CALTECH/MIT
VOTING TECHNOLOGY PROJECT

A multi-disciplinary, collaborative project of
the California Institute of Technology – Pasadena, California 91125 and
the Massachusetts Institute of Technology – Cambridge, Massachusetts 02139

PROCESSES CAN IMPROVE ELECTRONIC VOTING: A CASE STUDY OF AN ELECTION

Ted Selker
Media Arts & Sciences, MIT

VTP WORKING PAPER
October 2004

CALTECH-MIT/VOTING TECHNOLOGY PROJECT

Reno/Sparks, Nevada –September 7, 2004

Across the United States, I have personally watched hundreds of precincts vote since 2001. Most recently, I traveled to Reno/Sparks, Nevada to observe the rollout of the Sequoia Direct record electronic voting systems with verifiable paper trail printers on September 7, 2004. This experience was also enriched by the members of the Secretary of State of California's poll-watching effort, who invited me to join them to watch the election progress at eleven different polling places, which together represent almost forty different precincts. At each polling place I conducted interviews with poll workers and election officials as well as exit polls of voters. The California team, comprised of election officials from all over the state, who's broad experience provided helpful context and insights.

What I witnessed at the election both encouraged and horrified me. The paper "receipts" were less confusing than I feared they would be. Poll workers and voters alike showed an eagerness to "get it right," even when the new technology required them to endure some amount of initial frustration. However, things went gravely wrong when workers did not have adequate time to set up or test equipment; when, in the pressure of the moment, procedures were ignored or forgotten and, instead, solutions were improvised; or when no standard policy existed to guide election officials in proper protocol.

In my experience, such problems are not unique to that election, to the Sequoia electronic voting machines, or to the paper trail audit system. Indeed, the shortcomings I encountered in Nevada resemble those I have seen in precincts throughout the country and with every kind of voting system. Luckily, most of these problems can easily be solved if we focus on improving both training and process. We can learn from our mistakes. Toward that end, a detailed account of my day in Reno follows.

6:59 a.m. First Stop: Precincts 507, 516 517.

We arrived one minute before the polls were scheduled to open. Poll workers were busy setting up registration desks and voter materials. However, the screens on the voting machines were black. A distressed poll worker moaned, "This was terrible. We weren't able to get in here until 6:45 a.m." Washoe County Registrar Dan Burke later explained that it is often a problem to get into polling places at 6:00 a.m. on the day after a holiday, as was the case in this election.

Figure 1 A lone Poll worker documenting the state of a machine at the beginning of the day]

By 7:08 a.m. twelve people were lined up waiting to vote but the machines were still not ready. The Precinct President nervously admitted, "I assumed they would be turned on but…" She began switching on the machines. Two minutes later there were fifteen voters patiently waiting while she raced down the line of voting stations to get them operational as quickly as possible.

In her rush to make up for lost time, she recorded numbers from the front of the machines but did not check the odometer readings from the back She could have been taking numbers from the Paper Trail equipment or from the screen. Certainly, she was not verifying that these matched the odometers on the back.
The precinct president also worked by herself rather than in tandem with a co-worker, who should independently corroborate the readings to insure that the machines had a zero vote count at the beginning of the day. [Figure 1]. Instead, she alone marked numbers down. I hope her transcriptions were accurate.

These shortcuts could cause two serious problems:

(1) I observed no indication that the Reno machines had been tampered with, however, they had been set up on Saturday and Sunday and then left locked but unguarded in a church until the morning of the Tuesday election. (For this reason many jurisdictions do not allow elections to be held on the day after a holiday. But this, like most election protocol, is determined by local laws.) If, at the end of the election, a discrepancy had been discovered between the number of voters who checked in at registration and the total electronic vote count it would have been impossible to know which votes to subtract out, whether these were the accidental result of set-up testing or intentionally added by a vandal.

(2) Without a second set of eyes to separately record the odometer readings and serial numbers, a commonplace transcription error could frustrate election certification. Although these are easy to fix if caught in the moment, after the fact such mistakes cannot be remedied.

By 7:12 a.m. restless voters attempted to get the machines running on their own. One man stood for several minutes at a machine that was still turned off before someone came to his aid.

In the midst of this chaos, a low battery message on one machine's screen, indicating that it was not plugged in and had run down its battery power, was overlooked. This particular machine was placed near a window, which flooded the screen with sunlight and made it difficult to read. [Figure 2] The oversight might also have been possible because the green indicator light, which confirms that a machine is correctly plugged in, is located at the very back of the machine under a ledge [Figure 3].

Figure 2. Voting machine in poor lighting situation.

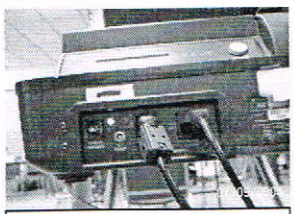

Figure 3. Back of voting machine with green indicator light.

It can only be seen if you bend down and over to look at it. However, in this polling place the voting machines were so close to the wall that it was probably impossible to check the power light easily. Indeed, a color band at the bottom of the touch screen appears when a machine is running on battery. This should have alerted election workers to the situation. Unfortunately, it went unnoticed.

This polling place (and all I visited in Reno) was spacious, well-lit, clean, and wheelchair accessible without even a curb to go over. I was, in fact, surprised at the sheer number of machines—twenty in every location I stopped at that day. Four or five would probably have been quite sufficient for the primary. Clearly, election officials really wanted to insure that they could adequately accommodate voters. Presumably, the high quantity of machines could ameliorate lines if there is high-turnout for the national election.

Unfortunately, their number taxed even large polling places. The physical arrangement of the numerous machines had not been worked out to protect voters' right to secrecy. At every polling place I visited, the machines were lined up in two long rows, typically with 15 feet between them, and with voters' backs facing each other. The machines were as close together as telephone booths (although an old-fashioned telephone booth with a door or screen would have been more private.) It was all too possible for any voter to watch how his neighbor was voting. Moreover, anyone standing in the center of the room, could closely observe every voter's actions.

The set-up I observed in Los Angeles, California for the recent gubernatorial election showed greater sensitivity to preserving privacy. There, voting machines were arranged in a square, each facing outward. I am told that before the polls opened, workers carefully stood at each screen to make sure that it was not possible to view the displays of the adjacent machines.

