review

Improving voting systems' user-friendliness, reliability, & security

Michael D. Byrne

abstract

About half of Americans have limited confidence that their vote will be properly counted. These fears have focused attention on voting system reliability, security, and usability. Over the last decade, substantial research on voting systems has demonstrated that many systems are less usable and secure than they should be. Producing truly reliable voting systems demands more than just following the federal guidelines enacted in 2005 (which, although well intentioned, have failed to substantially improve current systems) or simply updating voting systems to electronic voting computers using monies allocated by the 2002 Help America Vote Act (HAVA). In fact, HAVA has inadvertently led to the purchase of systems that may have actually increased the vote error rate. Key reforms needed to deliver reliable voting systems include substantial testing for usability, especially regarding the accurate capture of voter intent and the reduction of voter error rates, and measures to ensure the integrity of elections, such as election officials' ability to secure ballots.

Byrne, M. D. (2017). Improving voting systems' user-friendliness, reliability, & security. *Behavioral Science & Policy, 3*(1), 15–24.

Throughout the 2016 presidential election season, dark claims were floated about the election being rigged, and almost half of all Americans have limited confidence that their vote will be properly counted, according to an October 2016 survey.[1] These fears focus attention on the voting procedures and systems used in the United States. Are they, in fact, fair, and do they give all citizens a voice, as the Constitution requires? And in the wake of the vote recount efforts by Green Party candidate Jill Stein and the Clinton campaign, with both camps voicing concerns of potential computer hacking,[2] Americans may yet wonder: are their votes secure?

Core Findings

What is the issue?
After the Help America Vote Act passed in 2002, many states transitioned to direct recording electronic voting machines. While these improved usability for disabled voters, they did not improve error rates over traditional voting methods. Further usability and security testing is needed to improve the integrity of U.S. elections.

How can you act?
Selected interventions include:
1) Joining electoral machine research and design efforts in Travis County (Texas) and LA County (California), and across other jurisdictions, to pool resources
2) Embedding behavioral science research and insights in electoral design processes

Who should take the lead?
Behavioral science researchers, security experts, election officials

The voting process has been questioned before, particularly following the contested presidential election of 2000 and the infamous butterfly ballot (see the sidebar *The Butterfly Ballot From Palm Beach County, Florida, 2000*). Two years later, with strong bipartisan support, Congress passed legislation called the Help America Vote Act (HAVA) of 2002 to address election administration problems.

HAVA allocated billions of dollars to local jurisdictions to replace outdated voting equipment. But it turned out that many of the voting machines those jurisdictions rushed to purchase, most often voting computers known as *direct recording electronic* machines (DREs), offered little to no improvement. In fact, HAVA likely made usability worse for some voters in terms of preventing voter errors, because some of these replacement systems were measurably worse than traditional paper ballots, the best alternative then available.[3]

The fundamental problem with HAVA is that it put the need for purchases ahead of the science. The law imposed substantial pressure on county clerks to purchase new voting systems and granted them generous budgets to do so, yet it offered almost no scientific evidence to guide them on which systems were most usable and most secure. Commercial vendors, hungry for an allotment of the billions about to be spent, rushed in with poorly designed systems. These early systems were primarily DREs. They were not only scientifically unproven to enhance voting usability but also failed to follow industry best practices for both usability and computer security that had been established in the decades prior.[4]

However, there were positive consequences as well. The contested 2000 election spurred a wave of new research on many aspects of voting, including voting system usability, election administration practices and procedures, computer security, and statistical auditing methods.

For example, the Caltech/MIT Voting Technology Project (http://vote.caltech.edu/), an interdisciplinary research effort focused primarily on political science, has produced a substantial amount of valuable research on voting, particularly on election administration. For example, the idea of a residual vote—the difference between the total number of ballots received and the total number of votes cast in a particular race—came from this research and has now become a standard measure of the quality of voting systems.

The ACCURATE Center (http://www.accurate-voting.org/), a 6-year interdisciplinary research center funded by the National Science Foundation, focused instead on both computer security and voting system usability. The center is responsible for the vast majority of the research on voting system usability published since the center's inception in 2005. In addition, the center's research has yielded ideas that will likely be incorporated into the security and cryptography architectures of future voting systems.

