MODIFIED CIF TEMPLATE FOR VOTING SYSTEM TEST LABORATORIES (VSTLS)

INTENDED AUDIENCE

This document provides a template for the modified version of *ISO/IEC 25062:2006, the Common Industry Format (CIF)* usability test report. This modified version of the CIF has been specifically tailored for Voting System Test Laboratories (VSTLs), and their usability test administrator(s) and data logger(s).

This template has been created to enable VSTLs to effectively communicate the results of usability testing.

In addition to this template, a set of guidelines on how to complete the modified CIF template has been created to assist VSTL usability test administrator(s) and data logger(s) in developing a usability report.

INTENDED PURPOSE OF THIS DOCUMENT

This template has been prepared to help VSTLs meet the Voluntary Voting System Guidelines (VVSG) developed by the Technical Guidelines Development Committee (TGDC).

The VVSG requires that Voting Manufacturers submit their voting systems to one of the VSTLs for usability testing.

Section 3.2.1.1-D of the VVSG states that "The test lab shall report the metrics for the usability of the voting system as measured by the Voting Performance Protocol (VPP)."[1]

The following modified CIF template is intended to assist test laboratories in reporting the results of usability testing for each system tested. This template has been specifically tailored based on the VPP and the usability test requirements for voting systems.

[1] Voluntary Voting System Guidelines Recommendations to the Election Assistance Commission, Aug. 2007

STRUCTURE OF THIS DOCUMENT

This template includes sample content and placeholder for each section of the modified CIF. The content provided in this document is just a sample and is not intended to provide guidance or requirements. This template includes the following sections:

1.0 Executive Summary

2.0 Introduction

3.0 Method

4.0 Results

5.0 Appendices

In addition to these sections, the modified CIF shall also include a title page. A sample title page has been included on the following page.

HOW TO USE THIS DOCUMENT

This document is based on ISO/IEC 25062:2006 Common Industry Format (CIF), a format used to report the results of summative usability testing. Before using this document, you must become familiar with this standard. ISO/IEC 25062:2006 can be purchased from: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=43046.

It is important to note that the numbering format included in this template is identical to the numbering used in the document that provides guidelines and instructions for completing the modified CIF template for VSTLs.

When completing the modified CIF template, it is highly recommended that VSTLS, their usability test administrator(s) and their data logger(s) refer to the instructions and guidance in order to properly complete this template.

The data sample provided in this template is ***an example or placeholder of the types of content*** that may be useful in completing the modified CIF template. Gray background text (bounded in square brackets) needs to be replaced by VSTLs' supplied information. For detailed information about how to complete this template, please refer to the "Guidance on How to Complete the Modified CIF Template for Voting System Test Laboratories (VSTLs)".

Voting System Test Laboratory Usability Test Report of [NAME OF PRODUCT AND VERSION TESTED]

REPORT BASED ON ISO/IEC 25062:2006 COMMON INDUSTRY FORMAT FOR USABILITY TEST REPORTS

[FULL NAME OF PRODUCT AND VERSION TESTED]

DATE OF USABILITY TEST: [DATE USABILITY TEST WAS CONDUCTED]
DATE OF REPORT: [DATE REPORT WAS PREPARED]

REPORT PREPARED BY: [NAME OF SYSTEM TEST LABORATORY (VSTL)]
[VSTL CONTACT PERSON]
[VSTL PHONE NUMBER]
[VSTL EMAIL ADDRESS]
[VSTL MAILING ADDRESS]

TABLE OF CONTENTS

1.0 EXECUTIVE SUMMARY

A usability test of [name of product, version, and class of the system according to the overall class structure outlined in the VVSG] was conducted on [date] in [location] by [test laboratory]. The purpose of this test was to fulfill the requirements of the Voluntary Voting System Guidelines (VVSG) by measuring the performance of the voting system using the Voting Performance Protocol (VPP). The performance of the voting system under test (VSUT) was then compared to the benchmarks required by the VVSG to determine if the system met all of the benchmark metrics. The system [did], in fact, meet [and exceed] all of the performance benchmarks.

During the usability test, [XX] voters from the general population, matching the target demographic criteria outlined in the VPP, used the VSUT in a simulated election, while [XX] voters used the test method calibration system (henceforth referred to simply as the calibration system) [name of product, version, and class of the system according to the overall class structure outlined in the VVSG].

The ballot used was the NIST standard ballot, used from the VPP testing, consisting of one test ballot with twenty contests, including:

- Federal, state and local contests
- Partisan and nonpartisan contests
- Single member and multimember contests
- Retention races
- Constitutional amendments
- Referenda and ballot initiatives

This ballot included 28 tasks that model typical ballots from around the country, including:

- Voting for names at various locations within a list of names
- Voting a partial slate in a multimember contest
- Skipping elements of a ballot
- Write-in votes

During the usability test, participants were greeted upon arrival, asked to review and sign a consent/release form, and provided a written copy of the voting instructions. Participants were then escorted into the testing room and instructed to use one of two systems, either the VSUT or the calibration system. Participants were not told which system was being considered for certification.

Each participant worked alone and was not provided assistance from the test administrators. During the session, data logger(s) recorded whether or not the participant successfully cast a ballot (regardless of whether the ballot choices were correct) and the time taken to vote.

Following the conclusion of the testing, participants were asked to complete a post-test questionnaire and were compensated with a $ [50] for their time.

As part of the usability test, the usability test administrator(s) and data logger(s) also collected and analyzed the following types of data for each participant:

- Type of system used (VSUT and the calibration system)
- Number of tasks (or voting opportunities) successfully completed
- Time to complete the voting session
- Voter's assurance that he/she had tried to follow the instructions provided
- Voter's confidence that he/she had used the system correctly
- Voter's likeability ratings of the system

Various types of analysis, in accordance with the requirements set forth in the VPP, were used to score the ballots. The benchmark calculation Perl scripts were used to calculate the Mann-Whitney comparison of the current and nominal results of the calibration system, as well as the Total Completion Score for the calibration system. The script was also used to calculate the Total Completion Score, Perfect Ballot Index and Voter Inclusion Index for the VSUT. Time on Task, Average Voter Confidence and Average Voter Likeability were scored by the usability test administrator(s) and data logger(s) using a spreadsheet.

Following is a summary of the performance data collected on the VSUT and the calibration system.

Measure	Description	Minimum Benchmarks for VSUT	VSUT N=XX	Calibration System N=XX
Total Completion Score	*Percentage of test participants who were able to complete the process of voting and cast their ballots so that their ballot choices were recorded by the system.*	**98%**	**[Mean] [95% CI]**	**[Mean] [95% CI]**
Voter Inclusion Index	*Measurement that combines accuracy with variability in the level of accuracy among test participants.*	**.35**	**[Mean] [95% CI]**	**--**
Perfect Ballot Index	*Comparison of the number of participants who cast a ballot without any errors to those that had a least one error. It is used for detecting any systemic design problem that causes the same type of error by many test participants.*	**2.33**	**[Mean] [95% CI]**	**--**

Measure	Description	Minimum Benchmarks for VSUT	VSUT N=XX	Calibration System N=XX
Average Session Time	*Mean time taken per test participant to complete the process of activating, filing out and casting the ballot.*	--	**[Mean] [STDV]**	--
Average Voter Confidence	*Mean confidence level expressed by voters that they believed they voted correctly and the system successfully recorded their votes.*	--	**[Mean] [STDV]**	--
Average Voter Likeability	*Mean likeability level expressed by voters on how well they liked the system.*	--	**[Mean] [STDV]**	--

The results of the current test on the calibration system were compared to the nominal results, using a Mann-Whitney analysis. Please see the section on Usability Metrics for more on this analysis. Based on this analysis, this usability test is considered to be [valid].

In addition, the table above illustrates that the VSUT [not only meets all of the minimum benchmarks set by the VVSG, but exceeds the criteria and essentially passes the requirements].

2.0 INTRODUCTION

2.1 FULL PRODUCT DESCRIPTION

During the usability test, two systems were included in the test: the VSUT and the calibration system.

The calibration system used was [name of product, version and class].

The VSUT was [name of product, version and class]. Note: The version tested by the VSTL is *[or is not]* the same version of the VSUT tested by the voting system manufacturer in order to meet VVSG Requirements 3.2.1.2-A, 3.2.7-A.4, 3.2.8.1-B, 3.3.2-A, and 3.3.3-A.

Designed to present ballots to voters throughout the U.S. and collect voter responses, the [VSUT] consists of [description of system and how it is used]. It is typically used in federal, state and local elections and is set up by election volunteers in designated voting locations.

The usability testing attempted to simulate these environmental conditions and users' real-world context of use. To illustrate what the systems looked like at the time of testing, photos of both of the voting systems can be found in Appendix [J].

2.2 TEST OBJECTIVES

The primary goal of this usability test was to meet the requirements outlined in the VVSG and the VVP.

According to the VVSG, "It is not sufficient that the internal operation of voting systems be correct; in addition, voters and election officials shall be able to use them effectively and efficiently."

Therefore, the VVSG requires that voting systems be tested by test laboratories and that each system meet established benchmarks. Thus, the goals of this test were to fulfill the VVSG requirements, including:

- Measuring the performance of each voting system using the Voting Performance Protocol (VPP) (Requirements 3.2.1.1-D through 3.2.1.1-D.3)
- Comparing the performance of the voting system to the benchmarks specified by the VVSG (Requirements 3.2.1.1-A through 3.2.1.1-C)

Furthermore, the VVSG states, “The voting system should support a process that provides a high level of usability for all voters. The goal is for voters to be able to negotiate the process *effectively, efficiently, and comfortably*.”

To this end, metrics for effectiveness, efficiency and user satisfaction were captured during the usability testing.