Difficulties at this Reno polling place did not end with the poor layout and haphazard start. At 7:20 a.m. a voter complained that his smart card (needed to activate the voting machine) was invalid.

Another voter complained," It [the machine] will only let me vote for one race" Other voters began to realize that they, too, had only been prompted to vote for senator and had never had a chance to vote for the six other local races on the ballot.

One of the California observers commented, "There ain't no way in hell!...We are staying with OpScan."

By this time, twelve people were hanging about waiting to learn why they had only been permitted to vote on one race [Figure 4]. I cannot quantify how many people were deprived of the opportunity to vote for local officials. It was time for me to proceed to my next scheduled polling location.

Before leaving, I asked one of the exiting first voters, "Did you notice a second display?" trying to get him to tell me about the paper trail behind the glass. "No…," the man replied, "I just voted and reviewed my selections on the screen."

Figure 4. Voters who accidentally voted on provisional ballots.

8:00 a.m. Stop 2, Precincts 112, 126, 117.

Voting was already well underway when we arrived. I began by conducting short exit interviews of voters as they left the polling center. I asked them how this voting experience compared with previous elections; what kind of equipment they had used in the past; and whether they had reviewed their paper receipt.

Because the screens of the tightly packed machines faced me at locations throughout the day, I could actually see whether or not people looked at their printouts. Between the paper trail equipment's obvious placement; its noise and action; and the fact that the voting machine's screen went black when "receipts" printed, Sequoia did a good job getting voters to look at them. Most voters did glance at the "receipt," which was easily visible beneath the clear viewing area in the paper trail box. In fact, many people reached their hands over the plastic box trying to figure out where the paper came out because they are used to being able to keep things called "receipts". However, in the course of the day, I saw no more than 3 people spend even five seconds viewing the printout. Moreover, voters did this without being able to consult their onscreen selections. An improvement would be to allow voters to look at the onscreen "Review" display when they get to see and confirm the paper trail printout. Reno/Sparks voters that day had to rely entirely upon memory to make the comparison.

Nevertheless, people typically swore that their "receipts" exactly matched the votes they had entered on the machines. This might be the case. However, it more likely suggests that a number of them did not closely review the printouts because, on average, people make 1 —3% errors using Direct Electronic Register machines. This being the case, in a ballot of seven races, I would expect one in seven voters to discover a discrepancy between what the "receipt" printed and what he INTENDED to vote. Ideally, these voters could then ask to have their first vote invalidated so that they could fix their mistakes and enter their choices correctly.

The Reno/Sparks ballot that day had only seven races on it, far fewer than the number in many US elections. This may have been less confusing than longer ballots. Moreover, we had to look hard to find voters who did not appear to be have high literacy rates. Possibly, therefore, fewer errors occurred than might in the presidential election if more people with poor literacy skills choose to vote.

Many people brought their sample ballots with them. These were beautifully designed and extremely helpful. Unfortunately, there was nowhere for people to rest them while voting so they were forced to hold them as they used the machines.

Asked about the paper receipt, one voter stated, "I wasn't really interested because I had already made my choices." As a result of pre-election publicity most people did expect the paper trail printouts. I witnessed one person that day walk out in disgust grumbling, "How do I know that my vote counted?" It was clear that many people were not exactly sure what the "receipts" were for. I realized how important it was to phrase my questions carefully. The first time I asked another voter about the "receipt" he admitted that he hadn't reviewed it. However, with further prompting he began to say that he had used it to check his vote because he realized that this was what he was supposed to have done.

By far, however, the most disturbing exit interview that I conducted at this second location was with a voter who, when asked how difficult the technology was to use, replied, "I couldn't get the machine to work. It had a complicated display on it, so I pushed the button and then it let me vote." This answer implies that he had been allowed to proceed to a machine that was still in start-up mode and which had not yet had its odometer setting recorded. Not one to concede defeat, he managed to get the machine started on

his own. This meant that no record exists to confirm that that machine had a count of zero votes at the beginning of the election.

8:40 a.m. Third Stop: Precincts 120, 130, 206

When we arrived at the third polling place, a paper trail "receipt" printer stood open on a counter in front of us. "They told us to replace the paper if it jammed," a poll worker explained. The instructions provided for her included nothing about how to rethread the paper (we were told that the instructions at some polling places did and at others didn't) so she struggled to figure it out on her own. At first, she tore some paper off and tried to fit the end through the printer. Then, she grabbed a pair of scissors and cut more paper off the roll. As she did this, the paper trail "receipts" already recorded by the printer were cut off [Figure 5].

Figure 5. A poll worker holds paper trail "receipts" from that day's election in a

I photographed her rolling the paper trail printouts up to store them. Hopefully, she didn't add any extraneous scraps of paper or forget to include the piece she had separated from the roll. I could imagine her taping the roll to the inside of the printer ballot box and wondered if it would be found later. No second official helped, monitored, or certified the opening and closing of this paper trail equipment. When we left, the printer still lay open to view.

Through the course of my day in Reno, I was told about or observed printer jams in what I calculated to be more than one in every twenty printers. This was the equipment's public debut. Had it been thoroughly tested prior to the election? Had poll workers been properly trained on how to fix problems without compromising the security of the "receipts"? The answers to these questions were unclear from what I could observe or learn.

This location also had problems activating voters' smart cards. One couple explained that one of them had to have a card reprogrammed three times, the other, once, before they could vote. This may have been because the poll worker entered the precinct number incorrectly or because the activator or card had some connection problem.

Voters were nevertheless positive about the electronic voting machines. In my continuing exit interviews at this location, one described them as, "very much easier than punch cards." Another deemed them "very easy". Yet a third happily explained, "I looked at the review screen and noticed that the Supreme Court race that I care about was not selected so I went back and fixed it." This was the only person that day who told us that a review had helped fix an error. The correction in this instance was prompted by the on-screen display that summarizes voter selections rather than by the paper trail "receipt".

Responses to the paper printouts were somewhat vague, although by no means discontented. One voter merely said, "I checked." Another simply stated, "I heard it's new." It was unclear, however, if he understood what it was for. However, I was excited that voters in Reno/Sparks, unlike those in Wilton CT during the debut of Avante's paper trail, were not badly deterred by the printed "receipt."

