

Voting Systems Would Benefit From More Attention to Human Factors

In this interview with human factors/ergonomics voting systems experts, we target the human factors that are relevant to voting systems. Here, the experts explain how HF/E science can inform us about an issue of pressing importance to the general public. The objective is not to make a political statement but, instead, to report on the state of the science of voting systems to a broad spectrum of readers. HFES welcomes your feedback at info@hfes.org.

INTERVIEW QUESTIONS AND RESPONSES

In your opinion, what are some of the most important human factors issues that we faced in the United States four years ago with regard to the voting system?

Mike Byrne: In 2004, NIST (National Institute of Standards and Technology) released a comprehensive, 92-page report on human factors in voting systems. What is most striking about this report is that it references almost no data at all on the effects of different technologies on voting behavior. This is not because the authors failed to cite the relevant literature; there simply was no relevant literature to cite. There is no in principle reason that human factors research on voting systems could not have been done, it simply had not been.

Unfortunately, this lack of data had serious consequences. In 2002, Congress passed the Help America Vote Act, which provided funds to be used for updating voting systems. Private vendors rushed to provide products to fill this need. Ideally, these systems would be an improvement over the systems they replaced. However, without a precedent for usability evaluation, vendors were not required to provide any data about how well their systems performed. Even if they had, there were no data about the usability of extant system, such as punch cards and lever machines, against which to compare the new systems. Thus, there were no standards, no requirements, and no data available to inform local governments' decisions about what to purchase. As is the case in many other domains, the technology was deployed in advance of the conduct of fundamental science.

Tiffany Jastrzembki: In the 2004 presidential election, in light of the thousands of verified violations in the 2000 presidential election, disillusionment with the democratic tenet of having a vote accurately counted the way it was intended to be cast was still high. The 2000 election highlighted how poor ballot design (e.g., butterfly ballots) could lead to confusion as well as how specific voting devices (e.g., the punch card) themselves could lead to uncertainty for how the ballot was intended to be cast (e.g., hanging and pregnant chads). Both factors significantly contributed to spoiled ballots.

By 2004, record numbers of citizens registered to vote, record numbers of voters turned out at the polls, and record numbers of counties enacted sweeping changes to adopt new voting equipment and direct-recording electronic technologies.

Unfortunately, the usability of equipment and ballot design for electronic voting machines was poorly addressed, election officials were often poorly trained to handle the changes to their systems, too few machines were allotted to some precincts to handle the number of voters who were expected to turn out, and counties had the power to choose their own voting technologies and design their own ballots without any federal or state standardization or quality control. Those device and interface design choices had major implications for ease of use, verifiability, accuracy, and timeliness to cast a ballot (see Jastrzembki & Charness, 2007, for details).

Given these issues, it was no surprise that in places like Franklin County, Ohio, voters reported

standing in line for hours – in the rain – to cast their votes. Those circumstances deterred many citizens who showed up intending to vote; they just walked away. Voter registration rolls had increased 25% in that county, but the number of voting machines remained the same.

To compound the problem of too few machines, voters also reported a complicated, time-consuming ballot layout and design (Knapp, 2004), so movement through the polling station was slow and laborious. Both of these problems could have been avoided with good human factors analyses of their systems before election day finally came.

Douglas W. Jones: The first problem is ballot design. Regardless of the technology used, whether paper or electronic, we seem to be able to design ballots that are hard to use. On paper ballots, both 8 years ago in Florida and this year in Ohio, some counties broke the list of presidential candidates across two columns of the ballot – typically because unnecessarily long-winded instructions fill the top of the first ballot column. The predictable result is that some voters will see the presidential race as two separate races and cast a vote in each, invalidating their presidential vote.

Two years ago in Sarasota County, Florida, in designing a ballot for use with touch-screen voting machines, election administrators tucked the 13th Congressional District race above a banner near the top of the screen. Banner blindness – the common habit of ignoring information displayed above a banner or headline – almost certainly caused a significant fraction of voters to skip that race. Around 14% of voters expressed no preference at that time, but we cannot attribute this entirely to banner blindness because there were numerous reports of other problems that are difficult to reconcile with the banner blindness hypothesis.