3.0 METHOD

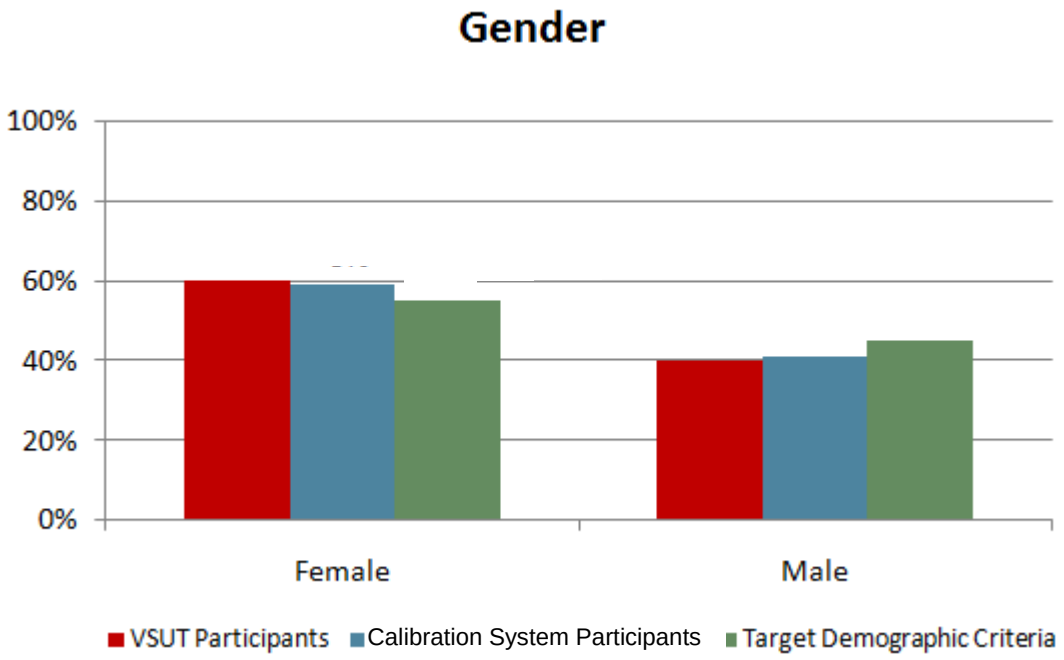
3.1 PARTICIPANTS

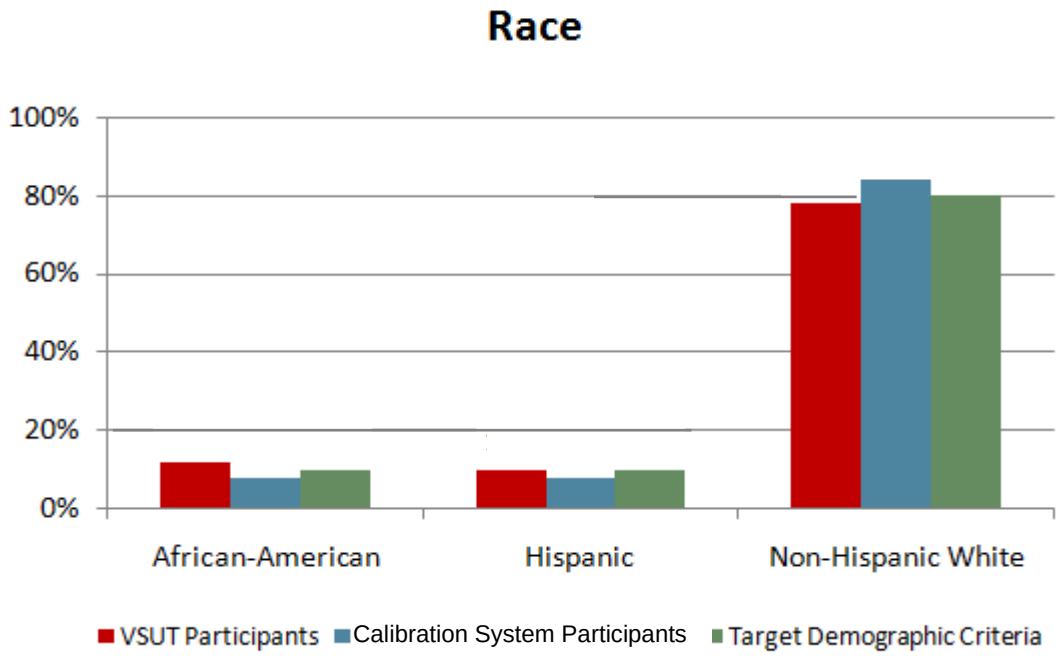
A total of [100] voters were tested on the VSUT, while [100] were tested on the calibration system. Both sets of participants had a varying mix of backgrounds and demographic characteristics. All participants in the usability test were U.S. citizens eligible to vote and literate in English. In addition, none of the participants had a disability, a significant connection to any manufacturer of voting systems, or a background in political science or computer science.

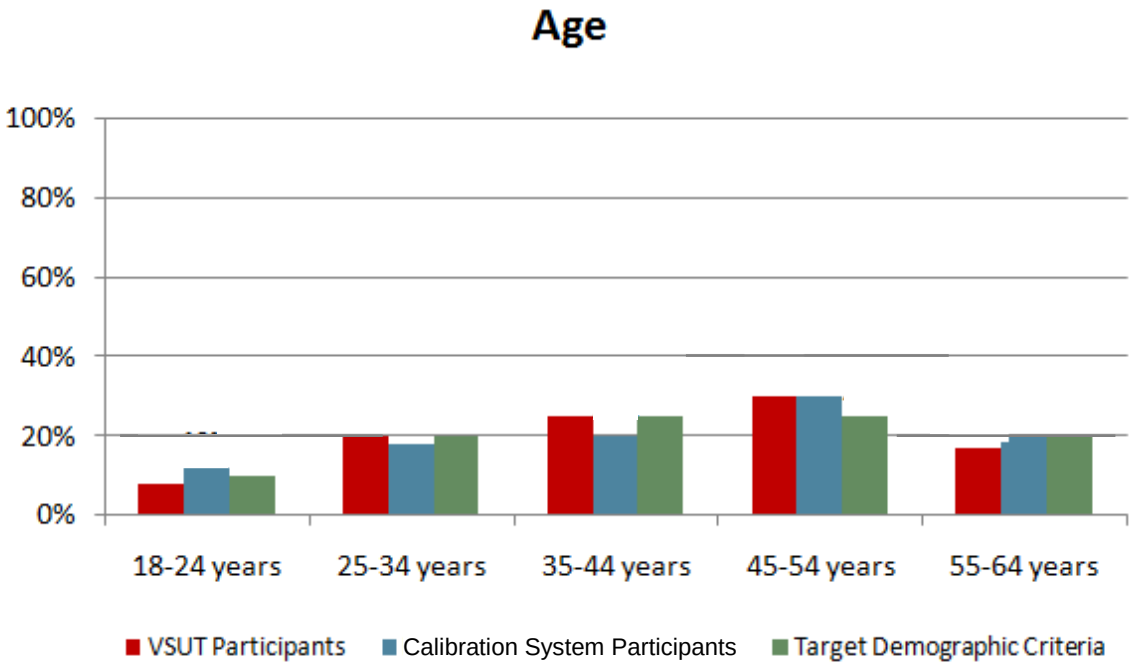
Participants were recruited by [name of recruiting firm] and were compensated $[50] for their time. Please see a copy of the recruiting screener, provided in Appendix [A].

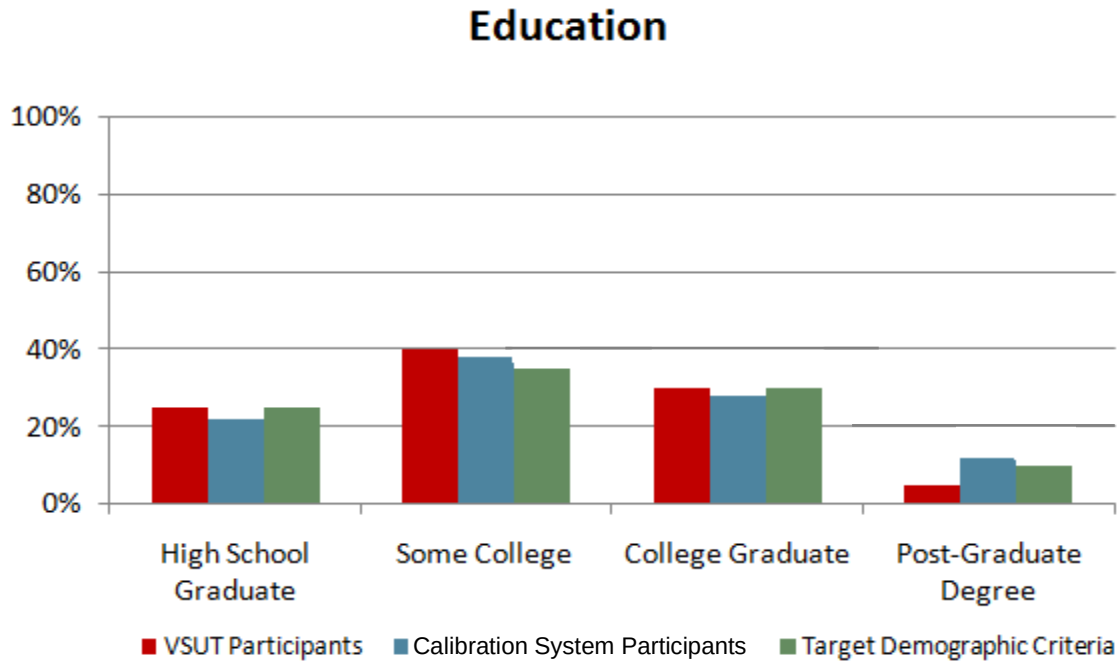
[The voters who were recruited and participated in the test closely matched the targets in the VPP. However, in some instances, the actual test population varied from the targets. Since the test results from the calibration system were considered satisfactory, the test is still considered to be valid].

Following is a graphical presentation of participant characteristics for the VSUT, the calibration system and the target demographic criteria:









Please see Appendix [B] for a full spreadsheet of participant demographics for the VSUT and the calibration system.

3.2 Context of Use in the Test

3.2.1 Tasks

During the usability test, participants were instructed to vote in a simulated election consisting of one test ballot with 20 contests, including:

- Federal, state and local contests
- Partisan and nonpartisan contests
- Single member and multimember contests
- Retention races
- Constitutional amendments
- Referenda and ballot initiatives

The test ballot used, also referred to as the NIST standard ballot specification, was the ballot recommended as part of the usability performance benchmarks. Please see Appendix [E] for a copy of the *implemented* test ballot.