9:00 a.m. Fourth Stop: Precincts 243, 215, 208

The stench of alcohol on the breath of the poll worker in charge of this location concerned me. She, however, was clearly used to it because she appeared coordinated and articulate. The location was running more smoothly than any I had yet visited.

As at the other polling places we visited, only one person seemed to have been trained to set up and close down the machines. She was it at this location. It had not occurred to her to have a second helper transcribing the odometer readings.

Voters here seemed pleased with the machines. One described the system as "quicker, clearer, with less room for error." Another stated, "It is obvious what you are voting for. The recap is nice." A third found it, "much easier," than previous voting techniques.

While most people accepted the systems and felt comfortable with them, here there was also a small amount of confusion about what the paper "receipts" were for. Many wondered why they weren't allowed to take them away. One remarked "Receipt for me or what? Not sure what they need to let us see that for."

10:00 a.m. –Return to Stop One
When I arrived back at the first location, a frazzled poll worker reported, "At 8:30 a.m. all the machines began shutting down here; the outlets were not active." Apparently, in the belated rush to turn on the machines, workers hadn't noticed that they had plugged them into dead outlets and had been running them all on battery power. A color bar on the bottom of the voting machines indicates when they are running on battery. However, this apparently was not noticed in the midst of an election. A future improvement might be to have the words "operating on battery: alert poll worker" in the red bar.

I noticed that the workers had daisy-chained ten machines to a single 50-foot extension cord, which was plugged into an outlet, which also had a microwave oven plugged into it. I couldn't stop myself from pointing out that they might unplug the microwave to avoid overloading the circuit. The poll worker explained that they would not use it. The plugs on the extension cord were warm to my touch. The other ten machines were similarly plugged into a second outlet.

At this point, the polling location manager blamed the problem of some people only being allowed to vote on the senatorial race, which we had seen earlier, on the fact that the card activator had been running on battery. This seemed strange because, in the meantime, election workers had figured out that these partially disenfranchised voters all belonged to the same precinct: 516.

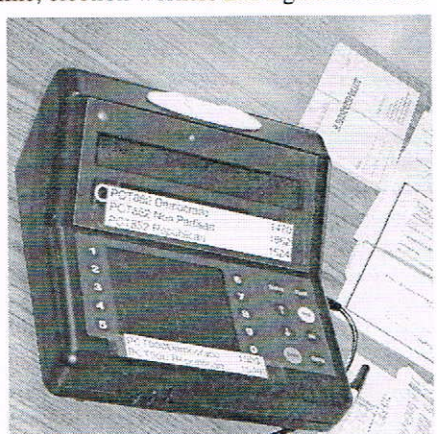

Figure 6. Activator box with paper labels for ballot codes.

A Sequoia representative had arrived in response to an earlier trouble-shooting call (laudably, he and other employees arrived within minutes whenever and wherever they were summoned). He denied that the machines operated differently on battery than when attached to an electrical outlet. He was convinced that the woman responsible for entering the codes for precinct 516 on voters' cards had accidentally been typing in the code for a provisional ballot. In Reno provisional ballots (pct 0000) only allow voters to make selections for races available to every precinct, in this case senator. It was an easy mistake to make because the provisional ballot code was the one closest to the front of the activation box. [Figure 6] The list of precinct codes was posted farther up the card activator on another scrap of paper above the buttons. The poll worker seemed to realize her error at about the same time that the inactive power plug problem was solved. She must have then begun typing in the correct numbers but never admitted her mistake. When we offered this explanation to the polling manager, she agreed that it sounded more plausible than hers. I do not know, however, how she documented these incidents.

10:35 A.M. Fifth Stop : Washoe County Government Campus**, Precincts 304, 413, 457 and Registrar's Office.**
At the Washoe County Government Campus I observed people being sent to the municipal Registrar's Office because they could not remember which precinct they lived in. Unfortunately, the early voting machines at the Registrar's Office had been turned off by law the previous Friday. Therefore, those sent there from a polling place to find out which precinct they were registered in had to subsequently travel to their assigned polling place in order to vote. Improvements would be to provide poll workers at every location with a means to help voters determine their correct precincts and to change the law to allow the early voting machines to stay active so that these people could vote at the Registrar's Office.

I also witnessed a technique for handling troublesome smart cards that resulted in unnecessarily shut down an entire polling location. A poll worker explained, "If the card doesn't work you just turn off the machine and put the card in a box not to be used." A policy of disabling machines every time a card didn't work could be used to advantage by someone who wanted to make it hard for others to vote in some part of town. Another way to disrupt a polling place, although no one did it that day, would be to claim that one's paper trail printout does not match one's vote.

10:45 a.m., Sixth Stop: Precincts 810, 854, 856, 976
In this location we watched one voter, with the aid of a supportive poll worker, struggle with her provisional ballot paperwork for ten long minutes before the poll worker suggested that she think about

where she lived and had registered previously. She then remembered where she voted at the last election and left to go there. County officials had told us that they hoped to discourage provisional ballots (which disenfranchise voting on local elections). This was the only person we saw that day who intentionally attempted to deposit a provisional ballot. As it turns out, the policy worked as planned in this case because the voter was sent to the precinct where she could vote on all races.

One voter complained that while using the electronic machine, "I didn't know how to end. It kept saying to return and kept going back." Others had some trouble finding the small buttons that say, "touch," "print," and "review" at the bottom of the screen. An onscreen display offering these functions in larger type would be an improvement.

1:40 p.m. Return to Polling Location 1,

We returned to our first polling place for a third time. Voting seemed to be progressing fairly normally. However, voters here were not shown the visual instructions on how to use the new machines. In any case, their sample ballots were missing the section illustrating the yellow slot where the smart card should properly be inserted. A poll worker told me that she had been verbally telling each voter how to use the machines. When I asked her to describe the directions she was giving, she explained that she told them to "put the card in the screen," which is not only incorrect but also physically impossible. However, I learned of only one person who attempted to do this that day.

Because of worries that finger grease would damage the computer screens (which neither I, nor a Sequoia engineer, whom I spoke with, consider problematic), poll workers had been directed to hand out unsharpened pencils and instruct voters to vote by applying the eraser end to the touchscreen. "Don't use your finger or fingernails," I overheard the same poll worked admonish, "They make our machines complain."