A second problem area involves the clumsy interfaces for election administrators. To the extent that human factors work has been done on voting systems, almost all work has been done on graphical user interfaces for voter interaction. The interfaces used by election administrators for ballot preparation, voting machine setup, and election results consolidation are generally awful. The result, predictably, is that errors are more frequent than they ought to be.

A third problem is that most human factors work on computer software has focused on train-

ing and the development of expertise. We know quite a bit about designing systems like Microsoft Word or Adobe Photoshop, in which the emphasis is on tools with long, gentle learning curves. Beginners can get started easily, and expertise develops steadily as users gain experience. Unfortunately, voting systems are in a different category. Typical voter interactions happen once every two or four years for a few minutes each. As a result, voters are not expected to gain any expertise, nor is it realistic to talk about training. To date, there is limited understanding about designing such systems. Even ATMs and bus ticket vending machines are poor starting points, and the number of people who have video players in which, years after they opened the box containing the product, the clock is still blinking is solid evidence of how bad we are at designing such systems.

Bill Killam: There are known performance issues with the current machines, which stem primarily from a lack of proper feedback and poor instructions. These problems differ among machines but have a measurable effect on user performance that varies by machine and by user demographic. And this reflects only the machine design.

Though there is no way of knowing the total effect, this certainly implies that voting performance can be affected by the specific machine used. Considering that each of the 5,000 or so voting districts in this country selects its own machine or machines, different elections are not consistent in their ability to represent voter intent. This discussion doesn't even include ballot design issues that are known to exist and have caused problems or issues associated with voluntary workers trying to run equipment that they may not know or understand, which is also a known problem.

Whitney Quesenbery: Whereas the butterfly ballot in 2000 was the most visible problem in recent elections, unfortunately, usability flaws in the design of ballots and voting systems is neither new nor isolated.

A recent report by the Brennan Center for Justice, *Better Ballots* (Norden, Kimball, Quesenbery, & Chen, 2008), documents cases of high residual vote rates (errors in which no vote is cast in a contest, either because the voter made no selection or selected too many candidates), many larger than the margin of victory. These problems affect both electronic and paper ballot design, so

we cannot just say that a return to paper ballots will solve all of the problems.

The most important human factors issues are as follows:

Ballot design and voting system interaction design. Many of the errors we see in ballot design are violations of basic principles of information design. We need better templates and more education for the people creating ballots. The U.S. Election Assistance Commission (EAC) has taken an important first step with its “Best Practices in Election Management,” which includes ballot design templates created by Design for Democracy.

A related issue is that many state laws include detailed design requirements that contradict good design practice. Human factors experts need to work to get those laws changed.

Accessibility. Federal law – the Help America Vote Act (HAVA) of 2002 – mandates that voting systems be accessible. Though progress has been slow, some initiatives have been effective.

At the EAC’s Technical Guidelines Development Committee, we took a universal usability approach to creating requirements. Our goal was to increase the accessibility of all systems as much as possible.

Accessibility does not refer only to “declared” disabilities such as blindness or deafness; it also refers to ensuring that ballots are accessible for people with low-vision disabilities, for those who do not read well, and for a full range of human ability. Disabilities associated with aging, such as vision and dexterity, are especially important.

Plain language. This may not sound like an elections issue, but many of the problems we see in using ballots come from poorly worded instructions. Voting is more complex than we think, especially when we consider all the local variations in elections laws and practices. In a highly mobile society, we cannot assume that everyone knows how to vote.

Plain language is also a critical consideration in ballot questions, when voters are asked to participate directly in passing amendments or other laws. Most of these questions are written in convoluted legal language and simply placed on the ballot as written. California is an exception: The Ballot Simplification Committee in San Francisco is the only example in the country of an official group that writes plain-language versions of ballot questions.

Support for election processes. NIST has just published guidelines for creating usable system

documentation and is conducting research on usability issues for poll workers. We also need to consider how to make audits and recounts more usable and accurate.

What human factors issues have been the focus of your work?