Using this ballot, participants were asked to perform 28 tasks that were selected to model typical voting tasks from around the country, as well as to thoroughly test the voting system's capabilities and usability, including:

- Voting for names at various locations within a list of names
- Voting a partial slate in a multimember contest
- Skipping elements of a ballot
- Write-in votes

Participants were provided with written instructions (see Appendix [D]) on how to vote and were asked to perform the tasks without assistance.

Once all participants had cast their ballots, each ballot was scored as recorded by the voting system (not as marked). That is, ballots were scored using the electronic record from the voting system, as opposed to the DRE screen or the paper ballot, etc.

Of the 28 tasks, each task was evaluated to determine if the voter was able to successfully cast a vote in a way that matched the instructions. Of the 20 contests, participants were asked to vote for one individual in 17 of the races, accounting for 17 points. In addition, there was one contest which included two votes, one contest with four votes, and one contest with five votes, accounting for the other 11 points. A perfect ballot received a score of 28.

Various types of analysis, in accordance with the requirements set forth in the VPP, were used to score the ballots. The benchmark calculation Perl scripts were used to calculate the Mann-Whitney comparison of the current and nominal results of the calibration system, as well as a comparison of the current and nominal Total Completion Score for the calibration system.

The script was also used to calculate the Total Completion Score, Perfect Ballot Index and Voter Inclusion Index for the VSUT. These results were then compared with the benchmarks set in section 3.2.1.1 of the VVSG.

In addition to the above metrics, time taken to cast the ballot was also recorded by an observer, while satisfaction and confidence data were recorded in a post-test questionnaire. All of these metrics are described in more detail in the Usability Metrics section of this report.

3.2.2 Test Location

The [VSUT] is intended to be used at designated polling locations across the U.S., including schools, libraries, churches and other public facilities large enough to house multiple voting stations.

In order to simulate this environment, the test was conducted at [location and description of location]. By using this location, our goal was to simulate a high quality, realistic polling station so that extraneous environmental factors would not contribute to errors in the test.

The test facility included a small waiting area and a testing room [(12' by 15' by 8 high)] with sufficient room for voters to maneuver and ample room for the two voting stations installed. Additionally, [indirect lighting was used to avoid the glare caused by overhead fixtures].

To ensure that the environment was comfortable for users, noise levels were kept to a minimum and the temperature was set to [XX] degrees F.

[Additional information such as, detailed emergency and evacuation plans were in place, and instructions were clearly posted throughout the facility to ensure a safe evacuation in case of an emergency.]

3.2.3 Voting Environment

During an actual election, voters are expected to use the voting system provided by the polling location, which may include systems that use mechanical levers, punch cards, touch screens, optical scans or paper ballots (that are hand counted). Voters may have experience with a wide-range of systems or may have experience with only one type of system.

During the usability test, all participants were instructed to use [VSUT] or the calibration system, just as if this system was implemented at their local polling location.

3.2.3.1 Display Devices

The calibration system used [description of the display including screen size, resolution and color settings. If print-based, include the media size and print resolution.]

The [VSUT] used [description of the display including screen size, resolution and color settings. If print-based, include the media size and print resolution.]

Both systems were set up by the test laboratory according to the manufacturers' documentation describing the system set-up and preparation.

Additionally, voters were instructed not to change any of the default system settings (such as control of font size).

3.2.3.2 Audio Devices

The calibration system used [insert description of audible cues if provided by the system.]

The [VSUT] used [insert description of audible cues if provided by the system.]

3.2.4 Test Administrator Tools

During the usability test, various tools were used to conduct the test sessions, including:

- [Informed Consent (See Appendix [C])]
- [Instructions for Participants (See Appendix [D])]
- [Post-test Questionnaire (See Appendix [F])]
- [Incentive Receipt and Acknowledgment Form (See Appendix [G])]

Participants' votes were recorded by the system, similar to a real-world election. Usability test data logger(s) used a stopwatch to time voter sessions. A note-taking form was used to capture time taken to vote and to record the number of participants who successfully cast a ballot for each system (See Appendix [H] for a copy of the note-taking form).

Lastly, a digital camera and video camera were used to document the testing sessions, testing environment, the voting system, and the calibration system. Copies of these photos can be found in Appendix [J].

Following the test, the team used the end-to end benchmark calculation perl scripts to analyze the data, which is described in more detail in Appendix [K].

3.3 Experimental Design

During the usability test, participants interacted with only one voting system, the VSUT or the calibration system. Each participant used the system in the same location, and was provided with the same ballot and written instructions. Participants were not told which system was being considered for certification.

The system was evaluated for effectiveness, efficiency and satisfaction. During the usability test, the testing team collected and analyzed the following types of data for each participant:

- Type of system used (VSUT or the calibration system)
- Number of tasks (or voting opportunities) successfully completed
- Time to complete the voting session
- Voter's assurance that he/she had tried to follow the instructions provided
- Voter's confidence that he/she had used the system correctly
- Voter's likeability ratings of the system

Additional information about the various measures and associated metrics can be found in the section on Usability Metrics.

3.3.1 Procedure

To ensure that at least 100 ballots were successfully cast on each system, [250] participants (matching the demographics in the section on Participants) were recruited and [210] participated in the usability test.

Participants were scheduled every [15] minutes to allow each participant adequate time to complete each voting session (which took approximately [10] minutes per participant). By scheduling the participants [15] minutes apart, the testing team was able to avoid excessive waiting by participants. A spreadsheet was used to keep track of the participant schedule, and included each participant's demographic characteristics as provided by the recruiting firm.

Upon arrival, participants were greeted and asked for their name. One test administrator(s) checked the participant's name to verify it matched the name on the participant schedule. Participants were then assigned a participant ID and asked to wait in a small area outside the testing environment.

Each participant was asked to review and sign a consent and release form (See Appendix [C]), which described their rights during the study. A test administrator witnessed each participant signing the form.

Participants were then provided with a written copy of the voting instructions (See Appendix [D]). No additional instruction was provided to the voters. Next, the participant was escorted

into the testing room and instructed to use one of the two systems, either the VSUT or the calibration system.

[Depending on the type of system, the test administrator may be required to start the voting process by enacting the role of the poll worker. Note any interactions here].

If the participant signaled for help during the testing (by raising his/her hand), the test administrator(s) replied:

> *"I'm sorry. I can't provide you with any help. Please do the best you can. If you are stuck and cannot continue, you can stop."*

[During the usability test, test administrator(s) and data logger(s) observed users' interactions (from a distance) and monitored each test session with a stop watch and recorded whether the participant successfully cast a ballot. Once the voter finished the test, he/she signaled to the test facilitator by raising his/her hand. The test facilitator recorded the time taken to vote and then escorted the participant back to the waiting area where he/she was asked to complete a post-test questionnaire (See Appendix [F]).]

After completing the post-test questionnaire, participants were thanked for their time and compensated $[XX]. Participants were asked to sign a receipt and acknowledgement form (See Appendix [G]) indicating that they had received the compensation.

[To ensure that the test ran smoothly, two staff members participated in this test, the usability test administrator and the data logger. One person greeted each participant as he/she arrived, administered the consent and release forms and gave the participant the voting instructions. A second person escorted participants to the voting system, timed the participant during the session, and then escorted the person back to the greeter, who gave voters the post-test questionnaire, compensated users for their time, and thanked each individual for their participation.]

The usability testing staff conducting the test were experienced usability practitioners. [Number of years of experience, educational backgrounds, and qualifications of the test administrator(s) and data logger(s)].

3.3.2 Participant General Instructions

During the usability sessions, participants were instructed to work alone and that the test facilitator would not be able to assist or answer any questions during the study. Participants were provided with three pages of written instructions.

The overall test instructions that were provided appear below:

> *Note: We are testing the voting system in a specific configuration. Though the system may allow you to change features such as font size and contrast, please do not adjust these settings for this test.*
>
> *In our mock election, we will be using fake names for candidates and colors for political party names. For example, you might see this:*
>
> *Joe Jones/Yellow Party*
>
> *Any similarity between names of candidates and real people or colors and real parties is purely coincidental.*
>
> *Please attempt to vote exactly as described on the following pages*
>
> *Once you start, we will not be able to help you.*
>
> *Please do the best you can. If you are stuck and cannot continue, inform the administrator.*
>
> *Thank you.*

A full copy of the instructions, in the exact format provided to participants, can be found in Appendix [D].

3.3.3 Participant Task Instructions

Following the overall test instructions, participants were provided with directions on how to vote in each of the 20 contests. A copy of the written instructions can be found in Appendix [D].

If participants signaled for help during the testing (by raising their hand), the test administrator replied:

> *"I'm sorry. I can't provide you with any help. Please do the best you can.*
> *If you are stuck and cannot continue, you can stop."*

3.4 Usability Metrics

According to the VVSG, voting systems "should support a process that provides a high level of usability for all voters. The goal is for voters to be able to negotiate the process effectively, efficiently, and comfortably."

To this end, metrics for effectiveness, efficiency and user satsifaction were captured during the usability testing. The goals of the test were to assess:

- Effectiveness of the [VSUT] by measuring the system's Total Completion Score, Voter Inclusion Index, and Perfect Ballot Index.
- Efficiency of the product by measuring the Average Session Time.
- Satisfaction of the system by measuring Average Voter Confidence and Average Voter Likeability.

The results of the Total Completion Score, Voter Inclusion Index, and Perfect Ballot Index were compared with the benchmarks set in section 3.2.1.1 of the VVSG. No benchmarks have been set for average session time, average voter confidence and average voter likeability.

