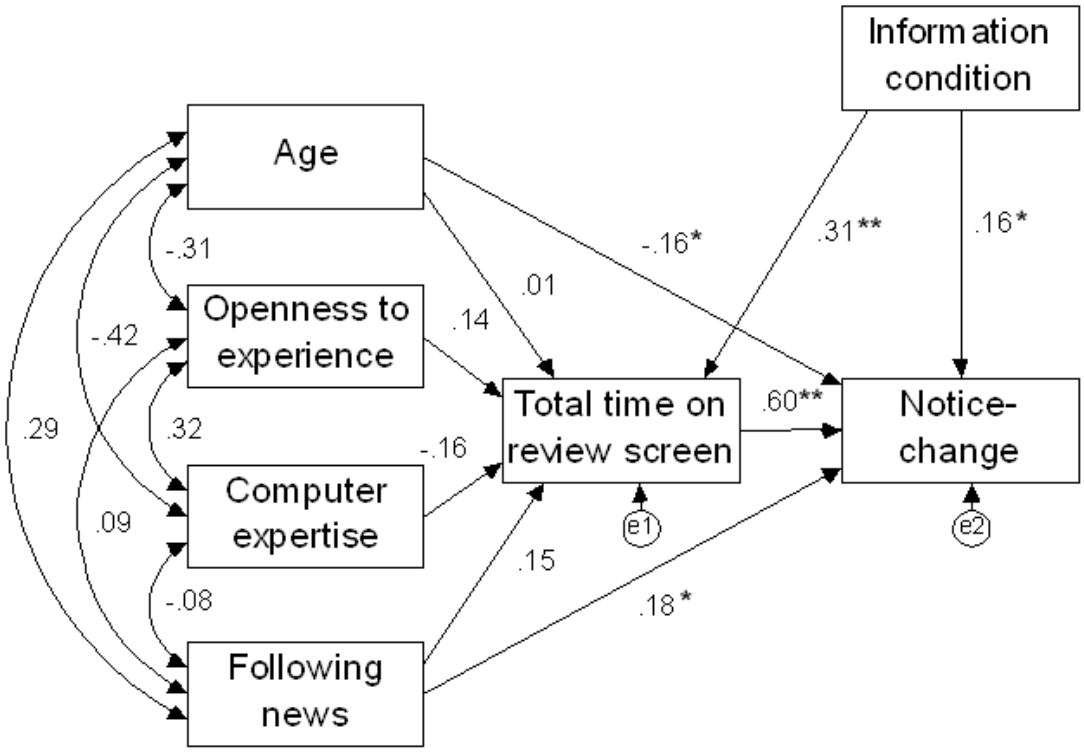


Figure 22. Path diagram and standardized parameter estimates for the Study 3 model. *$p < .05$. **$p < .01$.

As in Study 2, the total time on review screen was a strong predictor of whether voter would notice the changes (ß = .60, $p < .001$). Participants who spent longer on this screen were more likely to detect the vote flipping. Information condition also had both direct (ß = .16, $p = .03$) and indirect effects (ß = .31, $p < .001$) on the notice-change variable. Participants who were given more information and directed how they should vote were more likely to notice the changes, both directly and because these people spent more time on the review screen which made them more likely to notice the changes.

Two individual characteristics variables had significant direct paths to the notice-change variable. Age had a negative effect on noticing (ß = -.16, $p = .04$), which meant that older voters were less likely to notice that their votes had been changed. Also, how

carefully participants had been following the news about potential security problems with DREs impacted whether they would notice (ß = .18, p = .02). Those voters that had been following the news more carefully were more likely to notice the vote flipping.

DREs Versus Older Voting Methods

For the following results, outliers in the data were identified and replaced as in Studies 1 and 2. This resulted in the replacement of three DRE ballot completion times, one punch card ballot completion time, one DRE error rate, one lever machine error rate, one punch card error rate, and one DRE SUS score. For the error analysis, one participant was completely removed because of outlying error rates on both methods, and an additional five participants were removed because their ballot selections could not be read from the data files.

Time

Ballot completion times for the DRE, bubble ballot, lever machine, and punch card can be seen in Figure 23. Overall, the average ballot completion time was 281 seconds (*SD* = 147). Paired t-test revealed no reliable differences between the DRE and any of the other voting methods.

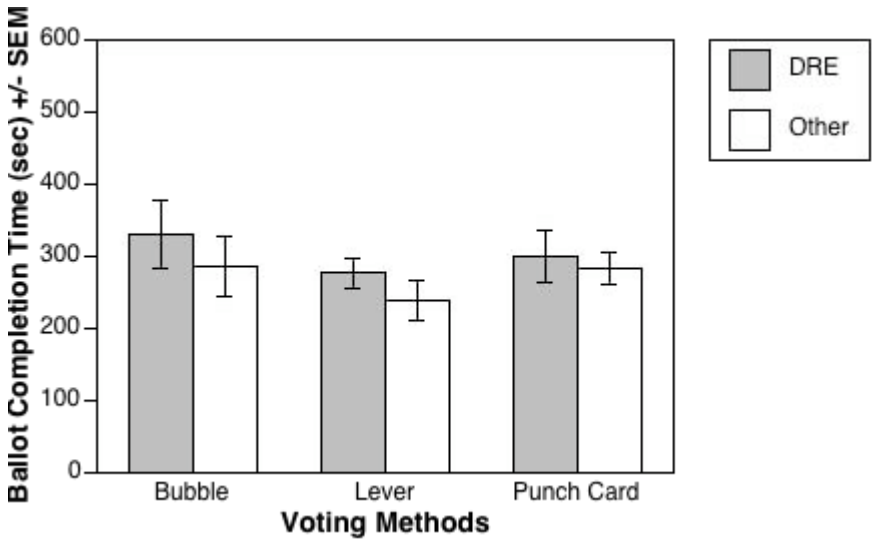


Figure 23. Pairs of mean ballot completion times (in sec) by voting method. Gray = DRE, white = other method (bubble, lever, or punchcard).

There were no reliable effects of information condition, which other method was used, or their interaction on the difference scores between the ballot completion times of the DRE and other methods, as shown by an ANOVA of these difference scores by information condition and other method.

A MANOVA examined the effect of information condition, other method, age, and education on the repeated variable of voting method (voting on the DRE and one additional method). There was a significant effect of age on ballot completion times ($F(1, 56) = 13.83$, $p < .001$), as can be seen in Figure 24 where a line was fitted to the ballot completion time by age data ($R^2 = .14$). There were no other reliable main effects or interactions.

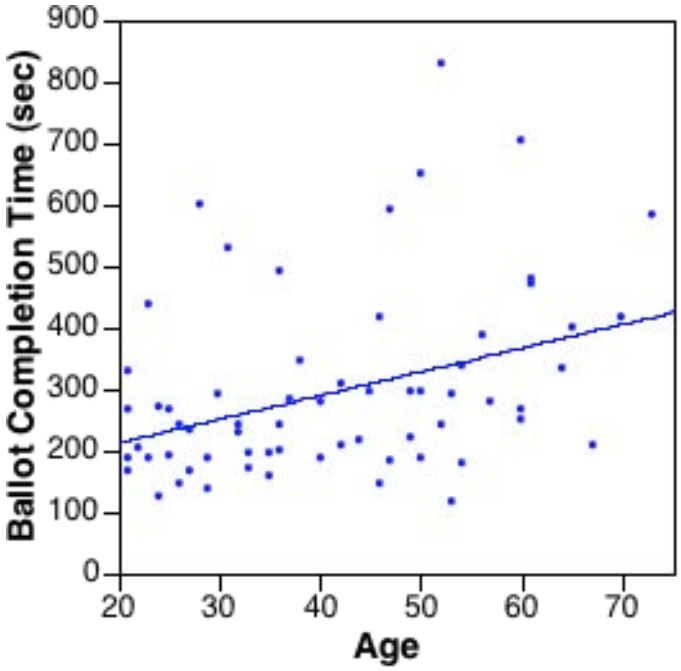


Figure 24. A line fitted to age by ballot completion time (in seconds).

Overall, ballot completion times on each voting method were unrelated to previous voting experience with the method. Time on the DRE was negatively correlated to self-rated computer expertise; voters completed their DRE ballots more quickly if they had more computer expertise (r = -0.28, p = 0.03).

Errors

Error rates per race for the four voting methods are shown in Figure 25. Across all voting methods, error rates were fairly high and variable (M = .03, SD = .06). There were no significant differences between the DRE and any other method as shown by paired t-tests.

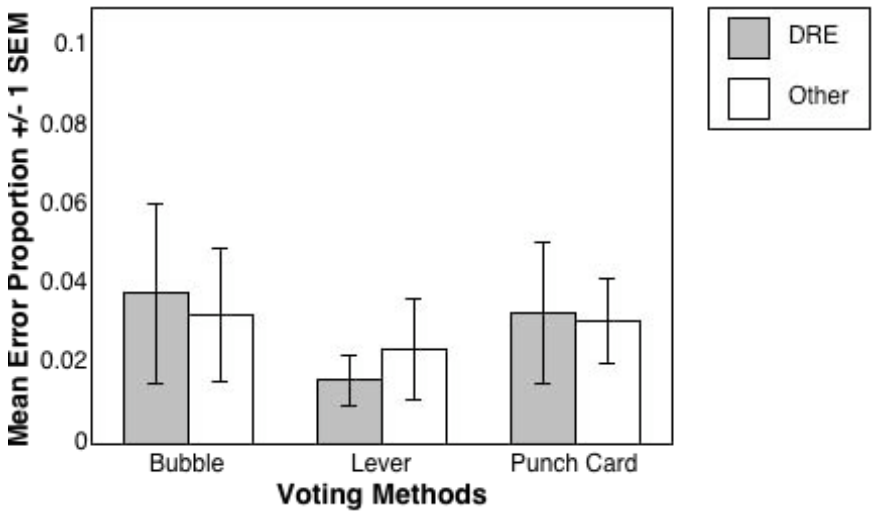


Figure 25. Pairs of mean error rates by voting method. Gray = DRE, white = other method (bubble, lever, or punchcard).

An ANOVA was performed on the difference scores between error rates on the DRE and other methods by information condition and other method. This showed no significant differences due to information condition, which other method was used, or their interaction.

A MANOVA was performed on the error rates on the DRE and other methods by age, education, information condition, and other method. There were no significant main effects or interactions of information condition, which other method the participant used, age, or education.

Error rates for each voting method were unrelated to previous voting experience on the method. Computer expertise was also not significantly related to the error rate on the DRE.

Errors per ballot can be seen in Table 30. Over 27% percent contained at least one error. A chi-square test showed that there were no differences in errors per ballot by voting method.

	Errors		
	Zero	At least 1	Total
DRE	44	15	59
Bubble ballot	11	4	15
Lever machine	19	5	24
Punch card	12	8	20
Total	86	32	118

Table 30. Number of ballots containing zero or at least 1 error by voting method.

In Study 3, data were also collected on post-completion errors on the DRE. These errors occurred when the voter left the voting station and reported to the experimenter that they had finished voting. In this study 6% of voters committed this type of error and walked away from the voting machine before actually casting their ballots. They believed they had finished the voting process, yet their ballots had not been cast.

Satisfaction

Satisfaction levels for each of the four voting methods can be seen in Figure 26. The SUS ratings for the DRE were higher than for any other method with an average of 91.8 (SD = 11.2). The bubble ballot produced an average SUS score of 76.8 (SD = 16.3), the lever machine average was 62.2 (SD = 27.0), and the punch card average was 57.8 (SD = 15.6). Paired t-test revealed that satisfaction levels for the DRE were reliably different that each of the other methods. For the DRE-bubble ballot pair, the t-test result

was $t(16) = 4.14$, $p = .001$. The t-test result for the DRE-lever machine pair was $t(25) = 4.39$, $p < .001$, and $t(19) = 8.46$, $p < .001$ for the DRE-punch card pair.

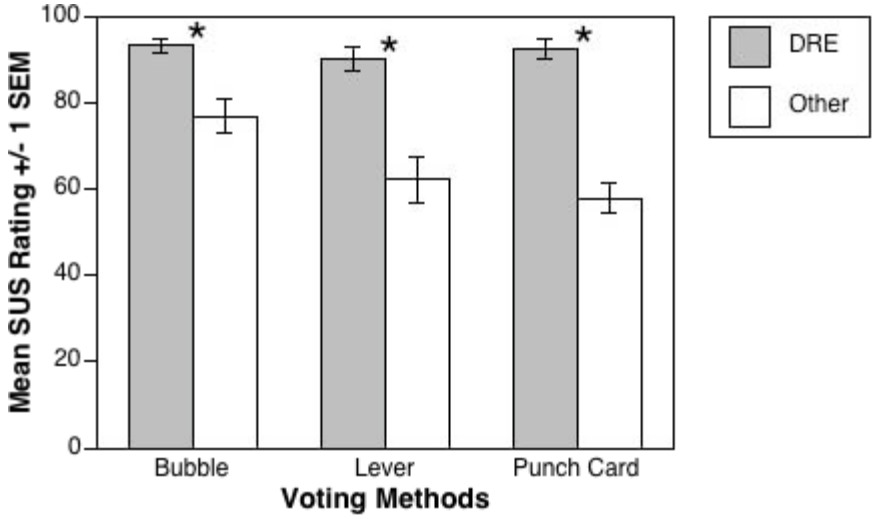


Figure 26. Pairs of mean SUS ratings by voting method. Gray = DRE, white = other method (bubble, lever, or punchcard). * $p < .05$.

An ANOVA was performed on the difference between the satisfaction levels on the DRE and other method that the participants used by information condition and other method. This did not reveal differences between other methods, information conditions, or the interaction of information condition and other method. However, the effect of which other method the participant used approached significance ($F(2, 54) = 2.54$, $p = .10$), meaning that the difference between the DRE and the bubble ballot was almost significantly smaller than the differences between the DRE and both the lever machine and punch card. Although not quite significant, this result shows that the DRE was better than all three methods, and was especially better than the lever machine and punch card.

A MANOVA examined the effect of the repeated variable voting method and also age, education, information condition, and which other method was used on SUS scores.

This revealed a main effect of which other method was used ($F(2, 56) = 4.48$, $p = .04$), and the repeated variable approached significance ($F(1, 56) = 2.64$, $p = .11$). The interaction of the repeated measure and which other method was used approached significance ($F(2, 56) = 2.41$, $p = .10$). Altogether this indicates that participants preferred the DRE over each of the other methods, especially over the lever machine and punch card, as also shown by the ANOVA above. The main effect of age also approached significance ($F(1, 56) = 3.9$, $p = .054$), with older adults giving higher SUS scores for both of the methods on which they voted. There were no other significant main effects or interactions on SUS scores.

Correlations revealed that previous voting experience was unrelated to satisfaction scores for each method. Computer expertise was unrelated to satisfaction scores on the DRE.

Discussion

The most striking result of Study 3 was that only 37% of participants noticed the vote flipping that occurred on their review screens. For each participant, the candidate choices for 1, 2, or 8 races had been changed and most voters did not detect this. The number, type, or location of the changes did not matter; voters either did or did not check the review screens carefully. This meant that anywhere from 4% - 30% (and likely more) of the ballot could be changed, mostly without detection. Also, because the location of the changes did not effect whether they would be noticed, it meant that important races such as US senator or state governor could be stolen just as easily as smaller, more local races further down the ballot such as sheriff or county tax assessor.

The type of the changes also was not important, indicating that votes could be changed to another candidate as well as simply changed to being an undervote. Removing a voter's selection in a race may be a more subtle method of vote flipping than changing it to another candidate. It may be more difficult to detect or prove fraud when a voter's selection is simply removed than when it is switched to a different candidate. This method of changing votes may make the problem seem more of a user error (e.g., the voter accidentally skipped a race) than a computer problem (e.g. fraud or a software bug).

In this study, the total time that participants spent on the review screen was again a very significant predictor of whether they would notice the changes. Voters who spent longer on the review screen were more likely to detect the vote flipping. As mentioned in Study 2, the direction of causality cannot be determined, but the relationship is certainly significant. Since the time voters spend on the review screen can be easily recorded by computers, a DRE could be programmed such that certain malicious software was only activated when the computer detected that the voter was not spending much time on the first page of review screens. This would mean that votes on the remaining review screen pages could be changed with a small chance of being noticed.

Another significant predictor of whether participants would notice the change was information condition. Participants who were more directed in who they should vote for and who were given a slate of choices to make were more likely to detect the changes both directly and indirectly. This meant that part, but not all, of the reason that these voters were more likely to notice was that they were spending longer on the review screen. As discussed above, this made participants much more likely to detect the vote flipping, although directional causality cannot be established. It may be the ease of

comparing the choices on a piece of paper to the choices displayed on the review screen that contributed to the predictive power of this variable. Participants did not have to remember for whom they had voted, they simply had to make visual comparisons between the screen and their slate.

Because voters are either not noticing the changes or are checking their review screens carefully and detecting the changes, it becomes important to identify differences between these types of voters. As for individual characteristics, the age of participants had a reliable effect on noticing the changes, and in this study age had a direct effect. In the Study 2, older voters were spending longer on the review screen, which in turn made it more likely that they would notice the changes. However, in the current study which had much higher power to detect effects, older voters were not spending any longer on the review screen than younger voters. The effect of age was not mediated by review screen time and so users were not more likely to detect the changes. In fact, age had a significant negative impact on the likelihood that voters would notice the changes.