An elderly man repeatedly attempted to use the unsharpened end of his pencil to vote with the onscreen prompts. He had to make four tries on one button to get his vote to register. Each time he jabbed the pencil into the screen I was afraid that he was going to break it. Allowing voters to use their fingers, as the machines were designed for, would have been easier and safer.

Nevertheless, when I conducted an exit interview with this same man he was enthusiastic about his voting experience, which he described as, "no problem." He added, "Next time will be a breeze."

Overall, people adapted to the electronic machines with few difficulties, especially given the fact that they had no training and only a few seconds of guidance on how to use them. One reason was probably the well-designed HAVA complaint instructions on the right of the screen [Figure 7] Most voters categorized the machines as "a lot easier" to use than the punch cards formerly utilized. One voter felt, "The electronic voting machine made it seem more official than the punch card." However, some voters here also complained about the hard-to-see buttons at the bottom of the screen.

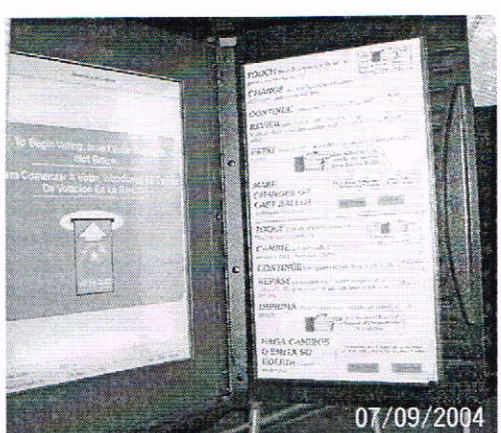

Figure 7. Clear voting instructions.

Many voters were disappointed that they did not get to keep their paper receipt. However, a poll worker noted, "People aren't looking at the paper."

2:20 p.m. Seventh Stop: Precincts 430, 456

As the day wore on, I began to suspect that I had been slightly set-up. Every location that I visited seemed homogeneously filled with educated, middle class, white voters. When I had asked poll workers whether they were set up for Spanish language voters they didn't know for sure. But, then again, I never saw anyone ask to vote in Spanish.

We went to this voting latest place hoping to see low income people vote. The place was very empty. We observed one person vote and then moved on in search of a place with Spanish language voters or with an

obviously different population mix in order to observe how other groups were adjusting to the new technology.

2:50 p.m. Eighth Stop: Precincts, 734, 758, 850, 851

Here, I met the first official Spanish interpreter that day. Actually, he informed me that he was simultaneously working two different locations, which he might dash back and forth from if needed. However, at three o'clock in the afternoon he had only received two requests for help.

I inquired about the audio capability built into the machines for use by blind voters although we witnessed no one who voted with earphones that day. The manager, who had worked on voting machine setup during the weekend, explained that no one where she had been had thought to test it beforehand. The audio systems could not be tested on the day of the election without actually testing live voting machines. Set up and testing of the audio ballot must be performed at polling places if we expect the disabled community to maintain its right to vote.

The printer malfunctions were even more worrisome. The same manager, who had set up seventy of voting stations, informed me that four printers had had to be serviced before the election even began. She then told me that two (out of the twenty) printers at this one polling place had not functioned since the poll had opened that morning.

I also learned there had been no guards or seals on the room after the machines were put in place. Better security for voting equipment should be put into place in order to be able to prove that it has not been used without authorization prior to the start of an election. It would not be unduly expensive or complicated to place numbered seals that tear when tampered with on the doors of polling places. A fancier portable security system possibly with surveillance camera that would store or transmit pictures of the polling places

"Where is my receipt," asked one voter. I continued to encounter many people who thought that the word "receipt" implied that the paper was something they could take with them. The term "record," which Sequoia uses, is less confusing.

3:20 p.m. Ninth Stop: Precincts 726, 786

Here, one voting machine had been shut down since the beginning of the election because the touchscreen wouldn't work. More voters complained that they couldn't keep the paper "receipt". "I think the voters should get a copy," one person affirmed. However, response to the machines was positive. An elderly woman using a walker progressed extremely slowly through her ballot. I was surprised, however, by her favorable impression of the new technology. She found the screen, "easier to read," and "a lot easier than the type that you fill in a circle." She was referring to an optical character recognition system.

4:00 p.m. Tenth Stop: Precincts 726, 732, 786, 852

This polling location was in a neighborhood with a lot of trailer homes and bars on windows.

For the first time all day, poll workers told me that two people had independently recorded serial numbers and odometers before opening the location. This was also the only location that reported that they had a spare card activator. Everyone else expected to rely upon trouble-shooters. Indeed, I witnessed both county officials and Sequoia personnel respond quickly to these requests.

In spite of well-trained workers, this polling site had its fair share of technical difficulties. One printer was out of service. I was also informed that ten smart cards had behaved erratically. Several of these could not be reprogrammed and had been put aside.

In anything, workers at this location were overzealous in their efforts to assist voters with the new technology. I watched one well-meaning helper literally stand over people to assist them as they voted. Any pretense of anonymity or secrecy was out the window.

One befuddled voter said, "The punch card was great," and wanted a copy of his paper receipt. On the

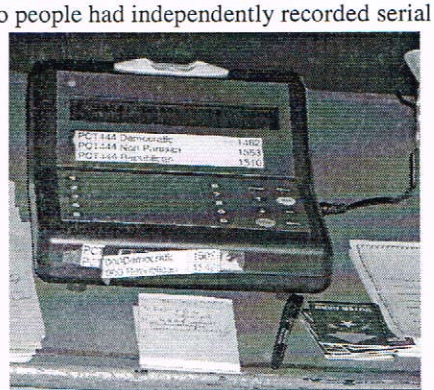

Figure 8. Card Activator labeling modified by poll worker to put provisional numbers out of the way.

whole, however, voters were generally positive about the electronic voting experience. An elderly person, who had been especially slow in voting, stated, "It takes time to learn but is o.k."

4:45 p.m. Eleventh Stop: Precincts 400, 444

My last stop before returning to the Registrar's Office was in a huge arena at the fairgrounds. It even contained an ATM machine, which one voter mistakenly attempted to shove his card into. (It gave his card back.) .