3.4.1 Effectiveness

To measure the effectiveness of the [VSUT], the testing team captured metrics relating to voters' completion rate and errors encountered. The following metrics were calculated for both the VSUT and calibration system:

Measure	Description
NPART	*Number of participants who attempted to vote on the system.*
NCAST	*Number of participants who successfully cast a ballot on the system.*
NPERFECT	*Number of participants who successfully cast a perfectly correct ballot on the system.*
Total Completion Score	*Percentage of test participants who were able to complete the process of voting and cast their ballots so that their ballot choices were recorded by the system. Failure to cast a ballot might involve problems such as a voter simply "giving up" during the voting session because of an inability to operate the system, or a mistaken belief that the casting has been successful.* ***VVSG Benchmark for Total Completion Score = 98%***

In addition, the following metrics for the VSUT were calculated:

Measure	Description
$\overline{BAS}$	*The Base Accuracy Score is the mean percentage of all ballot choices that are correctly cast by each of the test participants. Each voter is given 28 "voting opportunities" within the ballot. The number of these that are correctly performed divided by 28 yields that voter's base accuracy score.*
Voter Inclusion Index	*The base accuracy score is used to determine the Voter Inclusion Index (VII). A measure of overall voting accuracy that uses the Base Accuracy Score and the standard deviation.* $$VII = \frac{\overline{BAS} - LSL}{3S}$$ $\overline{BAS}$ = *Mean of Base Accuracy Scores* *LSL* = *Lower Specification Limit, which is set to 85%* *S* = *Standard Deviation* ***VVSG Benchmark for Voter Inclusion Index = .35***
Perfect Ballot Index	*The ratio of the number of cast ballots containing no errors to the number of ballot containing at least one error.* ***VVSG Benchmark for Perfect Ballot Index = 2.33***

3.4.2 Efficiency

To measure the efficiency of the [VSUT], the testing team measured voters' average time to complete the testing session.

3.4.2.1 Time on Task

To measure voters' efficiency with the [VSUT], he testing team used [a stopwatch] to record each voter's time to cast the ballot in seconds.

Measure	Description
Average Session Time	*Mean time taken per test participant (who successfully completed a ballot) to complete the process of activating, filing out and casting the ballot.* *The average time on task is calculated simply as the sum of times taken, in seconds, (TASKTIME-i) divided by the number of participants who successfully cast a ballot on the system (NCAST).* *The standard deviation for the TASKTIME-i is an optional metric that may also be reported.* ***Note: No benchmarks have been set for Average Session Time.***

3.4.3 SATISFACTION

To measure voters' confidence and likeability of the [VSUT], the testing team administered a post-test questionnaire which included three questions, one of which was used to establish user confidence with the system and one which was used to establish users' likeability with the VSUT.

3.4.3.1 SATISFACTION RATING

To measure voters' satisfaction with the [VSUT], the testing team analyzed average voter confidence and average voter likeability.

Measure	Description
Average Voter Confidence	*Mean confidence level expressed by voters that they believed they voted correctly and the system successfully recorded their votes.* ***Note: No benchmarks have been set for Average Voter Confidence.***
Average Voter Likeability	*Mean likeability level expressed by voters on how well they liked the system.* ***Note: No benchmarks have been set for Average Voter Likeability.***

4.0 RESULTS

4.1 DATA ANALYSIS

The results of the usability test were calculated using formulas specified in VPP (provided above in the Usability Metrics section). In addition to these analyses, which were performed to measure the effectiveness, efficiency and satisfaction with the VSUT, several additional analyses were performed on the calibration system to ensure a valid usability test and on the data sets to deal with any participants who did not follow instructions.

The following section will deal with four primary areas, including data collection, data scoring, data reduction and statistical analysis.

4.1.1 DATA COLLECTION

To collect the data for this usability test, a combination of methods was used:

- Participants' demographic information was provided by[the recruiting firm] based on participants' response to a set of recruiting questions. See Appendix [A] for a copy of the recruiting screener.
- Votes for each ballot were recorded using the VSUT and the calibration system.
- Session time, in seconds, was manually collected by one of the data logger(s), using [a stopwatch].
- Responses to a post-test questionnaire were completed by the participant using [a paper and pencil form]. See Appendix [F] for a copy of the questionnaire.

[Following the usability test, participants' demographic information, time on task, responses to the post-test questionnaire were recorded into an Excel spreadsheet, along with the participant's ID number.]

4.1.2 DATA SCORING:

The data was scored in a variety of ways. Following is a list of the ways that the data was scored for the calibration system and VSUT (including the scoring of ballots, time on task and post-test questionnaires).

Calibration System

To ensure that the test of the VSUT was a valid test, a test of the calibration system was conducted in parallel with the test of the VSUT. The purpose of this safeguard was to gather performance data on the calibration system in the same environment, with the same types of participants as those who tested the VSUT to ensure that the results from the calibration system matched previous performance established with this system in a previous test.

In order to ensure the validity of the testing procedure, the ***current*** results from the calibration system were compared to the nominal results (the nominal results were previously validated as truly representative of the performance of the calibration system).

To score the data from the calibration system, the benchmark calculation Perl scripts were used to calculate the Total Completion Score and 95% confidence intervals of the calibration system for both the current test and the nominal data. If the results of the current test did not fall within the 95% confidence intervals for the nominal data, than the test would have been considered invalid.

A Mann-Whitney test was also used to determine whether the procedure is valid or not. This was accomplished by comparing previous results of the calibration system scores against the current distribution of calibration system scores. If these were not sufficiently similar (if the z-score had been outside the normal 95% confidence interval), then the test procedure would have been rejected as invalid.

Based on the analysis of the calibration system, the Total Completion Score for the current system was [XX%], which fell between the confidence intervals for the nominal test of [XX%-XX%]. Additionally, the z-score from the Mann-Whitney test was [X.XX , Note: the z-score should be between 1.96 and -1.96].

[Both of these results indicate that the results of the current test are similar to the results of the nominal test, suggesting that the test was a valid test of both the calibration system and the VSUT].

VSUT

The following section will detail how each of the ballots were scored, how the time data was analyzed, and how the post-test questionnaires were scored.

- **Ballot Scoring**

  Ballots were scored as recorded by the voting system (not as marked). That is, ballots were scored as specified in VPP using the electronic record from the voting system, as opposed to the DRE screen.

  Of the 28 tasks, each task was evaluated to determine if the voter was able to successfully cast a vote in a way that matched the instructions. Of the 28 contests, participants were asked to vote for one individual in 17 of the races, accounting for 17 points. In addition, there was one contest which included two votes, one contest with four votes, and one contest with five votes, accounting for the other 11 points. A perfect ballot received a score of 28.

  Errors included missing votes, incorrect votes and unintended votes (votes in contests where the participants were instructed not to vote). For single member elections, retention races, constitutional amendments and ballot initiatives, only one error per task was counted. For multimember elections, the maximum number of errors was set to the number of candidates in the slate. Therefore, an incorrect vote in a multimember contest was counted as a single error (as opposed to being counted as two errors – one for failing to vote as intended and one for voting an unintended vote).

  Using this data, the test facilitator entered the data in the 'transcribe-ballots' script and used the 'tabulate-election' script to calculate the Total Completion Score, Perfect Ballot Index and the Voter Inclusion Index.

- **Time on Task**

  Time on task was manually entered, using the times captured by the data logger, in seconds. [Following the testing, the times were entered into the spreadsheet and matched to the appropriate participant ID. Once in Excel, the team calculated the mean time on task (for successful ballots) and the standard deviation.]

- **Post-Test Questionnaire**

  The post-test questionnaires were also manually scored. [After the testing, the data from the paper and pencil forms was transcribed into Excel and matched to the appropriate participant ID.] The mean scores for confidence and likeability were then calculated. Responses to the question about following the voting instructions were used to identify participants who did not follow the instructions.

4.1.3 Data Reduction

To identify participants who did not follow the instructions, or participants who intentionally did not follow the instructions provided to them, the first question on the post-test questionnaire asked voters if they tried to follow the instructions given to them.

The testing team used the procedures for dealing with participants who did not follow the instructions as outlined in the VPP, which is summarized below:

If more than five participants answered 'no' (that they did not follow instructions) for either the VSUT or calibration system, then the test would be deemed invalid (since there is an excessive number of participants who did not follow the instructions).

- indicates a serious problem with execution of the test.
- If there were between one and five participants who did not follow the instructions; and it is possible to identify them, then their data was to be removed, and the remaining data was analyzed.
- If there were between one and five participants who did not follow the instructions; and it is not possible to identify the participants, then:

- The original set of data for the VSUT would be analyzed. If the data passes the benchmarks for the Perfect Ballot Index and the Voter Inclusion Index, then the system is seen as having passed (even with the inclusion of these participants)

- If the VSUT does not pass using the original set of data, then the data shall be analyzed using a 'best-case' set of data, or the highest scores on the system. If the system still does not meet the benchmarks, then the system fails. If the system meets the benchmarks, using the 'best case' data but not the original set of data, then the results are indeterminate and the test shall be abandoned.

- Lastly, use the Mann-Whitney analysis for the calibration system using both the original and best-case set of ballots. If both analyses yield a z-score outside the normal 95% CI, then this run of the VPP test method is invalid and shall be abandoned.

4.1.4 STATISTICAL ANALYSIS

Once the test was shown to be valid and the participants who did not follow the instructions were discarded, the usability test administrator(s) and data logger(s) used the statistical analysis and procedures outlines in the VPP, along with the benchmark calculation Perl scripts provided, to calculate the Total Completion Score, Perfect Ballot Index and Voter Inclusion Index.

In addition, the team followed the criteria outlined in the VPP to calculate Average Time on Task, Average Confidence Score, and the Average Likeability Score.

A copy of the raw scores for both the VSUT and the calibration system can be found in Appendix [K].

4.2 Presentation of the Results

This section shall detail the results for both the calibration system and the VSUT.

Based on these minimum benchmarks set by the VVSG and the performance of the VSUT, [the VSUT meets or exceeds the criteria for all scores, including Total Completion Score, Perfect Ballot Index and Voter Inclusion Index].

The distribution of scores along with the raw data can be found in Appendix [K].