Another individual characteristic significantly predicted whether voters would detect the vote flipping on their review screens. This was how carefully participants had been following the news about potential security problems with electronic voting systems, and it had an even stronger effect than age. Participants who had been following the news more carefully were more likely to detect the review screen changes. It seems that being exposed to and perhaps educated about the potential problems with DREs raised the awareness of voters such that it caused a change in their behavior.

Another area of concern that could be targeted by voter education campaigns is the actual casting of the final ballot. As discussed in Chapter 3, leaving the voting station

while it is showing the review screen is a type of postcompletion error. These "fleeing voters" have completed their primary goal of selecting candidates, but they have not yet pressed the button to actually cast their vote. This phenomenon was reported in recent elections (Siegel & Doherty, 2006; electiononline.org, 2006) and appeared in Study 3 as well. In this study 6% of voters walked away from the voting machine before actually casting their ballots. They believed they were done and had finished the voting process, yet their ballots had not been cast. This is problematic because it is not clear how these problems are handled in real elections. Although some states have rules regarding this, poll workers are not highly trained and may not act consistently. In previous elections, it was reported that some poll workers pressed the cast ballot button to record the vote when they were confronted with the situation while others erased it. When a subsequent voter reported this problem, some poll workers cleared the system and let the new voter vote, while other cast the displayed ballot and denied the new voter a chance to vote. If a subsequent voter noticed this situation and did not report it, they could change the currently displayed vote to match their own beliefs and still vote their own ballot, resulting in stealing someone else's vote in addition to using their own.

In addition to evaluating review screen use by voters, Study 3 also measured the satisfaction that users felt with the DRE. Satisfaction ratings were only included in this analysis from voters who did not notice the changes to their review screens. Even though the machines were stealing their votes, participants felt very satisfied with their experiences using the DREs. These high satisfaction ratings for the DRE may have contributed to the low levels of vote-flipping detection. Because users liked the machines, they may have assumed that they worked properly and could be trusted.

Although DREs were preferred over each of the other three voting methods, voters were also fairly satisfied with the bubble ballot. This result has been reported previously (Byrne et al., 2007) and may not be surprising. By adulthood, most people have had a lot of experience in handling and responding to questions on paper. This experience seems to transfer to voting and helps users in figuring out and marking their paper ballots.

While voters were most satisfied with their voting experiences on the DRE, this was not due to improvements in error rates or ballot completion times. None of the methods help solve the problem of voter error, which this and previous studies (Everett et al., 2006; Greene et al., 2006; and Byrne et al., 2007) showed is a significant concern. In Study 3, over 27% of ballots contained at least one error, a figure that is much higher than would be desirable in situations like elections with such important outcomes. Furthermore, ballots in this study were scored by hand and participants were given credit for marks for which their intention could be inferred, even if the mark many not be counted by machine (e.g. when a participant circled the candidate's name instead of filling in the bubble next to the name). The actual error rates may be even higher and this study shows that the problem cannot be solved by changing voting technologies.

CHAPTER 8

GENERAL DISCUSSION

These studies find that in general voters are very highly satisfied with their experiences voting on the DRE. Almost all participants reported feeling that the review screen was useful and most felt it made them more confident that their vote would be recorded correctly, yet most participants did not use it to check the accuracy of their votes. It may be the fact that they liked it so much that made them less likely to notice the changes. Perhaps because voters liked the DREs so much, they assumed it had good security properties as well. This sort of finding is similar to the halo effects reported in the psychological literature. When people really like or think highly of one aspect of a person or object, they often believe that other aspects of the person or object must good as well. In this way, the "halo" from the first aspect casts its effect over the remaining aspects.

In Studies 2 and 3, voters who were given a full slate of candidates to choose were much more likely to notice the changes to their review screen. While this condition may not be realistic of situations in actual elections, the type of directed voting is likely to be used when the testing and certification of electronic voting machines is performed. These studies show that this type of testing may not reveal the full scope of potential problems with electronic voting machines. Since voters are more likely to check the accuracy of their ballots as shown on the review screen when they are directed how to vote and given a paper with candidates names to hold in their hand, any testing that is done with such directed conditions will significantly overestimate how many voters

check their ballots. It may be assumed that votes can be relied upon to verify their ballots selections before casting them, a false assumption with important security implications.

As mentioned in chapters 6 and 7, the number, type, or location of the changes did not affect whether voters would notice them. In some cases, eight races representing 30% of the ballot were removed, added, or changed and most voters did not notice. Participants fell into two groups: those who carefully checked their votes on the review screen and caught the changes, and those that did not. Examining the effects of performance and individual characteristics on who noticed the changes allowed for relatively accurate predictions of which voters noticed the changes. The ability to predict which voters will not notice the changes means that vote-flipping attacks could be targeted to avoid detection. The two variables that most likely could be used in elections to target attacks are age and how long voters spend on the review screen.

The finding that votes can be changed to another candidate or changed to an undervote equally easily is especially interesting given the problems in the 2006 election in Sarasota County, FL. In this county, there was a 14% undervote rate in the race for an open seat in the U.S. House. While this undervote rate may be due to poor ballot design or voters abstaining from selecting a candidate (electiononline.org, 2006), it is also possible that voters believed they had successfully chosen a candidate in the race. This possibility is supported by results from Studies 2 and 3 about how few people are using the review screen, and especially from study 3 about how voters did not notice that their votes had been changed to "None."

In Study 2, the personality trait openness to experience was positively related to noticing the changes. Although this effect did not replicate in the third study, it provides

preliminary evidence that personality traits may be able to predict specific behaviors in a voting context. While studies of personality have looked at the ability to predict outcomes in general categories such as job performance (Barrick & Mount, 1991) or training effectiveness (Gully et al., 2002), this study shows that it may be possible to predict how an individual will perform on a single, specified task.

The third study, which used a large sample size and participants who were representative of voters in general, showed that age had a negative direct impact on whether voters would notice the change. It is easy to obtain demographic information such as the average age of registered voters in a precinct. Because age is an important factor in whether voters will detect changes, malicious software attacks could be targeted to precincts with high numbers of older voters. This would mean that the vote flipping attacks could be executed selectively with a very low probably of detection.

Similarly, Studies 2 and 3 found that the longer participants spent on the review screen, the more likely they were to notice the changes. As mentioned in chapter 7, this information could be used by corrupt voting machines to steal votes without detection. If a voter spent a long time on the first page of multiple review screens, the DRE could act properly and not manipulate any votes. However, if the machine detected that a voter spent a very short amount of time on this first review screen page, it could make changes to the voter's choices on subsequent review screen pages with a very low probability that the changes would be caught. While this result may inform hackers on how to avoid detection, it also may point to potential solutions. If increased time spent on the review screen really leads to participants' verifying their ballots, voters could be strongly encouraged (or forced) to spend longer on this page. Of course, this may lead to other

problems such as voters leaving the voting booth while the review screen is displayed and before actually casting their ballots, or long wait times.

In addition to age and time spent on the review screen, one more variable significantly predicted whether voters would notice the changes to their review screens. This was how carefully they had been following the news on potential security problems with electronic voting machines. There may be an availability heuristic at work here, as some newspapers recently have been closely covering any potential problems with electronic voting machines. Given this coverage in the media, people who have been carefully following these stories may believe that security problems are much more prevalent than they actually have been proven to be so far. Although there have been allegations of fraud in recent elections, they have largely not been substantiated. If a voter has been following the news carefully, they seem to adopt a lower response criterion such that misses are less likely. As mentioned in chapter 3, studies have found that people can adjust their response criterions to reflect a conservative or risky strategy. In this way, the media are changing the behavior of voters, causing them to adopt a more conservative voting strategy and check their ballots more carefully.

While not useful from an attack-targeting standpoint, this has important implications on media effects and voter education campaigns. Voting is traditionally an area where training has not been used, as it is not feasible to train all voters on how to use a voting system. However, this study finds that mass media such as newspapers and news shows can impact voter behavior. Voter education campaigns have the potential to make a difference in how carefully voters check their ballots.

However, as the situation currently stands, voters cannot be depended upon to check the validity of their vote. Many security experts and election reform groups are calling for VVPATs to be required on all DREs and as of the 2006 elections, nearly half of the states mandated that their DREs have paper trails (electionline.org, 2006). However, these studies show that solutions to DRE security problems that require voter verification of their ballots may not solve vote-flipping problems. Users are not even checking their ballots on the review screen that is presented directly in front of them. In these studies, voters could not skip the review screen and had to look at it at least long enough to locate the button to continue to the next page. If voters are not checking their votes and detecting problems on this screen, they cannot be depended upon to verify their votes as shown on a VVPAT printout. On machines that are equipped with paper trails, the VVPATs are typically made by a thermal printer attached to the side of the voting machine. Voters who are not even verifying that the review screen accurately displays their votes will likely be much less inclined to verify their ballots on a separate device and with the quality of printing available on such devices. While there is not yet a definitive answer to how voters use VVPATs, these studies show that they are probably not a final solution to security problems. This is because as voters cannot be depended upon to verify that their choices have been recorded accurately.

This phenomenon likely applies to other domains where people are expected to review some information that is presented to them after completing their primary task. The effect may be seen in online shopping, where users select the items they would like to purchase and are then presented with a type of review screen, the shopping cart. Results from these studies suggest that shoppers may not check the contents of these

online shopping carts carefully before proceeding to the checkout. It is certainly important to consumers that their carts accurately reflect their desired purchases so that they can avoid buying extra products, making returns for unintended products, or paying for the shipment of a second order if an item was left out. In addition, the online businesses should be concerned that shoppers check the contents of their online shopping carts before placing an order. The satisfaction the customer feels with the company may be reduced if their order is not filled as intended and the company may have to process many returns if unintended items are purchased. However, the results of these studies show that shoppers may not be depended upon to verify the accuracy of the online shopping cart or review screen.

It is possible that improvements to the review screen itself could draw attention towards the information displayed there. For example, pictures of the product to be purchased or an icon representing a candidate's political party could be used on online shopping carts and DRE review screens. The use of this picture in addition to the name of the product or candidate could potentially improve the detection rates of problems. More specifically to voting systems, review screens could also highlight undervotes so that a user's attention was drawn to races where a vote had not been recorded. Also, these review screens could include the political party of the candidates chosen, a suggestion made by many participants. As many voters make their choices according to political party, it is much harder for them to check the accuracy of the review screen without this information. Including it would allow voters simply to scan the contents of the review screen to see if any votes had been recorded for the wrong party.

By studying how many voters check the accuracy of their votes on the review screen, these studies addressed potential security concerns of DREs. Another goal of these studies was to address the usability of DREs compared to older, more traditional voting methods. Previous research had evaluated different types of DREs, and the usability of older methods had been studied, but all these methods had not been included in a study that could make direct comparisons between them. It had been assumed that DREs would be better than the paper ballots, lever machines, and punch cards they have been replacing and Studies 1, 2, and 3 sought to examine that belief. Results from the studies reported here indicated that although performance on DREs in terms of efficiency and effectiveness is not any worse than on previous methods, it is also not any better. This finding replicated across all three studies and strongly shows that DREs do not lead to better voting performance as had been assumed. This was a robust effect and although there were some effects of age and education on ballot completion times, in general only computer expertise affected voters' levels of efficiency on the DRE.

No differences in error rates were seen between the DRE and the older voting methods, but the high frequency of ballots containing errors is cause for concern. In the second and third studies with the more representative participants, 27% of ballots contained at least one error. This closely matches the rate reported in previous work on older voting methods (Byrne et al., 2007). Because the outcomes of elections have such important and widespread impacts, this amount of error in electing officials is certainly higher than would be desired.

Although there were no differences in ballot completion times or error rates between the DRE and the other three methods, participants were most satisfied with their

voting experience on the electronic voting system. This was indicated by an average SUS score over 90 for this measure in the third study. According to suggestions made by Bangor, Kortum, and Miller (2007) from their work evaluating a range of technologies, scores above 90 indicate "truly superior products." Using this scale to interpret the SUS scores for the other voting methods shows that the bubble ballot has acceptable SUS scores, while the lever machine and punch card are marginal and could use improvement.

Even though participants with less computer expertise voted more slowly on the DRE, their satisfaction with it was not any less than those users with more computer expertise. The performance on and preference for DREs was not affected by age or education levels, and thus no evidence of a digital divide was seen. This may be surprising as the results of some studies, such as Roseman and Stephenson (2005) which found a decrease in voter turnout among older people when DREs were used, might have predicted a digital divide effect.

It is interesting that voters strongly prefer using the DREs even though their performance is not any better on them. This type of preference versus performance disassociation is not an uncommon finding, but has important implications in elections. Because of the controversy currently surrounding the use of DREs in elections, some groups are calling for a ban on DREs and at least one state (Florida) has proposed reverting to the use of paper ballots. Although participants do not like paper ballots as much as DREs, they are still much more satisfied with paper ballots than with lever machines and punch cards. As previous research has reported (Byrne et al, 2007), old-fashioned paper ballots actually work very well. Most voters can perform at reasonable

efficiency and effectiveness levels with paper ballots and are satisfied with the experience.

However, paper ballots are inaccessible for many groups of people. Paper ballots do not provide voters with disabilities such as low vision or manual dexterity impairments with the ability to vote an independent secret ballot. DREs have been touted as a solution to the accessibility problems of other voting methods, but Runyan (2007) reports that most DREs do not meet the disability requirements of HAVA or the American with Disabilities Act (ADA). Many systems do not allow for accessible input devices or adequate adjustments to screen displays. In addition, the audio component of many electronic voting systems does not provide efficient ballot navigation or synchronous audio and visual outputs. Finally, DREs that have been equipped with VVPATs to combat security problems do not allow independent verification of ballots for many people. Thus, even though DREs have the potential to improve accessibility to voting equipment over older methods, these improvements have not been realized.

Overall, the usability findings from the current studies show that the use of DREs does not lead to more efficient or effective voting performance, although voters are highly satisfied with these electronic systems. The high satisfaction participants feel with the DREs means that citizens may be unhappy about abandoning the new voting systems in the face of security concerns. Even though their performance on the DREs is not any better, voters may fight to keep the systems that they so strongly prefer.

Limitations and Future Directions

One general limitation of this work is that participants were not actually voting for real candidates in a real election. It is possible that the artificiality of the situation

could have affected the motivation and performance of participants. However, from observation, it was apparent that participants took these experiments very seriously and many commented on the importance of the work. In addition, error rates were fairly low, suggesting that voters were casting their ballots deliberately and carefully. Finally, if participants were less motivated to vote carefully in these experiments than in a real election and the results were exaggerated even by a factor of two, it still means that about a third of voters would not notice changes made to their ballots.

Other limitations of the current studies include that usability evaluations of the DREs and other methods used data from participants who were not in the control group. Data from participants who did not notice the changes to their review screens were also included. It is possible that these people might perform slightly differently from the users who did notice the changes. However, the results from the current study match with the trends found in previous work on traditional voting methods (Byrne et al., 2007), so it is unlikely that this is a problem.

The DRE used in these studies, VoteBox, is a prototype of an electronic voting system and may be representative of all DREs. Some systems may be better or worse than the VoteBox system and may return different usability measurements. Future studies could include different types of DREs or add features such as touchscreen input or multiple language and font size support to the existing VoteBox system.

One possible limitation of the DRE review screen work is that the review screen did not list the political parties of the candidates. This may make a difference, especially for people who vote a straight-party ticket. It is possible that adding this information to

the review screen would improve the frequency of vote-flipping detection and this could be tested in further work.

Another review screen manipulation that could be studied in future work is the degree to which attention is drawn to undervotes. Races in which the voter has not made a selection could be highlighted so that voters would be more likely to notice that they had undervoted. The type and amount of this highlighting could be varied in future experiments. Alternatively, in some studies voters could be explicitly warned that they were about to submit a ballot with undervotes.

Finally, from the design of the current studies, the cause and effect of how long voters spend on the review screen and whether they detect the changes cannot be determined. It could be that voters who have noticed the changes spend longer on the review screen trying to figure out the problem. Or it could be that voters who spend a long time on the review screen are carefully checking their votes and thus are more likely to notice the changes. Future work is needed to determine the direction of causality in this relationship.

Conclusions

This work evaluated the usability of DREs versus that of older, more traditional voting methods in terms of efficiency, effectiveness, and satisfaction. The DRE did not differ from the bubble ballot, lever machine, or punch card in terms of ballot completion times or error rates. However, DREs were strongly preferred over these other systems. The paper bubble ballot fared well in evaluations also, but voters were significantly more satisfied with their experience on voting on the DRE. This indicates that although participants' performance was not any better with the DREs, voters liked these electronic

voting systems and might react negatively to any attempt to ban or restrict their use in elections.

This satisfaction with DREs was observed even though voter's ballots on these machines could be changed largely without detection. Up to eight races were added, removed, or changed and less than 40% of voters noticed. Older voters, those who spent very little time on the review screen, and voters who had not been following the news on DREs were the least likely to notice the changes. A voter meeting these three criteria would be a worst-case scenario from a security standpoint, and this leads to the possibility of targeted attacks of vote flipping by malicious software on electronic voting machines. For many voters, their choices can be changed on the screen right in front of them without detection. This also has implications for VVPATs, which have been suggested as a partial solution to vote-flipping schemes. The findings here suggest that it is highly unlikely that voters will detect changes to their ballots on the VVPAT that prints out on a roll of paper next to the machine if they are not even noticing them on a screen presented directly in front of them.