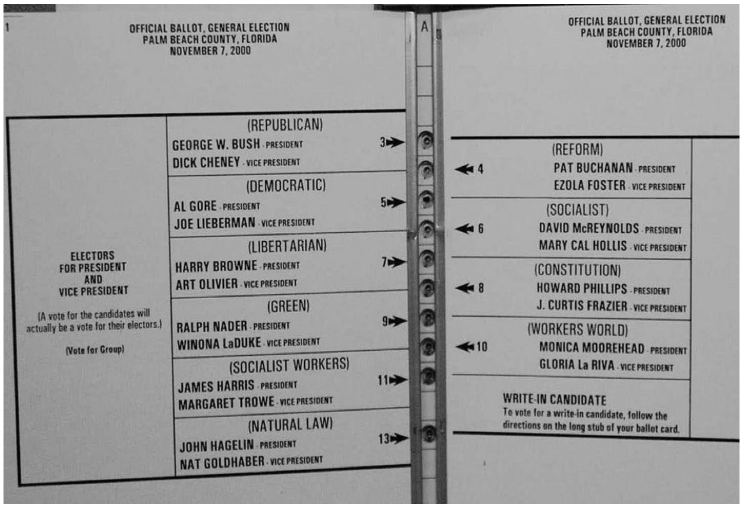
In this article, I focus on usability, but usability is not by any means the only important consideration. A truly successful voting system must address multiple factors, such as security, usability, accessibility, certification, ease of administration, cost, compliance with election laws, transparency, and auditability. Clearly, improving and updating the country's voting methods, practices, and administration is no simple task.

Human Factors of Voting Systems

Human factors is an academic discipline concerned with matching engineered systems to human capabilities. A human factors researcher,

The Butterfly Ballot From Palm Beach County, Florida, 2000

This confusing ballot played a key role in the 2000 U.S. presidential race. Many voters inadvertently chose Buchanan/Foster when intending to vote for Gore/Lieberman, and even more voters failed to cast a valid vote in the presidential race. A follow-up behavioral study using a paper ballot and Canadian participants (who were unaware of the Palm Beach ballot) showed that they made similar errors on 8% of the ballots and rated the butterfly ballot more confusing than a single-column ballot.[A]



A. Sinclair, R. C., Mark, M. M., Moore, S. E., Lavis, C. A., & Soldat, A. S. (2000, December 7). An electoral butterfly effect. *Nature, 408,* 665–666.

for example, might investigate how to design a website so its visitors can quickly and easily find what they are looking for. Most human factors academic programs are housed in psychology departments, although some are in industrial engineering programs or schools of information. Human factors started as a field of study in the 1950s, primarily in aviation, where practitioners investigated how to reduce deaths by designing cockpits and training pilots to match their capabilities with what the engineers who built the airplanes assumed they were.

Usability and human error are human factors problems. A 2004 report issued by the National Institutes for Standards and Technology (under the direction of the Federal Election Assistance Commission) acknowledged that there was a dearth of research data on voting system usability, noting that virtually no significant human factors studies of voting systems existed.[5]

An international standard for usability measurement provides a three-component definition for usability: effectiveness (the accuracy or completeness users achieve with specific goals in the target environment), efficiency (the resources expended by users in relation to the effect achieved), and satisfaction (a subjective rating of the system).[6] Table 1 applies these universal usability definitions specifically to the voting context.

Laboratory Studies on Legacy Voting Systems

In the mid-2000s, my colleagues and I did some of the first systematic studies designed to assess the usability of legacy voting technologies.[7,8] (For these studies and other research done by me and my colleagues, voting-age adults were recruited from the Houston, Texas, area through a mix of newspaper and Internet advertising. These samples were generally close to balanced on gender. They represented age ranges from roughly 20 to 80 years and contained people representative of a broad mix of other demographic variables, such as education and ethnicity. For details, consult the individual cited papers.)

