

Usability Testing of Voting Systems

October 2003

Federal Election Commission

999 E Street, NW
Washington, DC 20463

(800) 424-9530
(202) 694-1095



Audio Ballot

Smartcard

Introduction

The Federal Election Commission's Office of Election Administration (OEA) assists state and local election officials by responding to inquiries, publishing research, and conducting workshops on all matters related to election administration. Additionally, the OEA answers questions from the public and briefs foreign delegations on the U.S. election process, including voter registration and voting statistics.

In 2002, the OEA launched an effort to ensure the usability and accessibility of voting systems. This initiative generated a comprehensive set of human factors standards for voting systems. It also produced the three guides listed below that facilitate the development, usability testing, and procurement of user-centered voting systems:

- Developing a User-Centered Voting System
- Usability Testing of Voting Systems
- Procuring a User-Centered Voting System

This guide, titled *Usability Testing of Voting Systems*, is written for voting system manufacturers who may be developing new systems and for election officials who may be procuring new voting systems. It presents a rationale for conducting usability tests, outlines the basic steps in the testing process, and shares lessons learned from prior test experience.

Overview

Usability testing is an established technique for evaluating the quality of human interactions with various system components, including equipment, documents, architectural elements, environmental factors, and other people. As such, the technique is appropriate for judging the suitability of voting systems relative to the needs of both voters and election officials.

Usability testing is an important step in the user-centered design process (see the companion guide titled *Developing a User-Centered Voting System*). The process also includes conducting user research, formulating usability goals, and applying established design principles to generate an effective user interface design.

System developers will conduct a series of usability tests in order to ensure user interface quality. A test conducted during the system design process is called a formative test because user feedback from testing helps to “form” the system. A test conducted once the system design is virtually complete is called a summative test because it provides a final “summary” of the system’s usability. A rigorous, user-centered design process includes both formative and summative usability testing as part of an overall test and evaluation effort.

In a usability test of a voting system, test administrators conduct structured observations of representative users performing key tasks. An example of a task performed by voters might be to cast a ballot. However, depending on the test objectives and the stage of system development, this integrated task might be divided into subtasks (e.g., voting for a candidate in a particular race) in order to enable closer examination of the users’ thought processes and actions. Examples of representative tasks performed by election officials might be to set up and properly configure the system, create a ballot, and count the vote.

Test administrators can measure voting system performance in several ways. Objective measures, such as the time to perform a task and the number of use errors committed, require one to simply observe a user’s interaction with the system. Subjective measures, such as perceived ease of use, call for the test participant to express his or her opinion using rating scales, for example.

Test administrators can measure voting system performance in several ways.

Some usability testing approaches are quite formal while others are less so. Some tests involve many test participants while others involve just a few. Selecting the best approach depends on the test objectives, the stage of system development, and the resources available for the effort. Generally, usability testing becomes increasingly formal and rigorous as the system design progresses from the conceptual stage to the detailed engineering stage. Similarly, the cost and time associated with testing increase along with the test’s formality and rigor.

Sometimes, system developers wait until they have a near-final design before they conduct a usability test. As a consequence, they can become locked into an overall design solution, limiting what they can do to resolve usability problems that may become evident during testing. Testing earlier in the development process is the best strategy for catching user interface design problems when they are easier and less expensive to correct.

Meanwhile, it may be the customer – individual state or local election officials, for example – that initiates a test. In such cases, the system is usually finalized, subject to customization, to fulfill specific needs. This makes it both feasible and advisable to conduct a more comprehensive test at that point.

Testing earlier in the development process is the best strategy for catching user interface design problems when they are easier and less expensive to correct.

It takes one to two months to complete a typical usability test. A test may take considerably longer if one chooses to conduct test sessions at multiple locations, such as urban, suburban, and rural communities in several states, in order to evaluate the system's suitability to the needs of different populations. Moving to the other extreme, a "quick and dirty" usability test intended to resolve a specific design issue can take just an afternoon.