Mike Byrne: Our first order of business was to collect baseline usability data (efficiency, effectiveness, and satisfaction) for traditional voting technologies such as punch cards, lever machines, and paper ballots. Once we had these data – which show that, to the extent that these differ, paper is the better choice – we began to test voting computers, which are known in the industry as direct-recording electronic systems (DREs). In collaboration with colleagues in computer science, we have developed our own DRE software, modeled after commercial systems but observing standard usability guidelines. We then compared this DRE with more the traditional methods. Though the DRE offered no advantage in terms of objective measures of error rate and time taken to vote, it was substantially better in terms of subjective usability.

We have also looked at alternative methods for navigating through the ballot and have found that giving voters the ability to navigate to any arbitrary race at any time actually caused a substantial increase in errors.

We have also branched out beyond simple usability metrics for voting systems. We have looked at the rate at which voters detect malicious vote manipulation by the DRE software. About one third of voters notice, and they are not sensitive to the number of races tampered with, the method of tampering, or the location on the ballot of the altered race. We are currently evaluating whether or not simple user interface manipulations can increase this proportion.

We have also looked at the effect of different vote-recording technologies (i.e., thermal paper records produced by DREs, paper ballots) on the ability of individuals to audit or recount a race and are planning to look at the accuracy and speed of multiple-team counting procedures.

Tiffany Jastrzembki: The major thrusts of my research are twofold. First, I argue that designers and engineers need to produce more usable systems at the very earliest stages of design by analyzing the system using performance-based measures (e.g., accuracy and time to complete the

task) and by applying validated models of human processing to critical pathway analyses of competing designs. This technique helps elucidate bottlenecks and error-likeness to inform the design of better systems and reduces the need to build prototypes and run humans through multiple instantiations of each product.

Second, I argue that employing a gerontechnological approach to human factors design enables the designer to determine the most usable system for older adults' capabilities and limitations (users who may shed light on potential problems that may otherwise go undetected), and that the best design for the older adult often also corresponds to higher performance in younger users.

Douglas W. Jones: My work focuses on voting systems as a whole, and I recognize human factors as an important component. Right now, I am working on a small human factors study attempting to answer the question, What information can we routinely log from voter interactions that would allow us to identify human factors problems in the voting system without violating the voters' right to a secret ballot? We know that if we simply log every voter action, we would be able to retroactively reconstruct the difficulties they had, but that would violate the right to a secret ballot. Current machines do not log enough information, as was clearly demonstrated in the attempt to reconstruct what happened in Sarasota two years ago.

In the experiment, we will create known human factors problems (poor touch-screen calibration, insensitive touch-screen, banner blindness, etc.) and look for correlations between safe-to-log events and these problems.

Bill Killam: My firm has been working to develop a performance standard for voting systems that will eventually be part of the Federal Voluntary Voting Systems Guidelines (VVSG) that are used by many states as mandatory standards. These performance standards will serve to ensure adequate performance of voting systems. And if the performance requirements are raised over time, it will help push these systems to improve. However, using testing standards to improve voting systems is definitely a case of the tail trying to wag the dog.

Whitney Quesenbery: I serve on the federal advisory committee writing the VVSG and am chair of the subcommittee on human factors and

privacy (which includes usability, accessibility, privacy, and language support). In the VVSG 2005, we greatly expanded the accessibility section and wrote the first usability requirements for voting systems. We are currently in public review with an updated version. Among the new requirements we have proposed is a human performance test with pass-fail benchmarks for conformance with the VVSG. We believe this is the first conformance test of its kind in a standard.

The UPA Usability in Civic Life/Voting and Usability Project has partnered with The Brennan Center on work on ballot design. We have started a program to teach usability testing to election officials and conducted usability tests in six jurisdictions for the November 2008 elections. In many cases, we were able to help them improve their ballots and hope they will avoid the kinds of problems we have seen in Palm Beach, Sarasota, and other places.

There is no single solution to better voting systems – we need improvements at all points. The systems themselves, the ballot design and election management tools, the ballots themselves, and the procedures for running an election must all work together and must all be not only technically sound but designed for use by real voters and election workers.

What are the biggest challenges that you have encountered in terms of research and practice in this area?