REFERENCES

Alvarez, R. M. (2002). Ballot design options. Paper prepared for *Human factors research on voting machines and ballot designs: An exploratory study*. Retrieved October 31, 2006 from http://www.capc.umd.edu/rpts/MD_EVote_Alvarez.pdf

Andre, A. D., & Wickens, C. D. (1995). When users want what's not best for them. *Ergonomics in Design, 4*, 10 -14.

Bailey, R. W. (1993). Performance vs. preference. In *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting*, 282-286. Santa Monica, CA: Human Factors and Ergonomics Society.

Bangor, A., Kortum, P. T., & Miller, J. T. (2007). *An empirical evaluation of the System Usability Scale (SUS)*. Manuscript submitted for publication.

Bannet, J., Price, D. W., Rudys, A., Singer, J., & Wallach, D. S. (2004). Hack-a-Vote: Demonstrating security issues with electronic voting systems. *IEEE Security & Privacy Magazine, 2*, 32- 37.

Barrick, M. R., & Mount, M. M. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*, 1-26.

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

Berkun, M. M. (1964). Performance decrement under psychological stress. *Human Factors, 6*, 21-30.

Bremer, J. (n.d.). *Ballot design in an electronic environment - Lessons from the*

online market research industry. Retrieved October 31, 2006 from http://www.capc.umd.edu/rpts/MD_EVoteHuFac.html

Brennan Center for Justice. (2006). *The machinery of democracy: Voting system security, accessibility, usability, and cost.* Retrieved February 15, 2007 from http://www.brennancenter.org/stack_detail.asp?key=97&subkey=38150

Brooke, J. (1996) SUS: A "quick and dirty" usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & A. L. McClelland (Eds.), *Usability evaluation in industry* (pp. 189-194). London: Taylor and Francis.

Byrne, M. D., & Bovair, S. (1997). A working memory model of a common procedural error. *Cognitive Science, 21*, 31-61.

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

Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Chaum, D. (2004). Secret-ballot receipts: True voter-verifiable elections. *IEEE Security and Privacy, 2*(1), 38-47.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Cohen, S. B. (2005). *Auditing technology for electronic voting machines*. Unpublished manuscript, Massachusetts Institute of Technology: Cambridge, MA.

Conrad, F. G., Lewis, B., Peytcheva, E., Traugott, M., Hanmer, M. J., & Herrnson, P. S.

(2006). *The usability of electronic voting systems: Results from a laboratory study.* Paper presented at the Midwest Political Science Association, Chicago, IL. April 2006.

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual.* Odessa, FL: Psychological Assessment Resources.

Czaja, S. J., & Lee, C. C. (2003). Designing computer systems for older adults. In J. A. Jacko & A. Sears (Eds.), *The human-computer interaction handbook: Fundamentals, evolving technologies, and emerging applications* (p. 413-427). Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Di Franco, A., Petro, A., Shear, E., & Vladimirov, V. (2004). Small vote manipulations can swing elections. *Communications of the ACM, 47*(10), 43-45.

Dill, D. L. (2005). *Electronic voting: An overview of the problem*. Retrieved November 1, 2006 from http://www.verifiedvotingfoundation.org/article.php?id=5731

electiononline.org. (2006). *Election reform briefing 15: The 2006 election.* Retrieved March 25, 2007 from http://www.electionline.org/Publications/tabid/86/Default.aspx

Election Assistance Commission (2005). *Voluntary voting system guidelines*. Retrieved March 3, 2007 from http://guidelines.kennesaw.edu/vvsg/guide_toc.asp

Election Data Services Inc. (2006). *2006 voting equipment study*. Retrieved October 29, 2006 from www.electiondataservices.com/EDSInc_VEStudy2006.pdf

Ellis, L. B. M., Joo, H.-Y., & Gross, C. R. (1991). Use of a computer-based health risk appraisal by older adults. *Journal of Family Practice*, 33(4), 390-395.

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 Ergonomics Society 50th Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

Feldman, A. J., Halderman, J. A., & Felten, E. W. (2006). *Security analysis of the Diebold AccuVote-TS voting machine*. Retrieved November 13, 2006 from http://itpolicy.princeton.edu/voting

Fisk, A. D., Rogers, W. A., Charness, N., Czaja, S. J., & Sharit, J. (2004). *Designing for older adults: Principles and creative human factors approaches*. Boca Raton, FL: CRC Press.

Fitts, P. M. (1966). Cognitive aspects of information processing III: Set for speed versus accuracy. *Journal of Experimental Psychology, 71*, 849-857.

Frøkjær, E., Hertzum, M., and Hornbæk, K. (2000). Measuring usability: Are effectiveness, efficiency, and satisfaction really correlated? In *Human Factors in Computing Systems: Proceedings of CHI 2000.* New York: ACM.

Gerck, E. (2001). *Voting with witness voting: Qualified reliance on electronic voting*. Presented at the Workshop on Trustworthy Elections. Retrieved on February 21, 2007 from http://www.vote.caltech.edu/wote01/pdfs/gerck-witness.pdf

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., et al. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*, 84–96.

Goldberg, L. R., Sweeney, D., Merenda, P. F., & Hughes, J. E., Jr. (1998). Demographic variables and personality: The effects of gender, age, education, and ethnic/racial

status on self-descriptions of personality attributes. *Personality and Individual Differences, 24*(3), 393-403.

Gonggrijp, R., Hengeveld, W.-J., Bogk, A., Engling, D., Mehnert, H., Rieger, F., Scheffers, P., & Wels, B. (2006). *Nedap/Groenendaal ES3B voting computer: A security analysis*. Retrieved March 1, 2007 from http://www.wijvertrouwenstemcomputersniet.nl/Nedap-en

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: John Wiley and Sons.

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.

Gully, S. M., Payne, S. C., Koles, K. L. K., & Whiteman, J.-A. K. (2002). The impact of error training and individual differences on training outcomes: An attribute-treatment interaction perspective. *Journal of Applied Psychology, 87* (1), 143-155.

Herrnson, P. S., Niemi, R. G., Hanmer, M. J., Bederson, B. B., Conrad, F. G., & Traugott, M. (2006). *Voters' abilities to cast their votes as intended*. Paper presented at the Workshop on the Usability and Security of Electronic Voting System.

Herrnson, P. S., Niemi, R. G., Horowitz, J. M., & Richman, S. (2004). *Characteristics of contemporary voting machines*. Retrieved February 12, 2007 from www.capc.umd.edu/rpts/md_evote_ContempVotingMach.pdf

Herron, M. C., & Sekhon, J. S. (2003). Overvoting and representation: An examination of overvoted presidential ballots in Broward and Miami-Dade counties. *Electoral*

Studies, 22, 21-47.

Hoepman, J.-H., & Jacobs, B. (2007). Increased security through open source. *Communications of the ACM, 50*(1), 79-83.

Howell, W. C., & Kreidler, D. L. (1963). Information processing under contradictory instruction sets. *Journal of Experimental Psychology, 65,* 39-46.

Industry Usability Reporting Project. (2001). *Common industry format for usability test reports* (ANSI/INCITS 354-2001).

International Committee for Information Technology Standards. ISO 9241-11. (1998). *Ergonomic requirements for office work with visual display terminals (VDT)s - Part 11. Guidance on usability.*

International Personality Item Pool: A Scientific Collaboratory for the Development of Advanced Measures of Personality Traits and Other Individual Differences (http://ipip.ori.org/). Internet Web Site.

Jones, D. W. (2003). *A brief illustrated history of voting*. Retrieved September, 23, 2005 from http://www.cs.uiowa.edu/~jones/voting/pictures/#ballota

Jones, D. W. (2005). *Chain voting*. Retrieved March 5, 2007 from http://vote.nist.gov/threats/papers/ChainVoting.pdf

Jones, D. W. (2006). *Technologists as political reformers: Lessons from the early history of voting machines*. Presented at the Society for the History or Technology Annual Meeting. Las Vegas, October 13, 2006.

Kelsey, J. (2006). *Electronic IDV – Status report*. Presentation for the Technical Guidelines Development Committee. Retrieved March 3, 2007 from http://vote.nist.gov/ElectronicIDV-status-presentation.pdf

Kiayias, A., Michel, L., Russell. A., & Shvartsman, A. A. (2006) *Security assessment of the Diebold optical scan voting terminal*. Retrieved March 1, 2007 from http://voter.engr.uconn.edu/voter/Reports.html

Kimball, D. C. (2004). *Assessing voting methods in 2002*. Retrieved February 11, 2007 from www.umsl.edu/~kimballd/dkmpsa2.pdf.

Kimball, D. C. (2005). *Summary tables on voting technology and residual vote*. Retrieved March 5, 2007 from http://www.umsl.edu/~kimballd/dk_vote.htm

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

Kimball, D. C., & Kropf, M. (2006). *Ballot initiatives and residual ballots in the 2004 presidential election*. Paper presented at the Annual Meeting of the Southern Political Science Association. Atlanta, GA.

Kohno, T., Stubblefield, A., Rubin, A. D., & Wallach, D. S. (2004). Analysis of an electronic voting system. *Proceedings of the IEEE Symposium on Security and Privacy*. IEEE Computer Society Press.

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.

Mather, G. B. (1964). *Effects of the use of voting machines on total votes cast: Iowa – 1920-1960*. Institute of Public Affairs. The University of Iowa: Iowa City.

McCrae, R. R., & Costa, P. T., Jr. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*, 81-90.

McCrae, R. R., & Costa, P. T., Jr. (2003). *Personality in adulthood: A five-factor theory perspective* (2nd ed.). New York: The Guilford Press.

Mebane, Jr., W. R. (2004). The wrong man is president! Overvotes in the 2000 presidential election in Florida. *Perspectives on Politics, 2*(3), 525-535.

Mebane, Jr., W. R. (2006). *Voting machine allocation in Franklin County, Ohio, 2004: Response to U.S. Department of Justice letter of June 29, 2005*. Retrieved February 18, 2007 from macht.arts.cornell.edu/wrm1/franklin2.pdf

Merwin, D. H, & Wickens, C. D. (1993). Comparison of eight color and gray scales for displaying continuous 2D data. In *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting*, 1330-1334. Santa Monica, CA: Human Factors and Ergonomics Society.

National Institute of Standards and Technology. (2006). *Requiring software independence in VVSG 2007: STS recommendations for the TGDC*. Retrieved on February 27, 2007 from http://vote.nist.gov/DraftWhitePaperOnSIinVVSG2007-20061120.pdf

Nichols, S. M., & Strizek, G. A. (1995). Electronic voting machines and ballot roll-off. *American Politics Quarterly, 23*(3), 300-318.

Nielsen, J., & Levy, J. (1994). Measuring usability: Preference vs. performance. *Communications of the ACM, 37*(4), 66-76.

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

Pervin, L. A. (1996). *The science of personality*. New York: John Wiley & Sons, Inc.

Reason, J. (1990). *Human error*. Cambridge, UK: Cambridge University Press.

Rivest, R. L. (2006). *The ThreeBallot voting system*. Retrieved November 13, 2006 from http://theory.csail.mit.edu/~rivest/Rivest-TheThreeBallotVotingSystem.pdf

Roseman, G. H., Jr., & Stephenson, E. F. (2005). The effect of voting technology on voter turnout: Do computers scare the elderly? *Public Choice, 123*, 39-47.

Runyan, N. H. (2007). *Improving access to voting: A report on the technology for accessible voting systems*. Retrieved March 26, 2007 from http://voteraction.org/reports/nrreport/Improving_Access_to_Voting_.pdf

Salas, E., Driskell, J. E., & Hughes, S. (1996). Introduction: The study of stress and human performance. In J. E. Driskell & E. Salas (Eds.). *Stress and human performance*, 1-45. Mahwah, NJ: Lawrence Erlbaum Associates.

Saltman, R. G. (1988). Accuracy, integrity and security in computerized vote-tallying. *Communications of the ACM*, 31(10), 1184-1191 & 1218.

Saltman, R. G. (2006). *The history and politics of voting technology: In quest of integrity and public confidence*. New York: Palgrave Macmillan.

Siegel, S., & Doherty, O. (2006, September 6). Machines add features to help system, voters. *Gazette.Net*. Retrieved March 31, 2007 from http://www.gazette.net/stories/090606/montcou181637_31942.shtml

Simons, B. (2004). Electronic voting systems: The good, the bad, and the stupid. *ACM Queue, 2*(7), 20 – 26.

Sinclair, R. C., Mark, M. M., Moore, S, E., Lavis, C. A., & Soldat, A. S. (2000). An electoral butterfly effect. *Nature, 408*, 665-666.

Tokaji, D. P. (2005). The paperless chase: Electronic voting and democratic values. *Fordham Law Review*, 57.

Traugott, M. W., Hanmer, M. J., Park, W.-H., 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 conference of the Midwest Political Science Association, Chicago, IL, April 7-10.

Trewin, S. (2006). Physical usability and the mobile web. In *Proceedings of the 2006 International Cross-Disciplinary Workshop on Web Accessibility (W4a)*, 109-112. New York : ACM Press.

Verified Voting Foundation. (2004). *Resolution on electronic voting*. Retrieved February 27, 2007 from http://www.verifiedvotingfoundation.org/article.php?id=5028

VotersUnite (2006). *E-voting failures in the 2006 mid-term elections: A sampling of problems across the nation*. Retrieved March 1, 2007 from http://www.votersunite.org/info/E-VotingIn2006Mid-Term.pdf

VotersUnite (2007). *Election reform bill (Rush Holt's HR 811) urgently needs reform*. Retrieved March 6, 2007 from http://www.votersunite.org/info/2007ElectionReform.asp

Wand, J. N., Shotts, K. W., Sekhon, J. S., Mebane, W. R., Jr., Herron, M. C., & Brady, J. E. (2001). The butterfly did it: The aberrant vote for Buchanan in Palm Beach County, Florida. *American Political Science Review, 95*(4).

Welford, A. T. (1985). Changes of performance with age: An overview. In N. Charness (Ed.). *Aging and human performance*, 333-369. New York: John Wiley and Sans.

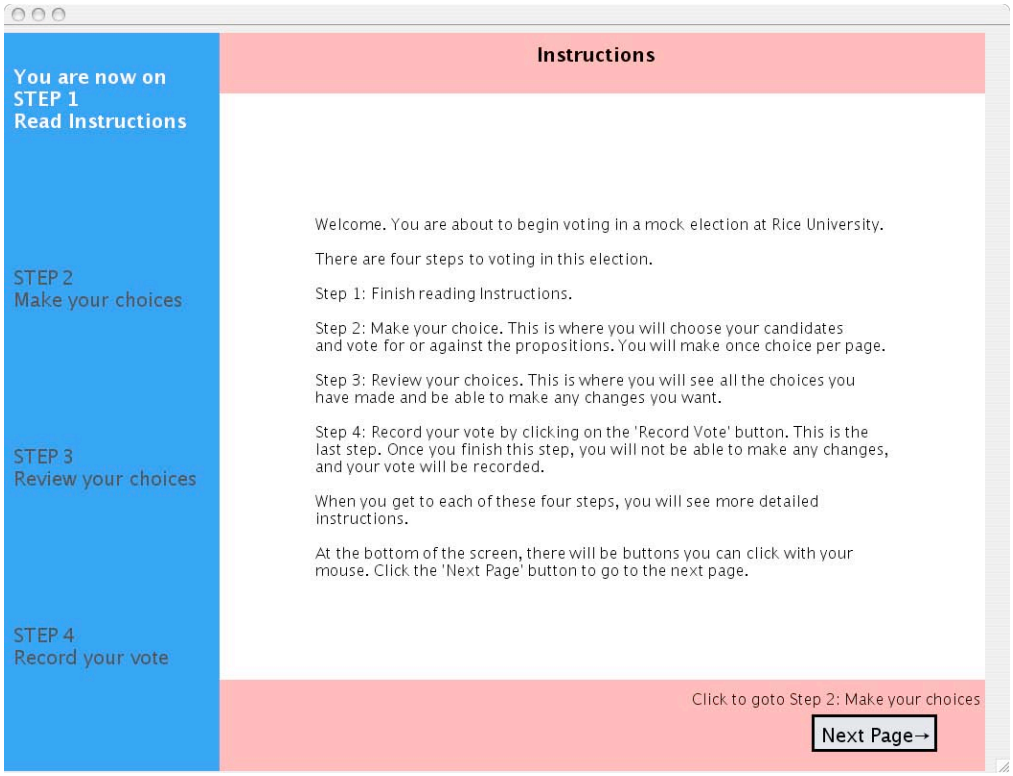
Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

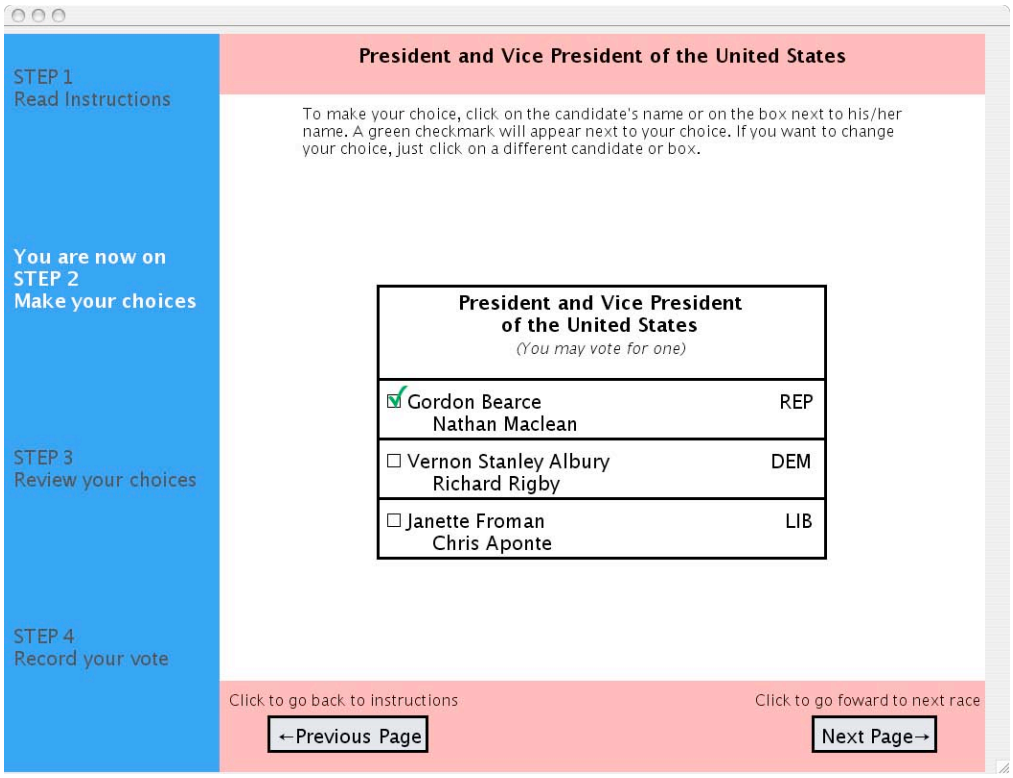
Wright, P. (1974). The harassed decision maker: Time pressures, distractions, and the use of evidence. *Journal of Marketing Research, 44*, 429-443.