First, we assessed voter error in each of three traditional voting systems: paper ballots, punch cards, and lever machines. Voter error occurs when a voter casts a vote for someone other than

Table 1. Usability of voting systems

Component	Definition	Note
Effectiveness	Voter error rate	Requires knowing the voter's actual intent. Therefore, for privacy reasons, the voter error rate cannot be studied in real elections; it can only be studied in the laboratory, where the voter's intent can be known.
Efficiency	Time needed to fill out and cast a ballot	
Satisfaction	Subjective user satisfaction	Generally measured by a subjective usability questionnaire such as the System Usability Scale (SUS).

“the voting technology used had no effect on how quickly voters cast their ballots”

the person whom the voter intended. Voter error also occurs when a voter fails to cast a vote when she or he intended to. These errors are impossible to assess in real elections because ballots are secret. In the laboratory, however, we were able to assemble diverse groups of voting-age adults and assign them to participate in a mock vote with each of the three voting systems.

We determined voter intent using one of two methods. We gave mock vote participants either (a) a list of candidates to vote for and measured errors as deviations from that list or (b) a booklet with names and descriptions of fictional candidates and then asked them to cast their ballots in all three systems but vote the same way on each ballot. When a voter cast a vote that did not match his or her other ballots, we counted that ballot as having an error.

We determined voter error rate, the measure of a voting system's effectiveness, by comparing each voter's intent with the actual vote that person cast. We measured a voting system's efficiency by tracking the time it took each person to cast a ballot. We also measured voter's satisfaction with a system using the System Usability Scale (SUS), a standard usability questionnaire.[9]

The key findings were straightforward: The voting technology used had no effect on how quickly voters cast their ballots, but it did affect their error rate and user satisfaction. Error rates with paper ballots averaged 1%–2%, which was markedly lower than the error rates produced by punch cards and lever machines (typically around 3%–4%, but sometimes even higher than that). Voters also indicated via the SUS questionnaire that they were somewhat more satisfied when voting with paper ballots than with punch cards and lever machines.

We also discovered that when voters have a list in hand of whom to vote for, they make fewer errors, regardless of the technology used. This is most likely because it is easier to work directly from a list and not from memory—an important distinction, because many voters do not bring lists into the voting booth. In fact, in some jurisdictions, it is illegal to do so.

Electronic (DRE) Voting Problems

As we were conducting our studies on the usability of legacy voting system, other research teams were investigating the new commercial DREs that flooded the market after HAVA become law.

A research team led by Paul Herrnson, a professor of political science now at the University of Connecticut, conducted a large study comparing the most popular commercial DREs available on the market.[3] They measured voter error by giving mock vote participants a list of candidates and measuring how often their actual votes diverged. They found that even the best commercial DREs were no better than paper ballots and most were worse, some substantially so.

HAVA mandated that voters with special needs be given an accessible way to vote. Commercial DREs are more accessible than paper ballots, punch cards, and lever machines, all of which are essentially impossible to use by voters with various disabilities, such as blindness or substantial motor impairments. The accessibility features (mostly audio-based presentation of the ballot) of these early DREs were quite poor by modern standards,[10] but they did allow jurisdictions to comply with HAVA's accessibility mandates. After HAVA, some jurisdictions combined paper ballots and DREs, whereas others moved entirely to DREs. These changes carried other costs. County clerks essentially had to become information technology managers, a new role for them. Furthermore, in some cases these changes likely led to more voter errors if paper ballots were replaced with DREs. Therefore, although DREs may make the physical act of casting a vote easier for people with certain disabilities, they are not necessarily better for the general voting population, at least compared with paper ballots.

In other words, voting systems became measurably less usable in jurisdictions that moved from paper voting to early DREs after HAVA became law in 2002. Jurisdictions that moved from punch cards or lever machines to DREs generally did not take as big a step backward, but overall, the change was not always progress.

"County clerks essentially had to become information technology managers"

Problems With Early DREs

To improve commercial DREs, it is first necessary to figure out what makes most of them so difficult to use. Some fail to conform to simple guidelines about text size and readability. Some require voters to follow novel and unusual procedures. Others have poor touchscreens, confusing instructions, or other complications. Almost every DRE on the market in 2008, when Herrnson's team conducted their study, has not one but multiple usability problems.

