AMERICAN INSTITUTES FOR RESEARCH®

NEW YORK STATE VOTER SYSTEM USER RATE ASSESSMENT STUDY

RESEARCH REPORT

DRAFT NOT FOR DISTRIBUTION

Dwayne G. Norris
Christine A. Paulsen

DECEMBER 11, 2006

Delivered to:

Anna E. Svizzero
Director of Election Operations
New York State Board of Elections
40 Steuben Street
Albany, New York 12207

Table of Contents

List of Exhibits

List of Appendices

Background

In March 2006, the Justice Department filed a lawsuit against the State of New York contending that the state had failed to comply with two of the requirements of the Help American Vote Act of 2002 (HAVA). Regarding one of these requirements, the lawsuit alleged that New York had failed to "adopt voting systems that are fully accessible by disabled voters and are capable of generating a permanent paper record that can be manually audited (U.S. Department of Justice, March 1, 2006)." Specifically in question was New York's use of the lever voting system, which was regarded as not being fully accessible to all voters with disabilities.

In order to be compliant with the HAVA requirements, New York State is expecting to replace the lever voting system with voting systems that are both accessible and capable of generating a voter verifiable paper audit trail. In addition to meeting the requirements under HAVA, any new voting systems purchased in the State of New York must also comply with Subtitle V of Title 9 of the Official Compilation of Codes, Rules, and Regulations of the State of New York. Section 6209.6 of the Code requires that:

> Vendors shall make available to the State Board, in a quantity to be determined by the State Board, voting systems for the purpose of conducting a usability test, which will establish (1) the minimum number of voting machines required in each polling place and (2) the maximum number of voters that can vote on one voting machine during the course of an ordinary 15-hour election day.[1]

The New York State Board of Election (NYSBOE) contracted with the American Institutes for Research (AIR)[2] in September 2006 to determine the Maximum Daily Rate (MDR) of voters for each voting system that is currently being considered for use in New York state elections (the second requirement under Section 6209.6). The first requirement, determining the minimum number of voting machines required in each polling place was not part of NYSBOE's charge to AIR—the primary goal of this study was to provide data on maximum daily rates in order to help NYSBOE in determining the minimum number of voting systems required.

[1] While the Code refers to this type of study as a usability study, this is a misnomer. In a typical usability study, a sample of users from the target population would attempt to vote using a system or set of systems with the purpose of determining ways that the interface may interfere or facilitate voters' ability to cast votes (Federal Election Commission, 2003; Dumas & Redish, 1999). Usability studies are not generally designed to compute outcomes like the maximum number of voters that can vote on a system in a specified period of time.

[2] The American Institutes for Research (AIR) is a not-for-profit organization that specializes in applied research and consulting in the behavioral and social sciences. AIR is headquartered in Washington, DC. More information about AIR can be found at www.air.org.

Study Purpose

For this study, the MDR is defined as the maximum number of voters a given voting system can accommodate in a 15-hour voting period at any registered polling place *all other things being equal*. By this, we mean that we have made no assumptions about other factors that will influence decisions about the minimum number of voting systems required at each polling place. There are many factors that determine how many voters can vote at a polling place during an election day and how many voting systems are required to accommodate them. Some of these factors include:

- Number of registered and active voters assigned to a given polling location;
- Length of the ballot;
- Size of the facility (whether it can accommodate large numbers of voters);
- Availability of trained poll workers and the extent of their training, including translators for voters with limited English proficiency;
- Availability of all materials needed for voting (e.g., ballots, pens, etc.);
- Variability in the number of voters who arrive at the polling place during different times of the day and the length of the resulting lines;
- Reliability and efficiency of the polling place procedures (e.g., checking voters in and out);
- Interpersonal differences among voters—for example, some voters will wait for long periods of time and some will choose not to wait in line, depending on many different conditions (e.g., weather, type of election, perceived importance of vote);
- Usability and accessibility of the voting systems; and
- Frequency of voting system failures and technical problems.

It was beyond the scope of work for this study to collect empirical data on these various factors. We strongly recommend that the NYSBOE take all such factors into consideration in making determinations about the type and number of voting systems to purchase in each district.

Voting Systems Included in Study

New York is considering several different types of voting systems to replace its lever system. These systems fall into two basic types: direct recording electronic systems (DREs) and optical scanning systems.

According to their website (http://www.fec.gov/pages/electpg.htm), the Federal Election Commission defines these as follows:

> **Direct recording electronic, or DREs,** are an electronic implementation of the old mechanical lever systems. As with the lever machines, there is no ballot; the possible choices are visible to the voter on the front of the machine. The voter directly enters choices into electronic storage with the use of a touch-screen, push-buttons, or similar device. An alphabetic keyboard is often provided with the entry device to allow for the possibility of write-in votes. The voter's choices are stored in these machines via a memory cartridge, diskette or smart-card and added to the choices of all other voters. (DREs can take several forms, including full-face paper ballot overlay, full-face touch screen, and paging touch screen.)
>
> **Optical scanning systems** employ a ballot card on which candidates and issue choices are preprinted next to an empty rectangle, circle, oval, or an incomplete arrow. Voters record their choices by filling in the rectangle, circle or oval, or by completing the arrow. After voting, the voters either place the ballot in a sealed box or feed it into a computer tabulating device at the precinct. The tabulating device reads the votes using "dark mark logic," whereby the computer selects the darkest mark within a given set as the correct choice or vote.

This study included several different types of voting systems from several different manufacturers:

Full-face Ballot Overlay DRE

- Liberty

Full-face Touch Screen DRE

- Avante
- Sequoia

Optical Scan

- Diebold

- ES&S
- Sequoia

In addition, NYSBOE asked us to include the lever machine in the study to provide some benchmarking data.

Each of the systems (except the lever and the Sequoia optical scan) were equipped with special features or, in the case of the optical scanners, were connected to systems (Automark) to enable them to be accessed by voters with disabilities. These features included audio, Braille keypads, "sip and puff" capability, large text display, high contrast displays, and language translations.

Study Design

This research study was comprised of five steps. Exhibit 1 below provides an overview of each of these steps and our objectives in carrying them out. Following Exhibit 1 is a description of the specific activities that we conducted during each research step.

Exhibit 1. Overview of Research Steps

Research Step	Objective
Review Voting Systems	To become familiar with the characteristics of each voting system, identify potential usability issues that may affect MDR calculations, and provide input to the development of study protocols and materials
Determine Population of Interest	To specify the composition of New York State eligible voters using U.S. Census data and guidance from NYSBOE staff
Prepare for Data Collection Sessions	To fully specify the procedures, instructions, locations/schedule, and desired composition of study participants, and to test the study procedures
Conduct Data Collection Sessions	To gather voting time estimates and other evaluative information to use in calculating a MDR for each voting system
Analyze Results	To compile all the information gathered in the study and use it to calculate a MDR for each voting system

Step 1: Review Voting Systems

The first step in this study was to become familiar with the voting systems for which MDR estimates were required. To accomplish this, the NYSBOE attempted to provide our research team with all available background information on each voting system. While we requested information such as vendor related descriptions, results from prior evaluations (formal or informal), and any vendor supported research, websites, or other informational sources, the information that

AIR received was limited and did not include all the voting systems under consideration. It was our understanding that the NYSBOE experienced difficulty in obtaining such information from the vendors. In fact, we did not know which voting systems were to be included in the study until very late in the process (after we conducted our in-person review of the few systems the NYSBOE did have in its possession).

During our in-person review, the following systems were available: Liberty DRE, Diebold DRE (withdrawn from study), Sequoia Optical Scan, and Sequoia DRE. In addition, an Automark system was available. None of these systems were programmed at the time and thus we could not conduct a usability assessment of them. Therefore, we performed our own limited research on those voting systems known to be included in the study at the time to help us refine our study design, protocols, and data recording tools. This research involved an internet search of reports, descriptions, news articles, and other background information on the voting systems. Thus, although the aggressive schedule for this study did not permit sufficient time for AIR to perform an intensive in-person review of the voting systems prior to the start of data collection, we were able to get basic background information on many of the voting systems included in the study.

Step 2: Determine Population of Interest

The second step in this study was to work with NYSBOE to fully specify the population of interest. This was a critical step because the calculation of an appropriate MDR must reflect the types of voters (in terms of their demographic profile) that will actually use the voting systems. Otherwise, the MDR may be viewed as an inappropriate estimate.

We collected and reviewed the New York Census data and discussed the most appropriate composition of the study sample. Based on these data, the NYSBOE was provided with guidance on proportions of voting-eligible individuals who should be included in the sample, based on key demographic variables such as sex, age, race and ethnicity, English language status, and disability.

Step 3: Prepare for Data Collection Sessions

The third step in the study was to prepare for the data collection sessions. This step involved resolving several critical study design issues, determining the locations and schedule for conducting the data collection sessions, and pilot testing the study processes.

Critical Study Design Issues

The first activity of this step of the study was to resolve some critical study design issues. A primary concern of this study was practice effects. That is, having a sample of voters vote on all seven systems could have been problematic in that the estimates of how long it takes a participant to vote on one system may be influenced by them having voted on another similar system. To the extent that this occurred, it would result in MDR estimates that do not reflect what will occur during actual voting (i.e., the MDR estimates may underestimate the time required to vote if practice effects occur). To minimize practice effects, we randomized the order in which individuals voted on the seven systems during the study sessions. Thus, there was no consistent pattern to the order in which study participants voted on each of the voting systems. Through randomization, practice effects, if they occurred at all, would be randomly distributed and thus would not bias the study results towards or against any given voting system.

Another concern of this study was determining how participants would vote during the data collection. The goal of any study design is to try to replicate the real world context as much as possible, while still be able to unambiguously analyze the phenomenon under study. Because this was not a real election, we felt it was not feasible to replicate the full range of voting choice behavior that occurs in an actual election (e.g., some vote for all offices, while others do not; some predetermine who they will vote for, while others decide when voting). After discussions with NYSBOE, it was decided that the general election ballot would serve as the template for this study and that we would provide fictitious names of candidates that study participants were to vote for in each of the offices that appear on New York's general election ballot.[3] Therefore, various different sample ballots were constructed by NYSBOE to represent some of the more "typical" voting patterns (e.g., straight Democratic ticket, straight Republican ticket, conservative ticket, liberal ticket) and participants were required to vote for the names that we provided on the sample ballot.[4]

In addition, because the number of propositions varies on any given election, it was decided to have participants vote for *only* 1 of 10 propositions on each of the seven different sample ballots prepared for the study. Unlike the offices, however, participants were instructed to read the proposition and vote as they thought appropriate. Furthermore, the propositions listed on each sample ballot were selected to be of approximately the same length in an attempt to standardize the

[3] The general election ballot was chosen as the template for this study in large part because it was regarded as the most comprehensive ballot used in New York elections.

[4] These sample ballots did not look like the actual ballots. Instead, the sample ballots simply listed the offices on the general election ballot along with fictitious names of candidates that study participants were to vote for in each office. The offices were listed on the sample ballot in the same order in which they appeared on the actual ballots used at each voting system.

time required to read it. We did not tell participants how to vote for propositions because we felt this would not adequately reflect the reading (and possibly decision time) required to respond to propositions. Only one proposition was placed on each sample ballot because we felt that study participants were unlikely to reread propositions when encountered on subsequent sample ballots throughout the data collection session.

The objective of constraining the voting choice behavior in this manner was to provide a consistent basis for interpreting the study results. That is, each participant voted for the same number of offices and propositions and thus the time it took participants to vote sample ballots on each voting system could be easily compared because the sample ballots were all the same length. Had we not done this, it would be impossible to interpret the resulting data because it would reflect the individual (and unknown) whims of the study participants (e.g., some might choose to vote all offices, while other may not). Appendix 1 provides an example of a sample ballot provided to study participants.

Study Locations

The second activity in this step of the study was to determine the locations and schedule for the data collection sessions. The most important consideration was access to individuals that fit our desired sample composition on key demographic characteristics. It was also important to secure space that allowed for seven voting systems to be located, for our research staff to observe and debrief participants, for participants to check-in and wait for the study to begin, and for easy access by disabled participants. AIR relied on NYSBOE to identify appropriate data collection locations.