Measure	Description	Minimum Benchmarks	VUST N=XX	Calibration System N=XX
NPART	*Number of participants who attempted to vote on the system.*	**At least 100**	**[XXX]**	**[XXX]**
NCAST	*Number of participants who successfully cast a ballot on the system.*	**--**	**[XX]**	**[XX]**
NPERFECT	*Number of participants who successfully cast a perfectly correct ballot on the system.*	**--**	**[XX]**	**--**
PCORRECT-i	*Proportion of voting opportunities taken by the i-th participant.*	**--**	**[Mean] [STDV]**	**--**
Z-Score from Mann-Whitney Test	*Z-score resulting from Mann-Whitney comparison of current and nominal distributions for the calibration system.*	**1.96 to -1.96**	**--**	**[XX]**

Measure	Description	Minimum Benchmarks	VUST N=XX	Calibration System N=XX
Total Completion Score	*Percentage of test participants who were able to complete the process of voting and cast their ballots so that their ballot choices were recorded by the system.*	**98%**	**[Mean] [95% CI]**	**[Mean] [95% CI]**
Voter Inclusion Index	*Measurement that combines accuracy with variability in the level of accuracy among test participants.*	**.35**	**[Mean] [95% CI]**	**--**
Perfect Ballot Index	*Comparison of the number of participants who cast a ballot without any errors to those that had a least one error and is used for detecting a systemic design problem that causes the same type of error by many test participants.*	**2.33**	**[Mean] [95% CI]**	**--**
Average Session Time	*Mean time taken per test participant to complete the process of activating, filing out and casting the ballot.*	**--**	**[Mean] [STDV]**	**--**
Average Voter Confidence	*Mean confidence level expressed by voters that they believed they voted correctly and the system successfully recorded their votes.*	**--**	**[Mean] [STDV]**	**--**
Average Voter Likeability	*Mean likeability level expressed by voters on how well they liked the system.*	**--**	**[Mean] [STDV]**	**--**

5.0 APPENDICES

The following appendices include supplemental data for this usability test report. Following is a list of the appendices provided:

- Appendix A: Recruiting screener
- Appendix B: Participant demographics
- Appendix C: Informed consent
- Appendix D: Instructions for participants
- Appendix E: Test ballot specification
- Appendix F: Post-test satisfaction questionnaire
- Appendix G: Incentive receipt and acknowledgment form
- Appendix H: Note-taking document
- Appendix I: Scoring tools
- Appendix J: Photos of the testing environment and voting systems
- Appendix K: Results

APPENDIX A: RECRUITING SCREENER

Following is a copy of the recruiting screener used to recruit voters for this usability test:

[NOTE: This recruiting screener is just a sample and is not meant to imply any requirements or guidelines.]

Instructions for the recruiting firm:

We request a total of [250] individuals who are defined as follows:

- All shall be eligible to vote in the US.
- We aim to match the target demographics below

[NOTE: These demographics match the target demographics outlined in the VPP].

Gender

Men	45%
Women	55%
TOTAL (participants)	[250]

Race:

African American	10%
Hispanic	10%
Non-Hispanic White	80%
TOTAL (participants)	[250]

Age

18-24	10%
25-34	20%
35-44	25%
45-54	25%
55-64	20%
TOTAL (participants)	[250]

Education

High School Graduate	25%
Some College	35%
College Graduate	30%
Post-Graduate Degree	10%
TOTAL (participants)	[250]

- Participants will be asked to vote on a voting machine and fill out some forms which will take about [10-15] minutes. However, individuals will be scheduled [six] per hour so that there may be some wait time incurred.
- Dates for testing are [date].

Recruiting Script for Recruiting Firm

Hello, my name is _____, calling from [Insert name of recruiting firm]. We are recruiting individuals with and without disabilities to participate in a usability study for a voting machine. We would like to ask you a few questions to see if you qualify and if would like to participate. This will only take a few minutes of your time and no one will attempt to sell you anything. This is strictly for research purposes. If you are interested and qualify for the study, you will be paid to participate. Can I ask you a few questions?

1. Do you or anyone in your household work for:

   ☐ An Internet or software development company
   ☐ A usability or market research business/company

   (If yes to any of the above, Dismiss)

2. When was the last time, if ever, you participated in a usability study?

   ☐ Within the last 6 months? (Dismiss)
   ☐ Over 6 months?
   ☐ Never

3. Are you qualified to vote in the US (a US citizen 18 years or older)?

   ☐ Yes
   ☐ No (dismiss)

4. Do you or any of your immediate family work in any of the following situations:

   ☐ Poll Worker or as a worker in another part of the voting process
   ☐ Voting Machine Manufacturing
   ☐ Voting Machine Development, Marketing, or Sales
   ☐ Any other position that is part of the voting process. Please describe

   (If yes to any of the above, Dismiss)

5. Do you consider yourself to be a person with disabilities?

   ☐ Yes
   ☐ No (go to question 7)

6. Please describe your disability. [Check as appropriate]

   ☐ In wheelchair
   ☐ Use a cane or walker
   ☐ Blind
   ☐ Low Vision -- What is your corrected vision?
   ☐ Movement impaired -- Please describe what technique you use to vote
   ☐ Hearing loss
   ☐ Cognitive Disability
   ☐ Other Please describe

   Also, please describe any adaptive equipment you use to vote

7. Which of the following describes your age?

   ☐ 18-24
   ☐ 25-34
   ☐ 35-44
   ☐ 45-54
   ☐ 55-64

8. Which of the following describes your education?

   ☐ High School Degree
   ☐ Some College
   ☐ College graduate
   ☐ Post Graduate

9. Do you have any difficulty reading or writing in English?

   ☐ Yes
   ☐ No (Dismiss)

10. Which of the following describes your race

- ☐ African American
- ☐ White
- ☐ Hispanic

11. Which of the following describes your gender?

- ☐ Male
- ☐ Female

12. Is the contact information that we used today, going to be appropriate in about a month?

- ☐ Yes
- ☐ No

13. Let me [verify/get] your contact information that will be current in a month

Name: _____
Contact phone: _____
Contact email: _____

Thank you for taking the time to answer my questions. We will be back in touch with you when a test is scheduled.

Confirming Participants

We would like to invite you to participate in a one-on-one interview session on [date] at [location]. You will be using a voting machine. The session will last approximately [20 minutes] and you will be compensated [$XX] as a token of our appreciation for taking the time to help us.

Again, this is strictly for research purposes and no one will attempt to sell you anything. May I schedule you to attend? We will mail you a confirmation letter and directions and we will call you the day before to confirm your attendance.

APPENDIX B: PARTICIPANT DEMOGRAPHICS

Following is a high-level overview of the voters who used the calibration system:

Gender

Men	[X]
Women	[X]
TOTAL (participants)	[X]

Race:

African American	[X]
Hispanic	[X]
Non-Hispanic White	[X]
TOTAL (participants)	[X]

Age

18-24	[X]
25-34	[X]
35-44	[X]
45-54	[X]
55-64	[X]
TOTAL (participants)	[X]

Education

High School Graduate	[X]
Some College	[X]
College Graduate	[X]
Post-Graduate Degree	[X]
TOTAL (participants)	[X]

Following is a full list of participant demographics of those voters who used the calibration system:

NOTE: These demographics are just a sample set of demographics and not meant to imply any requirements or guidelines.

Demographics				
Participant	**Gender**	**Race**	**Education**	**Age**
1	Female	African-American	Some College	25-34
2	Female	Non-Hispanic White	College Graduate	25-34
3	Male	Non-Hispanic White	College Graduate	45-54
4	Female	Hispanic	Post Graduate Degree	45-54
5	Female	Non-Hispanic White	High School Graduate	25-34
6	Male	Hispanic	Some College	45-54
7	Male	Non-Hispanic White	Post Graduate Degree	55-64
8	Male	African-American	College Graduate	45-54
9	Female	African-American	Some College	25-34
10	Female	Non-Hispanic White	College Graduate	25-34
11	Male	Non-Hispanic White	College Graduate	45-54
12	Female	Hispanic	Post Graduate Degree	45-54
13	Female	Non-Hispanic White	High School Graduate	25-34
14	Male	Hispanic	Some College	45-54
15	Male	Non-Hispanic White	Post Graduate Degree	55-64
16	Male	African-American	College Graduate	45-54
17	Female	African-American	Some College	25-34
18	Female	Non-Hispanic White	College Graduate	25-34
19	Male	Non-Hispanic White	College Graduate	45-54
20	Female	Hispanic	Post Graduate Degree	45-54
21	Female	Non-Hispanic White	High School Graduate	25-34
22	Male	Hispanic	Some College	45-54
23	Male	Non-Hispanic White	Post Graduate Degree	55-64
24	Male	African-American	College Graduate	45-54
25	Female	African-American	Some College	25-34
26	Female	Non-Hispanic White	College Graduate	25-34
27	Male	Non-Hispanic White	College Graduate	45-54
28	Female	Hispanic	Post Graduate Degree	45-54
29	Female	Non-Hispanic White	High School Graduate	25-34
30	Male	Hispanic	Some College	45-54
31	Male	Non-Hispanic White	Post Graduate Degree	55-64

Demographics				
Participant	**Gender**	**Race**	**Education**	**Age**
32	Male	African-American	College Graduate	45-54
33	Female	African-American	Some College	25-34
34	Female	Non-Hispanic White	College Graduate	25-34
35	Male	Non-Hispanic White	College Graduate	45-54
36	Female	Hispanic	Post Graduate Degree	45-54
37	Female	Non-Hispanic White	High School Graduate	25-34
38	Male	Hispanic	Some College	45-54
39	Male	Non-Hispanic White	Post Graduate Degree	55-64
40	Male	African-American	College Graduate	45-54
41	Female	African-American	Some College	25-34
42	Female	Non-Hispanic White	College Graduate	25-34
43	Male	Non-Hispanic White	College Graduate	45-54
44	Female	Hispanic	Post Graduate Degree	45-54
45	Female	Non-Hispanic White	High School Graduate	25-34
46	Male	Hispanic	Some College	45-54
47	Male	Non-Hispanic White	Post Graduate Degree	55-64
48	Male	African-American	College Graduate	45-54
49	Female	African-American	Some College	25-34
50	Female	Non-Hispanic White	College Graduate	25-34
51	Male	Non-Hispanic White	College Graduate	45-54
52	Female	Hispanic	Post Graduate Degree	45-54
53	Female	Non-Hispanic White	High School Graduate	25-34
54	Male	Hispanic	Some College	45-54
55	Male	Non-Hispanic White	Post Graduate Degree	55-64
56	Male	African-American	College Graduate	45-54
57	Female	African-American	Some College	25-34
58	Female	Non-Hispanic White	College Graduate	25-34
59	Male	Non-Hispanic White	College Graduate	45-54
60	Female	Hispanic	Post Graduate Degree	45-54
61	Female	Non-Hispanic White	High School Graduate	25-34
62	Male	Hispanic	Some College	45-54
63	Male	Non-Hispanic White	Post Graduate Degree	55-64
64	Male	African-American	College Graduate	45-54
65	Female	African-American	Some College	25-34
66	Female	Non-Hispanic White	College Graduate	25-34
67	Male	Non-Hispanic White	College Graduate	45-54