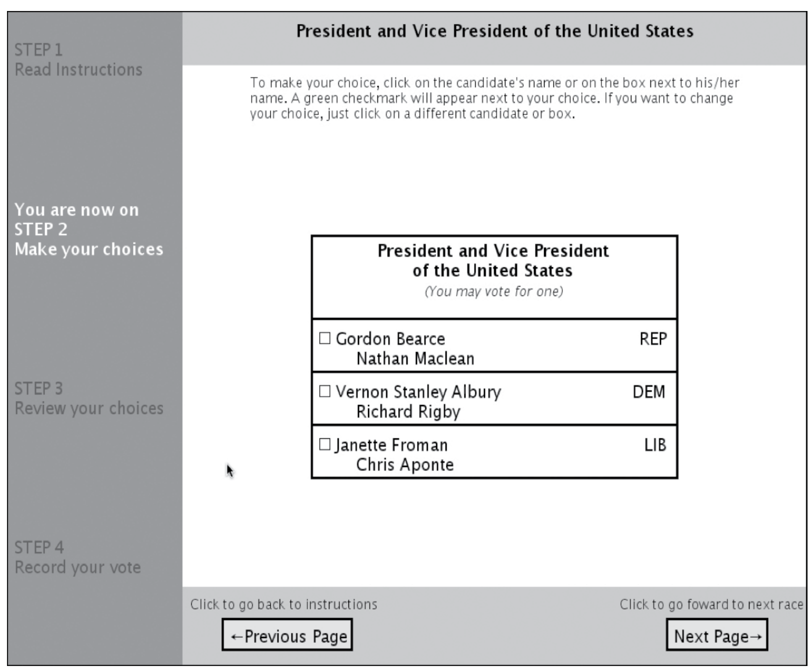
To understand why DREs are difficult to use, my colleagues and I constructed a DRE for research purposes. Like all other DREs, this one, called VoteBox, is a computer system. Unlike most other DREs, VoteBox is not a touchscreen; voters interact with it while either sitting or standing by clicking on buttons using a standard computer mouse. (See one of the VoteBox screens in Figure 1.)

VoteBox was intended to be a better DRE than many early commercial DREs, which did not conform to federal usability guidelines issued in 2005.[11] These guidelines, The Voluntary Voting System Guidelines, called for voting systems to meet basic criteria for usability and security. They also identified a minimum text size, a minimum contrast, and other features to help make the systems more usable.

The VoteBox DRE met these minimum standards, but it went no further. How did the VoteBox DRE compare with traditional systems? To examine the causes of differences in voting behaviors, we conducted a laboratory study in which we randomly assigned mock vote participants to vote on a medium-length ballot (27 races) twice: once using the VoteBox DRE and once using a traditional system (either a bubble-style paper ballot, a lever machine, or a punch card).[12] We also randomly assigned half of each group to vote from a list of preferred candidates, whereas the other half chose from a booklet with the names and descriptions of fictional candidates. We instructed those in the booklet group to choose candidates and vote the same way on multiple ballots. And to control for the order of voting (DRE vs. other system), we had half the voters vote first with the DRE and half vote first with the traditional system.

As it turned out, our in-house DRE was no more effective or efficient than the traditional systems. It took at least as long to vote on the DRE as it did on other technologies. (More educated voters vote slightly faster overall, regardless of technology—a result often seen in such studies, and here we found the same.)

Figure 1. Screenshot from the VoteBox research direct recording electronic voting system



The VoteBox DRE did not reduce the error rate compared with paper ballots. In fact, when we compared VoteBox or one of its variants with paper ballots in our subsequent studies, the two had a similar error rate of roughly 1.5%. The results demonstrate that simply following basic usability guidelines can help improve usability, but that alone is not enough to do better than the best legacy technology, paper.

Advantages of DREs

Paper ballots, although not very accessible, produce a record that is readable by humans, less vulnerable to malicious electronic tampering, and auditable later. Yet despite their drawbacks, DREs have some advantages over paper ballots. Even when voters make errors, interpreting the submitted ballot in a DRE is unambiguous, whereas interpreting a paper ballot is not. Consider the 2008 Senate election in Minnesota. A razor-thin margin of victory caused statewide recounts, and the two major political parties spent months contesting ambiguous paper ballots, such as the ones shown in Figure 2. (An excellent resource full of examples from this election can be found at http://minnesota.publicradio.org/features/2008/11/19_challenged_ballots/round1/.) Although DREs might not improve voter error rates, they also do not lead to such complications.

