A review of proposed

Voluntary Voting System Guidelines

[This page intentionally left blank]

Table of Contents

[This page intentionally left blank]

Executive Summary

In October 2007, the U.S. Election Assistance Commission (EAC) began to consider a new draft of standards – the Voluntary Voting System Guidelines (VVSG) – that future voting systems will have to meet in order to earn federal certification. Once fully implemented, only voting systems that have been federally certified can be purchased using federal funds, and so the VVSG will constitute a weighty statement of principle by the federal government as to how voting systems should be designed and implemented.

The current draft is an extensive document. It contains several hundred pages of detailed requirements meant to ensure voting system usability, accessibility, and security. Many of its provisions are the product of substantial research and experience.

This being said, we find the VVSG to be in need of extensive revision. In many cases it is excessively prescriptive, a characteristic we fear will stifle innovation. In other places, its provisions are based on faulty methodology, and require further testing and consideration.

Chief among our findings:

- In many instances, the VVSG prescribes design standards where performance standards would be sufficient. Excessive design standards can arrest manufacturers' ability to innovate new, better systems.

- Although the VVSG purports to be agnostic towards election processes, several provisions do in fact bear on process. These provisions exceed the document's mandate and could lead to burdensome requirements for administrators in the field.

- The VVSG's "software independence" requirement is, in effect, a requirement that direct recording electronic (DRE) systems produce a paper record, a provision we fear trades tangible losses in election facility for merely *perceived* gains in election security.

- The methodology through which the VVSG derives standards for voting system usability is flawed, failing to incorporate standard, easily implemented procedures for research in the social sciences.

- By focusing narrowly on provisions for voting system security, the VVSG distracts attention from topics that should be of at least as much concern to election administrators and the public, such as poll worker training, voter registration practices, and chain-of-custody procedures.

A number of the concerns we raise call for extensive revision to the VVSG. While such revision may be unpalatable, we strongly believe that the passage of time will only serve to make this important task more costly and difficult.

Introduction

What is the VVSG?

Although many changes to election administration in the United States receive vigorous scrutiny from an energetic monitoring community, vastly consequential federal involvement in this arena is now quietly proceeding with relatively little attention. In October 2007, the U.S. Election Assistance Commission (EAC) began to consider a new draft of standards – the Voluntary Voting System Guidelines (VVSG) – that future voting systems will have to meet in order to earn federal certification. Much broader in scope than any previous standards, the draft VVSG has potentially far-reaching implications for the future of voting technology and election administration, including controversial topics frequently the subject of spirited public debate.

Put another way, the VVSG,[1] once adopted, will constitute a weighty statement of principle by the federal government on a number of crucial election reform topics. In its current form, the VVSG delves into such important matters as the mandate of physical voting records, accessibility of voting technology, and protection of incentives for innovation. True, the standards are "voluntary" in that states purchase their own voting systems, which are not *required* to be federally certified. Still, in the future, federal funds will only be able to be spent on certified systems, and so even if some states opt out of the requirements, the VVSG's effects on the market, and through it the national landscape, will be immense. The document makes only a small exaggeration, therefore, when it claims that it can be "considered essentially as a mandatory standard" (1.1.2). Prudence tells us that the inconstancy of public opinion makes broad policies, once established, difficult and costly to change.[2] It is for this reason that we believe a thorough review of the draft VVSG before its adoption and implementation is essential.

The AEI-Brookings Task Force

In assembling a team of experts to review and comment on the draft VVSG, the AEI-Brookings Election Reform Project sought to attract distinguished scholars with diverse backgrounds and specializations. We did not consider individual viewpoints on specific policy matters as criteria for selection. Rather, we made a conscious effort to recruit individuals who have thus far kept a "low profile" in the frequently acrimonious sphere of election reform by abstaining from strident advocacy. We resolved to make only such recommendations as could enjoy broad – preferably unanimous – support from the entire group.

Members of the Task Force reviewed the VVSG and prepared extensive memoranda outlining their critiques of the document. The entire task force then met in

---

[1] Throughout this report, we use this abbreviation to refer to the *proposed* guidelines that are currently under review, not any past or future documents that might have the same name, unless otherwise noted.
[2] Anthony Downs, "Up and Down with Ecology – the 'Issue-Attention Cycle,'" *Public Interest*, Vol. 28 (Summer, 1972), pp. 38-50.