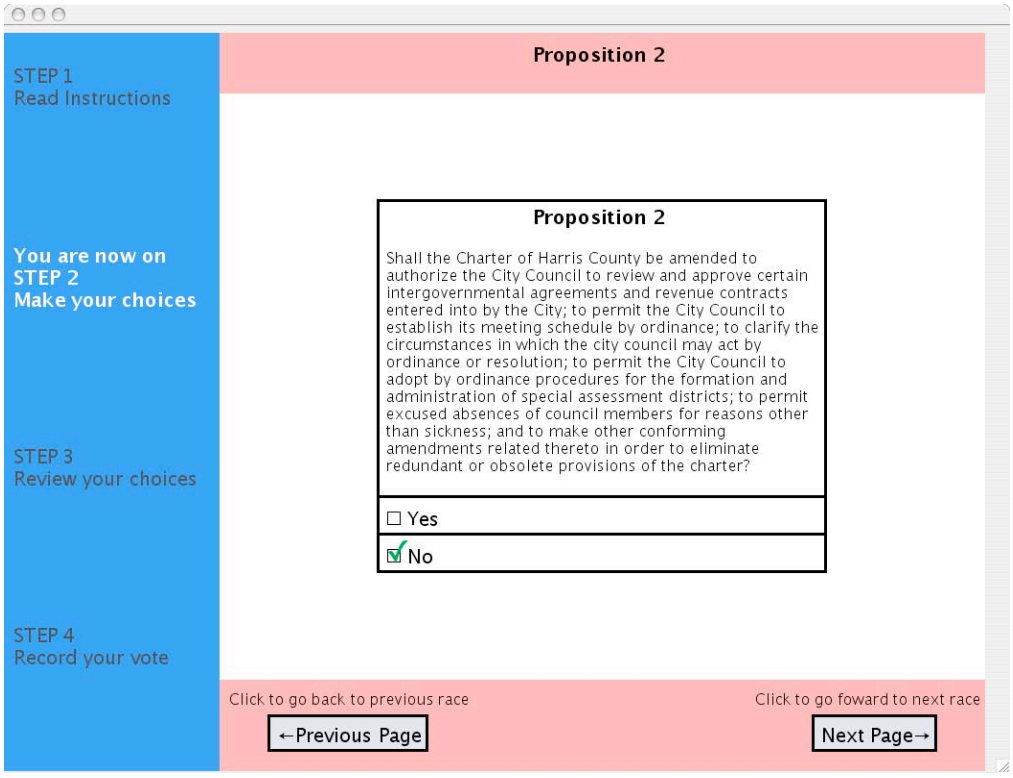
Yee, K.-P., Wagner, D., Hearst, M., & Bellovin, S. M. (2006). Prerendered user interfaces for higher-assurance electronic voting. In *Proceedings of the USENIX/ACCURATE Electronic Voting Technology Workshop*, Vancouver, BC.

APPENDIX A

This appendix contains screenshots from the DRE used in the studies.



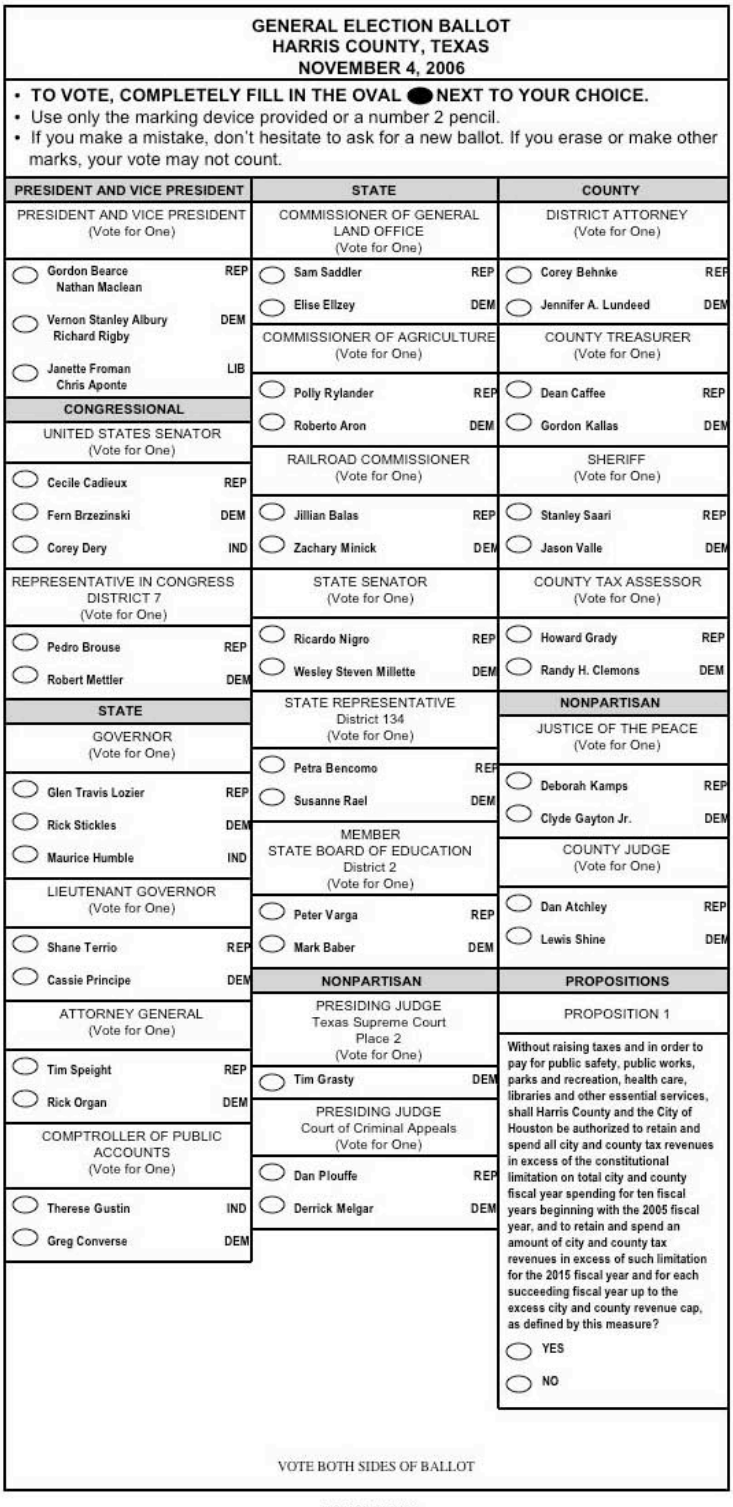

STEP 1
Read Instructions
You are now on
STEP 2
Make your choices
STEP 3
Review your choices
STEP 4
Record your vote
President and Vice President of the United States
To make your choice, click on the candidate's name or on the box next to his/her name. A green checkmark will appear next to your choice. If you want to change your choice, just click on a different candidate or box.
President and Vice President of the United States
(You may vote for one)
Gordon Bearce
Nathan Maclean
REP
Vernon Stanley Albury
Richard Rigby
DEM
Janette Froman
Chris Aponte
LIB
Click to go back to instructions
←Previous Page
Click to go foward to next race
Next Page→


STEP 1
Read Instructions
You are now on
STEP 2
Make your choices
STEP 3
Review your choices
STEP 4
Record your vote
Proposition 2
Proposition 2
Shall the Charter of Harris County be amended to authorize the City Council to review and approve certain intergovernmental agreements and revenue contracts entered into by the City; to permit the City Council to establish its meeting schedule by ordinance; to clarify the circumstances in which the city council may act by ordinance or resolution; to permit the City Council to adopt by ordinance procedures for the formation and administration of special assessment districts; to permit excused absences of council members for reasons other than sickness; and to make other conforming amendments related thereto in order to eliminate redundant or obsolete provisions of the charter?
Yes
No
Click to go back to previous race
←Previous Page
Click to go foward to next race
Next Page→


STEP 1
Read Instructions
STEP 2
Make your choices
STEP 3
Review your choices
STEP 4
Record your vote
Record Vote
You can not make any changes once you click the 'Record Vote' button. When you click the button, your vote will be officially recorded.
If you want to make changes, click the 'Previous Page' button to go back to the Review Screen.
Click to go back to Review Screen.
←Previous Page
Click to record your vote.
Record Vote
Thank you for voting!
Your vote has been recorded. You may now leave the voting booth.

APPENDIX B

This appendix contains the bubble ballot used in the studies (reduced to fit on the page).



GENERAL ELECTION BALLOT
HARRIS COUNTY, TEXAS
NOVEMBER 4, 2006

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

PRESIDENT AND VICE PRESIDENT

PRESIDENT AND VICE PRESIDENT
(Vote for One)

- ◯ Gordon Bearce / Nathan Maclean — REP
- ◯ Vernon Stanley Albury / Richard Rigby — DEM
- ◯ Janette Froman / Chris Aponte — LIB

CONGRESSIONAL

UNITED STATES SENATOR
(Vote for One)

- ◯ Cecile Cadieux — REP
- ◯ Fern Brzezinski — DEM
- ◯ Corey Dery — IND

REPRESENTATIVE IN CONGRESS
DISTRICT 7
(Vote for One)

- ◯ Pedro Brouse — REP
- ◯ Robert Mettler — DEM

STATE

GOVERNOR
(Vote for One)

- ◯ Glen Travis Lozier — REP
- ◯ Rick Stickles — DEM
- ◯ Maurice Humble — IND

LIEUTENANT GOVERNOR
(Vote for One)

- ◯ Shane Terrio — REP
- ◯ Cassie Principe — DEM

ATTORNEY GENERAL
(Vote for One)

- ◯ Tim Speight — REP
- ◯ Rick Organ — DEM

COMPTROLLER OF PUBLIC ACCOUNTS
(Vote for One)

- ◯ Therese Gustin — IND
- ◯ Greg Converse — DEM

STATE

COMMISSIONER OF GENERAL LAND OFFICE
(Vote for One)

- ◯ Sam Saddler — REP
- ◯ Elise Ellzey — DEM

COMMISSIONER OF AGRICULTURE
(Vote for One)

- ◯ Polly Rylander — REP
- ◯ Roberto Aron — DEM

RAILROAD COMMISSIONER
(Vote for One)

- ◯ Jillian Balas — REP
- ◯ Zachary Minick — DEM

STATE SENATOR
(Vote for One)

- ◯ Ricardo Nigro — REP
- ◯ Wesley Steven Millette — DEM

STATE REPRESENTATIVE
District 134
(Vote for One)

- ◯ Petra Bencomo — REP
- ◯ Susanne Rael — DEM

MEMBER
STATE BOARD OF EDUCATION
District 2
(Vote for One)

- ◯ Peter Varga — REP
- ◯ Mark Baber — DEM

NONPARTISAN

PRESIDING JUDGE
Texas Supreme Court
Place 2
(Vote for One)

- ◯ Tim Grasty — DEM

PRESIDING JUDGE
Court of Criminal Appeals
(Vote for One)

- ◯ Dan Plouffe — REP
- ◯ Derrick Melgar — DEM

COUNTY

DISTRICT ATTORNEY
(Vote for One)

- ◯ Corey Behnke — REP
- ◯ Jennifer A. Lundeed — DEM

COUNTY TREASURER
(Vote for One)

- ◯ Dean Caffee — REP
- ◯ Gordon Kallas — DEM

SHERIFF
(Vote for One)

- ◯ Stanley Saari — REP
- ◯ Jason Valle — DEM

COUNTY TAX ASSESSOR
(Vote for One)

- ◯ Howard Grady — REP
- ◯ Randy H. Clemons — DEM

NONPARTISAN

JUSTICE OF THE PEACE
(Vote for One)

- ◯ Deborah Kamps — REP
- ◯ Clyde Gayton Jr. — DEM

COUNTY JUDGE
(Vote for One)

- ◯ Dan Atchley — REP
- ◯ Lewis Shine — DEM

PROPOSITIONS

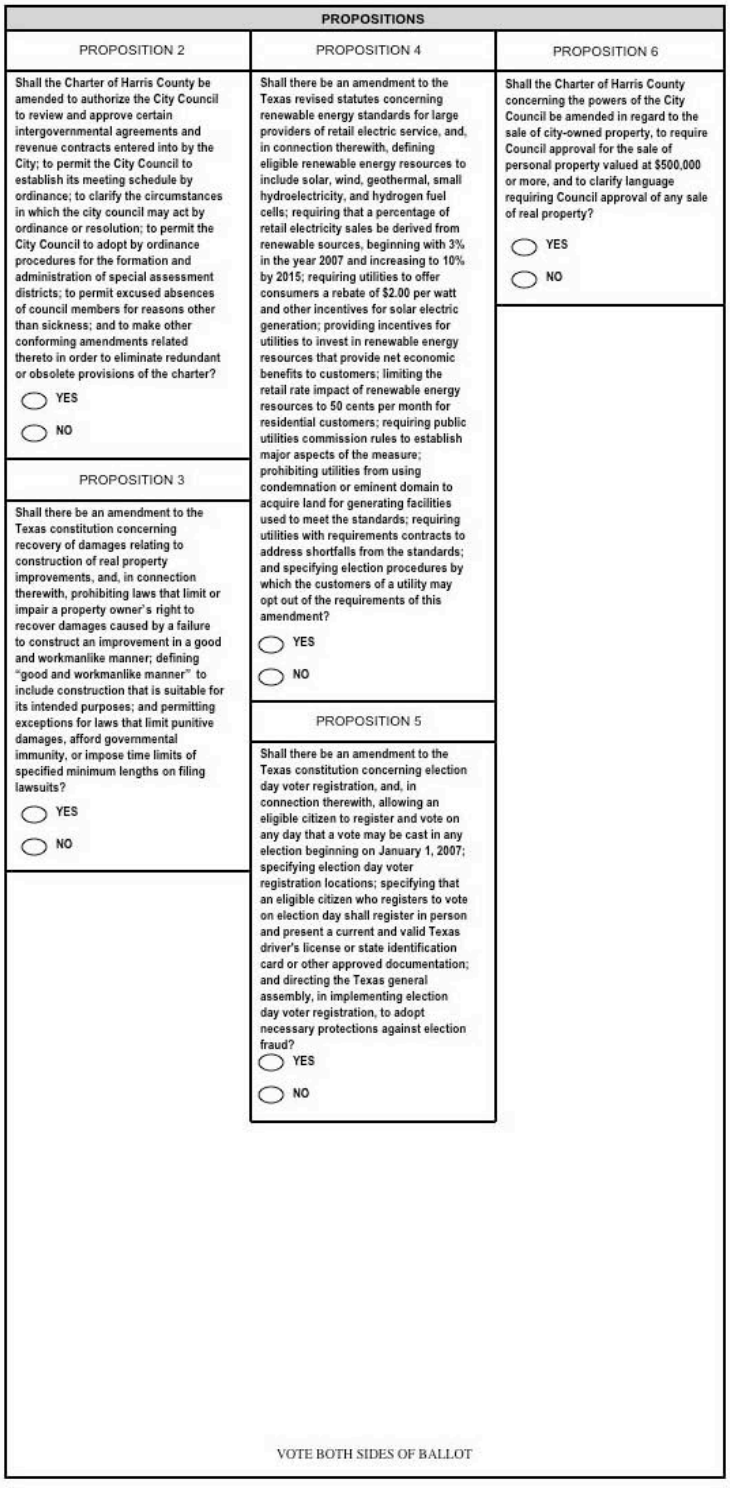
PROPOSITION 1

Without raising taxes and in order to pay for public safety, public works, parks and recreation, health care, libraries and other essential services, shall Harris County and the City of Houston be authorized to retain and spend all city and county tax revenues in excess of the constitutional limitation on total city and county fiscal year spending for ten fiscal years beginning with the 2005 fiscal year, and to retain and spend an amount of city and county tax revenues in excess of such limitation for the 2015 fiscal year and for each succeeding fiscal year up to the excess city and county revenue cap, as defined by this measure?