Development teams need to build time into their development schedules for testing. Ideally, testing will occur in parallel with development efforts so that it does not become an obstacle to the progression of other technical activities.

Similarly, customers should also allot time in their procurement process for testing.

A usability testing effort normally includes the following steps:

- Establish test objectives.
- Write a test plan.
- Recruit the test participants.
- Set up the test environment.
- Conduct one or more pilot test sessions.
- Conduct the test.
- Consolidate and analyze the data.
- Document the results.

The balance of this guide presents reasons to conduct a usability test and describes each step in more detail.

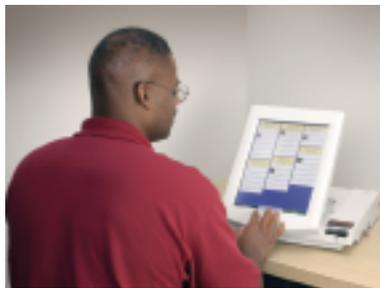
Rationale

SHOULD STATE AND LOCAL ELECTION OFFICIALS CONDUCT THEIR OWN USABILITY TESTS?

Perhaps. Election officials may choose to review the results of tests performed by vendors or by organizations responsible for voting equipment certification. Still, there are several reasons why election officials may choose to conduct their own usability tests.

First, they may be faced with a choice of several voting systems and wish to conduct their own comparative evaluation. Quantitative data regarding the amount of time required to cast a ballot and perceived ease of use can be factored into final procurement decisions.

Second, state and local election officials may have contracted with a vendor to develop a system with new capabilities that require validation. Equipment certification efforts may or may not fully address usability. Without adequate testing, significant usability problems could arise when the system is deployed.



Usability testing yields performance benchmarks enabling election officials to compare the performance of several available voting systems.

Third, they may want to evaluate the usability of one or more systems that they already have in use to see if there are any residual problems to resolve.

SHOULD VOTING SYSTEM DEVELOPERS CONDUCT USABILITY TESTS?

Yes. Usability testing is integral to design quality assurance and can generate usability improvements and associated marketing claims that give the developer a competitive advantage. So, there are compelling reasons to conduct usability tests.

Today, voting system developers must produce user-friendly solutions. Usability testing is essential to identifying user interface design shortcomings that could hinder users. Once identified, these shortcomings can be resolved. The discovery of usability problems late in the design process can be particularly problematic. Modifying the system design to alleviate a problem discovered at a late stage of development can be quite costly due to the need for expensive hardware and/or software changes. Such changes

can also delay an introduction to market.

Additionally, buyers can be quite sensitive to differences in usability. Therefore, the usability improvements that normally come from testing are essential to competitiveness, especially when many vendors offer systems with the same basic functionality.

Given the imperative to conduct usability tests, voting system developers need to decide whether they will utilize in-house staff or an independent vendor to conduct them. Using in-house staff to conduct the tests is usually expedient and may be less expensive, but using an independent vendor promises incrementally greater objectivity. Faced with the choice, some developers will ask in-house staff to conduct the formative tests, leaving the summative testing to outside experts.

Modifying the system design to alleviate a problem discovered at a late stage of development can be quite costly...

Test Objectives

As discussed elsewhere in this guide, there are many reasons to conduct a usability test, such as ensuring that the deployment of a new voting system goes smoothly. The general objectives of such tests might be to ensure that the voting system:

- Facilitates easy set up and configuration by system administrators.
- Simplifies the task of constructing suitable ballots.
- Guides voters through the entire voting process.
- Presents content clearly, regardless of whether the format is visual, tactile, or audible.
- Ensures that voters are able to cast their votes accurately and efficiently.
- Provides feedback during the voting process that enables voters to complete all required tasks and detect and correct any errors.
- Provides features that promote intuitive navigation among the content sections.
- Makes voters feel physically, intellectually, and emotionally comfortable and confident while participating in the process.