Washington, DC in December of 2007 to discuss each other's comments and arrive at consensus recommendations. Funding for the meeting was provided by the John S. and James L. Knight Foundation through a grant to the AEI-Brookings Election Reform Project.

Together, the eleven members of our team bring to the public discourse more than 100 years of experience in the academic and applied study of elections. We approach the task at hand from diverse perspectives in technology and the social sciences, including in our numbers two usability experts, a political methodologist, a human-computer interaction expert, and several public policy scholars. It is our hope that the recommendations of such a diverse group, speaking with one voice, will be of considerable value both to the public and the EAC as the nation contemplates how best to ensure the integrity of its electoral institutions.

Format and scope of recommendations

One inherent difficulty in analyzing a large document like the VVSG, a nearly 600-page behemoth, is choosing an appropriate level of scrutiny. We found ourselves torn between "broad stroke," thematic recommendations and line-by-line critiques that address specific provisions of the VVSG. It seemed that the broad stroke recommendations would most naturally take the form of simple prose, a style that held some appeal because it would be relatively easy for a wide audience of stakeholders, many of whom do not hold technical expertise, to digest. However, the EAC conducts its public comment process through a website where reviewers submit point-by-point recommendations germane to specific provisions of the document.

Some comments are suitable to only one style. For instance, we occasionally question the premise upon which a series of VVSG provisions is founded. In such a case, the EAC's public comment mechanism would require us to respond to each provision individually. We feel it more constructive to step back and explain our general concern.

We therefore adopt a two-pronged approach. This review is written in prose and provides general comments and recommendations about the VVSG. Most of its recommendations are not suitable to line-by-line edits. We hope it will be of use both to the EAC as it considers the VVSG and to media and laypersons seeking to increase their familiarity with this complex, but very important, document. In addition to this report, the product of collaboration among the members of our team, some individual members will also submit line-by-line recommendations to the EAC through the review website.

The VVSG in context

Chapter 3 of the VVSG's Introduction provides a full history of voting system standards in the United States. Readers interested in the full background of federal regulation in this area will find this resource useful. For the purposes of our review, we simply explain the genesis of the current draft standards so that a reader unfamiliar with their significance can understand them in context.

In October 2002, President Bush signed into law the Help America Vote Act (HAVA), a comprehensive reform of federal elections in the United States. Whereas previously

the federal government exercised little influence over the U.S. election system – the U.S. Constitution grants states considerable latitude in the conduct of elections – HAVA marked an effort by the federal government to increase oversight and regulation of the process. Largely a response to administrative and technological problems exhibited by Florida's conflict-ridden counting process during the 2000 presidential election, HAVA imposed administrative requirements such as provisional ballots; provided states with funds to update voting technology; required the establishment of computerized statewide registration lists; and created the Election Assistance Commission to review and coordinate election procedures.

Additionally, HAVA created the Technical Guidelines Development Committee, an organization composed of experts from the National Institute of Standards and Technology (NIST) and other bodies. The TGDC, in coordination with and approval from the EAC, was tasked with updating federal standards for certification of voting systems. HAVA mandated a short timeline for the adoption of the first set of standards; the TGDC was required to provide draft standards to the EAC only nine months after its formation. These standards, which are currently in effect, were finalized by the EAC in December 2005.

Because the TGDC undertook the 2005 standards on a short deadline, its members recommended a more comprehensive set of standards be developed in the near future. Whereas the 2005 standards built upon previously existing standards created under the supervision of the Federal Election Commission (FEC) before HAVA shifted authority over voting standards to the EAC, the current draft standards are a full overhaul meant "to meet today's more rigorous needs for electronic voting systems" (Introduction 3.4). In short, the current draft VVSG is the first comprehensive update to federal voting standards since before election administration came to the forefront of public attention in 2000.

A flawed document

Starting from scratch, the VVSG consists of a thorough, top-to-bottom review of voting system functionality that encompasses the diverse fields of computer science, electrical engineering, usability, political science, and others. The TGDC and NIST are to be congratulated for undertaking a monumental project.