The one way that VoteBox differed consistently from legacy systems in our experiments was in satisfaction: Voters repeatedly rated VoteBox as more satisfying to use than traditional systems. For example, using the SUS—the same standard usability questionnaire my colleagues and I used in all of our studies—our voting experiment participants rated VoteBox as substantially more satisfying to use than bubble-type paper ballots, lever machines, and punch cards. The satisfaction scores are, in fact, unusually high for engineered systems of any kind, and the results held for young and old voters, computer experts and computer novices, rich or poor, and similarly wide ranges on other demographic variables.

More critically, those who used VoteBox were more satisfied than were those who used traditional systems, regardless of which system allowed them to vote faster or make fewer errors (see Figure 3). This has two implications. First, just because voters like a system does not mean it generates lower error rates or allows people to vote in less time. Second, election officials who move away from DREs may find that their voters dislike the change.

Figure 2. Ballots from the 2008 Minnesota Senate race

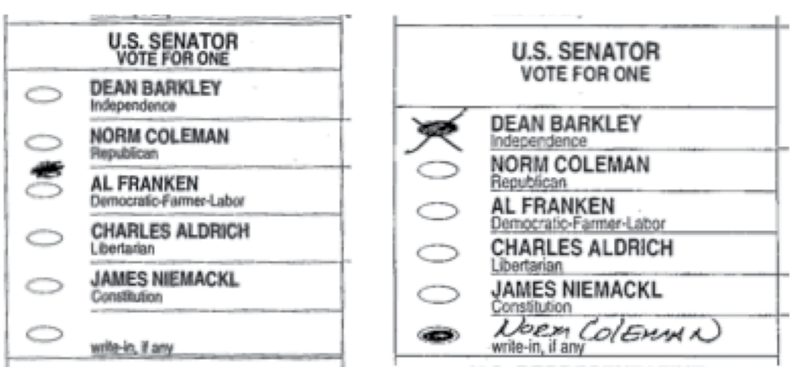
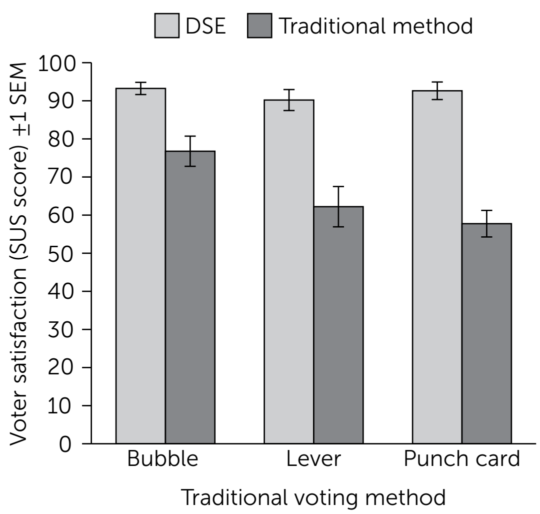


Figure 3. Voter satisfaction by voting system



Voters report being significantly more satisfied with the voting process when using direct recording electronic (DRE) voting machines than when using the following traditional methods: paper bubble ballots, lever machines, or punch cards. Voter satisfaction was measured by a standard survey called the System Usability Scale, which runs on a scale from 0 to 100, and the values shown are the mean across the voters. The error bars show the standard error of the mean—a statistical measure of how widely the results varied from voter to voter. Data come from "Electronic Voting Machines Versus Traditional Methods: Improved Preference, Similar Performance," by S. P. Everett, K. K. Greene, M. D. Byrne, D. S. Wallach, K. Derr, D. Sandler, and T. Torous, 2008, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: CHI 2008,* New York, NY: Association for Computing Machinery.

Ensuring DRE Security & Accuracy

Some voters have stated that they like voting on DREs because after they have navigated through the ballot, they can review those choices on the last screen before submitting their vote. **If the voting machine software was malfunctioning—or, worse, maliciously altered—would voters notice the altered votes on the review screen?** We have done multiple studies showing that most of the time, voters do not. In fact, **roughly two-thirds of voters failed to notice changes, even though the study used a permissive standard for what counted as noticing the change.**[13] When voters were asked if they noticed anything amiss on the review screen, they got credit for detection if they said that something was wrong, even if they could not articulate what it was. What this suggests is that security measures that depend on voters thoroughly checking their ballots are unlikely to be completely effective.

