

Notes to Reviewers
Volume Testing Supplement Version 1.0 for the VVSG-NI
April 1, 2009

This document represents supplemental guidance for the volume test method whose primary definition appears in Part 3 Requirement 5.2.3-D of the next iteration of the Voluntary Voting System Guidelines (VVSG-NI). When the VVSG-NI is approved by the Election Assistance Commission (EAC), the test suites will be available for use by voting system testing laboratories as a common basis for testing voting systems to determine conformance to the VVSG-NI.

Test suite reviewers are advised to first read and understand the VVSG-NI, especially the sections relevant to the test suites under review, before reviewing the test suites. The volume test requirement (and its subrequirements) is located in Part 3 Section 5, which can be found at:

<http://www.eac.gov/vvsg/part3/chapter05.php>. A complete version of the VVSG-NI in HTML, MS-Word, or PDF formats can be found at <http://www.eac.gov/vvsg>.

Commenting:

Please send comments on the test suites, by July 1, 2009, to: crt-software-test@nist.gov.

You may provide comments directly in your email and/or send attachments in MS-Word or PDF. If you wish, you may embed your comments within the PDF documentation using the instructions provided below. In general, please tell us the features you like and provide us with comments, corrections, and suggestions on how to improve the test suites. Please provide the following items:

- Test suite version number (found in the test suite documentation, currently Version 1.0)
- Your name and affiliation (include contact information if desired)
- Identification of the particular tests and requirements in the VVSG-NI for which your comment applies
- If including suggestions for changes to the tests, a description of the suggested change including an adequate justification for the change, or a draft replacement for the test including the justification and any other necessary documentation or commentary

All comments will be considered. After all comments have been received and incorporated into the test suites, a new version of the test suites will be posted on the NIST web site.

Embedding comments in PDF files:

If you wish to embed comments within the PDF documentation, you may do so using the free Adobe Reader software available from Adobe. The following detailed instructions for commenting the PDF file are current as of 2009-03-25 and Adobe Reader version 9.1.0. Versions 8.1.X are also usable.

1. Ensure that Adobe Reader is installed on your computer. Adobe Reader may be obtained from <http://get.adobe.com/reader/>.

2. Open the documentation PDF file in Adobe Reader.
3. There should be a menu on the toolbar labeled Comment or Review & Comment. Select Show Comment & Markup Toolbar from that menu to get a new toolbar that includes the Sticky Note tool, the Text Edits tool, and others. (These tools can also be accessed via Tools → Comment & Markup.)
4. To insert a comment someplace in the document, go to that page and use the Sticky Note tool. Once the text of the comment has been entered, the yellow note icon can be dragged to place it near the text in question.
5. To indicate desired textual changes, use the Text Edits tool to insert, delete, or replace text.
6. Save your changes using File → Save.

Volume Testing Supplemental for the VVSG-NI

Draft — Version 1.0 — April 1, 2009

This document and associated files have been prepared by the National Institute of Standards and Technology (NIST) and represent draft test materials for the Election Assistance Commission's next iteration of the VVSG. It is a preliminary draft and does not represent a consensus view or recommendation from NIST, nor does it represent any policy positions of NIST.

1 Introduction

This document provides additional requirements and guidance for the volume test method whose primary definition appears in Part 3 Requirement 5.2.3-D of the Voluntary Voting Systems Guidelines, draft 20070831 [1].

2 Background

By authorization of the 2002 Help America Vote Act (HAVA), NIST is assisting the Election Assistance Commission (EAC) with the implementation of Voluntary Voting System Guidelines (VVSG) for states and local governments conducting Federal elections. The EAC's Technical Guidelines Development Committee (TGDC) in collaboration with NIST researchers has developed a draft of the next iteration of the VVSG. The draft document is a set of detailed technical requirements addressing core requirements, human factors, privacy, security, and transparency of the next generation of voting systems. The EAC plans to issue the next VVSG after receiving and reviewing public comments.

NIST is developing a set of uniform public test suites to be used as part of the EAC's Testing and Certification Program. Test Labs will be able to use these freely available test suites to help determine that VVSG requirements are met by voting systems. The test suites address human factors, security and core functionality requirements for voting systems as specified in the VVSG. Use of the public test suites will produce consistent results and promote transparency of the testing process. The test suites can also assist manufacturers in the development of conforming products by providing precise test specifications. Also, they will help reduce the cost of testing since each test lab would no longer need to develop its own test suites. Finally, a uniform set of public test suites can increase election officials' and voters' confidence that voting systems conform to VVSG requirements.