Pilot Testing the Study Procedures

The third activity in this step of the study was to conduct a brief pilot test of the data collection sessions. This pilot test occurred prior to conducting the data collection sessions and served as a test of our procedures, data collection instruments, timing procedures, instructions, and other procedural requirements needed to conduct the data collection sessions. A total of 20 individuals participated in the pilot session which was held in Schenectady. These participants were then debriefed to get their reactions to the study procedures. Data from the pilot test were used to adjust our plan, procedures, instructions, and data collection materials prior to conducting the first data collection session in Rochester.

Step 4: Conduct Data Collection Sessions

Data collection sessions were held for two days in each of three locations: Rochester, Brooklyn, and Albany. At each location, participants were processed hourly in groups of approximately 14. Data collection ran from 8:00 am through 8:00 pm each day at each data collection location. The basic structure of each session is shown in Exhibit 2:

Exhibit 2. Structure of Each Data Collection Session

Time Allotted	Process Step	Activities
5 minutes	Staging	• Greeted participants and provided study overview • Participants read and signed consent forms • Participants assigned an ID number • Participants completed a demographic survey • Participants given specific voting instructions • Participants directed to the enter voting area
35 minutes	Voting	• Participants approached their first assigned voting station • Participants checked in with researcher at that voting station • Participants given a brief introduction to the system • Participants placed their vote • Participants completed voting reaction survey for that voting system • Participants moved to the next assigned voting station and repeat the process • Participants departed voting area and entered exit area
5 minutes	Exit	• Participants given debrief form • Participants asked if they have questions • Participants given cash for participation • Participants asked to sign a form indicating they received payment. Participants given carbon copy.

Pre-Session Planning

Prior to each session, a two-hour training session was provided to all staff involved in the data collection. Specifically, training was provided to inform all staff about the objectives of the study, review the protocol for data collection, including a review of the recording forms to be used to gather time information, and to provide staff with training on the appropriate ways to interact with study participants, including those with disabilities. Appendix 2 shows the agenda used in the two training sessions that occurred. As part of training, individuals that participated in the pilot test session were able to share lessons learned with other staff.

Conducting the Data Collection Sessions

During each data collection session, participants were first given an informed consent form (see Appendix 3) to read and sign prior to beginning the study. Once individuals completed the informed consent form, they were given additional background information about the goals of the study and asked to complete a participant background survey (see Appendix 4). This survey was used to gather demographic (e.g., gender, age) and background (e.g., voting frequency) information from all study participants.

Study materials were then given to participants and they received instructions on how to "vote" during the data collection session. Study materials included an instruction sheet that provided the order for voting on each voting system, as well as a way for us to track the voting systems each participant voted on (see Appendix 5 for a sample instruction sheet), and seven sample ballots (see Appendix 1 for a sample ballot). Both the voting system order and the sample ballots were randomly determined so that participants would not be using the same sample ballot on any voting system. With the study materials in hand, participants received oral instructions on how to proceed through the data collection session.

To begin the voting process, participants were instructed to go to the first voting system listed on their instruction sheet. Once at the voting system they were checked in by the research staff and given brief instructions on how to use the voting system. They then were to vote on the system using the sample ballots that were provided to them. That is, participants were asked to vote for the offices and fictitious candidates shown on their sample ballot. Once participants indicated that they understood the process they were allowed to begin. Research staff timed each participant on how long it took them to cast their vote. Staff also recorded any issues that arose during the voting process. When participants were done voting on a given voting system, they were given a seven item reaction survey to complete about their voting experience on that particular machine (see Appendix 6 for a copy of the reaction survey). They were also instructed that upon completing the survey, they were to go to the next voting system listed on their instruction sheet and vote the next sample ballot included in their packet of study materials. Participants proceeded in this fashion until the data collection session ended.[5]

As described below, specific criteria were used to record the time it took each participant to vote on each system.

[5] Although the voting machine order was randomly established prior to the data collection session, we often did instruct participants to skip a voting system on their list if the line to vote at that system was too long. Given the voting system order was random, we do not believe that moving folks around to open voting systems was problematic because participants were sent to any available voting system and then asked to proceed according to their instructions sheet if possible.

Recording Time Data for Optical Scan Systems with Paper Ballot

After providing instructions and answering participant questions, but before the participants were directed to the privacy booth,[6] AIR researchers asked participants if they were ready to begin. Once the participant said "yes," or otherwise indicated readiness, the researcher directed the participant to the privacy hut and said, "OK, please begin now." We started timing when the participant entered or arrived at the privacy hut. Then, we stopped timing when the participant stepped away from the privacy hut. This was recorded as Time 1 (time to complete paper ballot).

Next, we started timing again when the participant placed their ballot on the optical scan reader (and after any tearing of the ballot occurred, which was required for one type of ballot). This was the starting point for all optical scans regardless of whether someone had to push a button to feed the ballot into the machine.

All paper ballots were "kicked back" to the participants because all were purposely undervoted. This undervoting occurred because participants were instructed to vote for only 1 of 10 propositions listed on the ballot. When the optical scan reader kicked the ballot back, our researcher said "The system has kicked the ballot back because of an undervote, which we asked you to do. Would you like to review the ballot to make sure that you have completed the ballot as instructed? When you are satisfied that it is voted the way you intended to vote, you can place it back in." We stopped timing when the participants' completed and finalized ballots were confirmed by the optical scan reader or when the confirmed finalized ballot was accepted. We continued timing during any actions the participant took to resolve overvotes or undervotes, or otherwise review their ballot.

Recording Time Data for Voting Systems with Touch or Electronic Screens

After participant questions were answered, we asked the participant if he or she was ready to begin. Once the participant said "yes," or otherwise indicated readiness, the researcher said, "OK, please begin now."

If the participant had to take some behavioral action to initiate the process (e.g., push or touch button, activate joy stick, etc.) we started timing when he or she took that action. If there was no such action required and the ballot was open and ready for voting, we started timing immediately after telling the participant to begin. We stopped timing when the participant had correctly cast their ballot.

[6] Participants were given an option to vote in privacy booths similar to those used in actual elections or to simply sit in an open seating area.

In cases where participants attempted to leave the machine before the final ballot was submitted/confirmed, we reminded the participant that they had not finished voting. We did not tell them how to finish, but rather allowed the participants to figure out what must be done to finish.

Post-Session Activities

When participants had completed the voting process, they were given a short written debriefing (see Appendix 7), had their questions answered, and were thanked for their participation. Each participant was paid $30 for their participation and then dismissed from the data collection session.

Following each 45 minute session, the used study materials were stored away and the room was prepared for the next arriving group of participants. We also used the time between each 45 minute session to discuss any issues that arose, particularly those that would potentially affect the quality of timing participant's votes or that would affect the types of information we needed to record on the recording forms.

By using this strategy, we were able to make real-time adjustments to our procedures as opposed to waiting until the end of the full day or between data collections at different locations. As an example, it was clear after the initial session that some participants were a little confused about the requirement to vote for only the proposition listed on their sample ballot and not all 10 propositions listed on the actual ballot at each voting session. That is, our plans called for voters to intentionally undervote. Given this feedback, we revised the initial instructions provided to participants before they started the voting process to emphasize this requirement.

Step 5: Analyze Mock Election Results

The fifth step in this study was to compile the results from prior steps and analyze the data. There were four major sources of information stemming from the data collection sessions. The first type of information was the demographic profile of participants in the study. The second type of information was the time to vote (per system) for each study participant. The third type of information was the notes on our observations and issues that arose and may have a bearing on the calculation of the MDR. The final type of information was the survey results.

As Exhibit 3 shows, each of these sources of information was used to address different aspects of the research goals.

Exhibit 3. Sources of Information Gathered During Data Collection

Type of Information	Intended Use
Demographic data	To develop a profile of study participants on key demographic and background characteristics
Voting time data	To estimate the time it takes an individual to vote on each system and to estimate the MDR for each system
Data collection notes	To track any issues that arose during data collection for use in evaluating data quality
Reaction survey data	To gather participant opinions about aspects of the instructions, study, and voting systems

These data are presented in the Study Results section. Here, we briefly discussed how the data was evaluated and analyzed.

Data Quality

Prior to analyzing the data, it was entered into a database and evaluated for completeness and accuracy. This process involved double checking the data for data entry errors, evaluating hard copies of data to reconcile any anomalous values (where necessary), and checking that all background, time, and reaction survey values were within the expected range. For time data, we manually checked the recording forms whenever extreme times were found to ensure that they were accurately captured in the database.

As part of evaluating the data quality, it was necessary to review all the comments gathered from the research staff and make a determination about whether or not a particular participant's data were appropriate for us in calculating voting times and MDR estimates. The most common issue that surfaced in this review was that participants did not finish voting on a given voting system when time was called to end the data collection session. Because these participants did not finish, we could not use their time results in the analyses for the voting system on which they were stopped prior to casting their vote.

In addition to participants not having enough time to cast a vote on some systems, other issues precluded the use of some data in the analyses. These issues included participants failing to follow the directions, voting system breakdowns (e.g., paper jams in the optical scan reader), and individuals stopping the voting process on their own and not casting the ballot. With all these cases, we individually reviewed the comments on the recording forms and then made a decision on whether it was appropriate to include or exclude that data in the analyses. Furthermore, this was done on a system by system basis. Thus, it was possible that a participant's data was not appropriate for inclusion in calculations of voting

time for one voting system, whereas their data was appropriate for inclusion in calculations of voting time for the voting systems on which they successfully voted on during the study.

Data Analysis

Once the data were evaluated, we calculated various descriptive statistics to summarize the study findings. As the basis of estimating the time it takes an individual to vote on each machine and the MDR, we calculated the average/mean, trimmed mean, and median voting times. We also calculated the standard deviation of voting times on each voting system. The mean, trimmed mean, and median represent slightly different ways of estimating voting time, whereas the standard deviation provides an estimate of the variability in voting times. Below, we briefly discuss the features of each statistic we calculated, along with special considerations that should be exercised when interpreting the voting time estimates.

The first statistic calculated was the *average*, or *mean*, time to vote. While averages are readily understood, it is important to note that they can be dramatically affected by extreme values (i.e., outliers). In the present study, this generally occurred when one or more individuals required an extremely long time to vote in comparison to most other individuals voting on the same voting system. When this occurred, the mean voting time estimate may be higher than it would be if that individual's voting time was omitted from the calculation. This was a particular concern given the mix of individuals with and without disabilities, because those with certain disabilities that require the use of accessibility features on voting systems tended to take considerably more time to vote on a voting system then individuals without disabilities that did not require the use of accessibility features to vote.

The second statistic we calculated was the *trimmed mean*. The trimmed mean is calculated in precisely the same manner as the mean except that the extreme highest and lowest five percent of voting times are removed from the calculation. The trimmed mean is an attempt to eliminate the effects of extreme voting times—whether those times fall at the high or low end of the voting time range. Even with the trimmed mean, extreme times that are not removed can still have an influence over the final estimate.

The third statistic that we calculated was the *median*. The median represents the voting time at the middle of the distribution of all voting times. Thus, if we were to order the voting times for a given voting system from low to high, the median would represent the voting time in which 50 percent of the participants had lower voting times and 50 percent of the participants had higher voting times. Unlike the mean, the median estimate is less susceptible to extreme voting times.

The fourth statistic that we calculated was the *standard deviation* (SD) of voting times. The standard deviation provides a barometer for gauging the range of voting times around the mean. The larger the standard deviation relative to the range of times, the less representative the mean estimate is for the group on which it is calculated. For example, if five participants voting on a given system took 10 minutes each and another group of five participants took two minutes each, the mean voting time would be six minutes and the standard deviation would be 4.2. This standard deviation would be considered large and the thus the mean estimate of six minutes is less representative of how long it took any of the 10 participants in this example to vote. Now suppose that the first group of 5 participants each took 7 minutes to vote and the second group of 5 participants each took 5 minutes to vote. Here, the mean voting time would still be 6 minutes; however, the standard deviation would be 1.1. This smaller standard deviation reflects less variability around the mean estimate and thus the mean estimate is regarded as a better estimate of the voting time for these 10 participants.