One of the earliest proposals for improving the security of DREs was to have the machines also print out a paper record that voters could examine through glass.[14] These records are generally produced by inexpensive thermal printers—imagine a low-quality, light purple credit card receipt. If voters do not notice anomalies on the high-resolution display immediately prior to casting their vote, it seems highly unlikely that they would notice them under even worse visual conditions. This suggests that other security measures are necessary.

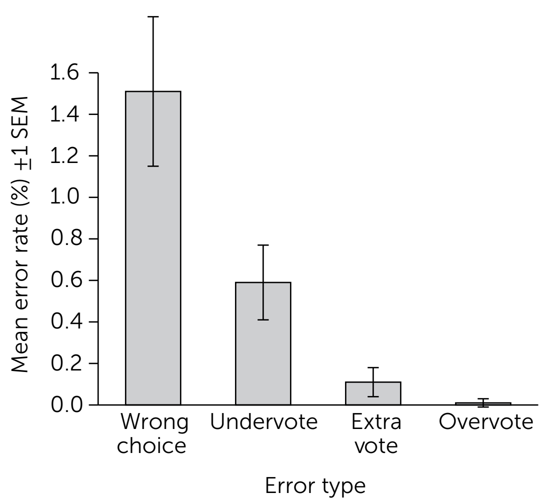
We tested two other interface manipulations in our experiments and found that there was little difference in error detection rates based on where the votes were on the ballot or the number of votes that were altered.[15] Changing the interface to highlight party affiliation and missing votes helped a bit, but even in the best case, this brought detection rates up to just 50%.

Instead of relying on voters to detect their own errors, sometimes errors can be detected using the residual vote rate.[16] Residual votes occur when voters fail to cast a vote or when they invalidate their vote, for example, by selecting two candidates in a contest where only one is allowed. When the residual vote rate is unusually high, this can alert election officials that something went wrong.

However, some residual votes do not indicate voter error. When a voter abstains on purpose—perhaps because he or she doesn't like any of the candidates—this also counts toward the residual vote rate. Say Mary is an avid voter but chooses not to vote in races in which she doesn't like any of the candidates or on propositions that she feels uninformed about. Mary's intentional abstentions would be counted toward the residual vote rate, despite not actually being errors.

There are also errors that do not show up in the residual vote rate—for example, if a voter meant to choose one candidate and instead selected another. Unfortunately, my colleagues and I have demonstrated that wrong-choice errors are much more common than other error types (for example, see Figure 4). This means that the residual vote rate is not necessarily a good indicator of a bad ballot. It also suggests that voting system designers cannot rely on the

Figure 4. Frequency of different voting error types



Overvote indicates that the voter selected too many candidates. *Omissions* indicate that a voter who intended to vote instead made no choice. *Extra vote* means a voter who intended to abstain accidentally selected a choice. The error bars refer to the standard error of the mean. Data are from "Now Do Voters Notice Review Screen Anomalies? A Look at Voting System Usability," by B. A. Campbell and M. D. Byrne, *Proceedings of the 2009 USENIX/ACCURATE Electronic Voting Technology Workshop/Workshop on Trustworthy Elections,* https://www.usenix.org/legacy/events/evtwote09/tech/full_-papers/campbell.pdf.

1-2%
voting error rates for paper ballots

55%
of Americans have only some, or little confidence that their vote will be properly counted

3-4%
voting error rates for punch cards and lever machines

“it is unlikely that a one-size-fits-all approach will be effective”

residual vote rate to indicate the true error rate and instead need to conduct laboratory usability studies that can verify voter intent. High residual vote rates can indeed indicate problems, but low residual vote rates do not necessarily mean that ballots were cast accurately.

Building Usable Voting Systems

Although there is still a great deal that is not known about voting system usability, the last decade has produced some key lessons:

- The most critical measure of a voting system's usability is the system's ability to accurately capture voter intent. The time it takes to cast a ballot is also important, but it is not particularly sensitive to design. Acceptable satisfaction with a voting system is relatively easy to achieve.