Even one poll worker here admitted that she had accidentally issued provisional ballots to people because it was the number closest to her on the card activator the same screw-up of not checking the ballot card I had witnessed at many locations throughout the day.

After realizing how easily this mistake could be made, a frustrated poll worker simply moved the scrap of paper with the programming numbers for provisional ballots out of the way. [Figure 8]. A better way of avoiding this frequent error would be to have two poll workers witness the card programming and then turn the activator display around to let the voter confirm it before the card is activated.

5:10 p.m. Return to Washoe County Government Campus, Precincts 304, 413, 457 and Registrar's Office

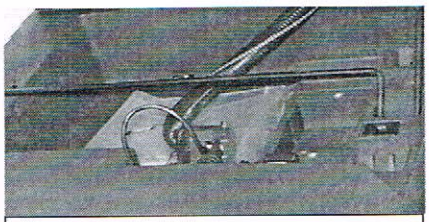

Figure 9. Unsecured Paper trail printer power supply cable.

I looked at the cables behind the voting machine and noticed that the data cable was secured with screws but the power cable for the paper trail printer was not. Simply unplugging it while a person voted would make it malfunction and make the printer not register that person's vote [Figure 9]. When it came time to shut down the polls, two women worked together. They intended to each write down the readings so that they could check them. However, when the time came they forgot to do this and merely stood with the machine between them so that crosschecking figures became physically impossible [Figure 10]. The process was unfamiliar and unpracticed. In fact, they packed the first machine up without removing its voting module. Realizing their mistake, they had to reopen the box to retrieve the memory module.

Figure 10. Official behind machine transcribes machine counts

Figure 11. counting room adjoining public lobby.

The small counting room had 5 doors and some irrelevant computers and other things in it. The doors were opened more than they should have been for optimum security [Figure 11.].

The official responsible for counting showed me around. When I asked him how he intended to back up the results and deliver the information, he admitted that he hadn't quite decided. I then inquired if he intended to burn a CD, in case something happened to the original results. He confessed that he had not made any particular plans to do so, but might.

I took interest in a memory module. He pointed out that it was just like the USB memory he had in his pocket. I noticed that the voting servers had slots for the memory cards and also extra USB connectors. I do not mean to imply that this well-meaning official had any nefarious intentions to defraud machines. However, the ease with which someone else might install a memory transfer device in the counting machines gave me a start. It was a particularly frightening end to a long day.

CONCLUSIONS:

Voting technology companies, election officials, outside research groups, and large segments of the public at large have worked hard since the 2000 election to improve the voting process. Numerous advances have been made. However, these changes will be meaningless if we do not give poll workers and election officials enough training and support. In order to operate polling places accurately, reliably and efficiently, those responsible must know how to both avoid problems and to resolve them when they arise.

Elections do not happen every day. The tasks that need to happen when they do take place are not, therefore, ingrained by force of habit and repetition. Poll workers and voters alike need instruction. All poll workers must have the chance to demonstrate that they can do their job on Election Day so that their learning process and mistakes do not compromise the integrity of actual elections.

In every instance that an election worker turns on a voting machine; copies down odometer numbers; opens a printer; programs a smart card; or views tally results on a backend computer he/she should be accompanied by someone to corroborate the correctness of these actions. Moreover, every person working at a polling place should be issued with a laminated checklist of instructions, which might even be worn around the neck. This would eliminate guesswork and the need to remember important but counterintuitive tasks (such as having two people independently recording odometer readings) because the instructions say to do so.

Most people working the polls in Reno had only received one hour of training. At all but one location only the manager had been specifically instructed about how to put paper in the new printers. However, she was never taught that a Paper Trail printer should be handled with the same care and security protocol. as a ballot box.

Voters should similarly have an opportunity to learn how to use new technology. Only one location that day offered all voters printed instructions with graphic illustrations. Another was careful to explain how the machines function. However, several thought that they had received no instructional materials suitable to offer voters. In most places, voters were left to walk up to the machines and read the instructions on them. People shouldn't have to figure out how to use voting machines on their own. Nor should they have to forfeit their right to secrecy by seeking assistance while in the act of voting.

As we become more dependent upon technology, it is imperative that we properly test this equipment for basic functionality as well as for security. Breakdowns occur and we must be ready for them. A standard method for handling such situations must be designed so that problems such as printer paper jams can be corrected without compromising the safety or secrecy of paper trail receipts.

Most importantly, we must create a checklist of operating procedure from start to finish. It does no good to improve the actual process of voting if machines are not secured before the election and if results are not responsibly transferred to election headquarters after the fact.

We have come a long way. The September 7 election in Reno shows us that we have a long way yet to go. If election officials in Reno immediately implement a system to insure that smart cards are correctly activated (for example, by having two workers and the voter corroborate the codes); if they improve the physical arrangement of the machines so that voter secrecy is respected; if they train precinct managers to always have a co-worker independently transcribe startup and shutdown readings; and if they plan a very early morning start for opening the polls they might run an accurate and secure election on November 2nd. The changes are small and not hard to implement, however, they could make a crucial difference.

JUS
JOURNAL OF USABILITY STUDIES

Vol. 2, Issue 1, November 2006, pp. 7-21

A Methodology for Testing Voting Systems

Ted Selker
Media Lab
Massachusetts Institute of Technology
20 Ames Street
Cambridge, MA 02139
selker@media.mit.edu

Elizabeth Rosenzweig
Bubble Mountain Consulting
62 Athelstane Road
Newton, MA 02459
eliz@bubblemtn.com

Anna Pandolfo
Media Lab
Massachusetts Institute of Technology
20 Ames Street
Cambridge, MA 02139
ap@media.mit.edu

Abstract

This paper compares the relative merit in realistic versus lab style experiments for testing voting technology. By analyzing three voting experiments, we describe the value of realistic settings in showing the enormous challenges for voting process control and consistent voting experiences.

The methodology developed for this type of experiment will help other researchers to test polling place protocols and administration. Comparing the results from laboratory experiments with voter verification and realistic voting experiments further validates the procedure of testing equipment in laboratory settings.

The methodology and protocol for testing voting systems can be applied to any voting technology. This protocol matches the real-world conditions of voting by replicating them for the experiment.

Keywords

Usability testing, voting technology testing, methods and tools, voting systems, audit trails, user-centered design.