Demographics				
Participant	**Gender**	**Race**	**Education**	**Age**
68	Female	Hispanic	Post Graduate Degree	45-54
69	Female	Non-Hispanic White	High School Graduate	25-34
70	Male	Hispanic	Some College	45-54
71	Male	Non-Hispanic White	Post Graduate Degree	55-64
72	Male	African-American	College Graduate	45-54
73	Female	African-American	Some College	25-34
74	Female	Non-Hispanic White	College Graduate	25-34
75	Male	Non-Hispanic White	College Graduate	45-54
76	Female	Hispanic	Post Graduate Degree	45-54
77	Female	Non-Hispanic White	High School Graduate	25-34
78	Male	Hispanic	Some College	45-54
79	Male	Non-Hispanic White	Post Graduate Degree	55-64
80	Male	African-American	College Graduate	45-54
81	Female	African-American	Some College	25-34
82	Female	Non-Hispanic White	College Graduate	25-34
83	Male	Non-Hispanic White	College Graduate	45-54
84	Female	Hispanic	Post Graduate Degree	45-54
85	Female	Non-Hispanic White	High School Graduate	25-34
86	Male	Hispanic	Some College	45-54
87	Male	Non-Hispanic White	Post Graduate Degree	55-64
88	Male	African-American	College Graduate	45-54
89	Female	African-American	Some College	25-34
90	Female	Non-Hispanic White	College Graduate	25-34
91	Male	Non-Hispanic White	College Graduate	45-54
92	Female	Hispanic	Post Graduate Degree	45-54
93	Female	Non-Hispanic White	High School Graduate	25-34
94	Male	Hispanic	Some College	45-54
95	Male	Non-Hispanic White	Post Graduate Degree	55-64
96	Male	African-American	College Graduate	45-54
97	Female	Non-Hispanic White	High School Graduate	25-34
98	Male	Hispanic	Some College	45-54
99	Male	Non-Hispanic White	Post Graduate Degree	55-64
100	Male	African-American	College Graduate	45-54

Following is a high-level overview of the voters who used the VSUT:

Gender

Men	[X]
Women	[X]
TOTAL (participants)	[X]

Race:

African American	[X]
Hispanic	[X]
Non-Hispanic White	[X]
TOTAL (participants)	[X]

Age

18-24	[X]
25-34	[X]
35-44	[X]
45-54	[X]
55-64	[X]
TOTAL (participants)	[X]

Education

High School Graduate	[X]
Some College	[X]
College Graduate	[X]
Post-Graduate Degree	[X]
TOTAL (participants)	[X]

Following is a full list of participant demographics of those voters who used the VSUT:

[NOTE: These demographics are just a sample set of demographics and are not meant to imply any requirements or guidelines.]

Demographics				
Participant	**Gender**	**Race**	**Education**	**Age**
1	Female	African-American	Some College	25-34
2	Female	Non-Hispanic White	College Graduate	25-34
3	Male	Non-Hispanic White	College Graduate	45-54
4	Female	Hispanic	Post Graduate Degree	45-54
5	Female	Non-Hispanic White	High School Graduate	25-34
6	Male	Hispanic	Some College	45-54
7	Male	Non-Hispanic White	Post Graduate Degree	55-64
8	Male	African-American	College Graduate	45-54
9	Female	African-American	Some College	25-34
10	Female	Non-Hispanic White	College Graduate	25-34
11	Male	Non-Hispanic White	College Graduate	45-54
12	Female	Hispanic	Post Graduate Degree	45-54
13	Female	Non-Hispanic White	High School Graduate	25-34
14	Male	Hispanic	Some College	45-54
15	Male	Non-Hispanic White	Post Graduate Degree	55-64
16	Male	African-American	College Graduate	45-54
17	Female	African-American	Some College	25-34
18	Female	Non-Hispanic White	College Graduate	25-34
19	Male	Non-Hispanic White	College Graduate	45-54
20	Female	Hispanic	Post Graduate Degree	45-54
21	Female	Non-Hispanic White	High School Graduate	25-34
22	Male	Hispanic	Some College	45-54
23	Male	Non-Hispanic White	Post Graduate Degree	55-64
24	Male	African-American	College Graduate	45-54
25	Female	African-American	Some College	25-34
26	Female	Non-Hispanic White	College Graduate	25-34
27	Male	Non-Hispanic White	College Graduate	45-54
28	Female	Hispanic	Post Graduate Degree	45-54
29	Female	Non-Hispanic White	High School Graduate	25-34
30	Male	Hispanic	Some College	45-54
31	Male	Non-Hispanic White	Post Graduate Degree	55-64

Demographics				
Participant	**Gender**	**Race**	**Education**	**Age**
32	Male	African-American	College Graduate	45-54
33	Female	African-American	Some College	25-34
34	Female	Non-Hispanic White	College Graduate	25-34
35	Male	Non-Hispanic White	College Graduate	45-54
36	Female	Hispanic	Post Graduate Degree	45-54
37	Female	Non-Hispanic White	High School Graduate	25-34
38	Male	Hispanic	Some College	45-54
39	Male	Non-Hispanic White	Post Graduate Degree	55-64
40	Male	African-American	College Graduate	45-54
41	Female	African-American	Some College	25-34
42	Female	Non-Hispanic White	College Graduate	25-34
43	Male	Non-Hispanic White	College Graduate	45-54
44	Female	Hispanic	Post Graduate Degree	45-54
45	Female	Non-Hispanic White	High School Graduate	25-34
46	Male	Hispanic	Some College	45-54
47	Male	Non-Hispanic White	Post Graduate Degree	55-64
48	Male	African-American	College Graduate	45-54
49	Female	African-American	Some College	25-34
50	Female	Non-Hispanic White	College Graduate	25-34
51	Male	Non-Hispanic White	College Graduate	45-54
52	Female	Hispanic	Post Graduate Degree	45-54
53	Female	Non-Hispanic White	High School Graduate	25-34
54	Male	Hispanic	Some College	45-54
55	Male	Non-Hispanic White	Post Graduate Degree	55-64
56	Male	African-American	College Graduate	45-54
57	Female	African-American	Some College	25-34
58	Female	Non-Hispanic White	College Graduate	25-34
59	Male	Non-Hispanic White	College Graduate	45-54
60	Female	Hispanic	Post Graduate Degree	45-54
61	Female	Non-Hispanic White	High School Graduate	25-34
62	Male	Hispanic	Some College	45-54
63	Male	Non-Hispanic White	Post Graduate Degree	55-64
64	Male	African-American	College Graduate	45-54
65	Female	African-American	Some College	25-34
66	Female	Non-Hispanic White	College Graduate	25-34
67	Male	Non-Hispanic White	College Graduate	45-54

Demographics				
Participant	**Gender**	**Race**	**Education**	**Age**
68	Female	Hispanic	Post Graduate Degree	45-54
69	Female	Non-Hispanic White	High School Graduate	25-34
70	Male	Hispanic	Some College	45-54
71	Male	Non-Hispanic White	Post Graduate Degree	55-64
72	Male	African-American	College Graduate	45-54
73	Female	African-American	Some College	25-34
74	Female	Non-Hispanic White	College Graduate	25-34
75	Male	Non-Hispanic White	College Graduate	45-54
76	Female	Hispanic	Post Graduate Degree	45-54
77	Female	Non-Hispanic White	High School Graduate	25-34
78	Male	Hispanic	Some College	45-54
79	Male	Non-Hispanic White	Post Graduate Degree	55-64
80	Male	African-American	College Graduate	45-54
81	Female	African-American	Some College	25-34
82	Female	Non-Hispanic White	College Graduate	25-34
83	Male	Non-Hispanic White	College Graduate	45-54
84	Female	Hispanic	Post Graduate Degree	45-54
85	Female	Non-Hispanic White	High School Graduate	25-34
86	Male	Hispanic	Some College	45-54
87	Male	Non-Hispanic White	Post Graduate Degree	55-64
88	Male	African-American	College Graduate	45-54
89	Female	African-American	Some College	25-34
90	Female	Non-Hispanic White	College Graduate	25-34
91	Male	Non-Hispanic White	College Graduate	45-54
92	Female	Hispanic	Post Graduate Degree	45-54
93	Female	Non-Hispanic White	High School Graduate	25-34
94	Male	Hispanic	Some College	45-54
95	Male	Non-Hispanic White	Post Graduate Degree	55-64
96	Male	African-American	College Graduate	45-54
97	Female	Non-Hispanic White	High School Graduate	25-34
98	Male	Hispanic	Some College	45-54
99	Male	Non-Hispanic White	Post Graduate Degree	55-64
100	Male	African-American	College Graduate	45-54

APPENDIX C: INFORMED CONSENT

Informed Consent

RESEARCH DESCRIPTION: The [name of test laboratory] is conducting a study to determine how easy it is for voters to use voting systems. Usability will be measured by determining the time it takes a voter to vote, the number of errors when the vote is cast, and voter satisfaction.

You will receive written instructions on how you as a voter "want to vote". You will be asked to vote as the paper instructs on a specific voting system. In addition to collecting your votes, there may be a camera focused on the system and your hands, but your face will not be photographed. After you cast your ballot, you will be asked for your opinion about the voting system and your voting experience. You will also be asked for demographic data to include age, gender, race, and education level. This process should take you no more than 30 minutes.

CONFIDENTIALITY: All the data collected will be anonymous. The data will be used by [name of test laboratory] to evaluate the usability of a voting system. The data will not be associated with any particular individual. All of the time and error data, demographic data, and voter experience and satisfaction data will be anonymous. All of the data will only be identified and linked together by a number, and will not be linked back to an individual in any way.

You are free to withdraw from the study at any time during the experiment. In total, we expect to have approximately [100] subjects complete the experiment. There are no risks involved in participating in this study, nor are there any immediate benefits. The long term benefits of this study should be improved voting systems.

CONTACT INFORMATION: For questions regarding this study, please contact [Contact name, phone number and email address].

"I have read the above description of this research project. I have also spoken to the usability test facilitator who answered any questions I had about this project. I acknowledge that I have received a personal copy of this form. I agree to participate in this research and I understand that I may withdraw at any time."

Signature: _____ Date: _____

Usability Researcher: _____

Signature Of Usability Researcher: _____ Date: _____

Witness: _____

Witness Signature: _____ Date: _____

APPENDIX D: VOTING INSTRUCTIONS

Following is a copy of the voting instructions in the exact format provided to voters during the usability test.