More specific test objectives may include:

- Evaluate the overall user interface design concept to determine if users are able to form a sufficiently clear understanding of their interactions with the system.
- Evaluate the layout of displays, controls, and on-screen information.
- Identify opportunities to enhance a preliminary design, such as changing terminology, shifting the order of information presentation and voter response, providing more or less feedback in response to voter inputs, increasing or decreasing the size of visual elements, etc.
- Measure the time required to perform tasks, recognizing that there are legal limits placed on the time allowed to vote in some jurisdictions.
- Observe what kinds of errors occur, including errors of omission and commission, and how well users are able to recover from them. For example, one may want to evaluate how well voters recover from an unintended undervote (i.e., failing to vote for a candidate in a particular race).
- Ensure that the voting system is not causing emotional or physical harm,

such as creating anxiety or feelings of inadequacy because users cannot understand how to perform a task or causing scratches due to contact with sharp edges on the equipment.

- Determine if the voting system effectively accommodates people with special needs. As a negative example, a voting system that uses a touchscreen may present touch targets that are too small for voters with dexterity limitations to touch easily. Or, as a positive example, a voting machine may incorporate assistive technology, such as synthesized speech output delivered via headphones, to present information to people with impaired vision or blindness.

The ultimate goal is to design a test that will effectively evaluate the voting system's key characteristics that influence user performance. For example, a test of a touchscreen-based voting system would pay close attention to the ability of users to reliably select one among an array of on-screen targets.

Test planners also need to establish evaluation criteria that are linked to design requirements, which in turn are linked to user needs and preferences. As discussed later, evaluation criteria may be based on objective (i.e., observable) measures, such as task times and task completion rates, or on subjective measures, such as ratings of display readability and overall intuitiveness.

Test Plan

A comprehensive plan helps usability test sessions go smoothly. It also helps generate visibility for usability initiatives within the development organization, thereby elevating the importance of the design assessment activity and ensuring that usability goals are achieved. A typical test plan might address the following topics:

- Objectives that may have a usability and marketing focus.
- Environment, including special furnishings, lighting, and equipment.
- Test administrators and their particular roles during the test.
- Participants, including the method of their recruitment and protection from emotional and physical harm.
- Test activities, which is essentially an agenda for the test session that describes steps such as greeting the test participant, directing the participant to attempt specific tasks, and conducting an exit interview.
- Tasks that may be quite broad or limited in order to focus on the voting system's salient interactive characteristics and aspects of concern.
- Performance measures such as task times, task completion rates, and ratings and rankings.
- Data analyses that may be limited to rudimentary statistical analyses or be extended to include more advanced statistical analyses.
- Reporting of test findings, which may be summarized in a bulleted list or presented more comprehensively in a detailed report.
- Schedule for conducting the test.

All interested parties, particularly individuals who will need to deal with proposed design changes derived from the test results, should review and comment on a draft test plan. The final test plan should reflect appropriate changes in response to comments. Additional changes may come in response to shifting test objectives, logistical constraints, and pilot test results.

Test Activities

Usability test sessions typically last one to two hours, even if specific user tasks take a matter of minutes or seconds. This is because test administrators spend a significant portion of the time orienting the test participants and soliciting their detailed feedback on their interactive experiences.

The typical order of activities is:

- Greet the test participant, thanking him or her for taking the time to participate in such an important activity.
- Administer a background questionnaire that solicits demographic information, such as the participant's age, education level, level of experience working with computers, and familiarity with specific voting systems.
- Orient the participant to the test environment and process, and reassure him or her that the assessment effort is focused on the system's capabilities rather than his or her own capabilities.



A typical usability test laboratory includes a test room and an adjacent observation room, connected by a one-way mirror.

- Ask the participant to review and sign an informed consent form that explains his or her rights as a participant.
- Train the test participant to “think aloud” and explain how to complete rating and ranking exercises. There is a tradeoff between asking the test participant to think aloud and asking the participant to behave more naturally (i.e., perform tasks without talking). The first approach – thinking aloud – provides better diagnostic information about the user interface design. The second approach – having the test participant work silently – yields more accurate task time data and may produce the

most realistic test participant behavior. Accordingly, formative tests conducted during system development usually include thinking aloud while summative tests aimed at determining final system performance typically do not.

