

Public Hearing
on Proposed Certification of the
Vote-PAD/DESI and Vote-PAD/Hart Systems

August 9, 2006

Testimony of Ellen Theisen,
President, Vote-PAD, Inc.

The Vote-PAD is a non-computerized device to assist people with a wide range of disabilities vote independently and privately. It neither counts nor tallies votes. It simply provides the means for people with disabilities to mark the same paper ballot used by other voters.

It has gone through the complete approval process in one state — Wisconsin. During that process, the State Election Board staff, headed by Kevin Kennedy, who is the current President of NASED, held a mock election and consulted with their advisory board, including members of the disabilities community. All the people with disabilities who tested the Vote-PAD in Wisconsin were able to use it independently. It was approved for use with hand-counted paper ballots by the Wisconsin State Election Board. We have made sales in Wisconsin and are now in the process of delivering orders.

Many people who have tested the Vote-PAD during our research and development process have expressed great satisfaction with the way it worked for them. Testimonial letters from 12 of these people were included in our application for certification in California. I have attached another letter we just received from the Vice-President of the Board of Directors of Bucks County Association for the Blind, in which he points out that he and 10 other members of his group were “all favorably impressed with the easy to understand training instructions and the independence we experienced in completing the voting form.”

The California certification testing for two systems using the Vote-PAD took place on July 19 and 20, and the Secretary of State’s staff has recommended against certification of both the Vote-PAD/DESI system and the Vote-PAD/Hart system. They are basing this recommendation on their evaluation of the results of that testing. I will address each of their reasons individually, but first I want to address the testing process.

As I wrote to the Secretary of State before the staff report was released, I have serious concerns about the testing process. While it was intended to verify the usability and effectiveness of Vote-PAD for use in an election, the test protocol was planned and implemented by people who have no experience conducting usability testing — particularly with respect to persons with disabilities — and without the oversight of a human factors expert. And, the results were evaluated by people who have no experience in human factors.

My primary concern is that the testing procedures failed to simulate a live election, so it’s difficult to know how the results would compare with the results of using the Vote-PAD in a live election. It is clear, however, that in spite of the artificial environment, and with qualitatively more difficult tasks than voters would perform, participants did very well. Here are four of the significant differences between a live election experience and the test experience:

- ◆ First, in a live election, voters would be introduced to the Vote-PAD and left alone so they could concentrate on voting their ballot. But in the testing, the monitors interrupted each participant no fewer than ten times, and some monitors interrupted them additional times to correct a mistake, offer a suggestion, or make a chatty remark. As the report from our usability consultant points out, this constant interruption no doubt disrupted the participants’ concentration and may have increased the error rate measured by the staff as compared to the rate that would be experienced in a real voting setting.

- ◆ Second, blind voters have told us that, in a live election, they would move through the contests in order, since the Vote-PAD instructions are set up in a linear fashion. But in the testing, two separate times the monitors asked each participant to skip a race, then later return to it, and THEN return to where they left off. This means that blind participants had to reorient themselves six times to where they were on the ballot and coordinate that with the audio or Braille instructions. Again, the disruptive protocol grossly distorted the error rate that would occur in an actual election, where blind voters would not engage this type of hop-sotch voting.
- ◆ In a live election, it is rare to vote more than one write in a particular election, and realistically, voters with disabilities are unlikely to vote a write in candidate unless it is very important to them, since it is a complicated process on any system. But in the testing, each participant was required to vote four write-in candidates — even after some of them expressed dismay at the second one.

As our usability consultant's report points out about the test script:

“One valid reason for doing this would be if an assessment found that most California voters, in a given voting session, vote for four write-ins, skip a race twice, continue voting and go back to the race.”

- ◆ In a live election, voters will never verify a ballot marked by someone else. But in the testing, blind participant were asked to do just that.

The staff has recommended against certification of the Vote-PAD for four reasons. I will address each one individually.

1) The demonstrated error rates for voters to cast a ballot using the Vote-PAD is unacceptably high – particularly for voters with visual impairments;

Unacceptably high. Compared to what? The certification testing chosen by the staff for the Vote-PAD – using people with disabilities who volunteered to test the system -- is unprecedented. Even though the staff has tested many other systems and components that are intended to be used by voters with disabilities, there are no results from other similar tests to provide a benchmark for comparison, since no other system has been tested in this way. This is true for the Hart system, which was tested just days before the Vote-PAD system. No explanation has ever been given as to why the same testing protocols were not used in testing the Hart system, or the ES&S AutoMark system. Accordingly, there is no way of knowing whether the error rate would have been even higher with these systems, had they been tested in the same manner.