- ◯ YES
- ◯ NO

VOTE BOTH SIDES OF BALLOT

FRONT Card 3 SEQ# 1



PROPOSITIONS

PROPOSITION 2

Shall the Charter of Harris County be amended to authorize the City Council to review and approve certain intergovernmental agreements and revenue contracts entered into by the City; to permit the City Council to establish its meeting schedule by ordinance; to clarify the circumstances in which the city council may act by ordinance or resolution; to permit the City Council to adopt by ordinance procedures for the formation and administration of special assessment districts; to permit excused absences of council members for reasons other than sickness; and to make other conforming amendments related thereto in order to eliminate redundant or obsolete provisions of the charter?

◯ YES

◯ NO

PROPOSITION 3

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

◯ YES

◯ NO

PROPOSITION 4

Shall there be an amendment to the Texas revised statutes concerning renewable energy standards for large providers of retail electric service, and, in connection therewith, defining eligible renewable energy resources to include solar, wind, geothermal, small hydroelectricity, and hydrogen fuel cells; requiring that a percentage of retail electricity sales be derived from renewable sources, beginning with 3% in the year 2007 and increasing to 10% by 2015; requiring utilities to offer consumers a rebate of $2.00 per watt and other incentives for solar electric generation; providing incentives for utilities to invest in renewable energy resources that provide net economic benefits to customers; limiting the retail rate impact of renewable energy resources to 50 cents per month for residential customers; requiring public utilities commission rules to establish major aspects of the measure; prohibiting utilities from using condemnation or eminent domain to acquire land for generating facilities used to meet the standards; requiring utilities with requirements contracts to address shortfalls from the standards; and specifying election procedures by which the customers of a utility may opt out of the requirements of this amendment?

◯ YES

◯ NO

PROPOSITION 5

Shall there be an amendment to the Texas constitution concerning election day voter registration, and, in connection therewith, allowing an eligible citizen to register and vote on any day that a vote may be cast in any election beginning on January 1, 2007; specifying election day voter registration locations; specifying that an eligible citizen who registers to vote on election day shall register in person and present a current and valid Texas driver's license or state identification card or other approved documentation; and directing the Texas general assembly, in implementing election day voter registration, to adopt necessary protections against election fraud?

◯ YES

◯ NO

PROPOSITION 6

Shall the Charter of Harris County concerning the powers of the City Council be amended in regard to the sale of city-owned property, to require Council approval for the sale of personal property valued at $500,000 or more, and to clarify language requiring Council approval of any sale of real property?

◯ YES

◯ NO

VOTE BOTH SIDES OF BALLOT

BACK Card 3 SB3 1

APPENDIX C

The races and propositions used in the study are included in this appendix.

President

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

US Senator
Cecile Cadieux (R)
Fern Brzezinski (D)
Corey Dery (I)

US Representative
District 7
Pedro Brouse (R)
Robert Mettler (D)

Governor
Glen Travis Lozier (R)
Rick Stickles (D)
Maurice Humble (I)

Lieutenant Governor
Shane Terrio (R)
Cassie Principe (D)

Attorney General
Tim Speight (R)
Rick Organ (D)

Comptroller of Public Accounts
Therese Gustin (I)
Greg Converse (D)

Commissioner of General Land Office
Sam Saddler (R)
Elise Ellzey (D)

Commissioner of Agriculture
Polly Rylander (R)
Roberto Aron (D)

Railroad Commissioner
Jillian Balas (R)

Zachary Minick (D)

State Senator
Ricardo Nigro (R)
Wesley Steven Millette (D)

State Representative(district 134)
Petra Bencomo (R)
Susanne Rael (D)

Member, State Board of Education, District 2
Peter Varga (R)
Mark Baber (D)

Presiding Judge, Texas Supreme Court, Place 2
Tim Grasty (D)

Presiding Judge, Court of Criminal Appeals
Dan Plouffe (R)
Derrick Melgar (D)

District Attorney
Corey Behnke (R)
Jennifer A. Lundeed (D)

County Treasurer
Dean Caffee (R)
Gordon Kallas (D)

Sheriff
Stanley Saari (R)
Jason Valle (D)

County Tax Assessor
Howard Grady (R)
Randy H. Clemons (D)

Justice of the Peace
Deborah Kamps (R)
Clyde Gayton Jr. (D)

County Judge
Dan Atchley (R)
Lewis Shine (D)

Proposition 1
Without raising taxes and in order to pay for public safety, public works, parks and recreation, health care, libraries and other essential services, shall Harris County and the City of Houston be authorized to retain and spend all city and county tax revenues in excess of the constitutional limitation on total city and county fiscal year spending for ten fiscal years beginning with the 2005 fiscal year, and to retain and spend an amount of city and county tax revenues in excess of such limitation for the 2015 fiscal year and for each succeeding fiscal year up to the excess city and county revenue cap, as defined by this measure?

Proposition 2
Shall the Charter of Harris County be amended to authorize the City Council to review and approve certain intergovernmental agreements and revenue contracts entered into by the City; to permit the City Council to establish its meeting schedule by ordinance; to clarify the circumstances in which the city council may act by ordinance or resolution; to permit the City Council to adopt by ordinance procedures for the formation and administration of special assessment districts; to permit excused absences of council members for reasons other than sickness; and to make other conforming amendments related thereto in order to eliminate redundant or obsolete provisions of the charter?

Proposition 3
Shall there be an amendment to the Texas constitution concerning recovery of damages relating to construction of real property improvements, and, in connection therewith, prohibiting laws that limit or impair a property owner's right to recover damages caused by a failure to construct an improvement in a good and workmanlike manner; defining "good and workmanlike manner" to include construction that is suitable for its intended purposes; and permitting exceptions for laws that limit punitive damages, afford governmental immunity, or impose time limits of specified minimum lengths on filing lawsuits?

Proposition 4
Shall there be an amendment to the Texas revised statutes concerning renewable energy standards for large providers of retail electric service, and, in connection therewith, defining eligible renewable energy resources to include solar, wind, geothermal, small hydroelectricity, and hydrogen fuel cells; requiring that a percentage of retail electricity sales be derived from renewable sources, beginning with 3% in the year 2007 and increasing to 10% by 2015; requiring utilities to offer consumers a rebate of $2.00 per watt and other incentives for solar electric generation; providing incentives for utilities to invest in renewable energy resources that provide net economic benefits to customers; limiting the retail rate impact of renewable energy resources to 50 cents per month for residential customers; requiring public utilities commission rules to establish major aspects of the measure; prohibiting utilities from using condemnation or eminent domain to acquire land for generating facilities used to meet the standards; requiring utilities with requirements contracts to address shortfalls from the standards; and specifying election

procedures by which the customers of a utility may opt out of the requirements of this amendment?

Proposition 5
Shall there be an amendment to the Texas constitution concerning election day voter registration, and, in connection therewith, allowing an eligible citizen to register and vote on any day that a vote may be cast in any election beginning on January 1, 2007; specifying election day voter registration locations; specifying that an eligible citizen who registers to vote on election day shall register in person and present a current and valid Texas driver's license or state identification card or other approved documentation; and directing the Texas general assembly, in implementing election day voter registration, to adopt necessary protections against election fraud?

Proposition 6
Shall the Charter of Harris County concerning the powers of the City Council be amended in regard to the sale of city-owned property, to require Council approval for the sale of personal property valued at $500,000 or more, and to clarify language requiring Council approval of any sale of real property?

APPENDIX D

This appendix contains the full voter guide used in the studies.

TEXAS
2006 GENERAL ELECTION
VOTER GUIDE

*****NOTE TO PARTICIPANTS: THIS VOTER GUIDE HAS BEEN COMPILED SOLELY FOR RESEARCH PURPOSES AND IS NOT INTENDED TO REFLECT THE VIEWS OF RICE UNIVERSITY OR OF THE RESEARCHERS ASSOCIATED WITH THIS STUDY. IT ALSO IS NOT INTENDED TO DEPICT REAL PEOPLE. *****

TEXAS GENERAL ELECTION
TUESDAY, NOVEMBER 4, 2006
POLLS OPEN 7:00a.m. to 7:00p.m.

CANDIDATES FOR PRESIDENT/ VICE PRESIDENT

Questions:
1. Do you support the 9-11 Commission's recommendations regarding reorganization of Congressional Intelligence Committees?
2. How do you propose to reduce the federal deficit?
3. What role should the federal government play in providing adequate health care for all Americans?

Gordon Bearce, Republican
(Nathan Maclean)

Background: BA Stanford University 1971; MBA Harvard University 1974. 14 years of service in the Navy Reserve. Governor of Missouri since 1992.

1. Yes. Currently, numerous committees oversee the many different areas of intelligence needed to make our homeland secure. We need to follow the 9-11 Commission's recommendations to help coordinate these committees and reorganize them so they can operate and pass along information efficiently.

2. In order to reduce the federal deficit, we must have long-term fiscal discipline. We should not borrow from Social Security or Medicare or any other programs, rather simply reduce government spending in other areas, such as areas where supplemental appropriations are routinely passed. I will not raise taxes, however tax cuts that solely benefited the wealthy should be repealed, as this simple act will help reduce the government deficit by over 1 trillion dollars.

3. We should take a strong step forward in helping all Americans get the adequate health care they deserve. We should expand Medicare and Medicaid to help cover those that are currently in need of quality health care, but aren't getting it. Everyone has a right to quality health care, and under my watch, I will ensure all Americans get what they deserve.

Vernon Stanley Albury, Democrat
(Richard Rigby)

Background: BA Princeton University 1967; JD Yale University 1970. District Attorney 1972-1982; Member of US House of Representatives since 1985.

1. Yes. We should restructure Congressional Intelligence Committees to help manage intelligence in a more expedient and precise manner, however complete reorganization will not help. We must restructure these committees and redefine their purposes, rather than simply removing some and adding power to others. This is a careful process that we need to take, allowing the Congressional Committees to cooperate and allow the Department of Homeland Security to oversee this intelligence.

2. I propose to reduce the federal deficit by controlling rampant supplemental appropriations bills. Too many congressmen and women are too concerned about giving their own districts money rather than looking out for the good of the entire nation. I will veto any supplemental appropriations bills that do not have a supermajority of the house behind them, and encourage fiscal discipline wherever I can.

3. The federal government should help provide adequate health care for all Americans. We should restructure Medicare and Medicaid so that they operate more efficiently and give Americans the medical coverage they need. Providing adequate health care does not mean

simply pumping money into these programs—we need to ensure that those in need get what they need, and not be short-changed by the bureaucracy.

Janette Froman, Libertarian
(Chris Aponte)
Background: BS Texas A&M 1980; JD University of Houston Law School 1984; Prior Candidate for Texas House of Representatives and Texas State Senate.

1. No. No amount of reorganization can fix the mess that the past few administrations have created. We need to rebuild our intelligence committees from the ground up—and establish term limits in the House so that those responsible for this disorganization are out of office.

2. I plan to fix the federal deficit by immediately cutting the Department of Defense's budget drastically. Their expenditures account for a large plurality of our government spending, and their rampant use of government funds needs to be curtailed.

3. None. It is not the federal government's responsibility to provide health care to all Americans. Too many take advantage of the system, and this problem has helped to exacerbate our federal deficit. Medicare and Medicaid ought to be repealed.

CANDIDATES FOR US SENATOR

Questions:
1. What changes, if any, need to be implemented in US free-trade policies?
2. A number of criticisms have been aimed at the Medicare prescription coverage program. What modifications, if any, would you support?
3. What, if anything, would you change about "No Child Left Behind"?

Cecile Cadieux, Republican
Background: JD, University of Texas 1985 – LLM in Taxation, University of Florida 1989, Authored or co-authored 14 professional articles; Married, one child

1. Chinese goods should be tariffed to cause their prices to be what they would be but for attachment of the yuan to the dollar. (China's currency has been attached to the dollar since 1995.) Attachment has prevented US manufacturers from being able to compete, thus causing loss of U.S. jobs.

2. The Program should be repealed and HHS should be directed to negotiate with the pharmaceutical companies to provide our seniors with the prices that are charged to western European and Canadian seniors. Catastrophic coverage should exist, but it should be funded by small Medicare Part A/B benefit reduction.

3. Test scores have not been improved since the federal Department of Education was created in 1979. Three levels of government is enough. Debts and unfunded liabilities of the federal government total $330,000 per full-time worker. I would dismantle the DOE.

Fern Brzezinski, Democrat
Background: I am a businesswoman, family woman, and public servant. I have been a business and political leader in Georgia for over 30 years, and I currently serve in the US House of Representatives. I am proud of my family, and I have 3 children and 4 grandchildren.

1. Our biggest challenge to our Free Trade Agreements is to make sure US Trade Representatives enforce the rights of US companies through the World Trade Organization.

2. The first phase of the Medicare Modernization Act has gone very well with the implementation of the Discount Drug Cards for seniors. The main provisions of the Act do not take effect until 2006. Any modification should only be considered after implementation in 2006.

3. As an original coauthor of NCLB, we are constantly monitoring its progress. We have already modified provision for testing of special education children and non-English speaking children. We must refine the "highly qualified teacher" provision, particularly in Special Education instruction.

Corey Dery, Independent
Background: I have a BA in Political Science from Yale University, and a JD from Duke University School of Law. I have served as a law clerk for the Texas Court of Appeals.

1. Trade agreements must guarantee that the US can act to protect workers from rapid changes in the international marketplace. I will carefully evaluate all trade agreements to ensure that they adequately protect the internationally recognized rights of workers including the right to organize and collectively bargain.

2. The Bush Administration's prescription drug plan must be changed so that our senior citizens can obtain prescription drugs at an affordable price. We should permit the government to negotiate with drug companies for fair prices for Medicare beneficiaries. We should also allow the re-importation of cheaper prescription drugs from other countries.

3. High quality education for our children is critical to the future of our economy and will give us a skilled and competitive workforce. As a member of the House Education and Workforce Committee, I have fought to fully fund Head Start, No Child Left Behind, and other important education initiatives.

CANDIDATES FOR US REPRESENTATIVE

Questions:
1. Do you support the 9-11 Commission's recommendations regarding reorganization of Congressional Intelligence Committees? Please explain.
2. What role should the federal government play in providing adequate health care for all Americans?
3. How would you address the growing federal deficit?:
4. What is your position on renewing and/or expanding the US Patriot Act?

Pedro Brouse, Republican
Background: Education: B.A. Accounting, University of Texas, Austin; Experience: Auditor/accountant-Texas Department of Public Welfare (1973-1977), US Navy (1979-1983), Initial Rentokil USA, Inc (1983-2004)

1. Congress should play a greater role in oversight.

2. I am very concerned about inadequate planning for seniors and veterans. Millions of Texas families are without health insurance…it is tragic that so many children are left out and so many Americans of the "greatest generation"—seniors and veterans, most of whom are over 80 years old----are left behind, when all of us in the younger generations owe the World War II generation so much.

3. Inadequate management of the budget and the economy has created this problem for our future. I am proposing a more responsible foreign/defense policy to address budgeting…and new legislation to address large/multi-nationals that "outsource" and go "offshore"…our renewed emphasis on economic development and lowering the tax burden on Americans who have the least income will help.

4. It should not have been renewed, but rather revised to accomplish cooperation within our US law enforcement system while respecting our cherished US Constitution and Bill of Rights…undermining our rights, liberties, and freedoms does not enhance security, it diminishes our great American democracy.

Robert Mettler, Democrat
Background: Education: Graduate, Senior Executive Fellows Program, Harvard University. J.D., St. Mary's Law School. B.S., Trinity University; Experience: Chief, Terrorism and National Security, US Attorney's Office; Bush-Cheney transition team member; Attorney General Greg Abbott transition team member; Deputy Attorney General for Criminal Justice under John Cornyn; Trial attorney, Public Integrity Section, US DOJ

1. I support the Commission's recommendations on Congressional Intelligence Committees. Today, Congressional Intelligence gathering is spread over several committees making it more difficult for Congressional leaders to address the key issues that will define and determine our success in the war on terror. By consolidating the Committee structure, we help create one area where key security issues can be fully and completely examined.

2. The best possible health care system will be driven by consumer choice; where patients and physicians can make decisions about appropriate care. Our current system, both public (Medicare/Medicaid) and private (HMOs), limits choice and drives up costs and must be reformed. Additionally, we must pass legislation to end runaway litigation that forces doctors to practice "defensive medicine," increasing costs and hampering development of cutting edge procedures and medicines while depriving Americans of the best health care possible.

3. Federal spending is driven by government bureaucracies and wasteful programs that are systematically funded, year after year, through massive "omnibus" spending bills which virtually no one actually reads, especially those in Congress. I strongly favor a Federal Agency "Sunset" Law so that each bureaucracy and every single funded program must justify its existence. This system in Texas has saved millions of dollars, and it is time we made Washington more closely account for every expenditure.

