

Testing Voting Systems

Part of [the Voting and Elections web pages](#)

by [Douglas W. Jones](#)

[THE UNIVERSITY OF IOWA Department of Computer Science](#)

Overview

Voting systems are subjected to many tests over their lifetimes, beginning with testing done by the manufacturer during development and ending on election day. These tests are summarized here, along with a brief description of the strengths and weaknesses of each test:

- [Internal testing at the vendor](#)
- [Independent testing authority certification](#)
- [State qualification tests](#)
- [Tests conducted during contract negotiation](#)
- [Acceptance testing as delivered](#)
- [Pre election \(logic and accuracy\) testing](#)
- [Testing as the polls are opened](#)
- [Parallel testing during an election](#)
- [Post election testing](#)

Internal testing at the vendor

All responsible product developers intensively test their products prior to allowing any outsiders to use or test them. The most responsible software development methodologies ask the system developers to develop suites of tests for each software component even before that component is developed. The greatest weakness of these tests is that they are developed by the system developers themselves, so they rarely contain surprises.

Independent testing authority certification

Starting with the 1990 FEC/NASED standards, independent testing authorities (ITAs) such as Wyle Labs have tested voting systems, certifying that these systems meet the letter of the "voluntary" standards set by the Federal Government and required, by law, in most states. Several states, such as Florida, that impose additional standards contract with the same labs to test to these stronger standards.

The ITA process has two primary weaknesses: First, the standards contain many specifics that are easy to test objectively (the software must contain no "naked constants" other than zero and one), and others that are vague or subjective (the software must be well documented). The ITAs are very good at testing to the specific objective requirements, but where subjective judgment or vague requirements are stated, the testing is frequently minimal.

Second, there are many requirements for voting systems that are obvious to observers in retrospect, but that are not explicitly written in the standards (Precinct 216 in Volusia County Florida reported -16,022 votes for Gore in 2000; prior to this, nobody thought to require that all vote totals be positive). The ITA cannot be expected to anticipate all such omissions from the standards!

Finally, the ITA tests are almost entirely predictable to the developers, as with the vendor's internal testing. Barring outright oversights or carelessness on the part of the vendor, and these do occur, and barring the vendor's decision to use the ITA process in lieu of an extensive internal testing program, the ITA testing can be almost pro forma. Catching carelessness on the part of vendors and offering a guarantee that minimal standards have been met is sufficiently important that the ITA process should not be dismissed out of hand.

State qualification tests

While some states allow any voting system to be offered for sale that has been certified to meet the "voluntary" federal standards, many states impose additional requirements. In these states, vendors must demonstrate that they have met these additional standards before offering their machines for sale in that state. Some states contract out to the ITAs to test to these additional standards, some states have their own testing labs, some states hire consultants, some states have boards of examiners that determine if state requirements are met.

In general, there is no point in having the state qualification tests duplicate the ITA tests! There is considerable virtue in having state tests that are unpredictable, allowing state examiners to use their judgment and knowledge of the shortcomings of the ITA testing to guide their tests. This is facilitated by state laws that give the board members the right to use their judgment instead of being limited to specific objective criteria. Generally, even when judgment calls are permitted, the board cannot reject a machine arbitrarily, but must show that it violates some provision required by state law.

State qualification testing should ideally include a demonstration that the voting machine can be configured for demonstration elections that exercises all of the distinctive features of that state's election law, for example, straight party voting, ballot rotation, correct handling of multi-seat races, and open or closed primaries, as the case may be. Enough ballots should be voted in these elections to verify that the required features are present.

Tests conducted during contract negotiation

When a jurisdiction puts out a request for bids, they will generally allow the finalists to bring in systems for demonstration and testing. It is noteworthy that federal certification and state qualification test whether a machine meets the legal requirements for sale, but they generally do not address any of the economic issues associated with voting system use, so it is at this time that economic issues must be evaluated.

In addition, the purchasing jurisdiction (usually the county) has an opportunity, at this point, to test the myriad practical features that are not legislated or written into any standards. As of 2004, neither the FEC/NASED standards nor the standards of most states address a broad range of issues related to usability, so it is imperative that local jurisdictions aggressively use the system, particularly in obscure modes of use such as those involving handicapped access (many blind voters have reported serious problems with audio ballots, for example).

It is extremely important at this stage to allow the local staff who will administer the election system to participate in demonstrations of the administrative side of the voting system, configuring machines for mock elections characteristic of the jurisdiction, performing pre-election tests, opening and closing the polls, and canvassing procedures. Generally, neither the

voting system standards, or state qualification tests address questions of how easy it is to administer elections on the various competing systems.

Acceptance testing as delivered

Each machine delivered by a vendor to the jurisdiction should be tested. Even if the vendor has some kind of quality control guarantees, these are of no value unless the customer detects failures at the time of delivery. At minimum, such tests should include power-on testing, basic user interface tests (do all the buttons work, does the touch screen sense touches at all extremes of its surface, do the paper-feed mechanisms work, does the uninterruptible power supply work).