At the same time, we fear the VVSG loses sight of what is perhaps its most important goal: to ensure that all vote collection systems accurately record and report voter intent. Although its requirements are extensive, we find significant problems in its treatment of security, usability and accessibility, matters at the very heart of HAVA's purpose. At some junctures, the VVSG would benefit from improved methodology and a greater familiarity with standing research in the social sciences. At other times, the VVSG potentially interferes with local prerogative over election conduct, prescribing solutions – such as physical records – that we do not find to be proper at all times nor in all places. In addition, the document could be made more wieldy, easy-to-use, and defined in its role.

Below, we expand on the reasons we hold these concerns and provide suggestions to improve the VVSG. We recognize that some of our suggestions will be unpalatable to those charged with preparing the final version of the VVSG. For instance, we find the document's treatment of usability standards to require substantial revision. Other points, such as our wariness of a paper trail mandate, run counter to passionate

beliefs held by well-intentioned advocacy groups.[3] The unpalatability or unpopularity of a proposition, however, are weak grounds for rejecting it. We have taken care to clearly explain the reasoning behind each of our recommendations. In return, if the reader finds himself in disagreement with our conclusions, we ask that he take as sincere our desire for a secure, reliable, and inclusive election system.

We especially hope that the EAC, NIST, the TGDC, and the general public will share our belief that the United States is at a crucial juncture with regard to election reform. The echo of the 2000 meltdown in Florida still resonates through the public consciousness, but will not do so forever. Moreover, states' commitment to certain administrative practices and technological options, though once pliable, grow more rigid with each passing day and each dollar spent. In short, it seems that *this* moment in time, more so than any other, is the easiest to undertake the arduous task of "getting it right," or as a notable former president once remarked, "Certainly it is not so easy to pay something as it is to pay nothing; but it is easier to pay a large sum than it is to pay a larger one."

First, Do No Harm

Haste makes waste

We sympathize with those who feel that the numerous election debacles since the presidential election of 2000 call for drastic corrective measures. While perhaps cliché, it is entirely accurate to call fair elections the *sine qua non* of any functional democracy, and so it is fitting to meet any threat to their legitimacy with dogged protectionism. The problems of the past eight years are no trifling matter, for even the suspicion of illegitimacy undermines trust in government and foments partisan rancor.

Nevertheless, it is possible for an improperly administered cure to be as bad as the disease. The haste with which policymakers have pursued election reforms has already led to a staggering amount of fiscal waste. For instance, election officials throughout the country rushed to replace punch card voting systems with modern touch screen voting computers, but in just a few years, these substitutes evinced their own shortcomings and many are now headed to the scrap pile.[4] Worse than wasted money, we fear that the continuation of a strong interventionist course could entrench ill-conceived solutions to complex problems. Unfortunately, the VVSG exhibits such a continuation in many ways.

Tempting though it may be to apply sweeping reforms to the conduct of elections, we adopt as a guiding principle the notion that states and localities – the entities traditionally entrusted with the conduct of elections – are generally capable and well-intentioned judges of what will and will not work within their borders. To override

---

[3] For a study of public opinion that reflects general ambivalence, however, see M. Glenn Newkirk, "Trends in American Trust in Voting Technology," InfoSENTRY Service, Inc., 17 March, 2008, available online at <http://www.infosentry.com/InfoSENTRY_Trends_Trust_VoteTech_2004-2008.pdf> (accessed 23 April, 2008).

[4] Gary Fineout, "Voting Change Irks Supervisors," *The Miami Herald*, 24 May, 2007, p. 20.

their authority is to assert that they would enact an ill-conceived policy choice if left to their own devices. It not only demonstrates an implausible degree of confidence in whatever reform is mandated, but also suggests that one policy size fits all. On the other hand, to encourage state and local prerogative is to allow for future innovations that exceed the boundaries for which imagination has heretofore provided. It is also to acknowledge the possibility that, just as touch-screen balloting has quickly fallen from grace as a cure to election quagmires, so might the apparent solution of today join the plentiful detritus with which the path to reform is littered.

Almost all public policies come with unanticipated consequences. This is one reason why Justice Louis Brandeis's entreaty that a locality can serve as a "laboratory; and try novel social and economic experiments without risk to the rest of the country,"[5] continues to resonate more than seventy years after he wrote it. We do not reject the notion that certain circumstances call for federal regulation. But we do feel that any intervention by the federal government into the sphere of election administration should be accompanied by a compelling explanation of why local authorities are incapable of tailoring their own policy. All things being equal, we believe the public interest in this area is best served when the federal government takes a minimalist approach.