4. No matter the threat, America must protect our civil liberties enshrined in the Bill of Rights. If we curtail civil liberties to fight terrorism, the terrorists win. However, our laws must keep up with the times, allowing us to investigate, disrupt and prosecute terrorists before they destroy critical

infrastructures. I support renewing the Patriot Act because it does just that: it takes existing legal principles and retrofits them to address the particular challenge of terrorism.

CANDIDATES FOR GOVERNOR

Questions:
1. What is your first priority as Governor?
2. How would your budget reflect support for environmental measures?
3. How would you improve and finance transportation?

Glen Travis Lozier, Republican

Biography: BA, Texas 1977; JD Georgetown 1980; As Attorney General, I have focused on the security of Texans, including domestic violence and protecting children. A former state and federal prosecutor, I have also served as Secretary of Public Safety.

1. As Governor, I want to create a Texas filled with opportunity. To do this, we must have better pay for better teachers so that our children get a better education. We must empower Texans to have more control over their healthcare options through health savings accounts and long term care incentives. And I will continue my efforts to combat domestic violence and gang activity.

2. As Governor, I will pursue responsible environmental policies to benefit future generations by employing a stewardship based model for governing our natural resources and environmental assets, emphasizing collaboration and citizen involvement; recommitting our state to pollution prevention; and creating an environmental enforcement team to target those who harm the environment through purposeful or grossly negligent actions.

3. As Governor, I will lead the way to innovative transportation solutions that empower Texans and work to reduce congestion by creating Regional Transportation Authorities to develop and implement solutions to regional transportation problems. I will use prioritize the use of technology on our roadways to make them less congested.

Rick Stickles, Democrat

Biography: BS, Rice 1975; JD Texas 1980; My life has been shaped by my parents, family, children, faith, and my community. Working in my father's firm, as a civil rights lawyer, and later as Mayor and Lt. Governor taught me to value strong communities, equal opportunity, hard work, fiscal discipline and finding common ground.

1. Education. Our teachers deserve better pay, and our schools can be made better simply by an emphasis on education in our state budget. I will raise standards and expect nothing less than excellence in the classroom and in recruiting the nation's best teachers.

2. We owe it to our children to leave them this beautiful state as we found it. Budget reform will allow us to make historic investments in environmental programs. We should value clean air and a clean environment, and through budget reform, we can achieve these.

3. We need a new approach to reduce traffic. We cannot simply tax and pave our way out of the problem. I will work to fix the hole in the transportation bucket by vetoing any diversion of Transportation funds. I will create incentives to better connect land-use and transportation decisions to reduce traffic and sprawl.

Maurice Humble, Independent
Biography: I have a BA in Economics and a JD from Texas; I am currently serving my fourth term in the Texas State Senate, and I chair the Education and Health Committee. I value my family and my three daughters, and the community I live and work in.

1. My first priority as governor would be to implement a comprehensive solution to the state's transportation problems. The state also has several other important issues that need to be addressed—including education, tax reform, and health care.

2. As a state senator, I have been a strong advocate for the environment. I have worked to provide $15 million each year for air quality improvement. I will continue to fight for environmental improvement across the great state of Texas.

3. We have a crisis on our hands that needs to be fixed—I am the only gubernatorial candidate willing to recognize this fact. We need a radical approach to fixing our transportation problems, including bolstering our transportation budget and tackling the issue at the state level, rather than with regional authorities.

CANDIDATES FOR LIEUTENANT GOVERNOR

Questions:
1. How do you see yourself functioning in the role of Lieutenant Governor?:
2. How would you influence the dynamics of the legislative process?:
3. What would you like the citizens of Texas to know about you?:

Shane Terrio, Republican
Biography: Occupation: Consultant with Riggs, Counselman, Michaels, and Downes, a Houston-based insurance agency. Education: BA, Political Science, Texas, 1979. Experience: Texas State Senate 1996-present

1. The Lieutenant Governor's statutory responsibilities include presiding over the Senate of Texas and chairing a number of state commissions. With ten years experience in the State Senate, I can easily fulfill these responsibilities. I also look forward to working with others to take a leadership role in a number of state programs, including efforts to reform Medicaid and make quality health care available to every Texan.

2. During my ten years in the State Senate I have built strong personal relationships with other legislators from both political parties. I have been recognized as one of the most effective members, and I have proven my ability to work with people who hold competing views on important issues and fashion sound public policies for Texas. I will continue to do that as Lieutenant Governor.

3. I have the background, knowledge, and experience in state government that is necessary to help lead Texas. I have also articulated a clear vision for the future of Texas—a vision that creates a pro-business environment and a commitment to invest the resources that economic growth generates in the core responsibilities of state government including transportation,

education, public safety, healthcare, and responsible efforts to protect our important natural resources.

Cassie Principe, Democrat

Biography: I've served Texas for 12 years in the legislature (both in the Senate and the House). I have a BA in Political Science from the University of Texas. I am a small business owner, and I am proud of my two grown children and my one grandchild.

1. The Lieutenant Governor presides over the Texas Senate. I will work closely with the Senate to continue the progress and build on the fiscal responsibility of the previous administration.

2. I believe that governing is not about finding fault but finding solutions. During my legislative career, I have proven the ability to reach out to those across the aisle to seek consensus on the important issues facing Texas, issues like education, transportation, the wise use of environmental resources, affordable health care insurance, and building a culture of freedom and personal responsibility.

3. I believe government must treat all its citizens with fairness, dignity and respect. My philosophy on government is that a representative has an obligation to listen, to have an open door for all people-including those who agree with and those who do not. I have fought for twenty years in Texas to build better communities, make our highways safer, provide tax relief and broaden educational opportunity, I have consistently been a voice for those who cannot afford to hire lobbyists; I consider myself "the people's lobbyist". This is how I approached my service on behalf of Texans at the federal, state and local level. I am eager to bring this effective experience to the job of being your Lieutenant Governor.

CANDIDATES FOR ATTORNEY GENERAL

Questions:
1. What do you want to accomplish as Attorney General?:
2. What potential do you view in this office?:

Tim Speight, Republican

Background: I am a retired U.S. Army officer, a former prosecutor, and a 14 year member of the Texas House of Representatives. I have earned degrees in Business, Management, Public Policy, and the Juris doctor.

1. I will crack down on violent sexual predators who target our children by enacting much tougher penalties for sex offenders, revamping the sex offender registry, requiring sex predators to register with State Police before being released from prison, monitoring sex offenders with GPS tracking systems, and other legal reforms. Other key priorities include strengthening efforts to protect Texans from identity theft, protecting Texas from terrorist threats, fighting drugs and gangs, implementing a family court system, protecting private property rights, and protecting Texas' pro-jobs environment by working to end lawsuit abuse and reducing regulations.

2. Our next Attorney General must have the experience to get the job done for our citizens from day one. As an army veteran who served in Europe during the Cold War, a local prosecutor who put murders, child molesters, and rapists behind bars, a proven legislator who played a key role

in abolishing parole for violent criminals and passing historic welfare reform, I bring the experience we need to this important office.

Rick Organ, Democrat
Background: BA, Texas, 1970; JD, Texas 1977; I have previously served the public as a District Attorney, and I have served in the Texas House of Representatives for 10 years.

1. In this post-9/11 world, I believe the next attorney general's top priority must be keeping Texas safe and secure. I will use the office to advocate for public safety and to pursue my security agenda. But the AG is also responsible for providing the best legal advice to the governor and legislature, and I believe that should be done promptly and without a partisan political agenda.

2. Texas needs an attorney general who is an advocate for all the people, not just the powerful. I believe the office can be a powerful force for reducing prescription drug prices, consumer fraud and identity theft. Also, I plan to work with the Department of Social Services to close the $2 billion child support gap.

CANDIDATES FOR COMPTROLLER OF PUBLIC ACCOUNTS

Question:
1. What will you do to "provide a window into Texas government"?

Therese Gustin, Independent
Training and Experience: I have a BA in Accounting from the University of Houston, and I am a Certified Public Accountant. I have worked in the Texas Comptroller's office for the past 15 years, and I am confident I can run this office better as the Comptroller.

1. If elected, I will work to audit and ensure that every Texas agency is spending money like it should and is being held accountable. I would make sure that government regulations are based on common-sense and that every agency is abiding by them.

Greg Converse, Democrat
Training and Experience: I am a Certified Public Account, and I received a BA in Accounting from the University of Texas, and an MBA from Rice University. I have worked for the Texas Treasury Department for the past 10 years.

1. The Comptroller's office should shed light on all the other bureaucracy and government in Texas, ensuring that everything is working properly. If elected, I will help the Texas government to run a smaller, more efficient operation, ensuring that no taxpayer's money is misused.

CANDIDATES FOR COMMISSIONER OF GENERAL LAND OFFICE

Question:
1. What will you do as Commissioner to uphold the General Land Office's responsibilities to protect natural resources?

Sam Saddler, Republican
Training and Experience: BS in Geology from Texas A&M in 1981. I have worked for the Texas General Land Office for the past 20 years. I am proud to work for the oldest state agency in Texas, and I have experience with all the intricacies of this office, therefore I believe I am qualified to be Commissioner.

1. One of the General Land Office's duties is to protect the natural resources that belong to our state. I will work closely with the Office of the Railroad Commission to ensure that our state's oil and gas deposits are taken care of. I will ensure that Texas' interests are at heart in these decisions, not local business interests.

Elise Ellzey, Democrat
Training and Experience: I have a BS in Petroleum Engineering from Louisiana State University. I have worked for Exxon as an engineer, and I have worked for the Texas Railroad Commission.

1. I will ensure that our natural resources are protected and that all the proper proceeds are given to the Permanent School Fund, to ensure that our children get the monies they deserve from drilling rights in this state. I will ensure that all contracts are handled appropriately.

CANDIDATES FOR COMMISSIONER OF AGRICULTURE

Question:
1. What can be done to revitalize Texas' agriculture industry?

Polly Rylander, Republican
Training and Experience: I have served two terms in the Texas House of Representatives, and I have a BA from the University of Houston, and an MBA from the University of Texas. I grew up on a farm, and I have worked within the agriculture industry for the past 10 years.

1. Marketing for Texas' agriculture products tops my list of priorities as Commissioner of Agriculture. If elected, I plan to help revitalize our extensive agriculture industry by promoting our products nationwide.

Roberto Aron, Democrat
Training and Experience: BS, Texas A&M 1975; MBA University of Houston, 1981; I have worked closely with the agriculture industry for the past 20 years, including working in New York in the financial markets.

1. With the Texas Department of Agriculture backing our state's industry, there is no need to revitalize it. Texas has one of the strongest agriculture exports of any state, and, if elected, I plan to help continue making Texas' agriculture industry successful.

CANDIDATES FOR RAILROAD COMMISSIONER

Questions:
1. How would you prioritize the goals of the Railroad Commission's Strategic Plan for 2005-09 in light of limited funding?
2. How do you propose to meet the Railroad Commission's stated responsibility for supporting research, education, training, and marketing of clean-burning alternative fuels?

Jillian Balas, Republican

Training and Experience: Geologist, petroleum geophysicist and energy attorney. Texas Railroad Commissioner since February 2003. Elected Chairman by colleagues. Former petroleum geophysicist for Amoco Production. Energy attorney at the General Land Office. Assistant Abilene city attorney; political science and legal studies instructor, Hardin-Simmons University. Elected Abilene City Councilman and Taylor County Judge.

1. The top goal of the Texas Railroad Commission is to strengthen the safety and productivity of the Texas energy industry. In this era of record high oil prices, we must reduce dependence on foreign oil, increase responsible energy production, and promote conservation and renewable energies such as wind, fuel cell and biomass energy. Since joining the Railroad Commission, I have helped reduce the agency budget, while improving safety and environmental quality in the energy sector.

2. As Chairman of the Texas Energy Planning Council, I worked hard to promote alternative energy sources. I have visited Texas wind farms and emerging technologies which promise to reduce dependence on foreign energy and improve environmental quality. My goal is to ensure emerging energy technologies are conceived and built in Texas, taking advantage of our vast expertise and infrastructure. The Railroad Commission also uses grants funds to promote cleaner burning fuels, such as propane.

Zachary Minick, Democrat

Training and Experience: Born and reared in west Texas. Degrees from Baylor, Southwestern Seminary, Yale, and the University of Illinois. Experienced in personal business development. Experienced in formulation, support, and implementation of public policy at the local, state, and national level. Experience in the negotiation and management of mineral properties.

1. The Commission's Strategic Plan for 2005-09 indicates it "does not expect significant changes in its mission, strategies, or goals during the next five years." The development of our oil and gas resources is primary. Safety and environmental concerns are secondary. Scant attention is given to alternative energy. No attention is given to monitoring intrastate natural gas transmission. The public's growing concern about the relationship between energy development and the environment needs a higher priority.

2. There may be an inherent conflict of interest in making a Commission devoted to the development of oil, gas and coal resources responsible for developing "clean-burning alternative fuels." A much broader range of knowledge, concern, and experience as well as a broader range of interests need to be involved. If this project is to remain the Commission's responsibility, it would have to greatly expand its knowledge base and staff.

CANDIDATES FOR STATE SENATOR

Questions:
1. What solutions would you propose to balance the state budget?
2. Should state funding for Public Education be expanded?
3. How do you propose to fund healthcare for the large number of uninsured in Texas?

Ricardo Nigro, Republican

Background: Education: B.B.A. from University of Texas-Austin, J.D. from South Texas College of Law; Experience: State Senator 2003-present; Travis County Commissioner 1998-2001; former Chief Clerk, Senate Committee on County Affairs; former Chief Clerk, Senate Joint Interim Committee on Regional Issues; former member of the Texas Open Records Steering Committee; former General Counsel for Senator Jeff Wentworth, and the Senate Interim Committee on Public Information.

1. I am a fiscal conservative and believe general government should be smaller and smarter. Last session we had a $10 billion budget deficit. The deficit was a spending problem, not a revenue problem. Citizens should not be asked to pay more in taxes due to the deficit. Government should do what families do: set priorities and live within a budget. That's why I helped pass a balanced budget without a tax increase.

2. Public Education is my top priority. State funding should be increased to improve educational standards and to abolish the need for the current Robin Hood school finance system. Even in the face of a $10 billion budget deficit last session, I supported $1.2 billion of additional investment in public schools. I also supported amendments to increase investment in textbooks, pre-kindergarten and kindergarten classes, and teacher retirement benefits.

3. It is important that the legislature create opportunities for more affordable and flexible market alternatives for health care coverage. Last session we created "Consumer Choice Health Plans" that will allow many currently uninsured Texas men, women and children to get the health care coverage that they could not afford prior to the passage of this legislation. Under this law, many small businesses will be able to provide coverage to employees and their families.

Wesley Steven Millette, Democrat

Background: Education: I have a Masters in Social Work and law degree from the University of Texas, and a B.A. in political science from Queens College. Experience: My experience includes seven terms in the Texas House, passing over 150 bills including the Landlord-Tenant Security Devices, Indoor Air Quality, Nursing Home Reform, and Mold Remediation Licensure acts. I served on the Public Health Committee, Human Services Committee, and Select Committee on Child Welfare and Foster.

1. To balance the budget, I'd close the loophole in the corporate franchise tax so limited liability partnerships pay their fair share; expand the sales tax base to include certain services; increase the cigarette tax, and/or amend the Texas Constitution t allow imposing a statewide property tax. I'd consider instituting a state income tax, if linked to restructuring our tax system so property and sales taxes are significantly reduced.

2. Yes. The state's contribution to public education has fallen below 40%, resulting in an increased reliance on local property taxes. This situation led Judge Dietz to rule that our system doesn't provide an "adequate" education, since almost half our school children under-perform. The ruling has been interpreted to mean that the state must come up with the substantial new money over and above the funds needed to offset a reduction in property taxes.

3. To fund health care for the large number of uninsured in Texas, I'd restore the cuts to the Children's Health Insurance Program and Medicaid, thus maximizing the receipt of federal matching funds. I'd institute a one dollar increase in the cigarette tax and dedicate the revenues to health services. I'd close the loophole in the corporate franchise tax so limited liability partnerships pay their fair share and dedicate a portion of the revenues to health care.

CANDIDATES FOR STATE REPRESENTATIVE

Questions:
1. Do you believe that changes or improvement should be made in the Texas public health care system?
2. Do you believe that additional revenue sources are needed to meet the needs of Texas residents? If so, please identify possible sources.
3. Given Texas' low national rating on education performance, what should be done to raise our standing?

Petra Bencomo, Republican

Qualifications: I received my B.A. from the University of Houston and J.D. from the University of Texas. I am an attorney at ConocoPhillips. I have worked three continuous legislative sessions (1998-2004). I have also served as Rep. Farrar's Chief of Staff and Rep. Moreno's campaign manager in the 2000 Democratic Primary.

1. We need increased funding for clinics that provide preventive healthcare. This would help relieve the overcrowding in emergency rooms and prevent hospital stays. We also need to increasing funding for children's healthcare programs, such as CHIP. Additionally, the state should use its purchasing power to reduce prescription costs.

2. Texas needs a fair, broad based business tax that reflects modern economy. We need to close the business tax loopholes and ensure that all companies pay equally. Additional revenue sources should not target those least able to pay, such as a regressive sales tax. We need a fair and equitable tax revenue system

3. Out Legislature needs to answer the funding needs highlighted by Judge Dietz. We need more funding for our schools to ensure that our students have the resources they need to learn and teachers have the resources they need to teach. We also need a teacher pay raise in order to recruit and retain qualified teachers.

