



March 9, 2007

Dear Mr. Flater,

Thank you very much for the opportunity to help the Core Requirements and Testing Subcommittee with its important mission. However, your letter and questionnaire pose considerable difficulty in crafting a responsible answer.

First, your letter says "the Core Requirements and Testing Subcommittee of the TGDC is preparing advice to the EAC..." My understanding of this process is that the sub-committees prepare items for consideration for the entire TGDC. The TGDC then meets and provides recommendations to the EAC. Perhaps the explanation in the letter is an abbreviated description of the process, but it is important to NASED that the entire TGDC has the opportunity to review any recommendations as important as error rates in voting equipment.

Secondly, the basic question being posed is a political one. As election administrators the critical answers are not technical. How much can we, as election officials, tolerate before errors are unacceptable? The 'correct' answers are:

1. How many device failures are acceptable in a given election?
None. Every vote must count.
2. How many counting errors are acceptable in a given election?
Fewer errors than the vote margin between first and second place candidates. The errors can be tolerated if they will not affect the outcome of the election.
If the error occurs in such a fashion that it is automatically detected and can be resolved and verified before reporting to the public in precinct polling place reports, election night reporting, and final canvass reporting, then a higher rate of errors would be tolerable. The error rate becomes a comfort zone criterion between where the local officials can tolerate the process to detect and correct and the system becomes unacceptable for further use in an election.
3. On average, how much volume is typically processed in an election?
This varies from state to state, jurisdiction to jurisdiction and election to election. There is no "typical" volume.

However much precision we demand in the ideal, the technical measures in the real world of machines and testing require technical standards. The difficulty in answering the questions as presented is that the questions themselves ignore the complexity of the issue the subcommittee is attempting to address.

Ideally, election administrators understand the limitations of their equipment and manage those limitations by having spare parts on-hand, extra machines to swap in, trouble-shooters roving the polling sites, maintenance procedures, and emergency supplies of paper ballots and disaster plans. Any failure that leads to even a single unrecoverable lost vote is unacceptable. In this area, no failures are permissible, so we design administrative processes to cope with the fact that perfection is not achievable.

A distinction needs to be made between the different designs of voting equipment and the effect that an equipment failure has on the voting process. For example, the failure of a polling place-based optical scan system usually means the loss of "second chance" voting but doesn't impede voting per se. On the other hand failure of a DRE device reduces the capacity of the polling place to process voters, unless either additional DREs or a back-up supply of printed ballots is available. Is the loss of "second chance" voting acceptable? For how long? With what frequency? Running out of paper ballots where that is the only method of voting, on the other hand, means that no one can vote until additional ballots are delivered to the polling place.

We need to distinguish among "failures" that can be fixed by poll-workers, vs. technicians, vs. machine replacement. The more difficult the solution, the lower the failure rate that can be tolerated. For example, if a VVPAT jams with some systems, poll-workers can be trained to fix this. If a touch-screen needs re-calibration, a technician must go to the polling place to correct the problem or replace the machine. Some equipment problems can only be corrected by replacing the machine.

Voting equipment also must be able to tolerate sitting around for long periods of time without being used and then work when it is turned on. Election administrators often find that a certain (usually small) percentage of machines fail at this point, so they have extra machines to cover this eventuality

For the purposes of developing testable standards, consider five categories of reliability of system components such as DREs and optical scan precinct count devices:

First, there is reliability based upon the design. Some designs are simply more prone to failure than others. This kind of design defect cannot be addressed in manufacturing quality control. Ultimately any manufacturing changes introduced to address the problem would really be design changes. In an ISO 9000 company the manufacturing group would not be allowed to make a change until the design was formally changed and documented. Other organizations with less formal structure might make such a change from the factory floor. Either way items in this category fail to meet expectations for reliability based upon design.

Second, perhaps properly part of design, and fairly new art is human factors design. How tolerant is the item of mishandling and operator error and to what extent does the design induce operator error. Will voters, poll workers, and system administrators be driven to improper actions by the design of the system and will those actions cause equipment failures?

Third, there is reliability based upon quality control in manufacturing. Reliability suffers from variations from design in the manufacturing process. A really good example discussed in public records is the defect that the Freeman, Craft McGregor Group found in the Sequoia Precinct scanners during volume testing performed for the City of Chicago and Cook County. (see: [http://electionupdates.caltech.edu/Chicago and Cook Report.pdf](http://electionupdates.caltech.edu/Chicago_and_Cook_Report.pdf)) In that instance a coat of Teflon paint on the ballot guide bars had been applied a few thousands of an inch thicker than the design. That was causing a gap between the guide bars and the mouth of the scanner. That gap combined with other issues brought a high rate of reported ballot jams in the Chicago/Cook 2006 Primary.

Fourth, there is reliability over the life cycle of the device. Virtually all components are going to have increased failure rates over time due to aging of components, the deterioration caused by use, and sometimes, lack of use. How reliable is the equipment given proper preventative and routine maintenance? Also, how reliable is it if it is poorly maintained?

Fifth, there is the reasonableness, affordability, and achievability of the tasks necessary to properly maintain the equipment and keep its reliability level high. Voting equipment is infrequently used, but the demand for reliability is constant.

The first three of these reliability categories can be evaluated with a high volume test using operators with a variety of experience, education, and intelligence. The California Volume Test Protocol (see: http://www.ss.ca.gov/elections/voting_systems/volume_test_protocol_final.pdf) does a pretty good job of finding Design, Quality Control and Human Factors issues with voting system equipment. It is an expensive test. If every state started doing it, it would drive the cost of systems sky high. This is a test that really needs to be conducted either at a national level or as a joint effort by a group of states with similar requirements.

The fourth category is somewhat evaluated by the shake and bake tests. However, unless there is some consistent monitoring and evaluation of equipment failures during elections, it is impossible to evaluate. One of the big problems with this issue is the inability of anyone to capture all of the election incidents, categorize them as to severity and sort them out to misunderstandings of fact, human error and confusion, poor or improper maintenance and real equipment failure. Then the failures need to be sorted into the five categories of

reliability. This is difficult to sort out because so much of it is dependent on the people observing failures, capturing evidence of failures and reporting failures.

Finally, it will be a challenge to provide any good hard standards for the Fifth category: reasonableness, affordability, and achievability of the tasks necessary to properly maintain the equipment and keep its reliability level high. One can measure maintenance requirements in dollars per year, hours per year, and/or the availability and maintenance of the required resources for maintenance. Up until now evaluating this has been left up to jurisdictions purchasing systems were looking at a 10-year cost of ownership when they made their purchase decisions. Many states have made disclosure of maintenance schedules and costs of service and parts a required element of the technical data packages and made it available to their counties. There were possibly some jurisdictions that did not do this analysis. There are two issues buried in this. One is the cost of proper maintenance. The other is the realistic achievability of proper maintenance. A particular risk here is a maintenance step that is poorly defined and for which the supervisor of the maintenance work cannot easily evaluate the quality or success of the work being performed by the workers.

In the development of the 2002 Voting System Standards, the FEC selected '1 in 10,000,000' as a compromise that was testable and affordable (higher standards require disproportionate amount of resources to evaluate). We find no reason to change that standard but rather recognize a need to review how the criteria are to be evaluated. A '1 in 10,000,000' standard is easily met for perfect test ballots but may not be achievable in an actual election.

Thank you very much for the opportunity to participate in this important discussion.

Sincerely,

Sandra J. Steinbach
Chairperson
NASED Voting Systems Board