The average/mean, trimmed mean, median, and standard deviation represent the basic statistics used to get an estimate of voting time. However, estimating the time to vote required more consideration than just the appropriate statistics to calculate. In addition, we also looked at these estimates of voting time for individuals that did and did not report having disabilities, and for individuals that did and did not report using accessibility features on the voting system. The voting estimates are reported for these two classifications of study participants for several reasons. First, it is a reasonable assumption that those who have a particular disability and/or who require the use of accessibility features to vote will take a longer time to vote on each voting system. While we do not argue that individuals with disabilities or individuals that require accessibility features of the voting system to vote should be excluded in calculations of voting times, we do believe it is informative to look at the calculations of voting times with and without individuals that fit this description. Second, as shown in the descriptions of study participants, the number of individuals that reported disabilities or reported using accessibility features varied somewhat across voting systems. This is somewhat expected because not all the voting systems were equipped with all possible accessibility features and thus individuals requiring special assistance could not vote on all the voting systems.

To calculate the MDR, we took each of the estimates of individual voting time discussed above (i.e., mean, trimmed mean, and median) and entered them into the following formula:

$$MDR = 54{,}000/IVT,$$

where 54,000 represents the total number of seconds in 15 hours (e.g., 60 seconds/minute X 60 minutes/hour X 15 hours/election) and IVT represents the estimate of how long it takes an individual to vote in seconds.

For the reaction survey results, we calculated the frequency of responses to the various options for each item.

Special Analysis Considerations for Optical Scan Voting Systems

There were three optical scan voting systems included in this study (i.e., Diebold, Sequoia, and ES&S). For each of these systems, two times were recorded for each participant that voted on them. The first recorded time indicated how long it took a participant to fill out the paper ballot. The second recorded time indicated how long it took a participant to correctly cast their ballot by feeding their completed paper ballot through the optical scan reader.

This feature of the optical scan voting systems brings up the issue of what is the appropriate time to report for the purposes of this study. However, because we separately recorded the times to complete the entire voting process on an optical scan voting system, we will present all the times. That is, we provide the voting time calculations for the optical scan voting systems in three ways—the time it takes to fill out the paper ballot, the time it takes to feed the paper ballot into the optical scan reader, and the combination of these two times. The combined time is calculated as the sum of the two component times (filling out the ballot and feeding the ballot into the optical scan reader) only and does not include any time that may occur between completing the ballot and casting the vote on the optical scan reader.

Study Participants

As noted in the study design section, participants were asked to complete a background survey as part of this study. Given the study design relied on a volunteer sample of participants, this survey was used to monitor the composition of study participants throughout the data collection phase of the study and then to described the composition of study participants. In this section, we present this description. Exhibit 4 contains a summary of the overall composition of study participants. Appendix 8 contains a summary of the composition of study participants that voted on each of the seven voting systems represented in this study. In all cases, the descriptions are based on the self-reported responses to the background survey. There was no attempt to verify the accuracy of these responses; however, given the anonymous collection of demographic and background information, there is no reason to expect that participants would not respond accurately to our background survey questions.

Across the three data collection sites, a total of 812 individuals participated in this study. The sample sizes for each of the three data collection sites were:

- Rochester = 286 participants
- New York City = 259 participants
- Albany = 267 participants

Of the 812 individuals in the study, 795 (99 percent) were residents of New York state. Most (89.7 percent) reported they were registered to vote in New York, 2.8 percent reported they were registered elsewhere, and 5.8 percent reported they were not registered to vote at all. We asked participants to indicate the frequency with which they vote. Most (69.7 percent) reported that they voted frequently, 14.5 percent reported that they voted occasionally, 7.4 percent reported that they rarely voted, and 7.1 percent reported that they had never voted before.

Exhibit 4. Comparison of Overall Study Sample to Target Population Characteristics

	N	Study Sample	Census
Gender	807		
Female		58.3%	52.6%
Male		41.1%	47.4%
Age	809		
18-24 years		18.2%	13.7%
25-44 years		28.2%	37.0%
45-64 years		38.7%	32.5%
65-74 years		11.1%	8.8%
75 years or older		3.4%	8.0%
Race / ethnicity	807		
American Indian / Alaska Native		1.6%	0.7%
Asian		3.8%	7.0%
Black / African American		33.8%	16.7%
Hispanic		6.2%	13.6%
Native Hawaiian / Pacific Islander		0.8%	0.1%
White		51.9%	65.0%
Other		4.2%	9.5%

Exhibit 4. Comparison of Overall Study Sample to Target Population Characteristics (Continued)

	N	Study Sample	Census
Education (Highest level completed)	807		
Some high school, no diploma		7.0%	15.7%
High school diploma or GED		22.0%	29.2%
Vocational /Trade/Business school		5.3%	NR
Some college, no degree		14.5%	15.6%
Associate's degree		10.1%	8.2%
Bachelor's degree		23.8%	17.9%
Graduate or professional degree (total)		13.9%	13.4%
Masters degree		10.3%	NR
Doctoral degree		1.4%	NR
Professional degree (JD, MD)		2.2%	NR
Other		2.7%	NR
First language	800		
English		93.0%	71.8%
Non-English (total)		6.1%	28.2%
Spanish		2.7%	NR
Korean		1.0%	NR
Mandarin		0.1%	NR
Cantonese		0.5%	NR
Other		1.8%	NR
Disability	782		
None		72.5%	76.9%
Disabled		28.7%	23.1%
Blind / visually impaired		3.8%	3.0%
Deafness / Hard of hearing		4.2%	11.0%
Mobility impairment / physical disability		10.3%	20.0%
Mental or cognitive disability		8.4%	6.0%
Other		2.0%	NR

Note: N = the number of participants who reported data for this variable. Percentages reported for race/ethnicity and disability may add up to more than 100% because some participants reported identifying with more than 1 category.
Sources: Sex, race/ethnicity, age: U.S. Census Bureau, Current Populations Survey, Nov. 2004 & 2005 Community Survey (Race only); Education, language: U.S. Census Bureau, 2005 American Community Survey; Disability: U.S. Census Bureau, 2005 American Community Survey; U.S. Census Bureau, Census 2000, Communication with several advocacy groups, including the Commission on Quality Care and Advocacy for Persons with Disabilities.
NR = Not reported.
*The existing Census data on the number and percent of adults who identify themselves as having disabilities is sometimes conflicting and of questional reliability (Andresen, E.M. and Fitch, C.A., 2000)

As Exhibit 4 shows, the study sample was fairly comparable to the actual voting eligible population in terms of the sex, age, and educational level of the adults in our sample.

The sample contained only slightly more females than the voting-eligible population in New York (58.3 percent versus 52.6 percent). In addition, the study sample contained somewhat fewer adults in the 25-44 year old and the "over 75" year old categories than the actual population in New York, while the study sample contained a slight overrepresentation of adults in the 18-24 year old, the 45-64 year old, and the 65-74 year old categories.

The study sample contained fewer adults who had not completed high school than expected (7.0 percent in the sample, versus 15.7 percent in the population).

Overall, discrepancies in participants' reported sex, age, and educational level were not dramatically different from the State population of adults in New York.

With respect to race and ethnicity, our sample contained a much greater proportion of Black/African American participants than expected, based on the Census data (33.8 percent versus 16.7 percent). Conversely, our sample contained fewer participants who reported they were Hispanic or White than expected.

The study sample also contained more native English speakers than one might expect by comparison to the Census data. However, it is important to note that vendors were not expected to provide ballots and other materials in every possible language. So, it was by design that our study contained a less representative sample of non-native English speakers. That said, we did encourage the NYSBOE to recruit participants who spoke as their first language Spanish, Mandarin, Cantonese, or Korean because these are the non-English languages that New York is required to support with its voting systems in certain locations. However, our sample included very few participants who spoke these languages.

With respect to disability, it was challenging to develop sample targets because the existing data on disability rates in the U.S. are conflicting and vary widely depending on factors like age. We consulted with experts and advocates from the State of New York and determined that the most effective approach would be to oversample from the population of individuals with disabilities in the hopes of attaining a good sized sample. With respect to participants who reported they did not have a disability, our numbers matched the expected numbers fairly closely (72.5 percent in the sample versus 76.9 percent in the population). Conversely, the proportion of study participants who identified themselves as disabled was 28.7 percent versus 23.1 percent in the population. With respect to type of disability, our sample underrepresented individuals who reported they were deaf/hard of

hearing or physically disabled and slightly overrepresented individuals with mental or cognitive disabilities.

Note that not all participants at a given location voted on all seven of the voting systems that were in the study. Appendix 8 summarizes the demographic characteristics of the participants that voted on the different voting systems. In general, the demographic composition of participants that voted on each voting system mirrored the demographic composition of the overall sample study participants.

Study Results

The study results centered on four central questions:

1. Were the voting times of study participants appropriately captured?
2. How long does it take to vote on each voting system?
3. What was the maximum daily rate (MDR) for each voting system?
4. What were the general reactions of study participants about both the study and each voting system on which they voted during the study?

Below we described the results as they pertain to these four research questions.

Were the Voting Times Appropriately Captured?

During each of the three data collection sessions, we conducted reliability checks by having a random sample of study participants independently timed by two staff members as their cast their vote. One person was the AIR staff member assigned to monitor a given voting system throughout the session; the other person was a staff person who had other non-timing responsibilities during the session. The second timer was instructed to randomly time a given participant without the knowledge of the primary timer. This was done to avoid influencing the timing behavior of the primary timer as a result of their knowledge that a second timing was occurring. These independent times were done on each voting system and at different points throughout each data collection session. We analyzed the two sets of timing data to assess the reliability (i.e., consistency) in timing between two different timers.

In the Rochester session, reliability checks occurred on four different occasions for each voting system. We evaluated these timing data by calculating the correlation between the two sets of times gathered across all the voting systems. The resulting correlation coefficient of 0.99 indicates that there was a high degree of consistency between the sets of times gathered by the primary and secondary timer. The

difference in times gathered by the primary and secondary timer ranged from a low of zero seconds to a high of eight seconds, with the average difference being three seconds.

In the New York City session, reliability checks occurred on six different occasions for each voting system. The resulting correlation between the times of the primary and secondary timer in New York City was also 0.99, with the difference between timers ranging from a low of zero seconds to a high of 30 seconds. The average difference in times for the New York City data collection session was three seconds.

In the Albany session, reliability checks occurred on three different occasions for each voting system. The resulting correlation between the times of the primary and secondary timer in Albany was 0.99, with differences between the timers ranging from a low of zero seconds to a high of 9 seconds. The average difference in times for Albany was two seconds.

Given the high correlations and the fact that when differences occurred they rarely varied by more than a few seconds demonstrates that the timing procedures as described above were being consistently followed by research staff assigned to each voting system. The few seconds difference that occasionally occurred during these reliability checks likely reflect the requirement for the independent timer to remain unobtrusive; it was not always possible for the secondary timer to hear the instructions to begin the voting process and still remain unobtrusive to the primary timer.

How Long Does It Take to Vote on Each Voting System?

Voting Time Overall Results

The second question to be answered by this study was "how long does it take to vote on each voting system?" As noted above, to address this question we calculated the mean, trimmed mean, and median times that it took to vote on each voting system. Exhibit 5 presents the voting estimates across all participants in the study.