Design vs. performance standards

At many junctures, in an effort to ensure voting system functionality, the TGDC, the committee that authored the draft VVSG, faced a choice between the prescription of design standards and performance standards. Design standards are precise specifications for the workings of a voting system while performance standards remain agnostic to process and only specify benchmarks for an acceptable level of functionality.[6]

We generally support well-conceived performance standards over design standards. As a matter of principle, the federal government should leave to vendors as much discretion as possible over how to design their products. Such meddling can stifle innovation and may bear partial responsibility for the paucity of new voting equipment brought to market in recent history.

A metaphor is illustrative. If, in 1935, the government wanted to ensure the safety of commercial aviation, it could have specified exacting design standards – measurements, position, composition, etc. – for plane engines. Such specifications might have achieved the short-term goal of ensuring that manufacturers adhered to the best practices of the day. But narrow constraints could also have forestalled the innovation of better alternatives that would have breeched old specifications, such as jet engines. A better alternative is to specify performance standards – the acceptable failure rate, etc. – and invite vendors to fulfill the requirements in whatever fashion they feel is most likely to satisfy consumer demand.

In a number of instances, the VVSG prescribes design standards that we feel are unnecessary. To cite two examples, Part 1 includes specifications for how

---

[5] *New State Ice Co. v. Liebmann* 285 U.S. 262 (1932).

[6] Maureen A. Breitenberg, "The ABC's of Standards-Related Activities in the United States," *National Institute of Standards and Technology* (May, 1987), available online at <http://ts.nist.gov/Standards/Conformity/stdpmr.cfm#Types> (accessed 24 March 2008).

manufacturers must design voting machine locks (Chapter 5.8.5) and ballot boxes (Chapter 6.1). While neither of the requirements seems unreasonable in and of itself, we worry that such extensive prescription could inhibit future innovation. Instead, it seems that the VVSG could simply require that voting machines and ballot boxes be resistant to tampering.

In other places, the problem is that the recommended design standards are not empirically connected to performance standards. There are few clear links between the goals for voting systems and the possible designs for achieving those goals. For example, it might be natural to expect that voter confidence will be enhanced with a physical voting record, but research in several venues shows this not to be the case.[7]

All things being equal, we recommend that the VVSG err on the side of self-restraint in the prescription of design standards so as to avoid "technology locks," whereby certain flawed approaches can become enshrined on account of regulatory inertia. In our view, the science of election administration and voting technology is insufficiently advanced to develop worthy design standards. Design standards should be reserved for instances in which there are clear and unambiguous empirical links between the goals for voting systems and the designs recommended, or in which it is essential to establish uniformity and interoperability across the industry. In this, we concur with a recent draft report from the NIST Security and Transparency Subcommittee.[8]

Of course, the substitution of performance standards for design standards would not prevent the VVSG from pointing to *examples* of technology worthy of emulation. Such examples could serve as role models and establish a conceptual benchmark for manufacturers to emulate while still making it easier for the voting technology industry to mature over time.

Mandating process

At its surface, the VVSG deals only with vote collection systems, the physical machinery of elections. The TGDC emphasizes at several points that the document is agnostic to process, calling administrative rules and procedures "outside the scope of the VVSG" (Part 1, Chapter 3.1.1). Such agnosticism is fitting, as the document is not meant to be a federal treatise on how to conduct an election, but rather a list of certification requirements for voting technology. The undertaking is daunting as it is.

Unfortunately, equipment and process do not exist in vacuums, but are inextricably linked. The VVSG could no more address technology while ignoring process than the Boston Red Sox could recruit hitters while ignoring their capabilities as fielders. Again, the VVSG should take a minimalist approach, considering as far as possible the process implications of its provisions and reserving for localities as much discretion as possible.

---

[7] Paul S. Herrnson, et al., *Voting Technology: The Not-So-Simple Act of Casting a Ballot* (Washington, D.C.: Brookings Institution Press, 2008); Robert Stein and Greg Vonnahme, "Voting Technology, election administration, and voter behavior," presented at the 2007 Midwest Political Science Association Meeting, Chicago, Il, April 1-3.