- Ask the test participant to explore the system’s user interface(s). Then, ask the test participant to describe his or her first reaction to the system, or even give it a letter grade (e.g., an A– versus a C+). This is a common approach to determining users’ initial impressions of a system they are seeing for the first time but not yet using to complete tasks. This step may be more appropriate for the assessment of system elements used by election officials who may “check out” the system before attempting specific tasks. It may not be an appropriate step for evaluating voter interactions with the system because voters are unlikely to explore the user interface before starting a task. Rather, they are likely to proceed directly through the voting process.
- Prompt the test participant to perform specific tasks. When testing the voter side of the system, the sole task may be to cast a ballot. While the user performs each task, the test administrator(s) as well as informed observers should take note of the test participant’s significant comments, requests for assistance, and observable errors.
- Ask the voter to rate completed tasks according to selected usability attributes, such as ease of use, task speed, ease of detecting and correcting any errors, confidence that he or she performed the task correctly, and overall satisfaction with the process. Some examples of ratings include:

How confident are you that you performed the task correctly?						
Not at all confident	1	2	3	4	5	Very confident
How satisfied are you with the overall voting process?						
Not at all satisfied	1	2	3	4	5	Very satisfied

- After the test participant completes each task, ask him or her to suggest ways to alter the user interface in order to make that particular task easier.
- Conduct an exit interview during which the test participant can offer his or her final assessment of the system and make additional suggestions on how to improve it.
- Complete the testing process by thanking the test participant for his or her time and feedback on the system.
- Compensate the participant for his or her time.

Test Participants

Most usability tests are conducted during the formative stage of system development. They usually involve a small number of test participants – often no more than a dozen people – because the overall test objective is to identify usability problems, rather than to determine their frequency of occurrence.

A summative usability test, conducted later in the development process and aimed at validating a system’s usability, may warrant a larger sample of perhaps 30 or more people. A larger sample enables more advanced statistical analyses, such as an analysis of variance. It also increases the chances of observing low-probability events, such as a particular use error, and lends greater credibility (i.e., face validity) to the test results.

If the test sample is small, it is better to recruit a diverse set of individuals rather than trying to proportionately match the characteristics of the general population of voters. If the test sample is larger, it makes sense to recruit a more proportionately representative sample. When running a single test on equipment intended for use by the general public, it is important to include individuals from populations with distinct characteristics, such as:

- Seniors (older individuals).
- People who are vision impaired.
- People who are hard of hearing.
- People who have mobility limitations.
- People who have learning and/or reading disabilities.
- Young adults who have never voted before.

Alternatively, one may conduct separate tests of the general purpose equipment and the equipment designed to accommodate people with special needs.

In accordance with the appropriate guidelines for human subjects protection, test administrators should take care to guard the test participants against physical and emotional harm. In fact, government and/or institutional rules and regulations may require test administrators to pass the test plan through an Internal Review Board before participant recruiting and testing can proceed.

Prior to recruiting participants, it helps to create a screener (i.e., recruiting script) that provides a list of questions for qualifying the prospects.



A visually impaired voter listens to her voting options through headphones.

It may be convenient to use “insiders,” such as company staff, as test participants. However, this practice is generally discouraged. The chance of positive or negative bias toward the company’s system is simply too great. Moreover, the use of insiders reduces the credibility of the test results, even if the chance of obtaining biased results is low.

Finding the right kind of participants outside a company’s walls requires a bit more work. There is the temptation to hedge on the principle of not recruiting insiders by recruiting their friends and family members instead. However, the best practice is to recruit individuals who have no affiliation with the development organization, thereby avoiding any conscious or unconscious bias. Common recruiting methods include cold contacting people listed in directories, intercepting people in retail locations, placing recruitment advertisements in newspapers, placing recruiting posters in community centers, and contacting organizations that serve the interests of special populations. Often, usability testing vendors will have a database containing the names of people interested in participating in usability tests, making recruiting more efficient.