3 Preparing for volume testing

3.1 *Equipment and materials*

Subrequirements of Part 3 Req. 5.2.3-D specify minimum numbers of certain devices for the volume test as well as minimum numbers of distinct ballots to be used for testing. For paper ballots, "distinct" means that two ballots are physically not the same ballot; i.e., the second ballot was not merely a recirculation of the first through the system again. For ballots cast electronically, "distinct" means that two ballots are cast in separate voting sessions as they normally would be in an election.

The test lab will need to prepare supplies of all materials needed to conduct an election as specified in the Voting Equipment User Documentation. In voting systems that use paper ballots, the VVSG requires a minimum of 10000 good test ballots, so testers will need 10000 blank ballots plus enough extras to make up for those that are spoiled due to tester error.

If the voting system under test does not include devices of the listed class, then the minimum for that class does not apply. This is reflected by the wording of the requirements, which have the form "For systems that include Xs, a minimum of N Xs SHALL be tested...."

The following table summarizes the numbers provided by subrequirements of Part 3 Req. 5.2.3-D. Where a number is not explicitly stated by a requirement, the implied minimum is given in square brackets.

Requirement	Device class	Minimum devices	Minimum distinct ballots, each device	Minimum distinct ballots, class total	Minimum ballot volume, class total
Part 3 Req. 5.2.3-D.1	VEBD	100	110	[11000]	[11000]
Part 3 Req. 5.2.3-D.2	Precinct tabulator	50	400	10000	[20000]
Part 3 Req. 5.2.3-D.3	Central tabulator	2	10000	10000	75000

A voting system under test is likely to include more than one of these device classes, and devices other than those listed will of course be needed to complete the voting system configuration as usual. The combination of different device classes may increase the total number of distinct test ballots and total ballot volume needed to satisfy all of the requirements simultaneously. For example, it is not possible to share ballots between a DRE and a central optical scan tabulator, so if a system includes both of these classes of devices, the total ballot volume processed by the system as a whole will have to be at least 86000 to satisfy both 5.2.3-D.1 and 5.2.3-D.3.

3.2 Location

The test lab may need to locate an extra-spacious testing venue to accommodate the equipment and test voters.

3.3 Test scenario

Part 3 Req. 5.2.3-D Functional testing, volume test

The test lab **SHALL** conduct a volume test in conditions approximating normal use in an election. The entire system **SHALL** be tested, from election definition through the reporting and auditing of final results.

The basic test suite that is provided in the Votetest distribution [2] contains 92 tests that exercise voting variations in small, simple scenarios to isolate the conditions under which failures occur. The basic test suite is intended to be complemented by a volume test that exercises all features together in a large, complex scenario where a significant volume of ballots is processed. The test lab should therefore design the volume test to exercise as many of the voting variations indicated in the implementation statement and as many of the system's interfaces (both user interfaces and device-to-device interfaces) as can productively be combined in a realistic testing scenario. The great variability of this set of voting variations and interfaces from one voting system to the next is why a single, standard volume testing scenario is not specified here.

Part 3 Req. 5.2.3-D and its subrequirements specify lower limits on some of the parameters of the volume test, such as the number of test ballots. In all other respects the test lab has flexibility. The test lab should design a test scenario that is defensible as a credible simulation of a real election while covering as much core functionality as is practical.

The scenario should be designed so that the ballot counts and vote totals that should appear in the vote data reports produced by the voting system at the conclusion of the volume test can be determined beforehand.

3.4 Personnel

The volume test may require additional test lab and manufacturer personnel beyond those who are involved throughout the operational testing of the voting system simply because there is more equipment to be managed. However, the primary staffing need is for test voters to cast ballots in the way that voters normally would in a real election.

For practical reasons, realism need not be carried as far as limiting each test voter to cast only one ballot. The California Volume Reliability Testing Protocol [3], on which the VVSG's volume test requirements were modelled, uses a ratio of 220 ballots per test voter in a 6-hour period when these ballots are cast exclusively by DRE. Test lab experience of the average ballot casting rate for different types of devices and paper versus electronic user interfaces can be used to estimate the optimal ratio for the voting system under test.