Exhibit 5. Estimates of Voting Time Based on All Participants

Voting System Name	N	Mean	SD	Trimmed Mean	Median
Avante DRE	363	04:08	02:11	03:53	03:39
Diebold Op Scan (total)	566	04:42	02:38	04:27	04:09
Diebold Op Scan (ballot marking)	566	04:09	02:25	03:56	03:43
Diebold Op Scan (ballot scanning)	566	00:34	01:26	00:23	00:21
ES&S Op Scan (total)	429	04:14	02:39	03:54	03:32
ES&S Op Scan (ballot marking)	429	03:46	02:37	03:26	03:06
ES&S Op Scan (ballot scanning)	429	00:28	00:29	00:24	00:21
Lever Machine	594	02:50	01:42	02:40	02:25
Liberty DRE	600	03:03	01:51	02:51	02:38
Sequoia DRE	463	04:20	02:41	04:03	03:42
Sequoia Op Scan (total)	609	04:05	02:57	03:47	03:30
Sequoia Op Scan (ballot marking)	609	03:27	01:50	03:16	03:03
Sequoia Op Scan (ballot scanning)	609	00:32	00:54	00:25	00:22

Note: N = the number of participants. SD = standard deviation. All estimates in minutes:seconds (e.g., 01:14) format.

As shown in Exhibit 5, the mean estimates for voting time where all fairly comparable for the optical scan systems. The one exception was that it appeared to take much more time to mark the Diebold optical scan ballot than the others. This pattern was consistent across the trimmed mean and median estimates of voter time.

For the DRE systems, the mean estimates for voting time ranged from 3:03 (Liberty) to 4:20 (Sequoia). The general pattern observed with the mean estimates remained consistent across the trimmed mean and median estimates of voter time on the DRE systems.

There are several features of these results that bear special attention. First, the standard deviations (see Exhibit 5) were they were moderate to large relative to the mean and reveal that there was considerable variability in how long it took to vote on all the voting systems. Second, as expected, the time estimates based on the mean are the largest and those based on the median are the smallest. The decrease in estimates of individual voting time as you go from the mean, to the trimmed mean, to the median reflects the effect that extreme times can have on a single estimate meant to reflect the "typical" time any given voter would take to vote. Fourth, the total time to cast a vote on the optical scan systems (i.e., marking ballot and scanning ballot) were comparable to the estimates of vote time on the DRE systems. Finally, in comparison to the vote time for the DRE systems or the total time for the optical scan systems, participants were able to vote faster on the

lever voting system. This finding is likely due, in part, to the familiarity of study participants with the lever voting system.

Voting Time Results for Individuals With and Without Reported Disabilities

Exhibit 6 shows the results for participants who did NOT report having any disabilities on the background questionnaire. In inspecting Exhibit 6, several observations stand out. First, the mean, trimmed mean, and median estimates generally were smaller, indicating that the group of participants without reported disabilities took less time to vote on each voting system. However, this trend did not hold up for the optical scan feeding times. In fact, for 6 of the 9 estimates for optical scan voting systems (i.e., mean, trimmed mean, and median for each of three voting systems), there was no change between times for all participants and those for participants who did not report disabilities. Second, although the standard deviations were still large, they did drop in all but two cases (ES&S and Sequoia ballot scanning times), indicating that generally those without self-identified disabilities were less variable in their voting times in comparison to the entire group of participants.

Exhibit 6. Estimates of Voting Time Based on Participants without Self-Identified Disabilities

Voting System Name	N	Mean	SD	Trimmed Mean	Median
Avante DRE	273	03:52	01:46	03:44	03:33
Diebold Op Scan (total)	432	04:20	01:50	04:13	04:00
Diebold Op Scan (ballot marking)	432	03:55	01:45	03:48	03:35
Diebold Op Scan (ballot scanning)	432	00:25	00:26	00:23	00:20
ES&S Op Scan (total)	344	03:54	01:58	03:41	03:25
ES&S Op Scan (ballot marking)	344	03:26	01:53	03:14	03:00
ES&S Op Scan (ballot scanning)	344	00:28	00:31	00:24	00:21
Lever Machine	453	02:45	01:38	02:34	02:20
Liberty DRE	462	02:54	01:36	02:44	02:32
Sequoia DRE	352	03:58	02:22	03:42	03:30
Sequoia Op Scan (total)	465	03:43	01:44	03:33	03:30
Sequoia Op Scan (ballot marking)	465	03:12	01:30	03:03	02:55
Sequoia Op Scan (ballot scanning)	465	00:32	00:59	00:24	00:22

Note: N = the number of participants. SD = standard deviation. All estimates in minutes:seconds (e.g., 01:14) format.

Exhibit 7 shows the results for participants who reported having a disability on the background questionnaire.[7] Inspection of Exhibit 7 shows that for the electronic

[7] All types of disabilities were included in this analysis. No attempt was made to evaluate these results by type of disability.

voting systems and for the lever voting system, participants with reported disabilities took longer to vote and were more variable in voting times in comparison to all participants or just participants with no reported disabilities. This was also true when considering the time it took to fill out paper ballots or the total time (ballot marking plus optical scan reading) for the optical scan voting systems. For the ES&S optical scan and Sequoia optical scan voting systems, the time to feed their ballots into the reader did not vary much between participants with and without disabilities.[8] There was greater variability between the participants with disabilities and those without disabilities on their time to feed ballots into the Diebold optical scan reader—participants with disabilities took longer.

Exhibit 7. Estimates of Voting Time Based on Participants with Self-Identified Disabilities

Voting System Name	N	Mean	SD	Trimmed Mean	Median
Avante DRE	80	05:08	03:07	04:44	04:16
Diebold Op Scan (total)	121	06:05	04:10	05:36	04:46
Diebold Op Scan (ballot marking)	121	04:58	03:53	04:31	04:00
Diebold Op Scan (ballot scanning)	121	01:07	02:55	00:31	00:23
ES&S Op Scan (total)	70	05:42	04:24	05:03	04:29
ES&S Op Scan (ballot marking)	70	05:15	04:26	04:36	03:55
ES&S Op Scan (ballot scanning)	70	00:27	00:17	00:25	00:22
Lever Machine	116	03:15	01:54	03:06	02:53
Liberty DRE	111	03:43	02:31	03:25	03:13
Sequoia DRE	93	05:31	03:12	05:13	04:36
Sequoia Op Scan (total)	117	04:45	02:27	04:31	04:06
Sequoia Op Scan (ballot marking)	117	04:15	02:23	04:02	03:35
Sequoia Op Scan (ballot scanning)	117	00:30	00:30	00:25	00:23

Note: N = the number of participants. SD = standard deviation. All estimates in minutes:seconds (e.g., 01:14) format.

Voting Time Results for Participants Based on Use of Accessibility Features

Exhibit 8 shows the results for participants who reported they did not use accessibility features on the voting systems and Exhibit 9 shows the results for participants who reported they did use accessibility features. Comparison of these two tables shows that regardless of which statistical estimate of time is used or regardless of which voting system is considered, the use of accessibility features almost always results in higher voting time estimates. Note, however, that given

[8] Note that some participants with disabilities asked for assistance in submitting their ballots into the optical scan reader from their own personal assistants or from AIR research staff.

the generally small number of individuals that report using accessibility features, the results of these numbers should be cautiously interpreted.[9] With fewer cases, it is plausible that these estimates reflect the instability in statistical estimates when based on only a limited number of data points.

Exhibit 8. Estimates of Voting Time Based on Participants that Self-Reported Not Using Accessibility Features

Voting System Name	N	Mean	SD	Trimmed Mean	Median
Avante DRE	331	03:57	01:50	03:46	03:36
Diebold Op Scan (total)	520	04:30	02:13	04:19	04:06
Diebold Op Scan (ballot marking)	520	04:00	02:00	03:52	03:39
Diebold Op Scan (ballot scanning)	520	00:30	01:12	00:23	00:20
ES&S Op Scan (total)	400	03:57	01:57	03:45	03:29
ES&S Op Scan (ballot marking)	400	03:30	01:53	03:18	03:03
ES&S Op Scan (ballot scanning)	400	00:27	00:30	00:23	00:21
Lever Machine	552	02:50	01:42	02:39	02:24
Liberty DRE	546	02:56	01:31	02:47	02:36
Sequoia DRE	420	04:10	02:29	03:54	03:39
Sequoia Op Scan (total)	560	03:51	01:53	03:39	03:26
Sequoia Op Scan (ballot marking)	560	03:21	01:43	03:11	03:01
Sequoia Op Scan (ballot scanning)	560	00:31	00:54	00:24	00:22

Note: N = the number of participants. SD = standard deviation. All estimates in minutes:seconds (e.g., 01:14) format.

[9] It is also important to point out that the Sequoia optical scan and lever voting systems did not offer accessibility features. It is unclear why some participants reported using accessibility features on these systems. This suggest that there was some confusion among study participants about what constitutes an accessibility feature and what does not.

Exhibit 9. Estimates of Voting Time Based on Participants that Self-Reported Using Accessibility Features

Voting System Name	N	Mean	SD	Trimmed Mean	Median
Avante DRE	14	08:26	04:47	08:10	06:08
Diebold Op Scan (total)	19	10:10	06:01	10:14	10:04
Diebold Op Scan (ballot marking)	19	07:58	06:47	07:51	04:51
Diebold Op Scan (ballot scanning)	19	02:13	04:07	01:45	00:30
ES&S Op Scan (total)	12	11:29	07:47	10:55	08:43
ES&S Op Scan (ballot marking)	12	10:53	07:57	10:17	07:53
ES&S Op Scan (ballot scanning)	12	00:36	00:24	00:35	00:32
Lever Machine	16	03:05	01:37	03:04	03:05
Liberty DRE	24	05:42	04:39	05:15	04:09
Sequoia DRE	22	07:05	03:46	06:44	05:56
Sequoia Op Scan (total)	11	04:52	02:37	04:48	04:35
Sequoia Op Scan (ballot marking)	11	04:14	02:10	04:07	04:17
Sequoia Op Scan (ballot scanning)	11	00:39	01:01	00:30	00:20

Note: N = the number of participants. SD = standard deviation. All estimates in minutes:seconds (e.g., 01:14) format.

What is the Estimated MDR?

The primary objective of this study was to calculate the maximum daily rate (MDR) of voting for each of the voting systems included in the study. As required by New York State law, the MDR indicates the maximum number of voters that each voting system could accommodate in a 15-hour voting period. As previously described, the calculation of the MDR was done without consideration to other factors that might affect the actual number of voters that can use a given voting machine in a 15 hour voting period, such as system breakdowns, volume of voters, and familiarity of voters with the voting machine.

With regard to optical scan systems, our decision to report three different times (time to mark ballot, time to scan ballot, and total time) presents some conceptual difficulties in calculating the MDR. Specifically, the optical scan ballots can be filled out simultaneously by more than one person even though only a single person can feed their optical scan ballot into a given optical scan reader at a time. Thus, an MDR that used the total time, for example, would likely underestimate the MDR in the typical voting situation where many more privacy booths than optical scan readers are provided. As the study results for all participants show (see Exhibit 5), the time it takes to fill out the paper ballot is considerable longer than the time it takes to read the ballot. As such, we do not use the optical scan total times to calculate a MDR. We do, however, calculate a pseudo-MDR for the

time it takes to fill out an optical scan ballot. This value can be interpreted as the maximum number of voters that can use a single ballot marking station/privacy booth in a 15 hour period.

MDR Overall Results

Exhibit 11 shows the MDR estimates based on all participants in the study. Given that the MDR results are calculated directly from the voting times, the results follow the same pattern as the voting time results. As shown, the MDR based on the mean voting time ranged from a low of 207 people (Sequoia DRE) to a high of 1,931 people (ES&S optical scan). However, when using the trimmed mean, it is the Diebold optical scan system with the highest MDR at 2,348 people. Finally, with the median voting time estimate, both the Diebold optical scan and ES&S optical scan systems both had the highest MDR at 2,571 people.

Exhibit 10. MDR Estimates Based on All Participants

		MDR Based On		
Voting System Name	N	Mean	Trimmed Mean	Median
Avante DRE	363	218	231	247
Diebold Op Scan (total)	N/A	N/A	N/A	N/A
Diebold Op Scan (ballot marking)	566	217	229	242
Diebold Op Scan (ballot scanning)	566	1588	2348	2571
ES&S Op Scan (total)	N/A	N/A	N/A	N/A
ES&S Op Scan (ballot marking)	429	239	262	290
ES&S Op Scan (ballot scanning)	429	1931	2246	2571
Lever Machine	594	317	337	372
Liberty DRE	600	295	317	342
Sequoia DRE	463	207	222	243
Sequoia Op Scan (total)	N/A	N/A	N/A	N/A
Sequoia Op Scan (ballot marking)	609	261	276	295
Sequoia Op Scan (ballot scanning)	609	1688	2185	2455

Note: MDR = maximum daily rate. N = the number of participants. These estimates do not reflect other factors that will influence the maximum number of voters who can vote during a 15-hour election day, such as system malfunctions, number of voters, etc.