Typically, test administrators compensate test participants for their time and travel expenses. Common forms of compensation include money and gift certificates. In keeping with human subjects protection goals, it is customary to ask participants to review and sign a statement of the test participant’s rights. Such statements provide an overview of the test process and instruct the participant that he or she has the right to withdraw from the test at any time without forfeiting compensation.

... the best practice is to recruit individuals who have no affiliation with the development organization, thereby avoiding any conscious or unconscious bias.

Test Environment

An effective usability test of a voting system can be conducted in all sorts of environments, ranging from an office to a conference room to an actual precinct. The goal is to choose (or create) a setting that is reasonably representative of the actual use environment – one that exposes the test participants to the various elements (e.g., noise and lighting conditions) that could influence their interactions with the voting system.

One can conduct an informal usability test – one intended to resolve specific design issues – virtually anywhere there is space for the necessary equipment. Therefore, an office or conference room would be quite adequate.

Usability testing laboratories and focus group facilities provide a convenient setting for tests focused on user interactions with the voting system’s hardware, software, and documents. Such facilities usually include a pair of rooms connected by a one-way mirror that facilitates video recording and unobtrusive observation by interested parties. They are well suited to comprehensive testing.



A view of a usability test from the observers’ perspective.

An actual voting center, such as a community meeting hall, is probably the ideal test setting in terms of maximizing environmental realism. However, such settings may make it harder to observe proceedings unobtrusively.

Test Sessions

Typically, one to two trained individuals can run a smooth usability test. Taking the two-person team approach, one administrator directs and solicits information from the test participant while the other one focuses on collecting data. However, some scenarios may require additional staff to play supporting roles in order to simulate an actual voting center.

It is common to video record the test sessions. Video recording enables people who were not able to witness a particular test session to watch it at a later time. Video recording also enables detailed analyses of significant user interactions that may happen too quickly to analyze in real time.

Conducting one or more pilot test sessions provides the means to validate the test plan and give the test administrators practice at running the test sessions smoothly. Pilot sessions may be scheduled to take place hours, days, or weeks in advance of the “official” sessions, depending on the test’s formality and complexity. Obviously, a longer period between pilot and formal testing provides more time to refine the test plan and associated materials (as needed).

In a usability test laboratory, the test administrators can stay with the test participant in the test room or direct the test from the observation room, communicating via an intercom. The preferred practice is to place the participant alone in the test room to perform tasks and respond to questions and rating exercises. This approach alleviates the tendency of test participants to ask the test administrator for assistance with tasks. Meanwhile, initial and follow-up interviews are more successful when a test administrator joins the participant in the same room.

Sometimes, test participants become frustrated while trying to perform a task. Ultimately, there is an appropriate time to either offer assistance or stop the task and move on to the next one. The need to provide assistance to the test participants indicates a possible user interface design problem and should be noted. Providing assistance when absolutely necessary enables the test to continue on to address other user interface elements. If it makes more sense to simply stop the task and move on to the next one, the test administrators should establish a time limit (five minutes, for example) for each task.

The preferred practice is to place the participant alone in the test room to perform tasks and respond to questions and rating exercises.



A test participant selects on-screen options based on audio prompts.

Test administrators will be busy collecting the following kinds of data throughout the test sessions:

- Responses to an initial questionnaire or interview focused on demographic factors, prior experience, and current opinion.
- Anecdotal remarks made while performing tasks.
- Task times.
- Ratings of the various user interface attributes, such as intuitiveness, ease of use, error prevention and recovery, and task speed.
- Initial and final grades of the system as a whole.
- Responses to a final questionnaire or interview focused on likes and dislikes, as well as suggestions for system improvement.

Data Analysis

The appropriate level of data analysis depends on the test objectives and the scale of the test.