- Almost all changes in the way people vote impact usability, from ballot layout to small choices in wording on instructions. So although guidelines are a good start and can help prevent certain classes of usability problems, they are insufficient to guarantee usable voting systems. Usability testing, both during the design process (usually multiple times) and after the design is finalized, is critical.

- DREs offer the best avenue to accessibility for those with a disability, but most DREs in use today produce untenably high error rates. Yet with careful usability testing, they can most likely be made more effective than legacy systems (even paper). Usability testing at multiple stages of development is a key requirement, one that no current commercially available system has met.

- Both security and usability must be considered early in the design of the system, and it is important to take great care not to compromise one for the other. This can be a difficult balance, but it is critical.

- Voting by mail is not an ideal solution. The vast majority of U.S. voters still vote in person at their designated polling place, but in some areas of the United States (predominantly on the West Coast), voting by mail has become popular. However, this approach is not favored by most voting security researchers because it offers essentially zero resistance to coercion and weak resistance to other forms of fraud. Voting by mail also usually relies on paper ballots, which can seriously limit accessibility. For these reasons, it is unclear whether voting by mail will continue to grow in popularity, and few researchers have investigated its usability.

(For additional voting system usability studies, see the online Supplemental Material.)

After more than a decade doing research on voting systems in collaboration with election officials, I have learned that elections are dramatically more complex and challenging to manage than most people realize. It is no easy task to maintain security and accessibility while also keeping things manageable for election officials, who have to navigate a maze of idiosyncratic voting laws and customs.

Designing usable voting systems requires more than just people with expertise in accessibility and usability. It requires collaboration between people with expertise in election administration; computer security; certification and legal compliance; auditing; and, of course, usability and accessibility. What's more, designing an effective system involves many trade-offs. Because of differences in election laws, local budgets, and demographics, it is unlikely that a one-size-fits-all approach will be effective. Instead, different jurisdictions will require different systems, so designing a usable voting system is a problem that will likely need to be solved multiple times.

Building on the research produced by Caltech/MIT, ACCURATE, and other groups, two collaborative efforts to build better voting machines are currently under way: the Los Angeles County (California) Voting Systems Assessment Project (VSAP; http://www.lavote.net/vsap/) and the Travis County (Texas) STAR-Vote project. (STAR stands for *secure, transparent, auditable,* and *reliable*.)[17] These two jurisdictions have different constraints in terms of election law, demographics, and resources. Nevertheless, both have brought election and voting system experts together to share their expertise, and the systems they are building share some major design features. Both will use a DRE user interface similar to the Center for Civic Design's Anywhere Ballot (http://civicdesign.org/projects/anywhere-ballot/) to support usability and accessibility, and both will produce a paper record to ensure the system is secure and auditable. Both projects are also committed to usability testing. Preliminary usability data from the VSAP project are available at http://www.lavote.net/vsap/research, and usability testing for the STAR-Vote project is under way at Rice University. If these systems ultimately prove successful, other jurisdictions may use Travis County's and Los Angeles's collaborative processes as models, and those with circumstances similar to those of Travis County and Los Angeles may adopt the systems themselves, although this is still years away.

Today, many of the DREs purchased in the early 2000s with HAVA funds are only a few years away from the end of their life cycle, and election officials are watching the Los Angeles and Travis County voting system development collaborations with keen interest. Further, many election officials are beginning to understand how behavioral science can help improve voting systems for their constituents.