Other MDR Results

We also calculated MDR estimates for participants with and without disabilities, and for participants who reported both using or not using accessibility features on each voting system. These results are provided in Appendix 9.

What Were the Reactions of Individuals to the Machines on Which They Voted?

The study design required participants to complete a reaction survey after voting on each voting system. This brief, seven item survey was used to gather additional information on whether participant's were familiar with each voting system, how they used the voting system during the study, and their general opinion about the ease of using each voting system. Specifically, the reaction survey required participants to respond "yes," "no," or "don't know" to the following items:

- I have voted on a voting system like this before.
- The voting system broke down.
- I used the accessibility options (e.g., the headphones, volume control, etc.)
- I voted on the system in a language other than English.

In addition, the survey required participants to rate their level of agreement with the following statements:

- The voting system was easy to use.
- The study instructions were easy to follow.
- I am confident that the voting system accurately recorded my vote.

The agreement ratings were gathered on the 5-point rating scale that ranged from "strongly agree" to "strongly disagree." Appendix 6 contains a copy of the reaction survey used for all voting systems.

Exhibit 12 shows the results of the first set of these questions. As shown, the responses were very consistent across voting systems with one exception—as expected, a large percentage (81 percent) of study participants reported being familiar with the lever voting system. For all other voting systems, 91 percent or more of the participants reported that they had not voted on a similar voting system before, that the system did not breakdown during the study, that they did not use the accessibility features, and that they voted in English.

Exhibit 11. Responses to Reaction Survey (Yes/No Items) Based on All Participants

	N	Percent (%) Yes	No	Don't Know
Avante DRE				
I have voted on a voting system like this before.	389	3.9	95.9	0.3
The voting system broke down.	386	4.7	92.5	2.8
I used the accessibility options.	385	3.6	95.8	0.5
I voted on the system in language other than English.	381	5.2	94.8	0.0
Diebold Op Scan				
I have voted on a voting system like this before.	571	8.2	91.8	0.0
The voting system broke down.	558	2.9	92.7	4.5
I used the accessibility options.	562	3.7	95.0	1.2
I voted on the system in language other than English.	559	4.7	94.8	0.5
ES&S Op Scan				
I have voted on a voting system like this before.	432	8.8	91.0	0.2
The voting system broke down.	423	1.9	95.0	3.1
I used the accessibility options.	423	3.3	95.5	1.2
I voted on the system in language other than English.	427	3.0	96.3	0.7
Liberty DRE				
I have voted on a voting system like this before.	607	5.9	93.6	0.5
The voting system broke down.	593	1.3	95.3	3.4
I used the accessibility options.	601	4.3	95.2	0.5
I voted on the system in language other than English.	599	6.0	93.8	0.2
Lever Machine				
I have voted on a voting system like this before.	610	81.3	18.0	0.7
The voting system broke down.	599	3.2	94.7	2.2
I used the accessibility options.	596	3.0	96.5	0.5
I voted on the system in language other than English.	601	6.5	93.0	0.5
Sequoia DRE				
I have voted on a voting system like this before.	469	5.5	94.2	0.2
The voting system broke down.	461	2.4	94.1	3.5
I used the accessibility options.	461	5.2	93.7	1.1
I voted on the system in language other than English.	461	6.9	92.8	0.2
Sequoia Op Scan				
I have voted on a voting system like this before.	602	6.0	94.0	0.0
The voting system broke down.	588	1.7	95.4	2.9
I used the accessibility options.	592	2.0	96.3	1.7
I voted on the system in language other than English.	589	4.8	94.6	0.7

Note: N = number of participants who responded to the reaction survey item.

Exhibit 13 shows the responses to the second set of reaction survey items. Unlike the first set of reaction survey items, there was a little more variability in the responses to the second set of items. The first item in Exhibit 13 deals with the ease of use of each voting system. As shown, 80 percent of participants agreed that the Liberty DRE voting system was easy to use. This represented the most favorable response in terms of ease of use. There is also a clear pattern by type of voting system, with participants giving more favorable ratings on ease of use to the direct recording electronic voting systems than to the optical scan voting systems.

The second item shown in Exhibit 13 addresses the study instructions. Regardless of which voting system participants voted on during the study, 80 percent or more of them indicated that they agreed that the study instructions were easy to follow. This question, in particular, provides additional support that the study design was clear and unambiguous to most participants.

Finally, in terms of how confident study participants were that the voting system accurately recorded their vote, there was also a clear pattern in that participants gave the direct recording electronic voting systems more favorable ratings than the optical scan voting systems. The most confidence was placed in the Avante DRE with 83 percent of participants agreeing that they were confident their vote was accurately recorded.

Exhibit 12. Response to Reaction Survey (Agreement Items) Based on All Participants

		Percent (%)		
	N	Agree	Neutral	Disagree
Avante DRE				
The voting system was easy to use.	388	76.5	12.9	10.6
The study instructions were easy to follow.	388	80.4	13.1	6.4
I am confident that the voting system accurately recorded my vote.	387	82.9	10.6	6.5
Diebold Op Scan				
The voting system was easy to use.	570	67.4	15.8	16.8
The study instructions were easy to follow.	567	80.4	12.9	6.7
I am confident that the voting system accurately recorded my vote.	566	67.8	19.1	13.1
ES&S Op Scan				
The voting system was easy to use.	432	63.9	15.5	20.6
The study instructions were easy to follow.	428	80.6	12.9	6.5
I am confident that the voting system accurately recorded my vote.	428	65.4	22.2	12.4

Exhibit 12. Response to Reaction Survey (Agreement Items) Based on All Participants (Continued)

	N	Percent (%) Agree	Neutral	Disagree
Liberty DRE				
The voting system was easy to use.	607	80.2	11.5	8.2
The study instructions were easy to follow.	603	86.1	9.6	4.3
I am confident that the voting system accurately recorded my vote.	605	76.5	16.4	7.1
Lever Machine				
The voting system was easy to use.	612	73.5	14.2	12.3
The study instructions were easy to follow.	607	83.0	13.3	3.6
I am confident that the voting system accurately recorded my vote.	611	78.7	15.2	6.1
Sequoia DRE				
The voting system was easy to use.	469	72.9	13.0	14.1
The study instructions were easy to follow.	466	82.6	12.7	4.7
I am confident that the voting system accurately recorded my vote.	469	79.7	12.6	7.7
Sequoia Op Scan				
The voting system was easy to use.	602	65.8	16.8	17.4
The study instructions were easy to follow.	599	82.0	11.5	6.5
I am confident that the voting system accurately recorded my vote.	601	60.2	21.1	18.6

Note: N = number of participants who responded to the reaction survey item.

Threats to the Validity of the Study

The findings reported in this document should be interpreted within the context in which this study was conducted. Specifically, there were several factors outside AIR's control that we believe may have impacted the study data, and ultimately, the study results.

One reason behind several of these threats is that the timeframe for conducting the study was very aggressive. The short timeframe to conduct the study resulted in less than optimal solutions to some of the critical components of our study design. For example, with more time, AIR would have performed all the outreach and recruitment efforts rather than relying on the NYSBOE to engage in outreach activities. A less aggressive schedule also would have enabled us to more carefully construct the target sample, as well as more carefully screen the study participants.

Another reason behind several of these threats is that voting systems represent a popular and heated issue for discussion in New York. Thus, we found that our

study generated a lot of interest from concerned citizens and the media. While AIR made every possible attempt to control the study environment and reduce the threat of external forces on the validity of the study, we still experienced several unexpected problems in conducting this study that could impact the credibility of the study findings. Each of these threats is discussed below.

Study Sample

NYSBOE appropriately identified voters as a "problematic entity" in the document entitled *Voting System User Rate Assessment Study* that was provided to AIR as background prior to the start of the study. In developing our work study design, we too regarded the issue of choosing the appropriate sample to participate in this study as the single biggest factor that would determine the success of this study. It is impossible to determine a MDR apart from the sample of voters on which the MDR is calculated. As a concrete example, it would be ill-advised to assume a MDR based on a sample of voters without disabilities generalizes to a sample of voters that includes individuals with disabilities, unless it can be shown that having a disability does not affect the time it takes to vote on a given voting system. Disability status is only one of the many demographic differences that may affect voting behavior and therefore must be accounted for in choosing an appropriate study sample.

Aside from simply accounting for various demographic characteristics in our sample of study participants, we were also concerned about the specific composition of study participants. It was AIR's goal that the demographic composition of the sample of voters used in the study mirror as much as possible the demographic composition of the population of voters expected to participate in New York elections. However, based on our conversations with NYSBOE staff, we understood that, at that time, there was no reliable source of information available about the composition of voters expected to vote in New York.

Not only does this reality complicate the process of defining the population of interest, it also leaves open the possibility that any MDR estimate could be regarded as not applicable to the true voting population. In the absence of data to profile voters, we suggested using New York Census data to determine the composition of voting-eligible people in New York.

It is important to understand the limitations of this approach. First, there may be dramatic differences between the composition of voting-eligible individuals as defined by Census data and either registered voters or active voters. Second, the composition of voters may vary as a function of voting district. However, the statutory requirement to calculate a single MDR for each voting system required that district level variations, to the extent that they occur, be ignored. Finally, by

using Census level data, this study calculates a MDR that most appropriately generalizes to voting-eligible adults.

To address this sampling issue, we provided the NYSBOE with guidance on the desired composition of the study sample. During execution of the study, we also built in provisions to continuously monitor the composition of study participants. After data collection occurred at the first location (Rochester), AIR notified the NYSBOE that the composition of the sample was heavily skewed toward certain demographic groups and that more targeted recruiting would need to take place at the remaining two data collection locations (Brooklyn and Albany) to ensure that the final composition of our study participants appropriately reflected the demographic composition of the population of interest.

Unfortunately, at the time of that discussion outreach efforts had already begun for New York City and therefore it was too late to implement remedial actions. After data collection occurred in New York City, AIR again informed the NYSBOE that the study sample was unbalanced (in comparison to the New York Census data) and that the distribution of participants across different demographic groups was too disparate from the distributions of demographic characteristics of the voting-eligible population in New York State. To address this problem, AIR took the lead on recruiting for the third and final data collection location, Albany. NYSBOE provided AIR with contacts in Albany who could assist with outreach to the specific populations we were lacking. In addition, AIR staff members screened each potential study participant to ensure that only persons from needed demographic groups were scheduled for the study. See Appendix 10 for a copy of the screening questionnaire used to select study participants for the Albany data collection.

Distractions: The Media and Other Observers

The NYSBOE informed AIR that, by law, the NYSBOE was obliged to notify citizens of New York that the study was taking place and, to the extent that it did not violate study participants' rights to privacy, that the NYSBOE was required to make the study observable to the public. Therefore, the NYSBOE alerted the media and citizen groups that the study was taking place through press releases and other means. In addition, one of the vendors issued a press release prior to data collection in Rochester inviting members of the media to attend a "demonstration" of their voting system during study hours.

At two of the data collection locations (Rochester and New York City) representatives from the media (print and television) were present. AIR allowed the media access to film the voting systems and to interview staff members from the NYSBOE, but we had an obligation to protect the privacy of the study

participants.[10] At no time did we give the media permission to film or interview study participants during the study. We did, however, advise study participants that media representatives were in the vicinity. And of course, study participants were free to talk with the media after they completed their participation in the study if they choose to do so.

At each of the data collection locations, the study proceedings were observed by citizen advocates, members of the media, and others who were interested in the study. At no time did AIR reveal any of the study participants' names to any observers. However, at times AIR staff members were distracted by the noise and activity of the observers. We addressed this issue with staff training and constant supervision of staff members to ensure they were able to perform their study responsibilities adequately. We also worked closely with the NYSBOE to keep distractions to a minimum. We do not believe that study participants were overly distracted by the noise and activity of the observers, but we have no way of knowing for sure.