Moreover, even if the tests were valid, and even had they been applied to other systems, to our knowledge, the Secretary of State has not adopted any regulations that specify what error rate is “acceptable.” It appears to be a completely subjective determination by the staff.

Interestingly, a study conducted by three Rice University psychology professors and presented at the USENIX/ ACCURATE Electronic Voting Technology Workshop on August

1, 2006 gives us an indicator of voter error rates in a more realistic situation.¹ The abstract contains this description of the study:

“The usability of three different voting methods was evaluated using the metrics recommended by the National Institute of Standards and Technology: efficiency, effectiveness, and satisfaction. Participants voted with two different types of paper-based ballots (one open-response ballot and one bubble ballot) and one mechanical lever machine.”

Thirty-six people participated; 13 were university students, and the rest were recruited from the general Houston population. All were sighted and fluent in English. They had voted in an average of 10 elections.

The study states: **“Overall, error rates were low but not negligible, with participants making errors on slightly less than 4% of the races.”**

To review: able-bodied, sighted college students and Houston residents who had previously voted in an average of 10 elections made errors on 4% of the contests.

Some participants were given a slate of candidates to vote for. Some were allowed to choose on their own. All were left alone while they voted their ballots. The error rate was 4%.

The scientists also checked the error rate “by ballot” and found that **“nearly 16% of the ballots contained at least one error.”**

Now, the 10.3% error rate recorded in the Vote-PAD testing can be seen in perspective.

The Secretary of State’s staff report tells us that 55% of the errors were made on the write-ins. That’s over half of the errors, and 21% of the total errors were because people didn’t fill in the write-in bubble. The staff admits, “While this is not uncommon with all paper ballots, **comparison rates for other voting systems are not available.**” That is precisely the problem. The staff is seeking to hold the Vote-PAD to a standard that has not been applied to any other system intended for used by voters with disabilities. And it is undisputed that the ability of blind voters to cast write-in votes is a recognized problem and a challenge no matter what system is involved.

It is very telling to examine what the error rate is when we take the write-in errors out.

If we remove the 124 write-in selections and the associated errors from the totals, participants had a 6.3% error rate on the selections they made on their own. And, this is a group of people made up of many who had never voted independently in a real election, many who were blind, several with development disabilities — and all of whom were subjected to at least ten interruptions while they were trying to concentrate on using a product that was completely new to them, in a situation that was entirely new to them and much more intimidating than a real election.

¹ *A Comparison of Usability Between Voting Methods*. Kristen K. Greene, Michael D. Byrne, and Sarah P. Everett; Department of Psychology, Rice University.
http://www.usenix.org/events/evt06/tech/full_papers/greene/greene.pdf

It is notable that in staff reports on other voting systems, some errors are called “artifacts of the testing process” and discounted entirely. For example, the staff report for the eSlate says:

“During the volume test, there were a noticeable number of instances of voter codes that were invalid. In several of these instances, we were unable to reconcile the statements of the voters, time stamps on the codes, and the status the system would show for the code. This issue is **probably** an artifact of the test methodology and operator error.”

Another staff report says that, of the seventy-eight errors logged during the Sequoia Edge I test:

“fifty-two recorded errors (66.67% of the recorded errors) were **attributed to human errors as artifacts of the testing process itself**. ... Most of these errors would not occur in a polling place on Election Day.”

However, none of the errors in the Vote-PAD report is discounted as an artifact of the testing process. I only received video tapes of the first day of testing last Friday afternoon, so I had time to examine the testing of only a few of the participants. But in the videos of two blind participants I watched, I counted three errors that I believe are artifacts of the testing process. Here’s one example:

The participant didn’t mark the write-in bubble when asked to write-in for Attorney General. After completing the write-in sheet, he was asked to skip back to the Secretary of State contest and vote, then return “to where he left off”. He couldn’t remember where he left off, so he used the verification wand to find his place. Since Attorney General didn’t have a mark in the bubble, he decided that was where he left off, and he voted for one of the listed candidates.

In a live election, of course, he would have remembered that he had decided to write in a candidate for Attorney General, and he would have corrected the write-in bubble, rather than voting for a different candidate. Nevertheless, this event was counted as an error by the monitor and not discounted as an “artifact of the testing process.”

Our usability consultant’s report points out that many additional opportunities for error, which in no way reflect the actual voting process, were introduced by the complicated tasks required of test participants. Realistically, those errors must be considered “artifacts of the testing process.” Therefore, the 10.3% error rate proposed by the staff is artificially high. In fact, since the staff did not evaluate the number of errors that fall into that category, we have no idea how much too high. How many errors recorded during the testing would simply not have occurred in a live election, where people were voting for their real choices? We won’t know until we have access to all of the tapes and compare them properly against documentation.