Voting Instructions

Note: We are testing the voting system in a specific configuration. Though the system may allow you to change features such as font size and contrast, please do not adjust these settings for this test.

In our mock election, we will be using fake names for candidates and colors for political party names. For example, you might see this:

Joe Jones/Yellow Party

Any similarity between names of candidates and real people or colors and real parties is purely coincidental.

Please attempt to vote exactly as described on the following pages

Once you start, we will not be able to help you.

Please do the best you can. If you are stuck and cannot continue, inform the administrator.

Thank you.

For President and Vice President of the United States, vote for
Adam Cramer and Greg Vuocolo

For Senator, vote for
David Platt

For Congress, vote for
Brad Schott

For Governor, vote for
Cathy Steele

Do not cast a vote for
Lieutenant Governor

For Registrar of Deeds, write in a vote for
Christopher Christopher

For State Senator, vote for
Edward Shiplett

For State Assemblyman, vote for
Amos Keller

For County Commissioners, vote for the following candidates:
Camille Argent
Mary Tawa
Joe Barry

and enter write in votes for:
Dorothy Johns
Charles Blank

For Court of Appeals Judge, vote for
Michael Marchesani

For Water Commissioner, vote for
Orville White
Gregory Seldon

For City Council, vote for the following candidates:
Randall Rupp
Carroll Shry
Donald Davis

For Chief Justice of the Supreme Court
Vote to keep Robert Demergue in office

For the question of retaining Justice of the Supreme Court Elmer Hull
Do not vote

For Proposed Constitutional Amendment C
Vote for this amendment

For Proposed Constitutional Amendment D
Vote for this amendment

For Proposed Constitutional Amendment H
Vote against this amendment

For Proposed Constitutional Amendment K
Vote against this amendment

For Ballot Measure 101: Open Primaries
Do not vote

For Ballot Measure 106: Limits on Private Enforcement of Unfair Business Competition Laws
Vote for the measure

Cast your ballot

APPENDIX E: TEST BALLOT SPECIFICATION

Following is a copy of the ballot specification.

INFORMATION APPLICABLE TO WHOLE BALLOT

Date and Time	2004-nov-02, 7:00 AM to 8:00 PM
State	Maryland
County	Madison
Party Line Voting Method	Enabled for partisan contests

INFORMATION APPLICABLE TO EVERY CONTEST

Full-term or partial-term election	Full-term
Voting Method	Simple vote for N candidate(s) - (i.e. no ranked voting)

CONTEST #0:

Title of Contest	Straight Party Vote
Partisanship	Partisan
Minimum Votes Allowed	0
Maximum Votes Allowed	1
Maximum Write-in Votes Allowed	0

- **Option #0.1:** Blue
- **Option #0.2:** Yellow
- **Option #0.3:** Purple
- **Option #0.4:** Orange
- **Option #0.5:** Pink
- **Option #0.6:** Gold
- **Option #0.7:** Gray
- **Option #0.8:** Aqua
- **Option #0.9:** Brown

CONTEST #1:

Title of Office	President and Vice-President of the United States
District of Office	United States
Partisanship	Partisan
Minimum Votes Allowed	0
Maximum Votes Allowed	1
Maximum Write-in Votes Allowed	0

- **Candidate #1.1:** Joseph Barchi and Joseph Hallaren / Blue
- **Candidate #1.2:** Adam Cramer and Greg Vuocolo / Yellow
- **Candidate #1.3:** Daniel Court and Amy Blumhardt / Purple

- **Candidate #1.4:** Alvin Boone and James Lian / Orange
- **Candidate #1.5:** Austin Hildebrand-MacDougall and James Garritty / Pink
- **Candidate #1.6:** Martin Patterson and Clay Lariviere / Gold
- **Candidate #1.7:** Elizabeth Harp and Antoine Jefferson / Gray
- **Candidate #1.8:** Charles Layne and Andrew Kowalski / Aqua
- **Candidate #1.9:** Marzena Pazgier and Welton Phelps / Brown

CONTEST #2:

Title of Office	US Senate
District of Office	Statewide
Partisanship	Partisan
Minimum Votes Allowed	0
Maximum Votes Allowed	1
Maximum Write-in Votes Allowed	1

- **Candidate #2.1:** Dennis Weiford / Blue
- **Candidate #2.2:** Lloyd Garriss / Yellow
- **Candidate #2.3:** Sylvia Wentworth-Farthington / Purple
- **Candidate #2.4:** John Hewetson / Orange
- **Candidate #2.5:** Victor Martinez / Pink
- **Candidate #2.6:** Heather Portier / Gold
- **Candidate #2.7:** David Platt / Gray

CONTEST #3:

Title of Office	US Representative
District of Office	6th Congressional District
Partisanship	Partisan
Minimum Votes Allowed	0
Maximum Votes Allowed	1
Maximum Write-in Votes Allowed	1

- **Candidate #3.1:** Brad Plunkard / Blue
- **Candidate #3.2:** Bruce Reeder / Yellow
- **Candidate #3.3:** Brad Schott / Purple
- **Candidate #3.4:** Glen Tawney / Orange
- **Candidate #3.5:** Carroll Forrest / Pink

CONTEST #4:

Title of Office	Governor
District of Office	Statewide
Partisanship	Partisan
Minimum Votes Allowed	0
Maximum Votes Allowed	1
Maximum Write-in Votes Allowed	1

- **Candidate #4.1:** Charlene Franz / Blue
- **Candidate #4.2:** Gerard Harris / Yellow
- **Candidate #4.3:** Linda Bargmann / Purple
- **Candidate #4.4:** Barbara Adcock / Orange
- **Candidate #4.5:** Carrie Steel-Loy / Pink
- **Candidate #4.6:** Frederick Sharp / Gold
- **Candidate #4.7:** Alex Wallace /Gray
- **Candidate #4.8:** Barbara Williams / Aqua
- **Candidate #4.9:** Althea Sharp / Brown
- **Candidate #4.10:** Douglas Alpern / Independent
- **Candidate #4.11:** Ann Windbeck / Independent
- **Candidate #4.12:** Mike Greher / Independent
- **Candidate #4.13:** Patricia Alexander / Independent
- **Candidate #4.14:** Kenneth Mitchell / Independent
- **Candidate #4.15:** Stan Lee / Independent
- **Candidate #4.16:** Henry Ash / Independent
- **Candidate #4.17:** Karen Kennedy / Independent
- **Candidate #4.18:** Van Jackson / Independent
- **Candidate #4.19:** Debbie Brown / Independent
- **Candidate #4.20:** Joseph Teller / Independent
- **Candidate #4.21:** Greg Ward / Independent
- **Candidate #4.22:** Lou Murphy / Independent
- **Candidate #4.23:** Jane Newman / Independent
- **Candidate #4.24:** Jack Callanann / Independent
- **Candidate #4.25:** Esther York / Independent
- **Candidate #4.26:** Glen Chandler / Independent
- **Candidate #4.27:** Marcia Colgate / Independent
- **Candidate #4.28:** Leslie Porter / Independent
- **Candidate #4.29:** Molly Dalton / Independent
- **Candidate #4.30:** David Davis / Independent
- **Candidate #4.31:** May Peterson / Independent
- **Candidate #4.32:** Patricia Dawkins / Independent
- **Candidate #4.33:** Suzanne Adams / Independent
- **Candidate #4.34:** Mary Miller / Independent
- **Candidate #4.35:** Rosalind Leigh / Independent
- **Candidate #4.36:** Elaine Henry / Independent
- **Candidate #4.37:** Gail Moses / Independent
- **Candidate #4.38:** Daniel Jones / Independent
- **Candidate #4.39:** Don Maybee / Independent
- **Candidate #4.40:** Lillian Cohen / Independent
- **Candidate #4.41:** Richard Mitchell / Independent
- **Candidate #4.42:** Pat York / Independent
- **Candidate #4.43:** Linda Rappaport / Independent
- **Candidate #4.44:** Mike Porter / Independent
- **Candidate #4.45:** Margaret Sharp / Independent
- **Candidate #4.46:** Cathy Steele / Independent
- **Candidate #4.47:** Lawrence Smith / Independent
- **Candidate #4.48:** Bill Kendrick / Independent
- **Candidate #4.49:** Fred Stein / Independent
- **Candidate #4.50:** Jerry Cole / Independent

CONTEST #5:

Title of Office	Lieutenant-Governor
District of Office	Statewide
Partisanship	Partisan
Minimum Votes Allowed	0
Maximum Votes Allowed	1
Maximum Write-in Votes Allowed	1

- **Candidate #5.1:** Chris Norberg / Blue
- **Candidate #5.2:** Anthony Parks / Yellow
- **Candidate #5.3:** Luis Garcia / Purple
- **Candidate #5.4:** Charles Qualey / Orange
- **Candidate #5.5:** George Hovis / Pink
- **Candidate #5.6:** Burt Zirkle / Gold
- **Candidate #5.7:** Brenda Davis / Gray
- **Candidate #5.8:** Edward Freeman / Aqua
- **Candidate #5.9:** Paul Swan / Brown

CONTEST #6:

Title of Office	Registrar of Deeds
District of Office	Countywide
Partisanship	Partisan
Minimum Votes Allowed	0
Maximum Votes Allowed	1
Maximum Write-in Votes Allowed	1

- **Candidate #6.1:** Laila Shamsi / Yellow

CONTEST #7:

Title of Office	State Senator
District of Office	31st District
Partisanship	Partisan
Minimum Votes Allowed	0
Maximum Votes Allowed	1
Maximum Write-in Votes Allowed	1

- **Candidate #7.1:** Edward Shiplett / Blue
- **Candidate #7.2:** Marty Talarico / Yellow

CONTEST #8:

Title of Office	State Assemblyman
District of Office	54th District
Partisanship	Partisan
Minimum Votes Allowed	0
Maximum Votes Allowed	1
Maximum Write-in Votes Allowed	1

- **Candidate #8.1:** Andrea Solis / Blue
- **Candidate #8.2:** Amos Keller / Yellow

CONTEST #9:

Title of Office	County Commissioners
District of Office	Countywide
Partisanship	Partisan
Minimum Votes Allowed	0
Maximum Votes Allowed	5
Maximum Write-in Votes Allowed	5