Biased Study Participants

During data collection at the first location (Rochester), AIR staff members noticed that three study participants had copies of the study materials in their handbags or in their pockets. When we inquired about these materials, we were informed that individuals from New York with access to study materials (not AIR) had provided these materials to the three study participants prior to their participation. In addition, we observed on at least two other occasions individuals asking for copies of the study materials from voting system vendors.[11] We do not know whether the individuals that had ask vendors for study materials were study participants or whether anyone else received copies of the materials from anyone outside AIR that was connected to the study.

Similarly, we observed a study participant at the second data collection location (New York City) with a stopwatch. AIR did not notice that she had a stopwatch until she dropped it and it broke. We can only assume that this study participant was timing herself as she used the voting systems. Given this individual appeared to be engaged in activities not called for by our protocol, it is reasonable to assume that the times associated with her voting may be distorted. Furthermore, we do not know how many other individuals that participated in our study attempted to simultaneously time their own or others' voting activity.

[10] As required by Federal regulations and internal AIR policies regarding the ethical treatment of human subjects in research studies, we promised study participants confidentiality as a condition of participation in this study. In addition to complying with Federal law and AIR policy, we believe that such a practice leads to more accurate study results because participants are not burdened by the prospect of having their participation in the study disclosed.

[11] All vendors were instructed to not provide copies of the ballot or to discuss the specific requirements of the study with anyone until after the completion of the study.

There were also observers at New York City and Albany with stopwatches. We do not know why observers were using stopwatches as they observed the study. We can only surmise that they were attempting to conduct their own data collection effort, but we do not know the nature of their study. We do not know what form of training these observers received and whether they were following the same methodology that AIR was using to record voting times. We did not train these individuals. In addition, voters marked their ballots from behind privacy curtains, so these observers would have no way of knowing when to start or stop their watches (and they would not know what criteria we used). Moreover, AIR never provided anyone with access to our time data, so there would be no way for these observers to compare their time data to ours or for them to make any credible claims about the reliability of their data.

Various sources have offered opinions about these events. Some individuals have accused citizen advocates of trying to skew the data in favor of a particular voting system. Other individuals believe that some participants were simply confused about the purpose of the study and viewed the study as a demonstration of voting system equipment, rather than fully understanding that a research study to determine the MDR of the voting systems was taking place. Still others claim that these incidents were limited to only a handful of cases and that there is no reason to be concerned about the behavior of study participants. Unfortunately, we are not in a position to speculate about the effect of these events or about the extent to which they occurred. We can, however, note that every attempt was made to fully explain the goals of the study, as well as the requirements for participating in it. To the extent that individuals did not follow the study protocol, it remains possible that the data gathered from them are inaccurate.

Vendor Participation

Our original study design called for a single vendor representative for each voting system to provide a standardized introduction on their respective voting system to each study participant who voted with their system. Appendix 11 contains the specific guidelines provided to vendors at the outset of the study. Vendors were required to submit a brief script and have it approved prior to data collection in Rochester. Beyond providing a brief introduction to their systems, vendors were instructed to not interfere with data collection or participants' voting behavior, unless specifically asked to by one of the AIR research staff.

During data collection at the first location, AIR staff members reported that some vendor representatives were occasionally straying from their scripts and providing more information than was necessary. In addition, on occasion vendors intervened when voters had questions during voting, or when it may have appeared they had a question but had not actually asked one.

Recognizing that the vendor representatives were not trained researchers and that they brought their own biases to the study (in favor of their own products), we decided to remove vendors from the voting rooms after collecting data at the first location.[12] For the balance of the study, AIR researchers delivered the standardized instructions to each study participant and vendors remained on stand-by in an adjacent area, in case questions or technical issues surfaced that AIR could not address. Although vendors remained present at the data collection sessions, they were strictly prohibited from interfering with the study.

Change in Study Design

Our original study design was based on the assumption that each study participant would vote on no more than five voting systems within a 45-minute period. However, the NYSBOE decided prior to the pilot test to include the lever voting system in the study. In addition, one vendor submitted an additional system just prior to data collection and NYSBOE decided to include this additional system in the study. Since outreach efforts had already begun and the study had already been publicized as 45 minutes in duration, we could not increase the duration of the data collection sessions. Therefore, study participants voted on a maximum of seven systems rather than five. The implication of this change is that most participants did not have the opportunity to vote on all seven systems in the 45 minute data collection session, which was part of our study design.

Questions Raised

The events summarized above have left AIR with several questions that are impossible to answer:

- Would the results presented in this report be the same with a more carefully selected sample of study participants?
- How many study participants received copies of study materials prior to participating in the study? Did study participants review the materials before participating? If so, what effect, if any, did that have on the tasks they performed in the study?
- Was there a concerted effort to tamper with the study or were these isolated incidents?
- Did some participants participate in the study with the goal of creating a favorable outcome for one voting system over another? If so, how did this affect their performance in the study? How many study participants engaged in this behavior?

[12] Vendors remained present during the entire data collection session in Rochester.

- Did vendor instructions to study participants that strayed from the pre-approved script have an impact on data collected at the first study location?

- Were participants more "practiced" because they used seven voting systems instead of five? Did the change in study design have any impact on participants' voting speed? Did participants try to "rush" in order to vote on all seven systems in the allotted time?

- Would more participants have been able to vote on all the systems if there were only five instead of seven?

AIR's Approach to Mitigating Threats to the Study

As AIR became aware of potential threats to the validity of the data collection efforts, we discussed the issues with the NYSBOE, the vendors, and study observers. In the interest of conducting a study that was objective, fair, and unbiased, AIR took the following actions to control the conditions under which the study was conducted.

First, when it became clear that the NYSBOE outreach efforts would not yield a sample of participants that were representative of the voting-eligible population of New York, AIR took over the recruiting and screening efforts. For the Albany data collection, we provided a toll-free number to potential participants and screened them by asking them a series of questions about their backgrounds to ensure that we recruited individuals whose backgrounds matched those we were lacking in the sample (See Appendix 10 for a copy of the screening questionnaire). To provide an additional incentive to participate, AIR randomly selected one participant from each day in Albany to receive $250.

Second, AIR actively controlled access to the study by observers and vendors. Observers and vendors watched the study proceedings from a designated area and we did not allow observers to interfere with the voting stations or to have access to data. We also did not allow observers to take copies of study materials. Some observers did ask for copies and we declined to share them until the study was complete. As indicated above, we also restricted media access to the study participants. In particular, the media were not allowed to film study participants. They were only allowed to film voting systems, the vendors, or NYSBOE staff members and this was restricted to when no study participants were in the study area.

Third, as soon as AIR became aware of study participants with access to study materials we addressed the issue with the NYSBOE. They were aware of the problem and told us they would address the problem with the parties that were involved. Another reason for AIR taking over the recruiting effort was to also

attempt to screen out participants who might have ulterior motives for participating in the study. The only way to do this was to screen out individuals who stated that they wanted to participate for reasons other than volunteering for a study. There were individuals whom we disqualified from the Albany data collection because they indicated to our staff member(s) that they had other reasons for being in the study or that they wanted to see a particular outcome (i.e., the success of one type of voting system over another).

Fourth, we analyzed the data multiple ways. For example, to eliminate any potential attempts to sabotage specific systems (i.e., biased participants who may have voted very slowly or very quickly on a system on purpose), our findings include a "trimmed mean" which removes outliers from the analysis (those data points that fall at the extreme 5 percent of the tail of the distribution of times for each system). This should eliminate from the findings data from participants who deliberately sought to interfere with the study.

Finally, as discussed above, we had already built features into our study design that also help mitigate these threats to the study. For example, AIR randomly assigned participants to voting systems in different orders. This meant that even if participants could not vote on all seven systems, there was no single system that was at a disadvantage—each one had an equal chance of being used in the study.[13]

Conclusions

The primary goal of this study was to estimate the MDR for voting systems that New York State is evaluating as replacements to the lever systems. To accomplish this, we designed and conducted a study whereby a sample of participants was timed as they voted on each of the systems. Like all research studies, there were pros and cons to our study design this report has attempted to point these features out so that individual's evaluating and using these study results can do so effectively. In this concluding section, we address some of the move salient features of the study that we believe deserve emphasis.

The Limits of MDR Estimates

This study only presents some of the factors that New York must consider in deciding how many and what types of voting systems to purchase. In conducting this study, we stuck to the original objective—estimating MDR. We have pointed

[13] We asked vendors to provide two of the same systems for each study location. In the pilot test, one of the two Avante DRE voting systems malfunctioned, so only one Avante DRE voting system was available for use throughout most of the entire study. Thus, the total sample size of voters who used this system is markedly lower than the sample size of voters who used the other systems. Similarly, the ES&S optical scan voting system was unavailable during the first day of data collection in Rochester because the ballot was incorrect; hence, fewer participants voted on this system then originally anticipated.

out many of the other obvious factors to consider throughout this report, such as system malfunctions, number and flow of voters, and the level of familiarity with a given voting system. The results of this study suggest some other, less obvious factors to consider. For example, while optical scan systems appeared to have the greatest potential to process the most voters in a day, participants had more confidence that their votes had been recorded accurately with the DREs than they did with the optical scan systems. Moreover, participants reported that they found the DRE systems easier to use than the optical scan systems.

Usability Testing to Inform MDR Estimates

Usability testing was outside the scope of this study, therefore we do not have data on the specific features of each system that may be unusable or problematic. We encourage New York to collect reports of any existing usability testing done by independent parties on the systems currently under consideration. Such data will add to the factors New York may consider in making decisions about voting system purchases.

The major reason for the difference in MDR between the optical scanning systems and the DREs is that the ballot marking and ballot casting process are combined so that each system is occupied and cannot process other voters while one voter is using it. Optical scanning systems involve two separate processes: while only one ballot at a time can be fed into the ballot scanner, multiple voters can mark their ballots simultaneously before scanning them in.

Potentially adding to the time difference between the DREs and the optical scanning systems is the time required for voters to verify their votes by reviewing the paper ballots that are printed for them after they mark their ballots.

There were also differences between the DREs themselves. Liberty's DRE was a full-face paper ballot overlay DRE, while the Sequoia and Avante DREs were full-face touch screen DRE systems. When AIR reviewed the Liberty DRE before the study began, we noticed that the paper printout was difficult to read and presented the voter with a series of coordinates, e.g., A1, B6, rather than actual names. In our expert opinion, this printout is less understandable to voters and may by one factor, among others, that accounts for the fact that voters spent less overall time on the Liberty DRE than the Sequoia DRE or the Avante DRE. However, this finding would need to be confirmed with user testing.

Variability in Voting Times

One of the most salient outcomes was the large variability in voting times on all the systems evaluated in this study. Because of time and resource constraints, we were not able to fully explore the nature of this variability.

Analysis of the sample composition showed that the overall sample, as well as the samples of participants voting on each system, was comparable to the composition of voting-eligible adults in New York State. However, there were some noteworthy exceptions (e.g., under representation of native non-English language participants). We analyzed the results across all participants and therefore can not definitively rule out that these exceptions might not alter the MDR estimates. We do believe that the randomization component of the study design does, however, mitigate any sampling effects biasing the MDR estimates in favor of one voting system over another.

Participants with Severe Disabilities

With seven voting systems to test, it was impossible to get a sizeable sample of participants with severe disabilities to vote on all the voting systems. Furthermore, two of the voting systems could not accommodate individuals with disabilities. We did report MDR estimates for participants with and without self-identified disabilities; however, most of the participants claiming disabilities reported that they had a mental or cognitive disability. Furthermore, the percentages of participants who self-reported the use of accessibility features were very small for all the voting systems.

Having a severe disability or using accessibility features impacts the time it takes to vote. To the extent that our volunteer sample lacked representation of individuals with certain types of disabilities and requiring the broad range of accessibility features on voting systems (e.g., sip and puff), the MDR estimates may be misleading.