The bottom line is that because of these problems with the design and implementation of the test, the error rate is distorted. Given this fact, and the failure of the staff to subject other systems to the same type of testing by persons with disabilities and arrive at an error rate for those systems, there is no basis for concluding that the error rate is “unacceptably high.” This is particularly true in the absence of any objective criteria established by the Secretary of State for determining acceptable voter error rates.

2) Voters with visual impairment cannot reliably cast a write-in ballot using this system;

Here's a brief and illuminating paragraph from a second voting system usability study presented at the workshop last August 1, 2006.² It will help put this issue, too, in perspective.

“Write-in candidates present a major challenge for voters. For a paper ballot/optical scan system, the problem is that they often forget to fill in the oval or complete the arrow that signals to the optical scanner that a write-in vote has been cast, a failure that could result in these ballots not being counted at all. With electronic systems, write-ins took more time than with the paper ballot, but other problems existed as well.”

In fact, significant problems are noted on three different electronic systems, ranging from the size and position of the write-in window, to confusing user interface, to difficult navigation.

Unfortunately, it appears that voters, with or without visual impairment, cannot reliably cast a write-in ballot using **any** system.

- ◆ As is true with all voting systems, writing in a candidate on the Vote-PAD is not easy for voters with visual impairments. However, many of the participants cast exactly the write-in they were told to cast. In fact, I watched a tape in which one blind participant accurately entered four write-in candidates' names on write-in sheets. Other blind participants legibly wrote candidate names directly onto the ballot through the write-in slots. Others legibly wrote the contest and name on the back of the write-in sheet. As Registrars will tell you, because the intent of these voters can be determined, these are all valid write in votes.

One problem with the write-in testing, which became apparent very quickly, was that the consultants and Secretary of State's staff were not familiar with Vote-PAD instructions and didn't understand the write-in process for the device. Because they did not know how it worked, they assigned write-in tasks to participants in a way that is counter to the Vote-PAD audio, Braille, large-print, and poll worker instructions. In fact, the staff report illuminates that the authors had not read the large-font instructions or had anyone check the Braille (highlighting added) and were simply making assumptions about the instructions:

“Once the ballot is completed, audio instruction (and **presumably** Braille and large-font instruction) guides the voter back through the write-in process, identifying each contest and its contest number in turn.”

If the monitors had understood this feature of the system going into the testing, we don't believe they would have given instructions that counter those given by the Vote-PAD device. In fact, one participant, on reaching the relevant instructions in the Braille book, said, ‘Doggone it. They have something in the back of the book that would have made it real easy.’”

² *The Importance of Usability Testing of Voting Systems*. Paul S. Herrnson, University of Maryland, Richard G. Niemi, University of Rochester, Michael J. Hanmer, Georgetown University, Benjamin B. Bederson, University of Maryland, Frederick G. Conrad, University of Michigan, Michael Traugott, University of Michigan. http://www.usenix.org/events/evt06/tech/full_papers/herrnson/herrnson.pdf

It is clear that the monitors didn't realize that the write-in process happens as the SECOND step to casting a ballot rather than being a simultaneous task – which is how they instructed participants to proceed. Perhaps the instructions that the poll workers give could be modified to make this more clear to voters in the polling place.

This lack of knowledge about the write-in instructions suggests that the staff may not have been sufficiently conversant with other instructions and procedures for use of the Vote-PAD, and their unfamiliarity could have led to additional mistakes and misperceptions about the Vote-PAD.

- ◆ It is notable that, in the staff report on the Vote-PAD, three of the eight pages discussing test results were devoted to write-in issues. However, in the staff report on the Hart eSlate system which was tested a week prior to the Vote-PAD, no write-in issues were raised, even though the consultant's report points out three severe problems with the write-in feature:
 - ◆ It is possible for the eSlate audio to report that one letter was entered when actually another was entered,
 - ◆ There is no way for a blind voter using the eSlate to review the name of a write-in candidate, and
 - ◆ The eSlate's audio feedback for the Clear/Backup function is confusing.

Since the eSlate can give inaccurate feedback on write in names, and – unlike the Vote-PAD – provides no way for voters to verify the name, it appears that voters with visual impairments can cast a write-in vote on the eSlate system even **less** reliably than on the Vote-PAD. It's difficult to understand why the staff didn't consider the eSlate write-in issues enough of a problem to include in their report.