Mike Byrne: The United States is the only industrialized Western democracy that does not have national laws that govern national elections. Instead, elections are regulated by states and counties, and there is enormous variance in laws, procedures, financial resources, equipment, poll worker training, and virtually anything else one can think of that might vary among districts. If there is more than one way to do something with respect to administering an election, it will be done in more than one way in the United States.

To take a simple example, consider the "simple" task of designing a ballot for use in a study. How long should the ballot be? In some jurisdictions, the number of local offices and ballot measures (e.g., California's propositions) is enormous and thus produces very long ballots. (My ballot in the 2006 election – an "off" year – was 46 races long.) Should the ballot include a straight-party option? A straight-party option can make voting extremely

fast for some voters and may also affect the rate of errors (both positively and negatively); this option is sure to have some impact on the usability of the ballot. Of course, this option is explicitly forbidden in some jurisdictions (e.g., California) and mandated by state law in others (e.g., Texas).

The other challenge is the variance in the voting population. There is virtually no other user interface that is intended to be used by such a wide audience. Even widely used technologies such as automobiles and personal computers do not have user bases that are as broad as the population that uses voting systems. Everyone over 18 (with a few exceptions) is guaranteed the right to vote and must be offered a reasonable opportunity to vote, regardless of age, cognitive ability, literacy, disability, and so on. We consciously try to recruit diverse subject populations for our studies but have only managed to reach the tip of the iceberg so far.

Tiffany Jastrzembki: The biggest challenge is having time to fully explore these applications in the voting domain. Applying validated, age-sensitive models to voting technologies and ballot design could help produce better systems in the real world that could translate to fewer spoiled ballots, less time spent at the polling station waiting in line, less time actually completing the ballot, and greater transparency that the ballot was cast as intended. Furthermore, analyses from our community must then reach government decision makers so that election standardization or regulation may be better ensured.

Douglas W. Jones: It is very difficult to create a simulated voting experience that is realistic for an experiment. Fictional candidates don't work the same as candidates the participants really care about. But if you use real candidate names, you need to be careful not to mislead voters into thinking they really voted and not to violate the privacy of voters by keeping records that connect them to how they voted.

There is a window of opportunity right after each general election for a study in which you can use the ballot from that election without risk of fooling people into thinking they have really voted. During that window, passions are still high, name recognition is still good, and you can run experiments that are, as a result, far more realistic.

Bill Killam: In our work on performance-based standards, the biggest issue is that this type

of work is new to many of the stakeholders. A lot of our work has been involved with convincing stakeholders at NIST and the EAC about validity, reliability, and so on – both how these must be applied and the limitations of their application (e.g., whether or not there is a need for generalization). For the basic issue of usability of voting systems, the biggest problems have been the lack of good, research-based design standards and lack of adherence to a user-centered design model.

Whitney Quesenbery: The biggest challenges are (a) disbelief that usability or human factors research is needed (“voting is easy, right?”); (b) lack of funding; (c) difficulty in gaining access to voting systems to conduct realistic research with actual systems (not just prototypes); and (d) methodological issues in conducting tests for an activity that should be private with systems that should not allow monitoring of user activity.

What has been done or can be done based on the science of human-technology interaction?

Mike Byrne: Besides beginning to understand the basic usability situation created by current and recent technologies, the most significant thing that has been done is that the federal government has recently proposed requiring systems to meet minimum standards on a suite of usability metrics, all centered on accuracy. Meeting these standards would almost certainly entail substantial improvements in the usability of current systems, so this is an important step in the right direction.

As for what could be done, an enormous amount of science could be carried out to better understand the problems and issues in this domain. Again, there are no in-principle barriers to the application of human factors science to voting systems; it simply has not been done to the extent that it needs to be. Given the importance of elections and the thin margins of victory that have become prevalent in the United States, this is becoming an increasingly important problem with relatively few people doing research in this area.

Tiffany Jastrzembki: As mentioned earlier, we need validated, age-sensitive models for design and engineering.

Douglas W. Jones: Unfortunately, the whole field of elections and election equipment design

has largely ignored human factors, except at a very intuitive level, for a very long time. Furthermore, intuition in human factors is frequently wrong, particularly when designing systems for use by novices. The problem is, the designers of the system are all experts in that system by the time they have any prototype to play with, and voters are never expected to develop expertise.