[8] NIST Security and Transparency Subcommittee, "Voting Systems Innovations Class," draft 10 March, 2007, available online at <http://vote.nist.gov/meeting-08172007/InnovationClass06262007.pdf> (accessed 22 January, 2008).

Consider a few ways in which some of the changes proposed by the VVSG – even marginal changes – can have a substantial effect on process:

- The VVSG requires voting systems to produce an "independent voter-verifiable record" (Part 1, Chapter 4.4). For all intents and purposes, this term is synonymous with a paper record, as we explain below. We also consider the general wisdom of a paper mandate below. Here, we simply point out that a paper mandate would have substantial effects on process. It would require the retraining of poll workers; the hiring of technicians with requisite expertise in the printer technology; the development of processes for validating the reliability of the paper trail; and the incorporation of procedures for transporting, storing, and securing paper records after an election. This is not all. Because paper trails lead to more time per vote, a paper mandate might require more machines to accommodate the long lines that develop in an important election and, in turn, more polling places to provide enough space for the new machines.

- The VVSG requires vote tabulators to be scanned for malicious software every 24 hours during operation (Part 1, Chapter 5.5). On its surface, the suggestion seems prudent, or at least innocuous, but we view it as an example of a decision best left to local authorities, who will have the best knowledge of their human resources, capabilities, and priorities.

- The VVSG's paper trail requirement could provide a disincentive for jurisdictions to employ vote centers, the consolidation of polling places that has been shown to lower costs, facilitate administration, and increase turnout, especially among groups that are typically politically unengaged.[9] Because the paper trails created by DREs in such a context will intermix ballots associated with different jurisdictions, the conduct of a meaningful recount would be a monumental undertaking that would require extensive sorting in order to filter out the pertinent ballots.

Software independence and paper records

Perhaps the most tangible change to come from the VVSG would be from its provision of "software independence" (Part 1, Chapter 2.7). On its surface, the software independence requirement is a provision designed to guard against the subversion of voting system software. It is, in effect, a requirement that voting systems produce a paper record such that digital records can be compared with something more tangible and, seemingly, trustworthy. While we strongly sympathize with the sentiments that have given rise to this proposal and recognize that software independence may indeed be a desirable feature in certain circumstances, we believe that a software independence *requirement* – or, indeed, a paper trail *requirement* –

---

[9] Robert M. Stein and Greg Vonnahme, "Engaging the unengaged voter: Vote centers and voter turnout," *Journal of Politics* (forthcoming, April, 2008); Robert M. Stein and Greg Vonnahme, "Voting place and its impact on voter participation," paper prepared for presentation at the annual meeting of the Midwest Political Science Association, Chicago, Illinois, April 3-5, 2008.

will be a long-run hindrance for voting technology. Because this position runs counter to some measures of popular opinion,[10] we take care to explain our reasoning.

Although seemingly an elusive concept, the VVSG's software independence requirement is in fact quite simple. The requirement states that

> Voting systems *SHALL* be software independent, that is, an undetected error or fault in the voting system's software *SHALL NOT* be capable of causing an undetectable change in election results (Part 1, Chapter 2.7-A, emphasis in original).

This statement can be simplified for a layperson. It calls for some sort of double check of any vote data stored by a computer. Voting machines should be designed such that it would be impossible to subvert an election through computer hacking because any alteration of computerized data would be revealed by an audit of whatever record served as a double check.

It is easy to see why the public and experts alike are suspicious of voting technology. In the years since the implementation of HAVA, newspapers have been replete with accounts of equipment faults and incompetence on the part of equipment manufacturers. Some reports have even left readers with concern over political infiltration in the upper levels of voting system manufacturers.[11] Since 2000, computer scientists have exposed numerous vulnerabilities in electronic voting machines.[12] Some of the companies that produce the machines have ineptly stumbled into one public relations disaster after another, such as when Diebold, Inc. inadvertently posted vulnerable computer code for its tabulation software to a public website.[13] Other revelations have come from experts who have vividly documented how a self-replicating virus might be installed in a Diebold AccuVote machine.[14] Similar vulnerabilities have been noted in other contexts.[15] Among the glaring shortcomings that have been documented is that the access ports on certain Diebold

---

[10] Tempress Nichols, "UGA Survey Finds Voters Favor Paper Trails From Electronic Voting," *NBC Augusta*, 16 August, 2007, available online at <http://www.nbcaugusta.com/news/local/4762281.html> (accessed 15 January 2008).