- **Candidate #9.1:** Camille Argent / Blue
- **Candidate #9.2:** Chloe Witherspoon / Blue
- **Candidate #9.3:** Clayton Bainbridge / Blue
- **Candidate #9.4:** Amanda Marracini / Yellow
- **Candidate #9.5:** Charlene Hennessey / Yellow
- **Candidate #9.6:** Eric Savoy / Yellow
- **Candidate #9.7:** Sheila Moskowitz / Purple
- **Candidate #9.8:** Mary Tawa / Purple
- **Candidate #9.9:** Damian Rangel / Purple
- **Candidate #9.10:** Valarie Altman / Orange
- **Candidate #9.11:** Helen Moore / Orange
- **Candidate #9.12:** John White / Orange
- **Candidate #9.13:** Joe Lee / Pink
- **Candidate #9.14:** Joe Barry / Pink
- **Candidate #9.15** Martin Schreiner / Gray

CONTEST #10:

Title of Office	Court of Appeals Judge
District of Office	Statewide, 4th seat
Partisanship	Non-partisan
Minimum Votes Allowed	0
Maximum Votes Allowed	1
Maximum Write-in Votes Allowed	1

- **Candidate #10.1:** Michael Marchesani

CONTEST #11:

Title of Office	Water Commissioners
District of Office	City of Springfield
Partisanship	Partisan
Minimum Votes Allowed	0
Maximum Votes Allowed	2
Maximum Write-in Votes Allowed	2

- **Candidate #11.1:** Orville White / Blue
- **Candidate #11.2:** Gregory Seldon / Yellow

CONTEST #12:

Title of Office	City Council
District of Office	City of Springfield
Partisanship	Partisan
Minimum Votes Allowed	0
Maximum Votes Allowed	4
Maximum Write-in Votes Allowed	4

- **Candidate #12.1:** Harvey Eagle / Blue
- **Candidate #12.2:** Randall Rupp / Blue
- **Candidate #12.3:** Carroll Shry / Blue
- **Candidate #12.4:** Beverly Barker / Yellow
- **Candidate #12.5:** Donald Davis / Yellow
- **Candidate #12.6:** Hugh Smith / Yellow
- **Candidate #12.7:** Reid Feister / Yellow

RETENTION QUESTION #1:

Wording of Question	Retain **Robert Demergue** as Chief Justice of the Supreme Court?

RETENTION QUESTION #2:

Wording of Question	Retain **Elmer Hull** as Associate Justice of the Supreme Court?

REFERENDUM #1:

Title of proposition	PROPOSED CONSTITUTIONAL AMENDMENT C
Wording of proposition	Shall there be amendments to the State constitution intended to have the collective effect of ensuring the separation of governmental power among the three branches of state government: the legislative branch, the executive branch and the judicial branch? a. Article III, Section 6 of the Constitution shall be amended to read as follows: Section 6. Holding of offices under other governments. - Senators and representatives not to hold other appointed offices under state government. --No person holding any office under the government of the United States, or of any other state or country, shall act as a general officer or as a member of the general assembly, unless at the time of taking such engagement that person shall have resigned the office under such government; and if any general officer, senator, representative, or judge shall, after election and engagement, accept any appointment under any other government, the office under this shall be immediately vacated; but this restriction shall not apply to any person appointed to take deposition or acknowledgement of deeds, or other legal instruments, by the authority of any other state or country. No senator or representative shall, during the time for which he or she was elected, be appointed to any state office, board, commission or other state or quasi-public entity exercising executive power under the laws of this state, and no person holding any executive office or serving as a member of any board, commission or other state or quasi-public entity exercising executive power under the laws of this state shall be a member of the senate or the house of representatives during his or her continuance in such office. b. Article V of the Constitution shall be amended to read as follows: The powers of the government shall be distributed into three (3) separate and distinct departments: the legislative, the executive and the judicial. c. Article VI, Section 10 of the Constitution shall be deleted in its entirety. d. Article IX, Section 5 of the Constitution shall be amended to read as follows: Section 5. Powers of appointment.- The governor shall, by and with the advice and consent of the senate, appoint all officers of the state whose appointment is not herein otherwise provided for and all members of any board, commission or other state or quasi-public entity which exercises executive power under the laws of this state; but the general assembly may by law vest the appointment of such inferior officers, as they deem proper, in the governor, or within their respective departments in the other general officers, the judiciary or in the heads of departments.

REFERENDUM #2:

Title of proposition	PROPOSED CONSTITUTIONAL AMENDMENT D
Wording of proposition	Shall there be an amendment to the State constitution concerning recovery of damages relating to construction of real property improvements, and, in connection therewith, prohibiting laws that limit or impair a property owner's right to recover damages caused by a failure to construct an improvement in a good and workmanlike manner; defining "good and workmanlike manner" to include construction that is suitable for its intended purposes; and permitting exceptions for laws that limit punitive damages, afford governmental immunity, or impose time limits of specified minimum lengths on filing lawsuits?

REFERENDUM #3:

Title of proposition	PROPOSED CONSTITUTIONAL AMENDMENT H
Wording of proposition	Shall there be an amendment to the State constitution allowing the State legislature to enact laws limiting the amount of damages for noneconomic loss that could be awarded for injury or death caused by a health care provider? "Noneconomic loss" generally includes, but is not limited to, losses such as pain and suffering, inconvenience, mental anguish, loss of capacity for enjoyment of life, loss of consortium, and other losses the claimant is entitled to recover as damages under general law. This amendment will not in any way affect the recovery of damages for ecomonic loss under State law. "Economic loss" generally includes, but is not limited to, monetary losses such as past and future medical expenses, loss of past and future earnings, loss of use of property, costs of repair or replacement, the economic value of domestic services, loss of employment or business opportunities. This amendment will not in any way affect the recovery of any additional damages known under State law as exemplary or punitive damages, which are damages allowed by law to punish a defendant and to deter persons from engaging in similar conduct in the future.

REFERENDUM #4:

Title of proposition	PROPOSED CONSTITUTIONAL AMENDMENT K
Wording of proposition	Shall there be an amendment to the State constitution authorizing Madison and Fromwit Counties to hold referenda on whether to authorize slot machines in existing, licensed parimutuel facilities (thoroughbred and harness racing, greyhound racing, and jai alai) that have conducted live racing or games in that county during each of the last two calendar years before effective date of this amendment? The Legislature may tax slot machine revenues, and any such taxes shall supplement public education funding statewide. Requires implementing legislation. This amendment alone has no fiscal impact on government. If slot machines are authorized in Madison or Fromwit counties, governmental costs associated with additional gambling will increase by an unknown amount and local sales tax-related revenues will be reduced by $5 million to $8 million annually. If the Legislature also chooses to tax slot machine revenues, state tax revenues from Madison and Fromwit counties combined would range from $200 million to $500 million annually.

REFERENDUM #5

Title of proposition	BALLOT MEASURE 101: Open Primaries
Wording of proposition	Requires primary elections where voters may vote for any state or federal candidate regardless of party registration of voter or candidate. The two primary-election candidates receiving most votes for an office, whether they are candidates with no party or members of same or different party, would be listed on general election ballot. Exempts presidential nominations. Fiscal Impact: No significant net fiscal effect on state and local governments.

REFERENDUM #6:

Title of proposition	BALLOT MEASURE 106: Limits on Private Enforcement of Unfair Business Competition Laws
Wording of proposition	Allows individual or class action "unfair business" lawsuits only if actual loss suffered; only government officials may enforce these laws on public's behalf. Fiscal Impact: Unknown state fiscal impact depending on whether the measure increases or decreases court workload and the extent to which diverted funds are replaced. Unknown potential costs to local governments, depending on the extent to which diverted funds are replaced.

APPENDIX F: POST-TEST SATISFACTION

Following is a copy of the post-test satisfaction questionnaire provided to participants to the conclusion of the usability test:

Participant Number ____ Machine Type: ☐ DRE ☐ Optscan

Post Test Questionnaires

Please complete the following questionnaire

1. To the best of my ability, I followed the instructions telling me how to vote.

☐ Yes
☐ No

2.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
	1	2	3	4	5
I was able to cast all of the votes in today's test exactly as instructed.					

3.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
	1	2	3	4	5
I would like to use this voting system in a real election.					

APPENDIX G: INCENTIVE RECEIPT AND ACKNOWLEDGMENT FORM

Following is a copy of the incentive receipt and acknowledgement form used during the usability test.

Acknowledgement of Receipt

"I hereby acknowledge receipt of $[XX] for my participation in a research study run by [test laboratory]."

Printed Name: _____

Address: _____

_____

Signature: _____ Date: _____

Usability Researcher: _____

Signature Of Usability Researcher: _____ Date: _____

Witness: _____

Witness Signature: _____ Date: _____

APPENDIX H: NOTE-TAKING DOCUMENTS

[Insert related note-taking documents]

APPENDIX I: SCORING TOOLS

[List all scoring tools and provide copies of the tools below, or links to tools, such as the end-to-end-benchmark calculation perl scripts.]

APPENDIX J: PHOTOS OF THE TESTING ENVIRONMENT AND VOTING SYSTEM

[Insert photos from the test environment and the calibration system and VSUT.]

APPENDIX K: RESULTS

[Insert raw data from the benchmark calibration Perl scripts, including the individual content scores and the frequency distribution. Also, and any detailed results, listed below, which were not included in the report above:]

- **Calibration system data - current results**
    - Identification (make and model) of the calibration system
    - NPART and NCAST
    - Measured TCS = NCAST / NPART
    - 95% CI for TCS
    - Distribution of raw scores
    - Z-score resulting from Mann-Whitney comparison of current and nominal distributions
- **Effectiveness Metrics for the VSUT**
    - NPART and NCAST
    - Measured TCS = NCAST / NPART
    - 95% CI for TCS
    - Distribution of raw scores
    - NPERFECT
    - Measured PBI = NPERFECT / (NCAST - NPERFECT)
    - 95% CI for PBI
    - Mean and standard deviation for the distribution of scaled scores (PCORRECT-i)
    - Measured VII
    - 95% CI for VII
- **Efficiency Metrics for the VSUT**
    - Mean for distribution of TASKTIME-i
    - Standard deviation for distribution of TASKTIME-i (optional)
- **Satisfaction Metrics for the VSUT**
    - Mean for distribution of confidence responses
    - Mean for distribution of likability responses