The demographics of the group of test voters **do not matter**. Unlike the usability test [4], the volume test does not treat errors made by test voters or test lab personnel as indicators of voting system performance. The volume test is strictly an assessment of the mechanical performance of the voting system when presented with the input that was determined in the test scenario. Deviations from the script by test voters or mistakes made by test lab personnel should be corrected and factored out of the results.

3.5 Procedures

Normal test lab procedures for recording observations, logging anomalies and resolving anomalies extend to the volume test; however, those procedures must be adapted for the less controlled environment of the volume test. The roles and responsibilities of test voters, test lab personnel, and manufacturer personnel in the volume test must be specified and made clear. The following table provides some guidelines to serve as a starting point.

Role	Responsibilities
Test voters	Cast ballots according to script Upon making a mistake or perceiving an anomaly (anything that differs from expectation), STOP, report the occurrence to test lab personnel and await instructions
Test lab personnel	Manage the volume test in general Fill all election official roles Provide test voters with scripts and any necessary materials (ballots, ballot markers and/or ballot activation credentials) Correct and compensate for test voter mistakes Categorize anomalies and follow up as appropriate Maintain complete record of all observations, mistakes, anomalies, and follow-up actions Collect data for reliability, accuracy, and misfeed rate Assign test verdict
Manufacturer personnel	Provide test lab with equipment and materials Assist test lab personnel with training, voting system operation, and resolution of anomalies as requested by test lab personnel If allowed, observe the volume test and maintain own records as needed to support quality assurance, corrective changes to the voting system and/or VVSG interpretation requests ¹

¹ Note: The EAC Laboratory Accreditation Program Manual, Version 1.0, effective July 2008, Section 2.11.1, prohibits manufacturers from observing the testing.

3.6 Training

Personnel **SHALL** be trained in the use of the voting system and in the volume testing procedures.

3.7 Set-up

The various devices in the voting system **SHALL** be prepared, calibrated and tested exactly as if they were being deployed for an actual election.

Testers **SHALL** follow the procedures specified by the manufacturer in the Voting System User Documentation.

It is an anomaly if for some reason those procedures cannot be followed. Investigation of the anomaly should determine whether the inability to follow the procedures results from a tester error, a nonconformity of the voting system, or a VVSG interpretation conflict.

4 Conducting the test

4.1 General test template

Like all end-to-end operational tests, the volume test follows the general test template specified in Part 3 Section 5.2.1.1 of the VVSG:

1. Establish initial state (clean out data from previous tests, verify resident software/firmware);
2. Program election and prepare ballots and/or ballot styles;
3. Generate pre-election audit reports;
4. Configure voting devices;
5. Run system readiness tests;
6. Generate system readiness audit reports;
7. Precinct count only:
 - A. Open poll;
 - B. Run precinct count test ballots; and
 - C. Close poll.
8. Run central count test ballots (central count / absentee ballots only);
9. Generate in-process audit reports;
10. Generate data reports for the specified reporting contexts;
11. Inspect ballot counters; and
12. Inspect reports.

4.2 Power outage

A two hour block of the volume test **SHALL** be run with the system disconnected from the normal power supply.

Part 1, Requirement 6.3.4.3-A.4 (Outages, Sags and Swells) requires that "All electronic voting systems shall be able to withstand, without disruption of normal operation or loss of data, a complete loss of power lasting two hours ..." To test this capability, a two hour block of the volume test is run with the system disconnected from the normal power supply. For testing details, see the electromagnetic compatibility test suite [5].

4.3 Imperfect marks and folds (optical scan only)

Part 3 Req. 5.2.3-D.4 Test imperfect marks and folds

The testing of MCOS *SHALL* include marks filled according to the recommended instructions to voters, imperfect marks as specified in Requirement Part1:7.7.5-D, and ballots with folds that do not intersect with voting targets.

Manual marking of paper ballots by test voters should produce a realistic mixture of near-perfect and imperfect marks. Ballots containing imperfect marks that still satisfy the manufacturer's specification of a reliably detectable mark *SHALL* be retained as valid test ballots.