Susanne Rael, Democrat
Qualifications: I will use my 35 years of legal, legislative and judicial experienced leadership and proven service as a former judge for city of Houston-Harris County, attorney, certified mediator and arbitrator, wife and mother, to make our schools better, our neighborhoods safer and improve our economy for families.

1. Every system should be reviewed constantly to maximize the resources being used to see how and where more efficiency for the delivery of services can be accomplished. I will continue to work with the legislature to ensure Texas' public health care system provides the care and services required by all Texans while recognizing the financial requirements of such a system.

2. My commitment is to the families of this District; to ensure everyone has an opportunity to receive a quality education, affordable healthcare, and to work to the fulfillment of the American Dream. As your State Representative, I will continue to seek the most effective and efficient manner to make these opportunities available to the families of this District.
3. In the next legislative session, I will continue to use my years of legislative experience to ensure all children have the resources necessary to receive a quality education at the highest level and our school teachers are paid a reasonable salary for the hard work. I will work with other legislators to ensure this effort is achieved.

CANDIDATES FOR MEMBER, STATE BOARD OF EDUCATION, DISTRICT 2

Questions:
1. How can schools effectively recruit and retain quality teachers?
2. What can be done about schools that have been rated "Academically Unacceptable"?

Peter Varga, Republican
Background: As a self-employed father of three, wife of a firefighter and Iraqi Freedom veteran, I am presently a UH Consumer Science/Teacher Certification applicant after earning an Associates Degree at HCCS. My 20 years of community service established the foundation for my commitment to a new direction for our schools.

1. Energetic recruitment and retention efforts should include an accelerated hiring timeline, active marketing campaigns, college and university partnerships, new teacher mentorship programs, professional development on classroom management, classroom routines and procedures, multicultural education, and lesson planning, paid summer orientations, maintain reduced classroom size, enforcement of disciplinary policies, placement of trained principals with management skills that promote teacher retention.

2. Student learning turns around all school ratings. Children's learning is promoted through the learning style of each child. A high teacher-student interaction can raise the level of learning. One cohesive team of the faculty, staff and principal as the instructional leader and manager who is supposed by strong parental and community groups can achieve a clearly defined shared vision of achievement.

Mark Baber, Democrat
Background: Director, Mayor's Citizens' Assistance Office (since 5/2000), Houston Parks and Recreation (1992-2000), Houston Community College (1987-1992); Precinct 105 Chairman, (since 1996), past president – Hawthorne Civic Club, Honors Diploma- Jeff Davis High School, A.A. Government – HCC, B.A. Political Science – University of Houston. Married seventeen years, father of three daughters.

1. Teachers are our most precious resource, so we must treat them as professionals and pay them like we are serious about quality education for our kids. We must both maintain standards and allow flexibility in teaching. We must let committed teachers teach what they know. We must provide quality environments where teachers want to teach and students want to study.

2. Schools with extraordinary challenges require extraordinary resources and commitment. We must provide special incentives to attract the most qualified and talented educators and to provide the best equipment and buildings. The community's stakeholders must also be actively engaged in helping to do their part. Parents, local community and business leaders, all of us, can and must turn our schools around.

CANDIDATES FOR PRESIDING JUDGE, TEXAS SUPREME COURT, PLACE 3

Questions:
1. What do you think the community can do to assist the judiciary in making decisions that protect women, their children and the community against family violence?
2. The U.S. Supreme Court has decided to hear a Minnesota dispute over whether judicial candidates can discuss their positions on issues that might come before their courts. Would you welcome a ruling that allowed you to freely comment on these issues?
3. How could we strengthen communications with the legal system when family is dealing with multiple courts and proceedings?

Tim Grasty, Democrat
Training and Experience: I have practiced trial law since 1981. I have never been sanctioned. I represent individuals, businesses, hospitals and educational institutions. I am active in delivering legal services to the poor. I am a mediator. I serve on a hospital board and volunteer through church, schools, and youth organizations.

1. Personal involvement with, and financial support of, prevention programs, assistance efforts and shelters is critical. Many such entities work with the courts. Citizens must press the legislature for appropriate action to address these problems. The court benefits when citizens willingly serve as jurors. The courts are open, be there.

2. No. Our government depends on objective, impartial and constitutionally constrained judges. Such a decision could overly politicize an already challenging selection process. Judges must decide each case on the facts and applicable law. The expression of opinions in the political context could suggest a predisposition or bias about certain cases.

3. The current presiding court system could be changed to allow a single court to handle a matter from filing to final disposition. Regardless, each file should be accurately documented as to activity and action. The courts provide forms which permit contemporaneous documentation. Judges should require attorneys to promptly complete filings.

CANDIDATES FOR Presiding Judge, Court of Criminal Appeals, Place 2

Questions:
1. Do you believe the composition of juries adequately and fairly reflects society at large? Why or why not?
2. What changes, if any would you support to assure that the rights of the legally indigent are adequately protected under current law and practice, particularly in death penalty cases?
3. While serving on the bench, do you believe you have a role in bringing important legal or judicial issues before the public or the legislature? Why or why not?

Dan Plouffe, Republican

Qualifications: Senior Judge, Texas Court of Criminal Appeals, 11 year member Associate Justice, Second Court of Appeals, 4 year member Board Certified in Criminal Law, Texas Board of Legal Specialization, Masters Degree-Judicial Process, University of Virginia School of Law Course Director- 2003 Advanced Criminal Law Seminar, State Bar of Texas

1. Since I have sat on the appellate bench for the past 16 years, I unfortunately have not had the experience to observe the jury selection process at the trial level. I do feel based upon the records on appeal involving jury selection that the trial courts are diligently enforcing the constitutional protections allotted to protect jurors.

2. In the last three sessions of the Texas Legislature, we have seen the enactment of the Texas Fair Defense Act and an amendment to the Texas Criminal Habeas Corpus Act to include Section 11.01, which covers representation of defendants in death penalty cases. I believe that both of these acts have gone a long way toward ensuring that indigent defendants are fairly and adequately represented, both at trial and on appeal.

3. Because the Texas Court of Criminal Appeals is in the best position to observe what are the current trends and issues affecting the criminal law, I feel that it is incumbent upon us to inform the legislature and the public of these matters and to hopefully help them fashion an adequate response.

Derrick Melgar, Democrat

Qualifications: I have practiced law for more than 20 years and have an extensive background in both civil and criminal trial work. As a part of my practice I have successfully argued cases before both The Supreme Court and The United States Court of Appeals for the Fifth Circuit.

1. Our right to a jury trial provides the most important protection we have against the abuse of power by the state. If the composition of the jury does not fairly reflect society, much of that protection is lost. Having picked many juries, I know that low income and minority Texans are not adequately represented in the jury pool. Remedying that requires both outreach to these communities and fair compensation for jury service.

2. Our state's failure to provide adequate representation to indigent defendants, particularly those in death penalty cases, is a national embarrassment. A statewide public defender's office should be established with adequate funding and competent attorneys to handle these cases. In addition Appellate Courts must be more aggressive in reviewing these cases to assure the defendant received adequate representation at trial.

3. While it is not a judge's job to legislate, they are in a unique position to recognize and advise on important legal and judicial issues facing the state. I would not hesitate to offer that expertise when appropriate and ethical.

CANDIDATES FOR DISTRICT ATTORNEY

Question:
1. What role should the District Attorney's office play in enforcing laws dealing with white-collar crime?

Corey Behnke, Republican

Training and Experience: District Attorney-present; Criminal District Judge 12 years; Assistant District Attorney 8 years; Private Practice 4 years; Board Certified Criminal Law; Co-chair Governor's Anti-Crime Commission; Member Texas Crime Victims Institute Advisory Council; National Council on Violence Against Women; Governor's Advisory Board on Juvenile Justice; University of Texas Law School.

1. I have prioritized white-collar crime prosecution. As law-enforcement's leader in pursuing this crime, my DA investigators and attorneys lead investigations & prosecutions. My efforts have resulted in millions being returned to victims and elderly individuals swindled of retirement money or scammed through home improvement and other frauds.

Jennifer A. Lundeed, Democrat

Training and Experience: BA, Texas, 1971. JD, Texas 1981. I have 20 years experience in criminal law. I am compassionate, rational and slow to anger. I will look at the big picture in making sure that justice is firm, fair, and serves the long-term interests of our community.

1. This office has a responsibility to protect the public from fraud whether by individuals, business or in cases involving public agencies. The DA has to enforce the law in a dignified manner. The DA must never serve the baser instincts of humanity such as envy, jealousy, or revenge.

CANDIDATES FOR COUNTY TREASURER

Question:
1. What do you hope to accomplish if elected to this office?

Dean Caffee, Republican
Training and Experience: BA in Accounting, Texas 1983. I have worked as a Certified Public Account in private practice for the past 18 years.

1. I hope to establish a transparent, smoothly run office. I will efficiently manage the staff of this office and ensure that the county's assets are handled properly and the county's budget is distributed as ordered.

Gordon Kallas, Democrat
Training and Experience: I am a Certified Public Account, and I hold certification as an elections administrator. I earned a BA in Accounting from the University of Oklahoma in 1979, and I have worked as a consultant for the local Area Development Partnership.

1. If elected, I hope to bring efficient management and vigor to make sure our county's monies are handled properly. With my experience, I will run a transparent and smooth county treasury office.

CANDIDATES FOR SHERIFF

Questions:
1. What is the impact of Homeland Security requirements on the Sheriff's Office?
2. What would you do to reduce juvenile crime in this County?
3. What would you do to improve relations between the Sheriff's office and the community?

Stanley Saari, Republican
Background: Education: BA in Social Work; Corrections Certificate; FBI, Secret Service Protection, and UT West Point Academies; Certified Public Manager; Police Senior Management Institute; 3809 hours CE; Experience: Manage $11 million budget and 211 employees at Austin Police Department; attained rank of Commander; 15 of 25 years in management; commanded Southwest & Southeast regions, SWAT Team, Investigations; managed Gang Suppression Unit, Homicide, Child Abuse, Sex Crimes, Robbery. Organized training conferences on gangs, criminal investigations and financial crimes.

1. Increased training and equipment for deputies who respond to WMD calls. Added security on high-risk terrorist targets. Increased calls for service on suspicious person' substance calls. Establishing an Intelligence Unit that provides potential threats. Screening information before public release to thwart false alarms. Educate the public on threats and providing instruction on how they can safeguard themselves against varied threats.

2. Work with private and public entities to expand programs such as Big Brothers/Big Sisters, mentoring, sports, scouting, and career development. Extra curricular activities keep at-risk kids and latch-key kids occupied and out of trouble. Expand the Juvenile First Offender Program to include other delinquent conduct cases. Use Juvenile Boot Camp for recidivists focusing on community service work. I would request additional bed space at Texas Youth Commission for serious habitual offenders.

3. Lead by example. Protecting and serving the community is a high calling and responsibility. Sheriff's deputies would interact with the community accordingly. Also, we would be more responsive to the community's needs. WE would determine what and where the needs are by reviewing citizen responses, internal affairs cases, crime statistics and data on hotspots of crime. We would also empanel a group of community representatives and sheriff's personnel to pinpoint additional issues and solutions.

Jason Valle, Democrat
Background: Education: BA in Criminal Justice, Southwest Texas State University, 1982 Graduate of Governor's Executive Development Program, University of Texas LBJ School of Public Affairs; Experience: Chief of Law Enforcement for Texas Alcohol Beverage Commission 1994-2004; 300 employees, 55 offices, budget of $15 million; National trainer for Department of Justice; Sheriff's Office (1985-1994) Corrections Officer, Mounted Patrol, DARE Officer, Deputy Sheriff Texas Department of Corrections (1985) Corrections Officer

1. Protecting our community and safeguarding the peace and welfare of all our citizens is a critical role of this office. We will do everything that we can to insure that our residents are informed, educated and prepared to respond to acts of bioterrorism and other threats. We will work tirelessly to partner with other local, regional and statewide groups to address preparedness, response and recovery efforts.

2. I believe that juvenile crime is something that we as a community must address. The sheriff's office, as an authority figure, must work to build a relationship with our youth. However, I believe everyone should be held accountable, without being condescending. When it comes to reducing juvenile crime, an ounce of prevention truly is worth a pound of cure. It is a countywide issue and will require countywide coordination and response.

3. The sheriff's office must begin to build relationships with the people whom we serve. Community policing refers to much more than the assignment of an officer to a certain community. We must knock down the walls of separation and build relationships on trust and respect with accountability and responsibility as our commitment to all we serve.

CANDIDATES FOR COUNTY TAX ASSESSOR

Questions:
1. What are the two biggest challenges facing the Tax Assessor-Collector office and how would you address them?
2. How can this office increase the number of registered voters in this County?

Howard Grady, Republican
Background: Education: B.A. degree, major-Economics, Texas Lutheran University; M.B.A. degree, Texas State University; Maintains certification as a Certified Internal Auditor; Experience: Deputy Clerk, Guadalupe County Clerk's Office; Caseworker/Eligibility specialist, Texas Department of Human Services; Assistant State Auditor, Texas State Auditor's Office; Field Monitor/Auditor, Contract Monitoring Department, Texas Workforce Commission; Self-employed auditor

1. The primary duty of the County Tax Collector is presenting accurate tax statements that are stated according to the properly assessed value of the property and the legal requirements. The County Tax Collector must ensure that voter rolls are accurate to ensure that everyone that is eligible to vote gets one voter's registration record. The County Tax Collector must confirm that all property statements and voting records are correct prior to mail-outs and issuance.

2. The office can send voter registration information in the mail-outs and the staff can routinely ask visitors to the tax offices if they are registered and would like to register. The County Tax Office can place voter registration materials at other county offices and various public places.

Randy H. Clemons, Democrat
Background: Education: B.A. degree in English, M.A. in Communications, Registered Texas Assessor-Collector (RTA). Certified by Texas Board of Tax Professional Examiners; Experience: Seventeen years' successful experience managing Tax Office operations, including property tax collections, current and delinquent; vehicle registration and titles; voter registration. Thirty-two years public service experience in federal, state, and local government.

1. a. Provide citizens with consistently superior service, (1) by decreasing their wait-time; (2) making services available at more convenient locations; (3) offering technological solutions to service delivery. Currently implementing all of these. B. reduce operating costs in the tax office (1) by using technology to our best advantage, (2) multi-tasking existing staff to reduce the need for more employees, (3) by creating and maintaining public/private partnerships for efficient, cost-effective service delivery.

2. By utilizing more than 2,500 Volunteer Deputy Registrars to register new voters. By focusing on voter registration year round, not just before major elections. By educating the public about voting, and using electronic media to help disseminate information. By raising awareness among younger voters. Travis County has 558,000 registered voters, which represents 90% of the population.

CANDIDATES FOR JUSTICE OF THE PEACE

Question:
1. In light of the recent US Supreme Court opinion recognizing the free speech rights of judicial candidates, what public policy issues, if any, will you raise in your judicial race?

Deborah Kamps, Republican
Training and Experience: I have worked for the Administrative Hearings Office for 10 years. My dedication, work ethic, and commitment to excellence in this office have qualified me for this position.

1. If elected, I will work closely with local schools in developing a pathway for truancy. This pathway will entail counseling and community service involvement. The old saying that "it takes a community to raise a child" can still be utilized today. This will help keep our community and our children successful.

Clyde Gayton Jr., Democrat
Training and Experience: BA, Texas 1987. I have worked as a clerk for the Administrative Hearings Office for the past 7 years. I have a wealth of knowledge regarding the intricacies of this office, and my dedication qualifies me for this office.

1. I plan to work closely with the community and other courts to help provide troubled youth a second chance in life. I would like to help establish extensive counseling services for youth entangled in drugs, and help them get back on a path to a successful life.

CANDIDATES FOR COUNTY JUDGE

Questions:
1. What would you do to ensure that indigent civil and criminal defendants have competent representation?
2. What can be done to alleviate the problem of overcrowded dockets in the courts?
3. Should judges recuse themselves from cases involving those who have contributed to their campaigns?

Dan Atchley, Republican
Background: Education: I graduated from the University of Texas and the Unveristy of Houston Law School, where I was on the staff of the Houston Law Review. Experience: I have been licensed from 27 years, and am board-certified in administrative law. I have served as Judge of the 353rd District Court since 1993, and was a trial attorney for 14 years prior to my election. I have experience in the range of cases heard by this court.

1. The district judges have adopted a plan for represenation of indigents in the criminal and juvenile system to insure that constitutional rights are protected. Appointed attorneys are required to complete continuing education and skilled attorneys are matched to the severity of the offense charged. The performance of the attorneys and aspects of the program are routinely evaluated. The plan contemplates that counsel appointed will meet with clients within 24 hours of incarceration.

2. Despite our explosive population growth, we have not had a new civil court since 1983, and the legislature approved one court last session. The county commissioners have supported hiring associate judges, who provide assistance with our family and juvenile dockets. We have one judge who hears many discovery matters to insure speed and consistency in those matters. Our central docket and ADR are major factors in helping us to reach cases timely for trial.

3. No. State and federal law are unanimous that campaign contributions alone do not require recusal. Lawyers on both sides of the docket contribute and are interested in fair judges. The state supreme court has long recognized the criticisms lodged at judicial campaign financing and suggested on several occasions that the legislature make changes to the system of judicial selection and campaign financing, but the legislature has not seen fit to adopt those recommendations.