[11] Paul Krugman, "Hack the Vote," *The New York Times*, 2 December, 2003, available online at <http://query.nytimes.com/gst/fullpage.html?res=940CE2DF1E3AF931A35751C1A9659C8B63> (accessed 15 January, 2008).

[12] Tadayoshi Kohno, Adam Stubblefield, Aviel D. Rubin, and Dan Wallach, "Analysis of an Electronic Voting System," *IEEE Symposium on Security and Privacy 2004*, IEEE Computer Society Press, May 2004, available online at <http://avirubin.com/vote.pdf> (accessed 15 January, 2008); Aviel Rubin, *Brave New Ballot* (Random House, 2006); Matt Bishop, "Overview of Red Team Reports," report prepared for the California Top-to-Bottom Review of Voting Machines, July, 2007, available online at <http://www.sos.ca.gov/elections/voting_systems/ttbr/red_overview.pdf> (accessed 15 January, 2008).

[13] John Schwartz, "Sharing Pits Copyright Against Free Speech," *The New York Times*, 3 November, 2003, p. C1.

[14] Ariel J. Feldman, J. Alex Halderman, and Edward W. Felten, "Security Analysis of the Diebold AccuVote-TS Voting Machine," Center for Information Technology Policy and Dept. of Computer Science, and Woodrow Wilson School of Public and International Affairs, Princeton University, draft 13 September, 2006, available online at <http://itpolicy.princeton.edu/voting/ts-paper.pdf> (accessed 15 January, 2008).

[15] University of California Red Team Reports, "Diebold Elections Systems, Inc.," available online at <http://www.sos.ca.gov/elections/voting_systems/ttbr/red_diebold.pdf> (accessed 11 April, 2008).

machines could be unlocked with the same keys used in hotel mini-bars, standardized keys that can be purchased cheaply on the Internet.[16]

We applaud the watchdog organizations such as Verified Voting, which has called for a high level of scrutiny for equipment manufacturers. This being said, we feel that public discourse on the topic of voting security suffers from a number of deeply rooted misunderstandings. First, while the security vulnerabilities that have been discovered thus far are serious, we believe the likelihood of a hacker actually subverting an election is very small. And while *security* measures should continue to be refined, the most serious liability for election administration – the challenge most likely to flummox an election and most deserving of attention – continues to be *usability*. Second, while many proponents seem to hold up a paper trail as a panacea capable of securing elections and ensuring public confidence in results, we see several ways in which the paper trails of today could plausibly harm election facility and perhaps even *undermine* confidence in the results. Third, the mandate of a paper trail is likely to stifle the development of better voting systems that could improve elections in ways that heretofore have not been considered.

While the public is right to take alarm at the numerous security vulnerabilities that researchers working in laboratories have discovered in the design and software of voting machines, both optical scan and Direct-Recording Electronic (DRE) machines, it is easy to overstate the facility with which a machine could be subverted under real-world circumstances. Many of the vulnerabilities that have been exposed could only be exploited by an "omniscient hacker" who had full knowledge of a system's vulnerabilities and could tamper with a machine without the constraints of other safeguards, such as a locked room or a watchful supervisor.[17] To our knowledge, there has not been a single case of even *attempted* criminal machine subversion in the United States.[18]

The paucity of real-world examples of machine tampering must be weighed against numerous examples of elections that have been flummoxed by *usability* issues.

---

[16] Dan Kaplan, "Princeton prof says mini-bar keys open Diebold voting machines," *SC Magazine*, 18 September 2006, available online at <http://www.scmagazineus.com/Princeton-prof-says-mini-bar-keys-open-Diebold-voting-machines/article/33915/> (accessed 15 January, 2008).

[17] Michael Ian Shamos, "Paper v. Electronic Voting Records – An Assessment," working paper, draft April, 2004, available online at <http://euro.ecom.cmu.edu/people/faculty/mshamos/paper.htm> (accessed 16 January, 2008).