When a test involves a dozen or fewer participants, data analyses are usually limited to descriptive statistical analyses, such as calculating means and standard deviations for the various quantitative measures. A considerably larger sample is normally needed to perform inferential statistical analyses, such as an analysis of variance.

Inferential statistical analyses help determine if the test results would hold up if the same kind of test was conducted with a larger population sample. However, the benefits of such analyses need to be weighed against the cost of extensive sampling. Moreover, simple statistical analyses are usually sufficient to identify and understand the major usability issues. Most system developers choose to fix or dismiss usability problems that arise in testing based on basic statistics and engineering judgment, rather than on complex estimations of problem frequency and severity.



A sample summary of rating data from a usability test of a voting system user interface.

As discussed earlier, most usability test professionals are comfortable running tests with relatively small population samples during system development. Their comfort stems from the practical view that a small-scale test will uncover most, if not all, major usability problems – the main point of conducting a formative test. For a summative test aimed at validating a system or making a comparison among systems, it may be appropriate to engage as many as 30 or more test participants in order to produce statistically significant findings.



A researcher interviews a test participant after she performed a series of hands-on tasks.

Anecdotal comments, such as “The ballot layout was really clean – organized looking” or “I didn’t know exactly what to do after I chose a candidate” can be just as valuable as quantitative data. Sometimes just a single anecdotal comment is enough to identify a possible usability problem. Certainly, many negative comments about the same usability problem are enough to confirm its persistence. As such, it is worthwhile to organize anecdotal comments into related groups and review them in order to identify unique issues and patterns.

...simple statistical analyses are usually sufficient to identify and understand the major usability issues.

Findings

Voting system developers in the early stages of design may seek detailed design recommendations. Such recommendations may address the overall conceptual framework or model, style of user interactions (e.g., using a touchscreen versus mechanical method of user input), or layout and appearance of the hardware, software, and paper-based user interfaces. At a later stage of design, developers may seek data that validates the usability of their system and provides the foundation for marketing claims.

Voting system customers, such as state and local election officials, may seek data describing a particular system's usability as compared to others. Or, they may seek confirmation that design changes made to accommodate their specific requirements did not introduce new usability problems.

Traditionally, usability test specialists document their findings in a detailed memorandum or report. Executive summaries, PowerPoint™-style presentations, and videos highlighting the more informative moments from various test sessions are also common deliverables.

It is common to produce a highlight video capturing the more interesting and informative moments from the test sessions.

Many test reports separate findings into general and specific ones. General findings address usability issues that apply to the user interface as a whole, such as problems finding one's way around an assortment of software screens. Specific findings address usability problems associated with a specific user interface element, such as an ambiguous icon or vague prompt. Particularly in the case of formative testing, it is helpful to present both types of problems in a format that names the problem, describes the problem in detail, recommends how to resolve the problem, and prioritizes the resolution (e.g., must fix, nice to fix, not important to fix).

Traditionally, usability test specialists document their findings in a detailed memorandum or report.

It is important to note that a formative test report should include positive as well as negative findings. Summative test reports often report the data only, excluding design recommendations because it is too late in the development schedule to make changes.

It is common to produce a 15-20 minute highlight video capturing the more interesting and informative moments from the test sessions. Such videos add a sense of credibility and gravity to written findings.

Additional guidance on how to report the results of a usability test are provided in ANSI NCITS 354-2001, Common Industry Format for Usability Test Reports (see

<http://zing.ncsl.nist.gov/iusr/>). The ANSI standard calls for a report that includes the following sections:

- Description of the product.
- Goals of the test.
- Test participants.
- Tasks the users were asked to perform.
- Experimental design of the test.
- Method or process by which the test was conducted.
- Usability measures and data collection methods.
- Numerical results.

One advantage of using the common report format is that it facilitates a comparison of test findings associated with several voting systems, particularly when different groups have performed the tests and prepared the reports.