Each optical scanner *SHALL* be fed at least 2 ballots that are marked to the specification given in Part 1 Req. 7.7.5-D and at least 2 ballots that are folded in a way that does not intersect voting targets. If in practice all of the absentee ballots processed by the system would be folded, then all of the simulated absentee ballots in the volume test scenario *SHALL* be folded accordingly.

5 Findings

A volume test can lead to a finding of nonconformity in three ways. As usual, the volume test is assigned a verdict of pass if and only if none of the conditions for failing apply:

Implicit Passing: Many test methods include a number of steps for each of which the system must perform correctly, or it fails. In general, it is easier to confirm that a system has *not* met a requirement, than that it has. If the test method is completed successfully without any failures, then the system passes.

5.1 Non-support of required functionality

As specified by Part 3 Req. 5.2.1.2-C:

Part 3 Req. 5.2.1.2-C Missing functionality

If the test lab is unable to execute a given test because the system does not support functionality that is required per the implementation statement or is required for all systems, the test verdict *SHALL* be Fail.

This condition applies if there is some feature of the voting system under test that causes the properly designed and scoped volume test to be a "non-starter;" e.g., the voting system supposedly supports voting variation X but there really is no way to do it.

5.2 Excessive failures, errors, and/or misfeeds

As part of the analysis of test results, the ballot counts and vote totals reported by the voting system under test are compared with the expected results as determined during the development of the test scenario. To the extent that these do not match, and it is determined that the input (i.e., the ballots cast) and the expected results were in fact correct, the voting system is assessed errors per the definition of report total error rate in Part 3 Section 5.3.4 of the VVSG. If it is found that the ballots cast were not all correct, as may be the case in a large volume test, the expected results must be adjusted to compensate before errors are attributed to the voting system.

If the voting system exhibits poor performance with respect to reliability, accuracy, or misfeed rate during the volume test, this may lead to a verdict of failure for conformance to the benchmarks specified in Part 1 Section 6.3 the VVSG. Conformity to these benchmarks is assessed at the completion of operational testing, so the voting system's behavior during the volume test could lead to an adverse finding on the benchmarks later on even if the volume test itself is assigned a passing verdict.

Because the core benchmarks of reliability, accuracy, and misfeed rate are based on the volume of ballots, votes, etc. processed by the voting system, data collected during the volume test contribute substantially to

the evaluation of these benchmarks. However, data collection during the volume test is the same as it is for all other applicable tests, as specified in Part 3 Section 5.3 of the VVSG.

For details, please refer to Part 1 Section 6.3, Part 2 Req. 6.1-J and its subrequirements, and Part 3 Section 5.3 of the VVSG.

5.3 Violation of any other VVSG requirement

As specified by Part 3 Req. 5.2.1.2-D:

Part 3 Req. 5.2.1.2-D Any demonstrable violation justifies an adverse opinion

A demonstrable violation of any applicable requirement of the VVSG during the execution of any test **SHALL** result in a test verdict of Fail.

DISCUSSION

The nonconformities observed during a particular test do not necessarily relate to the purpose of that test. This requirement clarifies that a nonconformity is a nonconformity, regardless of whether it relates to the test purpose.

In other words:

Serendipitous Detection of Failure: Although each test method is designed for specific requirements, it may also reveal violations of other requirements. These violations are to be noted by the tester and are counted as failures, just as if they had been the explicit purpose of the test.

While the purpose of a volume test could be argued to be broader than just evaluating reliability, accuracy, and misfeed rate, the outcome of this debate is irrelevant to the test results. "A nonconformity is a nonconformity."

6 References

- [1] Election Assistance Commission. *Voluntary Voting System Guidelines Recommendations to the Election Assistance Commission*, 2007-08-31. <http://vote.nist.gov/vvsg-report.htm>.
- [2] Votetest distribution, 2009-04-01. <http://vote.nist.gov/SystemTesting/reviewer-notes-votetest.htm>.
- [3] California Volume Reliability Testing Protocol, 2006-01-31. http://www.ss.ca.gov/elections/voting_systems/volume_test_protocol_final.pdf.
- [4] Human Factors Test Suite Version 1.0 for the Usability, Accessibility, and Privacy Requirements of the VVSG-NI, 2009-04-01. <http://vote.nist.gov/TestSuites/hfp-testsuite-version1-0.html>.
- [5] Electromagnetic Compatibility Test Suite, **to appear**.