Introduction

Voting is a difficult task. People don't vote very often, but when they do, they care very much about the result. A vote carries weight; it has great consequences, and therefore, voting machines must work as promised. Ballots often present the voter with a decision-making task where selections are made from a busy field of choices.

In 2002, the US Congress passed the Help America Vote Act. (HAVA, 2002), It included detailed guidelines for voting technology, such as voter verification and the ability to easily change a vote before it is cast and counted. This means that if the voter finds an error, he or she should be able to correct it. Voting systems today suffer many problems that result in lost votes, ranging from bad ballot design, usability issues around selection and casting of votes, to polling place problems, transportation, and counting procedures. These issues present obstacles to the accuracy, integrity, and security of valid vote counts.

Usability problems and registration issues are two of the biggest problems seen in recent elections (Ansolabehere 2004). Concern has been expressed regarding errors in voting, since such errors can result in a vote for the wrong candidate or in a residual vote, which renders the count invalid (Bowler, Donovan, and Happ, 1992), (Miller and Krosnick, 1998), (Wand, Shotts, Sekhon, and Mebane, 2002), (Sled, 2003), (Niemi and Herrnson, 2003).

Sled's *ex-post facto* study (Sled, 2003) on the evaluation of voting systems analyzed the California Governor Recall election of 2003. This election presented the voter with 138 selections for governor. The study showed that a vertical alignment effect caused 5% more selections for the adjacent three front-runners, on a punch card system than on an optical-scan or touch-screen system. This position is supported by results of other studies as well (Ansolabehere and Stephen, 2004).

The HAVA legislation created the Election Assistance Commission in 2002, which provides the first steps in developing guidelines for voting systems. As of this writing, only a few studies examine usability guidelines for voting, and none of them use the same methodologies. The value of our proposed methodology is that it is built on previous work for testing voting technologies.

The Caltech/MIT Voting Technology project developed several prototypes to demonstrate and test ballot design and verification auditing schemes for voting (Goler, Selker, and Wilde, 2006) and has run several studies to examine these issues. Two studies replicated the environment and condition of normal polling places. The first was the New York Reading Disabilities study, which compared full-faced ballots to standard Direct Record Electronic (DRE) voting machines for people with disabilities. The second study was the Arlington Voter Verification, which tested voter verification schemes in a realistic setting.

The Voting Technology project has also developed a simulator (Sullivan, 2005) to demonstrate and test ballot design and verification auditing schemes. This simulator can present a Voter Verifiable Paper trail as seen on a Sequoia or Diebold DRE, as well as experimental electronic ballots, such as the Lower Error Voting Interface (LEVI). The system can also present

an experimental, contemporaneous, Voter Verifiable Paper Audit Trail (VVPAT) with a paper trail that allows users to verify their vote after each selection, as opposed to waiting until the end of the voting process. Finally, the MIT simulator can present a contemporaneous Voter Verifiable Audio Audit Transcript Trail (VVAATT) (Goler, Selker, and Wilde, 2006). This simulator was also tested with the LEVI ballot design in the MIT LEVI Study.

The Voter Verification study compared four different verification approaches. Each of the following approaches produced an audit record of the vote:

- Audio with an audiotape record on an analog audiotape.
- A printed list of selections after all the selections are made
- A printed list of selections as the selections are being made
- An optical scan record.

Related Studies

This paper describes the general methodologies to be used in testing voting technology. It is based on three studies: the New York Reading Disabilities voting study, the MIT LEVI Study, and the Arlington Voter Verification study. The New York and Arlington studies included the use of a protocol created to match the real-world conditions, as well as investigate both usability and error detection in voting systems. The MIT study focused on two verification auditing systems.

New York Reading Disabilities Study

The Reading Disabilities study (Selker, Goler, and Wilde, 2005) was a previous effort to design a research protocol to create an ecologically valid voting experience. The hypothesis was that the pressure of a real voting experience affects the overall user experience.

A full-face voting system presented a voter with all races on the ballot simultaneously. This study sought to observe how full-faced voting interfaces compared to others that show less information simultaneously. In particular, we looked at how a full face voting interface affects the votes of people with reading disabilities.

This study used ESS iVotronic DRE and ESS V2000 paper over buttons full-faced DRE voting machine. This system employed the LS large LCD full faced voting machine.

This study was conducted in June of 2004 using a protocol of local registered voters with actual poll workers at the New York City west side YMCA polling place. Twenty-six percent of the subjects tested had a previously diagnosed reading disability.

A mock election was run, using poll workers who normally staff the polls on Election Day. These poll workers were given training for this experiment and were paid their standard rate. On the day of the mock election, 97 subjects went through the experiment during the hours of 7AM to 7PM. The subjects found themselves in lines 3 to 20 minutes long waiting at various stages of the voting process, just as they would on Election Day at many polling places around the United States. These conditions led to many problems,

and only data from 41 subjects were able to be collected.

When they first arrived, the voters were checked in at a Registration desk. Although normally voters register before they vote, this protocol differed from normal conditions by providing reading assistance to everyone while they were filling out the pretest questionnaire and consent form.

Participants were required to review actual voting materials from the 2002 Buffalo NY ballot. This review took place in an area with refreshments and tables with people providing reading support. In addition, there was a station staffed by a poll worker who explained how the voting machines worked.

To screen for reading disabilities, the protocol included the use of a battery of cognitive abilities tests (Nelson-Denny, WRAT-3, and WAIS-III vocabulary subtest). The screening process lasted for about 20 minutes, and the tests were administered in between each voting experience. In this case, testing had two main purposes; one was to assess reading disabilities and the other was to provide a distracting activity that would help to reduce any learning transference effect between trials.

At the end of the study, a post-test "exit poll" questionnaire was filled out with reading assistance.

The Reading Disabilities study found that people with reading disabilities took longer to vote than people without that disability. In addition, those diagnosed with reading disabilities experience the fall-off effect less on full-faced voting machines, but make more mistakes than people without reading disability.

MIT LEVI Study

This study tested verification auditing systems using the LEVI (Sullivan 2005) electronic ballot (Cohen 2005). The study had subjects vote in four elections on two different voting machines, one with a paper audit trail VVPAT, and one with an audio audit trail VVAATT. It was conducted in a quiet basement lobby area at MIT during the spring of 2005. The participants were drawn from the MIT community. It did not try to replicate real-world polling place conditions.