Bill Killam: Although there are ongoing efforts to develop standards for instructions and ballot design as well as our work on performance standards, much of this work has never been tested adequately. In addition, the Voluntary Voting System Guidelines contains interface design standards that have never been tested. We believe that the current document is neither necessary nor sufficient as currently written. More involvement by human factors engineers in developing and researching these guidelines, as well as adequate testing to ensure that the guidelines have the intended effect, is critical if these problems are to be solved in current voting systems and avoided in future systems.

SUMMARY

In summary, voting systems go beyond ballots and other voting devices to include polling places and systems used by election officials before, during, and after the vote. Importantly, there are numerous human factors issues surrounding these complex and varied systems. Ballots and devices are often poorly designed, election officials receive minimal training, and accessibility of the system to the wide range of eligible voters is limited. These problems are exacerbated by challenges that include the wide range of voters, the infrequency with which voters come in contact with the system, and the difficulties of conducting realistic voting experiments.

Some human factors work is ongoing, though it is relatively new and sparse, and the gaps in the research are large. There is a great deal of science that can be applied to improving the design of voting systems. Much of the science on human-computer interaction, human error, design for aging populations, human-systems integration, display design, and training can be found in the pages of *Human Factors: The Journal of the Human Factors and Ergonomics Society*, and other HF/E journals.

THE EXPERTS

Mike Byrne

Mike Byrne is an associate professor of psychology and computer science at Rice University. He received bachelor's degrees in psychology and engineering from the University of Michigan in 1991, a master's degree in computer science in 1995, and a Ph.D. in experimental psychology in 1996, both from the Georgia Institute of Technology. He is one of 10 principal investigators in the National Science Foundation-funded ACCURATE Project.

Tiffany Jastrzembski

Tiffany Jastrzembski completed her Ph.D. in cognitive psychology at Florida State University in 2006 and received her undergraduate education at Carnegie Mellon University. She is a cognitive research psychologist at the Air Force Research Laboratory and specializes in applied mathematical modeling of skill acquisition/decay, human factors engineering, expertise, and aging.

Douglas W. Jones

Douglas W. Jones is an associate professor of computer science at the University of Iowa. He has been working on voting systems since 1994, when he volunteered to serve on the Iowa Board of Examiners for Voting Machines and Electronic Voting Systems. In recent years, voting systems have been a priority in his research. He is one of 10 principal investigators in the National Science Foundation-funded ACCURATE project.

Bill Killam

Bill Killam is president and senior human factors consultant at User-Centered Design, Inc., in Ashburn, VA. He has been a member of the Human Factors and Ergonomics Society since 1985, holds degrees in engineering and psychology, and served as president of the HFES Potomac Chapter. He also teaches human factors engineering at George Mason University and the University of Maryland.

Whitney Quesenbery

Whitney Quesenbery is a user researcher and usability specialist with a focus on clear communication. She works with companies around the world to develop usable Web sites and applications. Whitney was president of the Usability Professionals' Association (UPA) and is a Fellow

of the Society for Technical Communication (STC). She serves on two federal advisory committees, working on voting system guidelines for the U.S. Election Assistance Commission and on updating U.S. accessibility regulations (Section 508) for the U.S. Access Board.