Using the Estimate MDR

Finally, we note that although the goal of study was to produce a MDR estimate for each system, we have in fact presented three estimates for each system (i.e., mean, trimmed mean, and median as shown in Exhibit 5). It is our suggestion that users of this report start with the mean as the best estimate of the MDR for each voting system, but consider the trimmed mean and median estimates of MDR in light of the other issues raised in this report.

References

Andresen, E.M. & Fitch, C.A. (2000). Reliability and validity of disability questions for US Census 2000. *American Journal of Public Health,* Vol. 90*(8)*, 1297-1299.

Dumas, J. & Redish, G. (1999). A Practical Guide to Usability Testing. Norwood, NJ: Ablex Publishing.

Federal Election Commission (October 2003). Usability Testing of Voting Systems. Prepared by American Institutes for Research under Contract No. FE-2-AC-0024.

U.S. Department of Justice (March 1, 2006). Justice Department sues New York State over voting rights: Press release. Washington, DC: Author.

APPENDICES

1. SAMPLE BALLOT

Ballot C

Governor and Lt. Governor contest
- Democratic candidate Stephen Massimilian for Governor
- Democratic candidate Gloria Dillon for Lt. Governor

Comptroller contest
- Socialist Worker candidate Eugene Ruff

Attorney General contest
- Democratic candidate Robert Squire

Associate Judge of the Court of Appeals contest
- Democratic candidate Stephen Harrick
- Democratic candidate Andre Decker

United States Senator contest
- Labor candidate Samuel Lawton

Justice of the Supreme Court contest
- Democratic candidate Robert O'Donnel
- Democratic candidate Carol Tubbs
- Democratic candidate Peter Tooley
- Democratic candidate Helena Sack

Representative in Congress contest
- Democratic candidate Lois Koss

State Senator contest
- Democratic candidate Theodore Kopp

Member of Assembly contest
- Liberal candidate Harry Sosses

County Judge contest
- Democratic candidate Donna Berke
- Democratic candidate Anders Culpo
- Democratic candidate Angela Pogoda
- Democratic candidate Gene Tillman

Judge of the Family Court contest
- Democratic candidate Eric Sheehy
- Democratic candidate Saul Currier
- Democratic candidate Lewis Tese
- Democratic candidate Peter Valle

District Attorney contest
- Liberal candidate Robert Hook

Councilman unexpired term contest
- Democratic candidate Gloria Castle

Ballot C

Ballot Questions

- Please skip all ballot questions except for the following two:
- Vote any way you like on Proposal Number Four, An Amendment, Continuation of the Services of a Judge or Justice After Retirement
- Vote any way you like on Proposal Number Six, An Amendment, Exchange of Certain Property Within the Adirondack Park

When you are done voting, please hand this form to the researcher

2. STAFF TRAINING AGENDA

Voting Machine Study
Training Agenda

- Project Background
- Project Overview
- Mock Vote Process
- Interacting with Participants
- Interacting with Vendors
- Interacting with Participants with Disabilities

3. CONSENT FORM

AMERICAN INSTITUTES FOR RESEARCH

Information about the Study

Please read this carefully.

Purpose

American Institutes for Research (AIR) has asked you to be part of a research study. This study will help us learn how actual voters use different types of voting systems. AIR wants to study the voting systems and NOT you or how well you can use the voting systems.

Procedures

This study is being conducted at three locations in the state of New York: Schenectady, New York City, and Rochester. First, we will ask you to complete a short survey about your background. Next, we will ask you to cast votes using several different voting systems. We will ask you to complete a short survey after you vote on each machine. We will assign you a unique identification number so that you do not have to put your name on any of the surveys. To make the process the same for everyone, we will tell you who to vote for. It is important that you follow the instructions provided to you today so that we can gather accurate information. The study will last approximately 45 minutes.

Confidentiality

Your name will only appear on this form (below) and your payment receipt. We will keep your name confidential and separated from all the survey information we collect from you during the study. Only your unique identification number will appear on the surveys that you complete today and there is no way for anyone to match your name with your identification number. Thus, you will complete the surveys anonymously. No one will ever know how you responded to the surveys.

Information Collected

We will collect information about your background as well as your impressions of the voting systems that you will use today. We will also record how long it takes you to vote using each system. We will not record your votes.

(Please turn over)

Benefits
Being in the study gives you a chance to share your thoughts about the usability of voting systems. Your role is important to helping us learn how actual voters use each of the voting systems. The information you provide today will help the election officials throughout New York State evaluate features, ease of use, and the time required to cast a vote using various voting systems that may be used in future elections.

Risks and Discomforts
Being in this study should not be risky or uncomfortable for you. You may take a break at any time you wish; just let an AIR staff person know that you would like to do so.

Voluntary Participation
Participation in this study is voluntary. You may stop participating in this study at any time and you may skip any questions you do not want to answer. If you choose to leave the study, we will still pay you the full amount we promised you.

More Information
If you have any questions or concerns about this research study, please feel free to contact the director of the research project, Dwayne Norris, Ph.D., at dnorris@air.org or 202-403-5129, or c/o AIR, Attn: Dwayne Norris, 1000 Thomas Jefferson Street, NW, Washington, DC 20007. If you have questions about your rights as a research participant, contact the IRB Chair at IRB@air.org or toll-free at 1-800-634-0797 or c/o AIR, IRB Chairperson, 1000 Thomas Jefferson Street, NW, Washington, DC 20007.

Informed Consent
By signing this form, you agree that you have read and understood the information described above and agree to participate in the study, as described.

_____ _____

Signature **Date**

_____

Please print name

You have the right to keep a copy of this form.
Please ask an AIR staff person to provide you with a copy.

4. BACKGROUND SURVEY

ID 3____

NY State Board of Elections
Study Participant Background Survey

1. Is your primary residence in the State of New York?
   - ❑ Yes
   - ❑ No

2. What is your gender?
   - ❑ Male
   - ❑ Female

3. How old are you?
   - ❑ 18-24 years of age
   - ❑ Between 25-34 years of age
   - ❑ Between 35-44 years of age
   - ❑ Between 45-54 years of age
   - ❑ Between 55-64 years of age
   - ❑ Between 65-74 years of age
   - ❑ 75 years of age or older

4. How would you best describe your racial/ethnic background? **Check all that apply:**
   - ❑ American Indian or Alaska Native
   - ❑ Asian
   - ❑ Black or African American
   - ❑ Hispanic
   - ❑ Native Hawaiian or Other Pacific Islander
   - ❑ White
   - ❑ Other (Please specify) _____

5. Select the highest level of education you have completed.
   - ❑ Some high school
   - ❑ GED or high school equivalency
   - ❑ High school diploma
   - ❑ Attended a vocational, trade, or business school
   - ❑ Less than two years of college
   - ❑ Associate's degree
   - ❑ Bachelor's degree
   - ❑ Master's degree
   - ❑ Doctorate (e.g., Ph.D. or Ed.D.)
   - ❑ First professional degree (J.D., M.D.)
   - ❑ Other (Please specify) _____

Please Complete Questions on the Back

6. What is your **first** language? **Check one:**
   - ❑ English
   - ❑ Spanish
   - ❑ Korean
   - ❑ Mandarin
   - ❑ Cantonese
   - ❑ Other (Please specify) _____

7. Do you identify yourself as having any of the following disabilities? **Check all that apply:**
   - ❑ Blindness or visual impairment
   - ❑ Deafness or hard of hearing
   - ❑ Mobility impairment (a long-lasting condition that substantially limits your physical activities such as walking, climbing stairs, reaching, lifting, or carrying)
   - ❑ Mental or cognitive disability (a physical, mental, or emotional condition lasting 6 months or more that makes it difficult to learn, remember, or concentrate)
   - ❑ Other (Please specify) _____
   - ❑ None of the above

8. How frequently do you vote in political elections?
   - ❑ Frequently. I vote in every, or nearly every, political election.
   - ❑ Occasionally. I try to vote in every political election, but do not always vote.
   - ❑ Rarely. I rarely, if ever, vote in political elections.
   - ❑ I am a first time voter.

9. Are you registered to vote?
   - ❑ Yes, I'm registered in New York
   - ❑ Yes, I'm registered elsewhere
   - ❑ No

Thank you

5. INSTRUCTION SHEET

2001

INSTRUCTIONS

Please vote in the following order:

Voting Station 7 _____
Voting Station 1 _____
Voting Station 2 _____
Voting Station 3 _____
Voting Station 4 _____
Voting Station 5 _____
Voting Station 6 _____

6. REACTION SURVEY

ID _____ Voting Station ____6____

Please indicate your level of agreement with the following statements about the voting system you just used.

	Yes	No	Don't Know
I have voted on a voting system like this before.	○	○	○
The voting system broke down.	○	○	○
I used the accessibility options (e.g., the headphones, volume control, etc.).	○	○	○
I voted on the system in a language other than English.	○	○	○

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
The voting system was easy to use.	○	○	○	○	○
The study instructions were easy to follow.	○	○	○	○	○
I am confident that the voting system accurately recorded my vote.	○	○	○	○	○

7. DEBRIEFING FORM

New York State Board of Elections
Research Study Debriefing Form

Thank you for participating in today's study. The information you provided today will help the election officials throughout New York State evaluate features, ease of use, and the time required to cast a vote using various voting systems that may be used in future elections.

The answers you provided today are anonymous and confidential. As we did not record your name on any study documents, any reports we generate on the basis of these data cannot be linked to you specifically. All reports will contain group-level (e.g., age group, voting experience) not individual-level findings.

Thank you!

8. Composition of Study Participants for each Voting System

Composition of Study Participants for Each Voting System

	Avante	Diebold	ES&S	Lever	Liberty	Sequoia DRE	Sequoia Op Scan
Gender							
N	360	564	428	588	588	460	600
Female	57.5%	60.6%	62.9%	57.3%	58.3%	59.6%	59.2%
Male	42.5%	39.4%	37.1%	42.7%	41.7%	40.0%	40.8%
Age							
N	360	565	428	589	590	462	602
18-24 years	22.5%	21.1%	22.2%	19.2%	20.2%	18.8%	20.1%
25-34 years	14.7%	15.0%	16.8%	14.1%	14.6%	13.9%	14.3%
35-44 years	13.6%	14.5%	12.1%	13.6%	14.7%	16.7%	15.1%
45-54 years	20.6%	20.0%	20.6%	21.2%	21.4%	20.8%	21.3%
55-64 years	15.3%	16.6%	15.0%	18.5%	16.1%	16.2%	16.4%
65-74 years	10.3%	10.3%	10.0%	9.7%	9.3%	11.3%	10.0%
75 years or older	3.1%	2.5%	3.3%	3.7%	3.7%	2.4%	2.8%
Race / ethnicity							
N	359	563	428	587	589	461	602
American Indian / Alaska Native	1.1%	1.2%	1.6%	1.4%	1.5%	1.1%	1.5%
Asian	5.0%	3.9%	4.4%	4.4%	3.9%	3.9%	4.2%
Black / African American	34.3%	32.1%	27.1%	32.9%	32.9%	34.9%	32.4%
Hispanic	5.8%	4.8%	5.1%	5.1%	5.4%	4.6%	5.5%
Native Hawaiian / Pacific Islander	0.8%	0.9%	1.2%	1.0%	0.5%	0.7%	0.8%
White	49.3%	53.6%	57.2%	51.6%	52.5%	51.2%	53.3%
Other	3.6%	3.4%	3.3%	3.6%	3.2%	3.7%	2.3%
Education (Highest level completed)							
N	360	565	427	588	588	460	600
Some high school	5.3%	4.2%	3.3%	5.6%	5.4%	5.2%	5.5%
GED or HS diploma equivalent	5.8%	6.2%	4.4%	7.3%	6.6%	7.6%	7.7%
High school diploma	11.9%	14.2%	14.5%	15.3%	14.8%	14.8%	13.0%
Vocational /Trade/Business school	4.4%	4.1%	4.7%	4.9%	4.9%	5.7%	5.3%
Less than 2 years of college	16.1%	15.4%	16.4%	15.3%	14.3%	13.7%	14.0%
Associates degree	10.3%	12.0%	9.8%	10.9%	10.4%	11.5%	11.5%
Bachelors degree	26.9%	25.3%	26.2%	23.8%	25.5%	24.1%	25.0%
Masters degree	11.1%	11.0%	12.2%	9.9%	10.2%	10.0%	10.3%
Doctorate degree	1.4%	1.8%	2.3%	1.5%	1.7%	1.7%	1.5%
Professional degree (JD, MD)	2.8%	2.7%	2.8%	2.6%	2.6%	3.0%	2.7%
Other	3.9%	3.2%	3.3%	2.9%	3.6%	2.6%	3.5%