The protocol included the introduction of errors on the ballot that tested whether the verification system was helping voters find the errors. The results showed that, of 106 ballots, only 14% of people with audio verification reported the errors, and only one person reported finding an error with paper trail verification. Even more striking was the difficulty that people had in reporting those problems. When asked if they saw any errors, 85% of audio verification people reported that they found an error and 7% reported found an error on the verification record. To test if these results could be replicated in a naturalistic setting, we deployed a second study, the results of which are also reported here.

	NY Reading Disabilities	MIT LEVI	Arlington Voter Verification
System and Ballot design	ESS full-face LCD prototype, mixed media product and screen by screen DRE	LEVI software prototype	Diebold software simulation
Environment	YMCA polling place	MIT Media Lab basement lobby	City hall polling place
Subjects	Registered voters, some with documented learning disabilities	MIT Community	Registered voters
Errors	No errors introduced	Introduced Errors	Introduced errors

Figure 1. Comparison of the Three Studies

Arlington Voter Verification Study

The Voter Verification study examined auditing schemes to help voters verify their vote before it was cast and counted.

The Voter Verification study built on the New York Reading Disabilities Study by creating an ecologically valid voting experience. This study employed a protocol using registered voters voting in a mock election in the town's regular polling place administered by real poll workers. To provide voters with information about the election, voting materials were acquired from a local newspaper and offered to the study participants for review. The voters were given voting cards with proposed selections for the races.

The hypothesis was that the effectiveness of verifying ballots might vary depending on the technology used. Four voting technologies were tested:

- Standard Diebold Optical Scanned paper ballot in which voters write the votes, which are then scanned.
- The MIT voting simulator with a Diebold-like ballot and a Voter Verified Paper Audit Trail (VVPAT), which prints a paper trail once the voting is completed.
- The MIT voting simulator with a Diebold like ballot and a Voter Verified Paper Audit Trail (VVPAT) with contemporaneous paper trail, which prints the vote each time a selection is made on the ballot.
- The MIT voting simulator with a Diebold-like ballot with aVoter Verified Audio Audit Trail Technology (VVAATT), which provided contemporaneous audio feedback each time a selection was made on the ballot. The audio feedback used a male computer-generated voice that spoke as each selection was made.

The protocol had each subject vote on four different machines. Before they voted, the subjects signed a consent form and filled out a pretest questionnaire. After the participants had finished voting on each system, they filled out a post-task questionnaire. Once they had completed all the voting systems, the participants filled out a post-test questionnaire.

The assignment of each subject to a specific protocol was partially counterbalanced to avoid order effects. An additional procedure was added to avoid some of the possible carry-over effects. This procedure was inserted between voting experiences for each subject. Scrabble® was chosen as a word-oriented syntactic experience to match, and hopefully, to mask the kind of activity that people performed when voting. Each time a subject changed machines, they were asked to create a few words on the Scrabble® board. Although initial results look good, further work needs to be done to examine the effects of that inserted distracter activity.

Experiment Goals

The Arlington Voter Verification experiment had several goals:

- To understand the ability of voters to recognize and correct errors that were introduced by fraud (experimental manipulation in our case) or errors made by the voter's wrong selection across the various proposed technologies.
- To observe whether voters verify their selections before, during, and after casting their votes. There were two kinds of errors purposely introduced into the ballot for the VVPAT system only.
- To observe the effect that the different verification methods have on the ability to change a vote before it is cast and counted.
- To improve the experimental protocol in an effort to ensure ecological validity by mimicking actual conditions of an election.

Subjects

The study included 48 participants, all registered to vote and currently residents of the town of Arlington, Massachusetts. The subjects had the following demographic characteristics of the registered voting subjects:

- Age ranged between 30 and 82 years old, with an average of 50 years old
- Five percent of the subjects didn't own any computer equipment, and had never had experience using one; 80% of the subjects reported using the computer daily
- 92% of the subjects agreed that voting process in general is an easy one.
- When asked the frequency with which the subjects vote, 64% reported to always vote during elections, and 30% reported to voting frequently

Overall Test Environments

The study was conducted in the city hall in Arlington, Massachusetts, in a multipurpose auditorium with a stage. The room is used for town meetings, elections, and many other events. It has balconies that hang over the areas used in the polling place, which create some darkened areas that are often used for voting booths. In this study, only the registration table was under the balconies. All other activities occurred in the well-lit areas.

The Ballot

The ballot used was a combination of two previous elections in the town of Arlington. Those two elections were November 2004 and a previous town election in September 2003.

The reasoning for using familiar ballots was that the experiment would exactly simulate the voting situation

by providing subjects with elections with which they were already familiar.

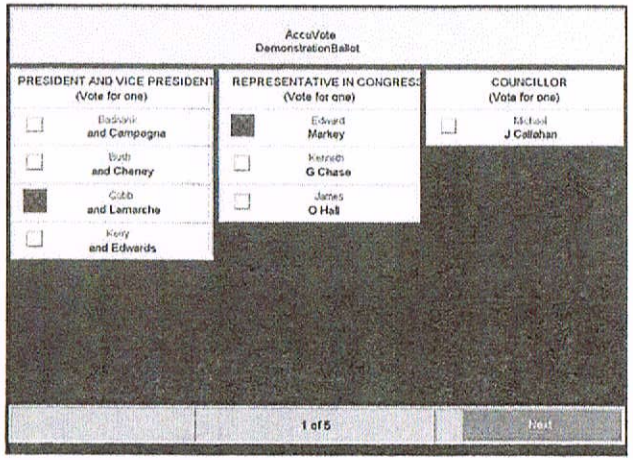

Figure 2. Diebold-style voting interface

The ballot used in the Voter Verification study had 13 races. Eight of the races had several candidates to choose from, while five races had only one listed candidate. Subjects had pre-assigned candidates to vote for in nine races, and were allowed to choose their own candidate in four races.

ERRORS INTRODUCED TO VOTING PROCESS

Since verifications was a goal of this study, two kinds of errors were purposely introduced into the ballot for the VVPAT system:

- Error 1: Changing a vote to a different candidate.
- Error 2: Removing a vote from a candidate. The vote is omitted in the audit prints.