Lewis Shine, Democrat

Background: Education: UT Austin, BA, School of Social and behavioral Sciences, 1977 TSU, Thurgood Marshall School of Law, 1983; Experience: 12 Years 10 Months Judicial Experience, Associate District Court Judge, Third Administrative Judicial Region (central Texas). Presided over 100,000 family law cases including contempt of court/jail cases. 5 years 6 months Attorney, Private Practice with criminal defense emphasis. 2 years Hearings Examiner, Parole Revocation, Texas Youth Commission.

1. Evaluate defense Attorneys according to the Fair Defense Act. Should an Attorney not meet the set standard, a specific plan for continuing legal education can be required before placing the Attorney on the appointment for indigent defendants list. Formal complaints to the State Bar of Texas Grievance Process may be necessary.

2. Judges can require Defense Attorneys to appear in Court and set expectations that Defense Attorneys obtain discovery about the case prior to their client's day in court. The Court can set several of the Defense Attorney's clients' cases on a specific day. In civil cases, the Court can require Attorneys to talk on the telephone or by personal meeting prior to setting the case for trial.

3. Judge's rules for recusal are currently in place and the State Commission on Judicial Conduct is active in determining any violations in this area. I am in favor of the rules for recusal. Currently Judges seek lawyers' campaign contributions by necessity and any financial relief that would change this action would be welcomed by any judge.

APPENDIX E

This appendix contains an example of a survey packet used in Study 2. The SUS and Big 5 assessment are not included here.

Voting Study Survey Packet

Experiment # 295
2006/2007

Participant Number _____

Date _____

Background and General Voting Survey

1. Age: _____

2. Gender: ____ Male ____ Female

3. Year or Position at Rice:

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

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

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

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

7. Are you a native English speaker? ____ No ____ Yes

If no, what is your native language? _____

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

____ No ____ Yes

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

____ Humanities
____ Social Sciences
____ Natural Sciences
____ Engineering
____ Architecture
____ Music

10. How many hours per week do you use a computer?

____ less than 5 hours
____ between 5 and 10 hours
____ between 10 and 20 hours
____ between 20 and 30 hours
____ between 30 and 40 hours
____ over 40 hours

11. Please rate your level of computer expertise (1 = novice, 10 = expert)

1 2 3 4 5 6 7 8 9 10

12. Which of these activities do you use a computer for? Check all that apply.

____ Word Processing (e.g. Microsoft Word)
____ Programming (e.g. Java, C++, Scheme)
____ Web design
____ Graphic Design (e.g. Adobe Photoshop, Illustrator)
____ Video Editing
____ Personal Finance (e.g. Quicken, Turbo Tax)
____ Games
____ Music
____ Multimedia (e.g encyclopedias; interactive CDs)
____ Spreadsheet management (e.g. Microsoft Excel)
____ Data Analysis (e.g. SAS, SPSS)

13. What is your political affiliation?

____ Republican
____ Democrat
____ Libertarian
____ Independent
____ Other, please specify: _____

14. How many **national** elections have you voted in? _____

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

_____

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

---

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

_____

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

18. Do you typically cast your vote on an absentee ballot?

____ No ____ Yes

19. Please indicate **how many times** you have used each type of technology or ballot to cast your vote in any election.

____ Fill in the bubble (or box)
____ Connect the arrows (or lines)
____ Open response
____ Lever machines
____ Punchcards
____ Electronic – touchscreen
____ Electronic – other
____ Don't know
____ Other, please specify:_____

20. Have you ever felt worried about figuring out how to use the ballot or technology to cast your vote?

____ No ____ Yes

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

____ No ____ Yes

22. Did this prevent you from voting?

____ No ____ Yes

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

____ No ____ Yes

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

____ No ____ Yes

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

**** End of previous voting experience section****

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

____ No, not at all

____ Yes, somewhat

____ Yes, very closely

27. If you have been following the news about computer voting and security, has it affected your trust of these systems?

____ No ____ Yes ____ Does not apply to me

Why or why not? _____

_____

_____

_____

28. How often do you use an ATM (Automated Teller Machine) to get money or complete other transactions at a bank, grocery store, or other location?

____ never
____ very infrequently
____ occasionally (for example 1-4 times a year)
____ often (for example once a month)
____ frequently (for example once a week or more)

Additional Demographic Survey

1. What is your current occupation? _____

2. Please indicate the highest level of education you have completed.

____ Some high school
____ High school or G.E.D.
____ Some college or Associate's degree
____ Bachelor's degree or equivalent
____ Postgraduate degree (such as M.A., Ph.D., M.D., J.D.)

3. Are you:

____ African American
____ American Indian
____ Asian American
____ Caucasian
____ Mexican American or Chicano
____ Other Hispanic or Latino (please specify) _____
____ Multiracial (please specify) _____
____ Other (please specify) _____

4. Which of the following income ranges best describes your yearly wages?

____ below $20,000
____ $20,000 to $40,000
____ $40,000 to $60,000
____ $60,000 to $80,000
____ Above $80,000

5. If you are retired, which of the following income ranges best describes you maximum yearly wages while you were working full-time?

____ below $20,000
____ $20,000 to $40,000
____ $40,000 to $60,000
____ $60,000 to $80,000
____ Above $80,000

Voting Method Comparison Survey

1. Of the two voting methods you used in this study, which was your **favorite**? (Please circle one.)

Bubble ballot **Computer**

2. Why was this your favorite type? _____

_____

3. What, if anything, did you **not** like about the other voting method (the kind that was not your favorite)? _____

_____

4. Have you ever voted on another type of ballot or voting equipment that you liked better than these? (Please circle one.)

No **Yes**

If yes, please describe the other ballot. _____

_____

_____

Accuracy Comparison Survey

1. Please consider the **computer voting method**. How accurate did you feel this voting method was? In other words, how confident were you that the voting method recorded the vote that you intended? Please indicate your response on a scale of 1 to 5, with 1 being least confident and 5 being most confident.

Not at all confident				Very confident
1	2	3	4	5

2. Please consider the **bubble ballot voting method**. How accurate did you feel this voting method was? In other words, how confident were you that the voting method recorded the vote that you intended? Please indicate your response on a scale of 1 to 5, with 1 being least confident and 5 being most confident.

Not at all confident				Very confident
1	2	3	4	5

Security Comparison Survey

1. Please consider the **computer voting method**. How secure did you feel this voting method was? In other words, how confident did you feel that your vote could not be changed after the fact? Please indicate your response on a scale of 1 to 5, with 1 being least confident and 5 being most confident.

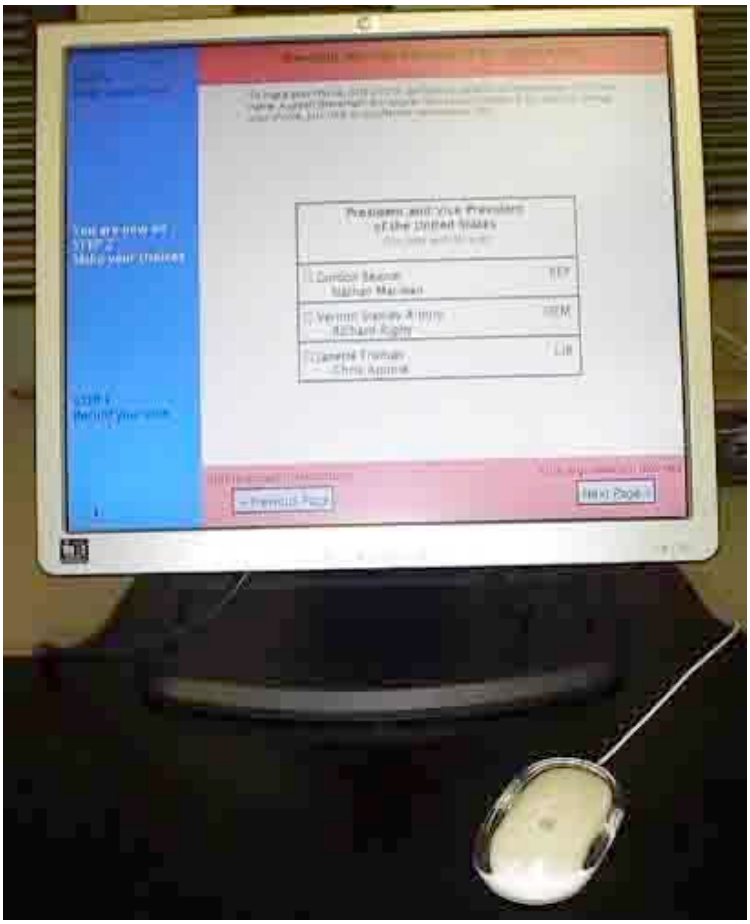
Not at all confident				Very confident
1	2	3	4	5

2. Please consider the **bubble ballot voting method**. How secure did you feel this voting method was? In other words, how confident did you feel that your vote could not be changed after the fact? Please indicate your response on a scale of 1 to 5, with 1 being least confident and 5 being most confident.

Not at all confident				Very confident
1	2	3	4	5

Did you feel that the instructions on the following voting method were easy to understand? (Please circle one.)

No **Yes**



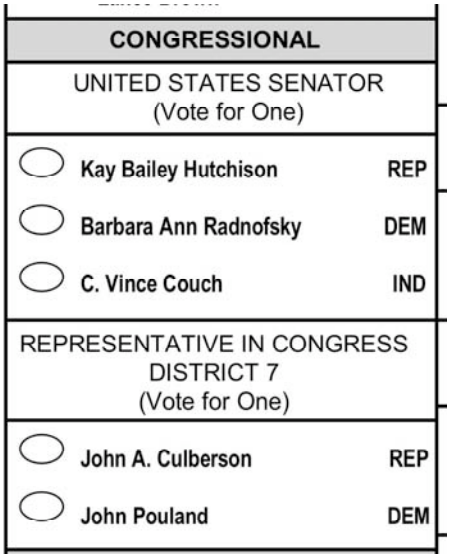
If no, what was unclear? _____

_____

_____

Did you feel that the instructions on the following voting method were easy to understand? (Please circle one.)

No **Yes**



CONGRESSIONAL

UNITED STATES SENATOR
(Vote for One)

- Kay Bailey Hutchison — REP
- Barbara Ann Radnofsky — DEM
- C. Vince Couch — IND

REPRESENTATIVE IN CONGRESS
DISTRICT 7
(Vote for One)

- John A. Culberson — REP
- John Pouland — DEM

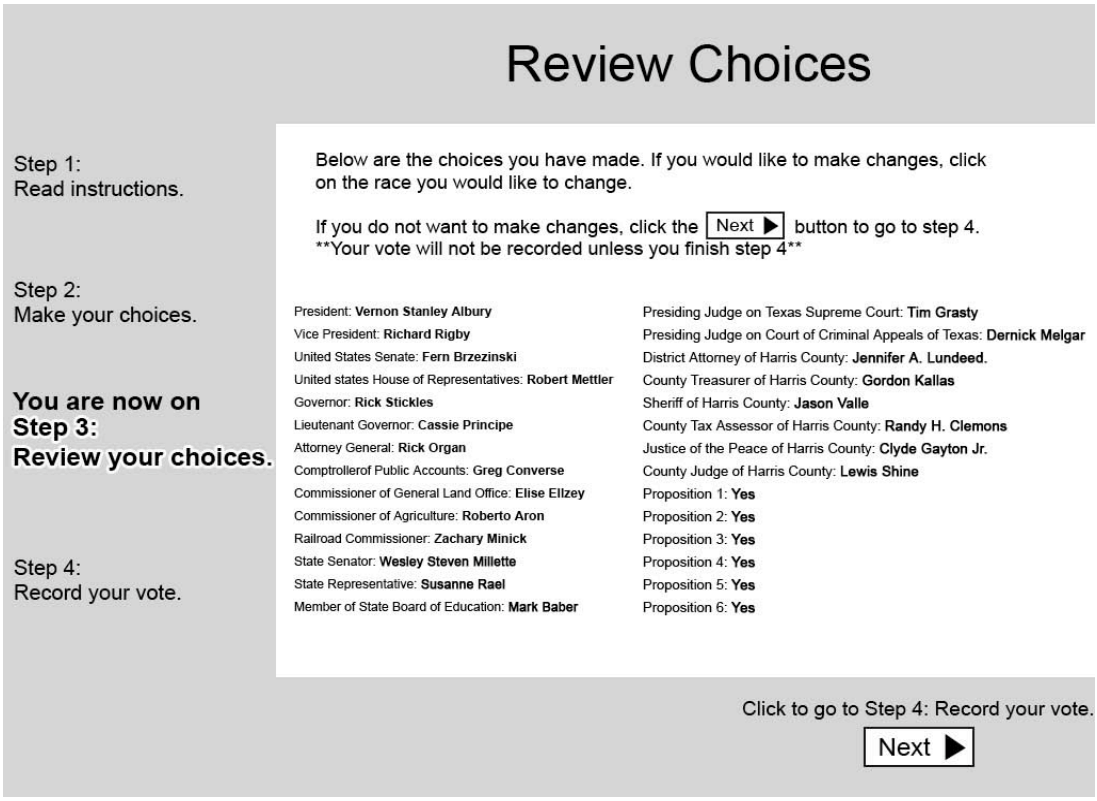
If no, what was unclear? _____

_____

_____

Review Screen Survey

When you voted with the computer voting system, you were shown a summary screen (see picture below) before casting your ballot. This screen displayed the names of the races and what selection you made.



1. Did you feel that having a summary screen was useful?

____ No ____ Yes

2. Did the summary screen perform as you expected it to?

____ No ____ Yes

3. Please describe why the summary screen either did or did not perform as you expected.

_____

_____

_____

4. Did you change any of your choices after viewing the summary screen?

____ No ____ Yes

5. If you made a change after viewing the summary screen, please describe why you made a change.

_____

_____

_____

6. If you made a change after viewing the summary screen, were you satisfied with the ability to change your selection?

____ No ____ Yes

7. Did you feel that having a summary screen made you feel confident that your vote would be counted correctly?

____ No ____ Yes

Why or why not?

_____

_____

_____

8. Although the other voting method you used, the bubble ballot, did not provide you with method to review your vote like a summary screen, did you check each of your choices to make sure they were correct?

____ No ____ Yes

Why or why not?

_____

_____

_____

9. Did you feel checking your vote with the computer or the bubble ballot was better, or were they the same? Why?

_____

_____

_____

10. Other comments:

_____

_____

_____

_____

11. The summary screen you saw today did not behave as it should have. Instead of showing you exactly which races you saw and which (if any) candidates you selected, it removed two races. This means that you actually were given the opportunity to vote in 27 races, but only 25 races were shown on the summary screen. Did you notice that there were races missing?

____ No ____ Yes

12. If you noticed, what did you think when you saw the summary screen?

_____

_____

_____

_____

Voter Guide Survey

In the study today, you were offered a voter guide. Please answer the following questions about it.

1. Which of the following best describes how much you used the voter guide? (Please circle one.)

Did not look at it

Briefly glanced at it

Read it carefully

2. Did you look back at or use the voter guide when marking on the ballots?

No Yes

3. In real-life voting situations, how often do you use a voter guide?

Never

Hardly ever

Sometimes

Most of the time

Always

4. If you use voter guides in real-life voting situations, please describe how you use it (For example, do you read it at home, while waiting in line, while voting? Do you mark your choices in advance? Do you bring it with you to look at when you vote?

_____

_____

_____

APPENDIX F

The new offices and propositions used in Study 3 are included in this appendix.

Clerk of the Supreme Court
Carolina Delao (R)
Trudy Millsap (D)

State Auditor and Inspector
Susanne Guillemette (R)
Caitlin Setzer (D)

Commissioner of Labor
Hector Hanneman (R)
Fay Garfinkel (D)

Insurance Commissioner
Ursula Hilderbrand (R)
Dustin Stroman (D)

Corporation Commissioner
Alisa Thoreson (R)
Jorge Reaux (D)

Associate District Judge
Glen Strock (R)
Shane Kelsch (D)

Sheriff/Coroner
Wesley Pomeroy (R)
Dean Smathers (D)

Property Valuation Administrator
Shane Croskey (R)
Tim Voorhis (D)

Extra Proposition 1
Shall there be an amendment to the Texas constitution concerning initiative and referendum petitions, and, in connection therewith, changing petition rights and procedures; allowing petitions to be submitted at all levels of Texas government; limiting initiative ballot titles to 75 words; changing single-subject requirements and procedures; limiting the annual number of new laws that governments may exclude from possible referendum petitions; establishing standards for review of filed petitions; specifying that petitions may be voted on at any November election; limiting the use of government resources to discuss a petition; requiring voter approval for future petition laws and rules and for changes to certain voter-approved petitions; and authorizing measures to enforce the amendment?

Extra Proposition 2
Shall the Texas Constitution be amended to create a Healthy Future Trust Fund which will: 1) be used to reduce and prevent tobacco use, to increase funding for healthcare access and treatment for eligible low-income individuals and Medicaid recipients, and cover administrative costs; 2) be funded by a tax of four cents per cigarette and twenty percent on other tobacco products; and 3) be kept separate from general revenue and annually audited?

Extra Proposition 3
Shall the municipality of the City of Houston impose a sales tax of one-eighth of one percent (1/8%) for the purpose of providing funding for local parks for the municipality and specifically, funding for the construction and maintenance of new and existing recreation centers and creation programs in parks, including but not limited to programs for children and seniors?