By necessity, when hundreds or even thousands of machines are being delivered, these tests must be brief, but they should also include checks on the software versions installed (as self-reported), checks to see that electronic records of the serial numbers match the serial numbers affixed to the outside of the machine, and so on.

It is equally important to perform these acceptance tests when machines are upgraded or repaired as it is to perform them when the machines are delivered new, and the tests are equally important after in-house servicing as they are after machines are returned from the vendor's premises.

Finally, when large numbers of machines are involved, it is reasonable to perform more intensive tests on some of them, tests comparable to the tests that ought to be performed during qualification testing or contract negotiation.

Pre election (logic and accuracy) testing

Before each election, every voting machine should be subject to public testing. This is frequently described as logic and accuracy testing or simply L&A testing, a term that is more appropriate in the realm of punched-card and mark-sense ballot tabulating machines than in the realm of direct-recording electronic systems, but the term is used widely and in many states, it is enshrined in state law.

The laws or administrative rules governing this testing vary considerably from state to state. Generally, central-count paper ballot tabulating machinery can be subject to more extensive tests than voting machines, simply because each county needs only a few such machines. Similarly, precinct-count paper ballot tabulating machinery, with one machine per precinct, can be tested more intensively than voting machines, which may number in the tens per precinct.

An effective test should verify all of the conditions tested in acceptance testing, since some failures may have occurred since the systems arrived in the warehouse. In addition, the tests should verify that the machines are correctly configured for the specifics of this election, with the correct ballot information loaded, including the names of all applicable candidates, races and contests.

The tabulation system should be tested by recording test votes on each machine, verifying that it is possible to vote for each candidate on the ballot and that these votes are tabulated correctly all the way through to the canvass; this can be done, for example, by casting a different number of votes for each candidate or issue position in each race or contest on the ballot.

When multiple machines are configured identically, this part of the test need only be performed in full and manually on one of the identical machines, while

on the others, it is reasonable to simplify the testing by verifying that the other machines are indeed configured identically and then using some combination of automated self-test scripts and simplified manual testing.

For mark-sense voting systems, it is important to test the sensor calibration, verifying that the vote detection threshold is appropriately set between a blank spot on the ballot and a dark pencil mark. The calibration should be tested in terms of pencil marks even in jurisdictions that use black markers, because it is inevitable that some voters will use pencils, particularly when markers go dry in voting booths or when ballots are voted by mail. One way to judge the appropriateness of the threshold setting is to see that the system distinguishes between hesitation marks (single dots made by accidentally resting the pencil tip on a voting target) and X or check-marks, since the former are common accidents not intended as votes, and most state laws allow an X or check to be counted as a vote even though such minimal marks are never recommended.

For touch-screen voting systems, it is important to test the touch-screen calibration, verifying that the machine can sense and track touches over the entire surface of the touch screen. Typical touch screen machines have a calibration mode in which they either display targets and ask the tester to touch them with a stylus, or they display a target that follows the point of the stylus as it is slid around the screen.

For voting systems with audio interfaces, this should be checked by casting at least some of the test ballots using this interface. While doing this, the volume control should be adjusted over its full range to verify that it works. Similarly, where multiple display magnifications are supported, at least one test ballot should be voted for each ballot style using each level of magnification. Neither of these tests can be meaningfully performed using automatic self-testing scripts!

The final step of the pre-election test is to clear the voting machinery, setting all vote totals to zero and emptying the physical or electronic ballot boxes, and then sealing the systems prior to their official use for the election.

Ideally, each jurisdiction should design a pre-election test that, between all tested machines, not only casts at least one vote per candidate on each machine, but also produces an overall vote total arranged so that each candidate and each yes-no choice in the entire election receives a different total. Designing the test this way verifies that votes for each candidate are correctly reported as being for that candidate and not switched to other candidates. This will require voting additional test ballots on some of the machines under test!

Pre-election testing should be a public process! This means that the details and rationale of the tests must be disclosed, the testers should make themselves available for questioning prior to and after each testing session, representatives of the parties and campaigns must be invited, and an effort must be made to make space for additional members of the public who may wish to observe. This requires that testing be conducted in facilities that offer both adequate viewing areas and some degree of security.

It is important to assure that the voting machine configuration tested in the pre-election tests is the same configuration used on election day. Loading new software or replacing hardware components on a voting machine generally requires the repetition of those parts of the pre-election tests that could possibly depend on the particular hardware or software updates that were made.

Testing as the polls are opened

Prior to opening the polls, every voting machine and vote tabulation system should be checked to see that it is still configured for the correct election, including the correct precinct, ballot style, and other applicable details. This is usually determined from a startup report that is displayed or printed when the system is powered up.