RESULTS FROM VOTER VERIFICATION STUDY

Out of the 48 participants who ran this study, a rich data set was obtained for 35 participants..

ERRORS FOUND BY PARTICIPANTS

Although there were two fraudulent errors introduced on the VVPAT system, 83% of the participants did not report finding any errors. Though voters were encouraged to check their work on the paper ballot Optical Scan system, no one reported an error, even though 17% of the participants did, in fact, make a selection error on the paper ballot.

Subjects voting with the VVPAT contemporaneous system were presented a printout of each selection as they made it. In this process, they made 50% more candidate selection errors than VVPAT, the Optical scan, or VVAATT audio verification. The contemporaneous printout did allow them to find and report 60% of those erroneous selections during the verification process.

The VVAATT contemporaneous system used delayed computer-generated audio. Voting took longer with this system than with the other voting systems. The results from the VVAATT system showed that participants made half as many errors as on the VVPAT contemporaneous system. However, the results were very similar to the number of selection errors on the Opt Scan system. In the case of VVAATT, 60% of the errors were also recognized and reported during verification.

We find verification process valuable because it allows for a reduction of the total number of errors. Voters reduced their errors by 30% to 60% in systems that

provided verification compared to a 0% reduction in errors on systems that don't provide any means for verification. These results, although preliminary, corroborate previous findings by Cohen and Selker (2005) that point toward the possibility of a more accurate alternative audit technologies, other than paper trails alone.

Ease of Use

Subjects' perceived the Optical Scan paper ballots and VVPAT systems easier to use than the VVAATT and VVPAT contemporaneous systems. Since the Optical Scan was the technology they were used to, we believe that supported their perception it was easier to use. Furthermore, the lack of feedback for the mistakes on the Optical Scan gave them the perception that there were no problems.

One way to interpret this result is that the immediate feedback provided by the VVPAT contemporaneous and VVAATT might translate into a cognitive load that is higher than a system that does not interrupt the voter while they are voting. The fact that the contemporaneous feedback forces them to think about their vote again might explain the higher number of errors found in these systems. Furthermore, the interruption was disruptive and could explain the perception of a more difficult-to-use system.

From the descriptive measures, we observe that, for Optical Scan and VVPAT systems, 94% of the total subjects agree that these systems were easy to understand, followed by the VVPAT Contemporaneous system, with 83%, and finally, the VVAATT system with 77%.

The multivariate tests ($F(3,31)$ $p<0.001$) and the alternative univariate test ($F(3,99)$ $p<0.001$) are significant.

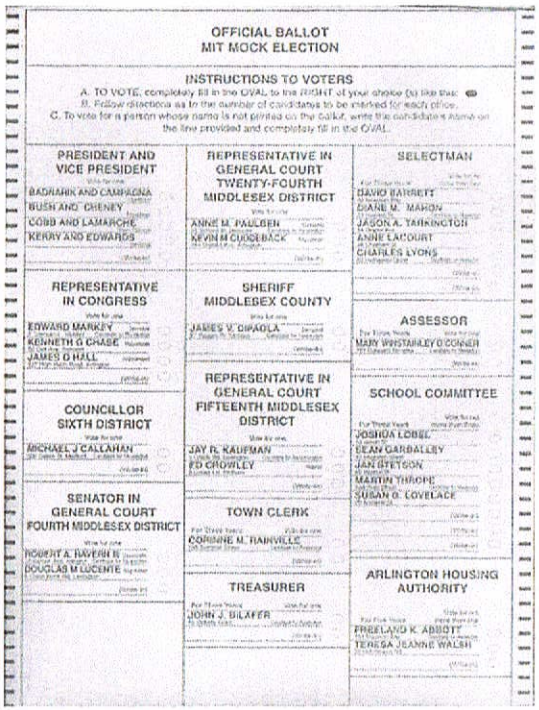

Figure 3.. Paper Ballot used in conjunction with Optical Scanner

Discussion

The methodology used in this study shows that real-world problems in voting places add vast burdens to the process. The data for comparing verification results in a laboratory versus ecological experiments are

consistently comparable. They show that the surrounding experience does not change the problems of ballot comprehension or verification results. The yield of valid data from the set of experimental subjects was affected greatly by the complexity of the protocol. In the Reading Disabilities study various problems with the protocol and environment meant that only 41 of 97 subject data were retrievable and useful. In the Voter Verification, the yield was a bit better with 35 out of 48 subject of the participant data being useful.

Many of the difficulties in the Voter Verification study had to do with voting machine testing and set up, which can be easily solved with more solid prototypes. Another set of problems stemmed from variation in protocol based on experimenter and poll worker confusion. These issues can be solved, in part, by training experimenters and poll workers ahead of time.

Finally, the use of a local ballot from the same town that the participants come from is very controversial for the subjects. In New York Reading Disabilities study, the two-year separation of ballot and experiment, and geographical distance from Buffalo may have helped them follow a voting card, since they were using a ballot that was not meaningful to them. In the Arlington Voter Verification study, a few voters purposely decided not to follow the instructions and wouldn't vote for the candidates pre-assigned on the voting card if these were not people they would normally vote for. Some subjects in Arlington and in New York even voted randomly to avoid giving away their selections.

We believe it is crucial that voters mark their voting cards themselves and make selections from those cards. This protocol would allow more control over tracking votes and errors.

This real-world protocol produced data that points to real-world problems, such as distraction when voting, confusion with new voting technologies, as well as poll workers playing a key role in voter confidence and security.

A method for testing voting machines that purposely introduces errors provides a solid approach for examining error detection. By knowingly introducing errors into the experimental design, researchers are provided with simple metrics to study the issues that are important in determining if voting verification is working.

The data pointed to key issues:

- Poll workers and polling place conditions may have as much to do with the usability of voting as the machines themselves. Poll workers may not provide consistent in directions or protocol. In both the Reading Disabilities study and the Voter Verification study, this was true, and many of pollworkers did not want to read the script or the protocol. This factor indicates a larger problem that must be studied concerning poll worker training. In any case, we recommend using trained experimenters to act as poll workers in voting studies that don't concern evaluation of poll workers.
- Voters in Arlington who had used Optical Scan ballots in their elections said they were comfortable with them. However Optical Scan ballots were the only ballots in which the voters were never successful at finding an error. This could be because the voters