It is the job of security experts and election administrators to worry about keeping ballots safe; it is partly up to behavioral scientists to ensure that what is recorded on those ballots accurately matches voters' intent. Without that, all the security machinery in the world does not guarantee the integrity of American elections. For citizens to trust in the elections, they have to be able to trust that voting systems are user-friendly. Behavioral science has a key role to play in ensuring that they are—thus securing the integrity of U.S. elections.

author affiliation

Byrne: Rice University. Corresponding author's e-mail: byrne@rice.edu.

supplemental material

- https://behavioralpolicy.org/publications/
- additional references

references

1. Cooper, B., Cox, D., Lienesch, R., & Jones, R. P. (2016). *The divide over America's future: 1950 or 2050? Findings from the 2016 American Values Survey* [Research report]. Retrieved from PPRI website: http://www.prri.org/research/poll-1950s-2050-divided-nations-direction-post-election/
2. Davidson, A. (2016, November 28). The recount road to nowhere. *New Yorker*. Retrieved from http://www.newyorker.com
3. Herrnson, P. S., Niemi, R. G., Hanmer, M. J., Bederson, B. B., Conrad, F. C., & Traugott, M. W. (2008). *Voting technology: The not-so-simple act of casting a ballot*. Washington, DC: Brookings Institution Press.
4. California Secretary of State. (2007). Top-to-bottom review. Retrieved from http://www.sos.ca.gov/elections/voting-systems/oversight/top-bottom-review
5. Laskowski, S. J., Autry, M., Cugini, J., Killam, W., & Yen, J. (2004). *Improving the usability and accessibility of voting systems and products* (NIST Special Publication 500-256). Washington, DC: U.S. Government Printing Office.
6. International Committee for Information Technology Standards. (1998). *ISO 9241-11:1998. Ergonomic requirements for office work with visual display terminals (VDT)s—Part 11: Guidance on usability.* Geneva, Switzerland: Author.
7. Greene, K. K., Byrne, M. D., & Everett, S. P. (2006). A comparison of usability between voting methods. In *Proceedings of the 2006 USENIX/ACCURATE Electronic Voting Technology Workshop.* Retrieved from https://www.usenix.org/legacy/events/evt06/tech/full_papers/greene/greene.pdf?CFID=9073719468CFTOKEN=91391177
8. Byrne, M. D., Greene, K. K., & Everett, S. P. (2007). Usability of voting systems: Baseline data for paper, punch cards, and lever machines. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: CHI 2007* (pp. 171–180). New York, NY: Association for Computing Machinery. doi:10.1145/1240624.1240653
9. Brooke, J. (1996). SUS: A "quick and dirty" usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & A. L. McClelland (Eds.), *Usability evaluation in industry* (pp. 189–194). London, United Kingdom: Taylor & Francis.
10. Runyan, N. H. (2007). *Improving access to voting: A report on the technology for accessible voting systems*. Retrieved from Dēmos website: http://www.demos.org/sites/default/files/publications/improving_access%20(1)_0.pdf
11. United States Election Assistance Commission. (2005). Voluntary voting system guidelines. Retrieved from http://archives.eac.gov/program-areas/voting-systems/docs/vvsgvolumei.pdf/attachment_download/file
12. Everett, S. P., Greene, K. K., Byrne, M. D., Wallach, D. S., Derr, K., Sandler, D., & Torous, T. (2008). Electronic voting machines versus traditional methods: Improved preference, similar performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: CHI 2008* (pp. 883–892). New York, NY: Association for Computing Machinery. doi:10.1145/1357054.1357195
13. Everett, S. P. (2007). *The usability of electronic voting machines and how votes can be changed without detection* (Doctoral dissertation). Rice University, Houston, TX.
14. Mercuri, R. T. (2005). *Physical verifiability of computer systems*. Retrieved from http://www.notablesoftware.com/Papers/PhysVerify.pdf
15. Campbell, B. A., & Byrne, M. D. (2009). Now do voters notice review screen anomalies? A look at voting system usability. In *Proceedings of the 2009 USENIX/ACCURATE Electronic Voting Technology Workshop/Workshop on Trustworthy Elections*. Retrieved from https://www.usenix.org/legacy/events/evtwote09/tech/full_papers/campbell.pdf
16. Caltech/MIT Voting Technology Project. (2001). *Residual votes attributable to technology: An assessment of the reliability of existing voting equipment* [Report, Version 2]. Retrieved from http://www.hss.caltech.edu/~voting/CalTech_MIT_Report_Version2.pdf
17. Bell, S., Benaloh, J., Byrne, M. D., DeBeauvior, D., Eakin, B., Fisher, G., . . . Winn, M. (2013). STAR-Vote: A secure, transparent, auditable, and reliable voting system. *Journal of Election Technology and Systems, 1,* 18–37.