In addition, the final step before opening the polls should be to verify that the ballot box (whether physical or virtual) is empty, and that the ballot tabulation system has all zeros. Typically, this is done by printing a zeros report from the machinery. Ideally, this zeros report should be produced by identically the same software and procedures as are used to close the polls, but unfortunately, outside observers without access to the actual software can only verify that the report itself looks like a poll closing report with all vote totals set to zero.

Some elements of the acceptance tests will necessarily be duplicated as the polls are opened, since most computerized voting systems perform some kind of power-on self-test. In some jurisdictions, significant elements of the pre-election test have long been conducted at the polling place.

Observers, both partisan observers and members of the public, must be able to observe all polling place procedures, including the procedures for opening the polls.

Parallel testing during an election

Parallel testing, also known as election-day testing, involves selecting voting machines at random and testing them as realistically as possible during the period that votes are being cast. The fundamental question addressed by such tests arise from the fact that pre-election testing is almost always done using a special test mode in the voting system, and corrupt software could potentially arrange to perform honestly while in test mode while performing dishonestly during a real election.

Parallel testing is particularly valuable to address some of the security questions that have been raised about Direct-Recording Electronic voting machines (for example, touch-screen voting machines), but it is potentially applicable to all electronic vote counting systems. It is better to test optical mark-sense scanners, however, by randomly selecting precincts for hand recount after each election, as in California.

It is fairly easy to enumerate a long list of conditions that corrupt election software could check in order to distinguish between testing and real elections. It could check the date, for example, misbehaving only on the first Tuesday after the first Monday of November in even numbered years, and it could test the length of time the polls had been open, misbehaving only if the polls were open for at least 6 hours, and it could test the number of ballots cast, misbehaving only if at least 75 were encountered, or it could test the distribution of votes over the candidates, misbehaving only if most of the votes go to a small number of the candidates in the vote-for-one races or only if many voters abstain from most of the races at the tail of the ballot.

Pre-set vote scripts that guarantee at least one vote for each candidate or that guarantee that each candidate receive a different number of votes can be detected by dishonest software! Therefore, parallel testing is best done either by using a random distribution of test votes generated from polling data representative of the electorate, or by asking real voters to volunteer to help

test the system (perhaps asking each to flip a coin to decide secretly whether they'll vote for the candidates they like or for the candidates they think their neighbor likes).

It is important to avoid the possibility of communicating to the system under test any information that could allow the most corrupt possible software to learn that it is being tested. Ideally, this requires that the particular machines to be tested be selected at the last possible moment and then opened for voting at the normal time for opening the polls and closed at the normal time for closing the polls. In addition, mechanical vote entry should not be used, but real people should vote each test ballot, with at least two observers noting either that the test script is followed exactly or noting the choices made. (A video record of the screen might be helpful.)

Parallel testing at the polling place is a possibility. This maximizes exposure of the testing to public observation and possibly to public participation, an important consideration because the entire purpose of these tests is to build public confidence in the accuracy of the voting system.

However parallel testing is conducted, it is important to guard against any possibility of contamination of the official canvass with ballot data from voting machines that were subject to parallel testing. By their very nature, these votes are indistinguishable from real votes, excepting the fact that they came from a machine under test. Therefore, physical quarantine of the vote totals from the parallel testing is essential. Use of a different color for paper in the printer under test, use of distinctively colored data cartridges, warning streamers attached to cartridges, and similar measures may all be helpful. In addition, if the serial number of the voting machine is tied to its votes through the canvass, a check to make sure that the serial numbers of the machines under parallel testing do not appear in the canvass is obviously appropriate.

If polling places are so small that there is no room to select one machine from the machines that were delivered to that polling place, it is possible to conduct parallel testing elsewhere, pulling machines for testing immediately prior to delivery to the polling place and setting them aside for testing. In that case, it is appropriate to publish the location of the testing and invite public observation. Casual drop-in observation can be maximized by conducting the tests near a polling place and advertising to the voters at that polling place that they can stop by after voting to watch or perhaps participate.

Post election testing

Some jurisdictions require routine post-election testing of some of the voting machinery, to make sure that, after the canvassing process was completed, the machinery is still working as well as it did before the election. Generally, these tests are very similar to pre-election or logic and accuracy testing.

Clearly, where the machines themselves hold the evidence of the vote count, as with mechanical lever voting machines or direct-recording electronic voting machines, this evidence must not be destroyed until law and prudence agree that it is no longer relevant to any potential legal challenge to the election.

In the event of a recount, all of the pre-election tests that do not involve possible destruction of the votes being recounted must be repeated in order to assure that the machinery used in the recount is operating correctly.

Reprinted, with small revisions, as *Appendix E, Voting Machine Testing* of [The Machinery of Democracy: Protecting Elections in an Electronic World](#), part of the Voting Rights & Elections series published by the [Brennan Center for Justice](#) at the NYU School of Law, June 27, 2006.