And because the staff chose not to test the eSlate by having voters with visual impairments actually vote on the system – as they did with Vote-PAD – there is no way of knowing how high the write-in error rate would have been, or how reliably blind voters could have cast write in votes on those systems. And again, because to our knowledge the Secretary of State has no established standard for objectively measuring such reliability, it is difficult to understand what subjective standard was used to reach the staff's conclusion.

Certainly write-in candidates present a major challenge for voters using the Vote-PAD, but there is no evidence to suggest that the challenge is greater than it is on other systems, since this stringent testing by people with disabilities has not occurred on any other system certified in California.

3) There is no reliable method for voters with visual disabilities to independently verify their complete ballot, or even that they have been issued the correct ballot and that it has been properly aligned in the Vote-PAD booklet;

This assertion has three parts: no reliable method for blind voters to verify a ballot, no ability to reliably verify that they have been issued the correct ballot, and no ability to verify the alignment of the ballot. I'll address them separately.

- 1) During the testing, some participants with visual impairments used the wand to verify their votes as they moved through the ballot. Some found an error they had made. Others used it with great success to find their place after being asked to skip around the contests. Few if any, however, went through the entire verification process at the end of voting their ballots. The video tapes shows that, for some, this may have been because the monitors told them it was time to cast their ballots before they got to that part of the audio or Braille. One expressed concern that he had already taken up too much of the monitor's time.

Since the first part of the testing fails to show whether blind voters can reliably verify their ballots, perhaps the staff is basing their conclusion on the results of the Phase II testing, in which participants with visual impairments were asked to verify a ballot and write-in sheet they had not marked. This, of course, would never happen in a real election and provides no basis whatsoever for determining how reliably a blind voter could verify a ballot that they had personally voted.

Amazingly, some of the participants in Phase II were asked to verify the name on a write-in sheet without being told the name that was written in. It is absurd to ask a blind participant to search through 780 cells to find marks they didn't even make – like needles in a haystack. It isn't surprising that many refused to try and that not one of them was completely successful.

Furthermore, it is unreasonable to expect that participants, exhausted and frustrated from skipping around the ballot in Phase I, would be able to, or even motivated to, succeed at such a harsh and meaningless task, unrelated to a live voting experience.

- 2) Now for the second part of the staff's assertion – that voters with visual impairments cannot reliably verify that they have been issued the correct ballot. The staff report suggests a simple solution, with which we agree. Allow a second poll worker – or a friend who might have accompanied them to the polling place – verify that they have the correct ballot. Easy, reliable, and it would take only a minute or two.

It's certainly simpler to solve this problem on the Vote-PAD than it is on systems that require poll workers to correctly program the voter access card – a problem that often occurs with the electronic systems.

- ◆ For example, 7000 voters in Orange County, California were given the wrong ballot on a new eSlate system in the March 2004 primary. The LA times estimated that 1500 voters cast the wrong ballot.
 - ◆ Just last month, in Tennessee, some Smyrna City voters found that the City elections were missing from the last screens of their electronic ballots. Other Tennessee voters were given electronic ballots with the wrong State House contests. It took poll workers 40 minutes to correct the problem.
- 3) Finally, the third part of the staff's assertion – voters with visual impairments cannot verify that the ballot has been properly aligned in the Vote-PAD booklet. Here again, it is apparent that the staff is not fully familiar with the Vote-PAD system they were testing.

The Vote-PAD is designed with alignment holes carefully positioned at the bottom of the sleeve for precisely the purpose of allowing blind voters to verify the alignment of their ballots. During testing this issue did not arise, but if voters have questions about the alignment in a live election, poll workers will demonstrate how to use those holes for verification before the ballot is ever voted. The process is much simpler and more reliable than calibrating a touch screen.

4) The length of time involved for voters with visual impairments to cast a ballot is excessive and likely to have a negative impact on voter participation.

Excessive? Compare to what standard?

Again, the staff has never subjected any other system – including the AutoMark or eSlate – to testing by numerous persons with a range of physical and developmental disabilities. Only the Vote-PAD has been singled out for that type of testing. Accordingly, there is no way of knowing how much additional time would be required on the AutoMark for participants to find their place in the instructions and on the ballot after skipping a contest, or how much additional time would be required for returning to where they left off.

Nor is there any way of knowing how long it would take blind voters to vote a full primary ballot if they were directed to engage in the time-consuming task of voting four write-in candidates. So there is simply no data to which we can compare the voting times on the Vote-PAD versus other systems. Accordingly, it appears that the staff has merely made an assumption that voting would be faster on other systems, and that therefore the amount of time to vote on the Vote-PAD is comparatively excessive, with no comparative, empirical data to support that speculation.