SUGGESTED READINGS

- ACCURATE Research Group – A Center for Correct, Usable, Reliable, Auditable, and Transparent Elections. Publications accessible online from <http://accurate-voting.org/pubs/>
- Bederson, B. B., Lee, B., Sherman, R. M., Herrnson, P. S., & Niemi, R. G. (2003). Electronic voting system usability issues. In *Human Factors in Computing Systems: Proceedings of CHI 2003* (pp. 145–152). New York: ACM.
- Brennan, L. (project director). *The Machinery of Democracy*. http://www.brennancenter.org/content/resource/the_machinery_of_democracy_voting_system_security_accessibility_usability_a/
- Byrne, M. D., Greene, K. K., & Everett, S. P. (2007). Usability of voting systems: Baseline data for paper, punch cards, and lever machines. In *Human Factors in Computing Systems: Proceedings of CHI 2007* (pp. 171–180). New York: ACM.
- Caltech-MIT/Voting Technology Project. (2001). Voting – What is, what could be. Retrieved October 3, 2008, from <http://www.vote.caltech.edu/2001report.htm>
- Election Assistance Commission. (2005). *Voluntary Voting System 2005 Guidelines* (VVSG). Available at <http://www.eac.gov/voting%20systems/voting-system-certification/2005-vvsg/>
- Everett, S. P. (2007). *The usability of electronic voting machines and how votes can be changed without detection*. Doctoral dissertation, Rice University, Houston, TX.
- Everett, S. P., Byrne, M. D., & Greene, K. K. (2006). Measuring the usability of paper ballots: Efficiency, effectiveness, and satisfaction. In *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* (pp. 2547–2551). Santa Monica, CA: Human Factors and Ergonomics Society.
- Everett, S. P., Greene, K. K., Byrne, M. D., Wallach, D., Derr, K., Sandler, D., et al. (2008). Electronic voting machines versus traditional methods: Improved preference, similar performance. In *Human Factors in Computing Systems: Proceedings of CHI 2008* (pp. 883–892). New York: ACM.
- Greene, K. K., Byrne, M. D., & Everett, S. P. (2006, August). A comparison of usability between voting methods. In *Proceedings of the 2006 USENIX/ACCURATE Electronic Voting Technology Workshop*, Vancouver, BC, Canada. Berkeley, CA: USENIX Association. Available from <http://portal.acm.org/citation.cfm?id=1251003.1251005&coll=GUIDE&d1=GUIDE&CFID=5682855&CFTOKEN=81039522>
- 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.
- Jastrzembski, T., & Charness, N. (2007). What older adults can teach us about designing better ballots. *Ergonomics in Design*, 15(2), 6–11.
- Knapp, J. (2004). Effect of voting-machine allocations on the 2004 election – Franklin County, Ohio. Retrieved October 6, 2008, from <http://www.copperas.com/machinery>.
- 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.
- Laskowski, S. J., & Redish, J. (2006). *Guidelines for writing clear instructions for voters and poll workers*. Retrieved October 8, 2008, from <http://vote.nist.gov/032906PlainLanguageRpt.pdf>
- Mebane, W. R., Jr. (2004). The wrong man is president! Overvotes in the 2000 presidential election in Florida. *Perspectives on Politics*, 2(3), 525–535.
- Mercuri, Rebecca. Research and work, found at <http://notablessoftware.com>
- Nichols, S. M., & Strizek, G. A. (1995). Electronic voting machines and ballot roll-off. *American Politics Quarterly*, 23(3), 300–318.
- Norden, L., Kimball, D., Quesenbery, W., & Chen, M. (2008). *Better ballots*. Available at http://www.brennancenter.org/content/resource/better_ballots/
- Runyan, N. (2007). *Improving access to voting: A report on the technology for accessible voting systems*. Retrieved October 7, 2008, from http://demos.org/pubs/improving_access.pdf
- Runyan, N., & Tobias, J. (2007). *California accessibility review for the top-to-bottom voting systems review*. Retrieved October 8, 2008, from http://www.sos.ca.gov/elections/voting_systems/ttbr/accessibility_review_report_california_ttbr_absolute_final_version_16.pdf
- UPA Web site lists references: http://www.usabilityprofessionals.org/civillife/voting/usability_research.html
- Usability Testing of Voting systems (UPA 2004 Workshop Report). http://www.usabilityprofessionals.org/civillife/voting/documents/voting_summative_test.pdf
- User-Based Conformance Test for the Usability of Voting Equipment*. (2006). <http://vote.nist.gov/032906User-BasedConfTesting3-10-06.doc>
- VVSG Guidelines available at <http://vote.nist.gov>
- Voluntary Voting System 2005 Guidelines* (VVSG). <http://www.eac.gov/voting%20systems/voting-system-certification/2005-vvsg/?searchterm=vvsg>
- Wand, J. N., Shotts, K. W., Sekhon, J. S., Mebane, W. R., Jr., Herron, M. C., & Brady, H. E. (2001). The butterfly did it: The aberrant vote for Buchanan in Palm Beach County, Florida. *American Political Science Review*, 95(4), 793–810.