	Avante	Diebold	ES&S	Lever	Liberty	Sequoia DRE	Sequoia Op Scan
First language							
N	359	562	423	583	586	457	596
English	92.8%	96.1%	94.8%	94.5%	95.1%	94.5%	95.3%
Spanish	2.8%	1.4%	2.1%	2.1%	1.7%	1.8%	1.8%
Korean	1.1%	0.9%	0.9%	1.0%	0.7%	1.1%	1.2%
Mandarin	0.3%	0.0%	0.2%	0.0%	0.2%	0.2%	0.0%
Cantonese	1.1%	0.5%	0.2%	0.7%	0.7%	0.7%	0.3%
Other	1.9%	1.1%	1.7%	1.7%	1.7%	1.8%	1.3%
Disability							
N	353	553	414	569	573	445	582
None	77.3%	78.1%	83.1%	79.6%	80.6%	79.1%	79.9%
Blind / visually impaired	4.0%	3.3%	2.7%	2.3%	2.4%	2.0%	1.4%
Deafness / Hard of hearing	4.0%	3.8%	3.1%	3.9%	4.4%	4.3%	3.4%
Mobility impairment	7.4%	9.0%	6.0%	7.6%	6.3%	7.4%	8.4%
Mental or cognitive disability	5.4%	4.9%	4.1%	5.8%	5.2%	6.1%	5.5%
Other	2.0%	0.9%	1.0%	0.9%	1.0%	1.1%	1.4%
Voting frequency							
N	357	561	422	581	584	455	595
First time voter	6.4%	6.6%	4.0%	7.6%	6.7%	6.8%	8.6%
Rarely	4.8%	6.6%	6.2%	6.9%	8.2%	8.1%	7.6%
Occasionally	18.2%	15.9%	16.8%	15.7%	16.3%	14.9%	15.0%
Frequently	70.6%	70.9%	73.0%	69.9%	68.8%	70.1%	68.9%
Registered voter							
N	358	561	422	583	584	456	595
Yes, in NY	92.7%	91.1%	93.6%	91.6%	90.9%	90.8%	90.1%
Yes, elsewhere	2.5%	3.0%	3.3%	2.4%	3.3%	2.4%	3.0%
No	4.7%	5.9%	3.1%	6.0%	5.8%	6.8%	6.9%

Note: N = the number of participants who reported data for this variable. Percentages reported for race/ethnicity and disability may add up to more than 100% because some participants reported identifying with more than 1 category.

9. MDR Results for Participants Using and not Using Accessibility Features

MDR Estimates for Participants Reporting the Use of Accessibility Features

Voting System Name	N	MDR Based On Mean	Trimmed Mean	Median
Avante DRE	331	228	239	250
Diebold Op Scan (total)	N/A	N/A	N/A	N/A
Diebold Op Scan (ballot marking)	520	225	233	247
Diebold Op Scan (ballot scanning)	520	1800	2348	2700
ES&S Op Scan (total)	N/A	N/A	N/A	N/A
ES&S Op Scan (ballot marking)	400	257	273	295
ES&S Op Scan (ballot scanning)	400	2000	2348	2571
Lever Machine	552	318	340	375
Liberty DRE	546	307	323	346
Sequoia DRE	420	216	231	247
Sequoia Op Scan (total)	N/A	N/A	N/A	N/A
Sequoia Op Scan (ballot marking)	560	269	283	298
Sequoia Op Scan (ballot scanning)	560	1742	2250	2455

Note: MDR = maximum daily rate. N = the number of participants. These estimates do not reflect other factors that will influence the maximum number of voters who can vote during a 15-hour election day, such as system malfunctions, number of voters, etc.

MDR Estimates for Participants Not Reporting the Use of Accessibility Features

Voting System Name	N	MDR Based On Mean	Trimmed Mean	Median
Avante DRE	14	107	110	147
Diebold Op Scan (total)	N/A	N/A	N/A	N/A
Diebold Op Scan (ballot marking)	19	113	115	186
Diebold Op Scan (ballot scanning)	19	406	514	1800
ES&S Op Scan (total)	N/A	N/A	N/A	N/A
ES&S Op Scan (ballot marking)	12	83	88	114
ES&S Op Scan (ballot scanning)	12	1500	1543	1688
Lever Machine	16	292	293	292
Liberty DRE	24	158	171	217
Sequoia DRE	22	127	134	152
Sequoia Op Scan (total)	N/A	N/A	N/A	N/A
Sequoia Op Scan (ballot marking)	11	213	219	210
Sequoia Op Scan (ballot scanning)	11	1385	1800	2700

Note: MDR = maximum daily rate. N = the number of participants. These estimates do not reflect other factors that will influence the maximum number of voters who can vote during a 15-hour election day, such as system malfunctions, number of voters, etc.

10. SCREENING PROTOCOL AND QUESTIONS

NYSBOE Recruiting Screener

Name: _____

Phone: _____ Email: _____

Date Scheduled: _____ Confirmation Sent: _____

We need to schedule 14 participants, plus 2 wait-list participants, per block for the following blocks:

Thursday, November 16

Block 1: 8:00am to 8:45am
Block 2: 9:00am to 9:45am
Block 3: 10:00am to 10:45am
Block 4: 11:00am to 11:45am
Block 5: 1:00pm to 1:45pm
Block 6: 2:00pm to 2:45pm
Block 7: 3:00pm to 3:45pm
Block 8: 5:00pm to 5:45pm
Block 9: 6:00pm to 6:45pm

Friday, November 17

Block 10: 8:00am to 8:45am
Block 11: 9:00am to 9:45am
Block 12: 10:00am to 10:45am
Block 13: 11:00am to 11:45am
Block 14: 1:00pm to 1:45pm
Block 15: 2:00pm to 2:45pm
Block 16: 3:00pm to 3:45pm
Block 17: 5:00pm to 5:45pm
Block 18: 6:00pm to 6:45pm

Potential participants will be disqualified if:

- They are under 18 years of age, or
- They do not live in New York State.

Introduction

Hello, my name is _____. I'm calling on behalf of the American Institutes for Research (AIR). AIR is an independent, not-for-profit organization that is hired by other companies and agencies to make their products easier to use. Currently, we are recruiting people to participate in an important study of different voting systems that might be used in the future in NY elections. This is not a sales presentation and you will not be asked to purchase anything.

The study will be held in ?? and will last approximately about 45 minutes. If you participate, we will give you $30. Does this sound like something you might be interested in?

Great! May I take a few moments to ask you some questions to determine if your background matches the profile we are looking for? These questions are voluntary; you can skip the ones you don't want to answer. All of the information will be kept confidential.

1. Is your primary residence in the State of New York?
   - ❑ Yes
   - ❑ No **(Disqualify, see script at end)**

2. How old are you?
   - ❑ Younger than 18 years of age **(Disqualify, see script at end)**
   - ❑ 18-24 years of age
   - ❑ Between 25-34 years of age
   - ❑ Between 35-44 years of age
   - ❑ Between 45-54 years of age
   - ❑ Between 55-64 years of age
   - ❑ Between 65-74 years of age
   - ❑ 75 years of age or older

3. Are you a voting inspector or poll worker?
   - ❑ Yes **(Disqualify, see script at end)**
   - ❑ No

4. What is your gender?
   - ❑ Male
   - ❑ Female

5. How would you best describe your racial/ethnic background? **Check all that apply:**
   - ❑ American Indian or Alaska Native
   - ❑ Asian
   - ❑ Black or African American
   - ❑ Hispanic
   - ❑ Native Hawaiian or Other Pacific Islander
   - ❑ White
   - ❑ Other (Please specify) _____

6. What is the highest level of education you have completed:
   - ❑ Some high school
   - ❑ GED or high school equivalency
   - ❑ High school diploma
   - ❑ Attended a vocational, trade, or business school
   - ❑ Less than two years of college
   - ❑ Associate's degree
   - ❑ Bachelor's degree
   - ❑ Master's degree
   - ❑ Doctorate (e.g., Ph.D. or Ed.D.)
   - ❑ First professional degree (J.D., M.D.)
   - ❑ Other (Please specify) _____

7. What is your **first** language (the one you understand the best)? **Check one:**
   - ❑ English
   - ❑ Spanish
   - ❑ Korean
   - ❑ Mandarin
   - ❑ Cantonese
   - ❑ Other (Please specify) _____

8. Do you identify yourself as having any of the following disabilities? **Check all that apply:**
   - ❑ Blindness or visual impairment
   - ❑ Deafness or hard of hearing
   - ❑ Mobility impairment (a long-lasting condition that substantially limits your physical activities such as walking, climbing stairs, reaching, lifting, or carrying)
   - ❑ Mental or cognitive disability (a physical, mental, or emotional condition lasting 6 months or more that makes it difficult to learn, remember, or concentrate)
   - ❑ Other (Please specify) _____
   - ❑ None of the above

9. How frequently do you vote in political elections?
    - ❑ Frequently. I vote in every, or nearly every, political election.
    - ❑ Occasionally. I try to vote in every political election, but do not always vote.
    - ❑ Rarely. I rarely, if ever, vote in political elections.
    - ❑ I am a first time voter.

10. Are you registered to vote?
    - ❑ Yes, I'm registered in New York
    - ❑ Yes, I'm registered elsewhere
    - ❑ No

Thank you

If the person qualifies:

- Record their contact information.
- Schedule them into one of the blocks.
- Let them know that we will be starting on the hour, but that we'd like them to show up 15 minutes before their scheduled time so we can ensure their participation. We will not allow anyone into the study room after the study has begun.
- If they are waitlisted, let them know that they should still show up 15 minutes ahead of time. We will make every effort to include them, but if we cannot, we will still pay them.
- Send a confirmation email with directions.

If the person does not qualify:

Say, "Thank you for calling today, but the group that you fall into is full at the moment. If an opening becomes available may we keep your contact information on file to call you back?"

11. VENDOR GUIDELINES

Summary of Vendors Roles in NYSBOE Voting System Study

1. Vendors are required to provide AIR with a written script so that we may ensure the instructions for each system are provided in the most consistent manner possible. This script will serve as your brief introduction to participants on how to use your system. You may not demonstrate the actual voting process. Please only provide an overview of how to vote and the most important features voters need to know for voting. This introduction should be no longer than 1 minute. You should be prepared to discuss accessibility features, too, when needed.

2. You will not be allowed to hand out or display additional promotional or educational items because we cannot ensure that all participants will see these on an equal basis.

3. After your brief introduction, participants will have an opportunity to ask you additional questions or seek clarification. You must limit your response to the specific question and are not allowed to add any extraneous information because we do not want to bias any participants. If a researcher feels that your answer does not pertain to the direct question, you will be asked to stop your explanation.

4. After you have answered the participant's questions, please step to the side or back of the system. We do not want the participants to have added pressure from anyone or have any interference that would degrade the quality of this study. We've assured participants that their voting process will be private, so please respect their privacy.

5. Once we begin timing, you will no longer have the chance to help the participant or answer any questions unless called upon by the researcher to answer questions about the system's functionality. This helps us ensure that the times recorded are as objective as possible.

6. Vendors should provide two of each system. Vendors should provide only one operator per unique system. We understand that this may be challenging to provide instructions at two systems, so we will do our best to control the flow of people.

7. Please note that any deviation from these procedures will be noted by a researcher and the data related to the deviation will need to be dropped to make sure the study is as objective and credible as possible. Thank you for your cooperation.