Scaling the Testing Effort

Voting system development is a complex process that involves many tradeoffs, including deciding how much usability testing is sufficient to ensure a high-quality solution. Unfortunately, there is no definitive answer regarding how much to test and what to test. To some extent, the decision to conduct a greater or lesser number of usability tests is a matter of organizational confidence in the evolving design and the organization's philosophical commitment to user-based testing.

Nominally, developers should conduct at least several cycles of usability testing. Early testing may focus on alternative design concepts, which may be rendered as relatively primitive, "paper and pencil" mockups, or preliminary, computer-simulated prototypes. Later testing may focus on identifying residual usability problems associated with a working prototype.

However, there is a considerable jump in realism between "paper and pencil" mockups and computer-based prototypes, never mind a working prototype being prepared for field testing. That is why developers who seek to reduce their financial and technical risk often choose to conduct several rounds of usability testing, increasing the number of opportunities for users to point the user interface design in the right direction. The extra testing represents an upfront investment that should pay off in the long run.

A strong testing program might include the following tests:

TEST	GOAL	PARTICIPANTS
#1	Evaluate a set of preliminary, conceptual designs on paper to identify the preferred concept or an appropriate hybrid of multiple concepts.	6 to 12
#2	Evaluate the preliminary design, which may take the form of a computer simulation and/or physical mock-up.	6 to 12
#3	Evaluate the refined design, which may take the form of a working prototype.	12 or more
#4	Evaluate the near-final design, which may take the form of a prototype that is ready for field testing.	30 or more

Tips

Here are several tips drawn from experienced usability test specialists:

- Given limited resources, it is better to conduct several modest-size usability tests throughout the system development process rather than a single, larger usability test near the end of development. However, voting system developers may still be compelled by certification and marketing goals to conduct a large, summative usability test.
- Ideally, usability tests should be performed by usability specialists. While it is possible for non-specialists to run a test, they are less likely to diagnose certain usability issues that may be evident to experienced individuals. That said, usability tests conducted by non-specialists are still worth while.
- The differences between individuals tend to outweigh differences attributable to geographic factors. Therefore, it can be just as effective and more efficient to test in a single geographic location. However, this approach may lose credibility among individuals accustomed to broad-based sampling.
- Formative usability tests conducted early in the design process tend to produce dramatic results within the first few test sessions. However, caution should be taken to avoid generalizing from the results of the first few test sessions.
- Sometimes, people question the credibility of results drawn from a test involving just a dozen or so test participants. Tests involving 30 or more test participants tend to have more credibility among non-specialists.
- Usability tests focused on developing marketing claims – among other goals – should be conducted by independent experts.

The following books provide additional tips on usability testing:

Dumas, J. & Redish, J., (1999), *A Practical Guide to Usability Testing*. Portland, OR: Intellect.

Rubin, J., (1994), *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests*. NY: John Wiley & Sons.

ACKNOWLEDGEMENTS

American Institutes for Research (www.air.org) prepared this guide for the Federal Election Commission under the terms of Contract No. FE-2-AC-0024.

American Institutes for Research wishes to acknowledge the following individuals for their contributions to the guide and the project as a whole.

Project staff:

Michael Wiklund, American Institutes for Research (project director and guide author)
John O'Hara (consultant)
Stephen Berger, TEM Consulting (consultant)

Federal Election Commission staff:

Penelope Bonsall (contract technical monitor)
Peggy Sims
Brian Hancock
Bryan Whitener

Advisory Board members:

David Baquis, U.S. Access Board
Barry Beith, HumanCentric Technologies
Pauline Brunelli, Federal Voting Assistance Program
David Capozzi, U.S. Access Board
Paul Craft, Department of State, Florida
Lance Gough, Chicago Board of Elections
Linda Lamone, State Board of Elections, Maryland
Sharon J. Laskowski, National Institute of Standards and Technology
Susan Palmer, Cingular Wireless
Whitney Quesenbery, Whitney Interactive Design
Hans A. von Spakovsky, U.S. Department of Justice
Scott Wiedmann, Federal Voting Assistance Program