In conclusion, there seems to be a staff assumption that when electronic systems are used by people with disabilities, they will have an acceptable error rate, will not take an excessive length of time, will provide reliable write-in capabilities, will be easy to use, and will allow the voters to verify their intent. Therefore, the accessibility testing that was conducted on the Vote-PAD was not conducted on other systems that will be used by voters with disabilities, not even on the Hart InterCivic eSlate, which was tested just one week before the Vote-PAD.

On the other hand, there appears to be an assumption that those advantages are not present in the Vote-PAD, so stringent, unprecedented testing was done. I maintain that, in spite of the inappropriate testing which inadvertently set up the participants for failure and frustration, their remarkable successes showed the Vote-PAD to be at least as accessible as, if not more than, other systems certified for use in California.

Accordingly, the certification of the Vote-PAD for the November 2006 election is fully warranted, on condition of certain modifications in procedures and instructions. This conditional certification approach has many precedents in previous certifications.

For example, according to the staff report on the AutoMark:

“The AutoMARK is relatively slow. ... Longer ballots and employment of the alternate audio or binary interfaces could significantly lengthen this time, making it prohibitive for use by the majority of voters and raising potential issues with Election Code §14224 ...”

“... there is no ability on the AutoMARK to pause the audio instruction.”

“During testing, staff was able to induce the AutoMARK to crash by rapidly and repeatedly inputting instructions.”

And then:

“It should be noted that none of these issues are significant enough to delay certification of the AutoMARK VAT. This is a first-generation device and it is expected that future versions would bring improvements in most of these issues.”

Certainly for new systems, the staff finds problems that need remedying. Rather than denying certification, the normal procedure is to recommend solutions. In most cases, the staff recommends certification, along with modification of polling place procedures. But in the case of Vote-PAD, they recommended denying certification, even though improved polling place procedures would provide solutions to all their concerns, except those caused by the test situation itself.

We submit that, contrary to the staff recommendation, there is sufficient evidence to justify the certification of the Vote-PAD. To ensure its successful implementation in polling places, we recommend the following conditions, and would be agreeable to other reasonable conditions for using the Vote-PAD in the November 2006 election:

- ◆ One additional poll worker will be employed in each polling place specifically trained on the Vote-PAD and dedicated to instructing voters in its use.
- ◆ Poll workers will emphasize that voters who wish to use the write-in sheet should wait until they have completed their ballots, so they have access to the contest identification scheme which is at the end of the guide to contest choices. Vote-PAD documentation will be modified accordingly.
- ◆ Poll workers will tell voters that they may turn the write-in sheet over and use the back to write in a contest and candidate if the write-in slot is too small and they don't want to use the write-in grid. Vote-PAD documentation will be modified accordingly.
- ◆ Poll workers will emphasize that voters using the audio or Braille should move through the ballot in order, so as to avoid confusion. Vote-PAD documentation will be modified accordingly.
- ◆ Once a ballot has been inserted into the ballot sleeve and aligned, and before the ballot is voted, a second poll worker or friend of the voter will check that it is the correct ballot and that it has been properly aligned.
- ◆ With the cover partially closed to protect the privacy of the previous voter, each sleeve will be cleaned after use with rubbing alcohol to remove stray marks and other residue. Vote-PAD documentation will be modified accordingly.
- ◆ The pens will be black, but not indelible ink. Vote-PAD documentation will be modified accordingly.
- ◆ Wet wipes will be provided for voters to remove ink from their hands. Vote-PAD documentation will be modified accordingly.

- ◆ Write-in sheets will be stapled, not paper-clipped, to the ballot. Vote-PAD documentation will be modified accordingly.
- ◆ The Secretary of State will assign polling place monitors in a number he determines appropriate to observe the use of the Vote-PAD in the November 2006 election.

As a final point: Assuming a system is accessible, just because it is computerized, is an injustice to the disabilities community.

California owes it to its disabilities community to conduct appropriate accessibility testing on all systems that it certifies for use by people with disabilities. We urge the Secretary of State to conduct an appropriate test of the Vote-PAD for full certification in the near future, and to include appropriate usability testing in its test protocol template for all systems.

Attachments:

Letter from Alvin Blazik, Bucks County Association for the Blind

Comments on the Vote-PAD testing process by Valerie J. Berg Rice, Usability Consultant to Vote-PAD, Inc., including qualifications.

The Importance Of Usability Testing Of Voting Systems, Herrnson, et. al.

A Comparison of Usability Between Voting Methods. Greene, et. al.

California Secretary of State's Staff Report on the AutoMark system.