[18] Technology skeptics are most likely to point to the election results of Florida's Sarasota County in the 2006 midterm elections as a possible example of tampering. In this case, voting machines reported an abnormally high number of undervotes for a congressional race between Christine Jennings and Vern Buchanan. Widespread disagreement about the reason for this discrepancy remains. It is the unanimous consensus of our task force that the most likely cause was a poorly designed ballot. This is also the conclusion of a recent analysis by the Government Accountability Office, Nabajyoti Barkakati, "Results of GAO's Testing of Voting Systems Used in Sarasota County in Florida's 13th Congressional District," *GAO Reports*, 8 February, 2008, available online at <http://www.gao.gov/new.items/d08425t.pdf> (accessed 11 February, 2008). Also see Lauren Frisina, Michael C. Herron, James Honaker, and Jeffrey B. Lewis, "Ballot Formats, Touchscreens, and Undervotes: A Study of the 2006 Midterm Elections in Florida," working paper, draft 23 November, 2006, available online at <http://www.dartmouth.edu/~herron/> (accessed 16 January, 2008). But in disagreement, see Walter Mebane and David Dill, "Factors Associated with the Excessive CD-13 Undervote in the 2006 General Election in Sarasota County, Florida," working paper, draft 18 January, 2007, available online at <http://www.umich.edu/~wmebane/smachines1.pdf> (accessed 16 January, 2008).

Consider the following instances, which include the most significant election debacles of the last decade:

- The 2000 presidential election in Florida drew attention to the difficulty some voters have with certain ballot formats. The confusing butterfly ballot in Palm Beach County led a number of Al Gore's supporters to accidentally vote for Pat Buchanan, as well as a number of accidental "overvotes," sinking the contest into a dramatic month-long legal battle.[19]

- In Cuyahoga County's 2006 primary, Ohio election officials dealt with a variety of printing problems, including failure of printers to print, paper jams, and paper spooling issues. Since Ohio law requires that the paper printout be considered the official ballot, such incidents could threaten reliability and accuracy of votes.[20]

- In the 2006 midterm elections, the vote totals for Florida's Sarasota County reported an implausible 13% undervote rate in a congressional election between Christine Jennings and Vern Buchanan. The most likely explanation for this discrepancy is a poorly designed ballot, where Florida's gubernatorial race was displayed more prominently on the same screen as the congressional race.[21]

- In a 2006 primary, voters in Montgomery County, Maryland, including one of our authors, arrived at certain polling locations to discover that the voting machines had been delivered that morning without requisite access cards. Other polling sites had insufficient numbers of paper ballots for Democratic or Republican elections, forcing voters to either return later or not vote at all.[22]

- In California's 2008 presidential primary, an unclear ballot design and insufficient poll worker training led approximately 49,500 voters to unknowingly spoil their ballots. The voters did not realize that it was necessary to complete a circle indicating in which party's primary election they wanted their ballot to count.[23]

Of course, there is not a zero-sum relationship between usability and security. We certainly hope to achieve both, and indeed can imagine many ways in which they might support each other. However, the VVSG's software independence requirement

---

[19] Jonathan Wand, Kenneth Shotts, Walter R. Mebane, Jasjeet S. Sekhon, Michael Herron, and Henry E. Brady, "The Butterfly Did It: The Aberrant Vote for Buchanan in Palm Beach County, Florida," *American Political Science Review*, Vol. 95, No, 4 (Dec., 2001), pp. 793-810.

[20] "DRE Analysis for May 2006 Primary, Cuyahoga County, Ohio," Election Science Institute, August 2006, available online at <http://bocc.cuyahogacounty.us/GSC/pdf/esi_cuyahoga_final.pdf> (accessed 23 January, 2008).

[21] See f.n. 16.

[22] Olivia Doherty and C. Benjamin Ford, "Card snafu fixed by late morning, but voters angry," *The Gazette*, 13 September, 2006, available online at <http://www.gazette.net/stories/091306/montcou173827_31942.shtml> (accessed 15 January, 2008).

[23] Richard C. Paddock, "Some independent votes won't count in L.A. County," February 12, 2008, available online at <http://www.latimes.com/news/local/la-me-vote12feb12,0,6970893.story> (accessed 14 April, 2008).

would, as we continue to explain below, trade *real* sacrifices in usability for mere *perceived* gains in security.

The objective of the VVSG's software independence requirement is a double check for the electronic tallies that have become the subject of so much scrutiny. Optical scan systems, where voters receive a paper ballot and darken circles to select candidates, much like on a standardized test, have such a double check built in. Ballots are counted by electronic tabulators, either in the polling place or at some central location, but the original ballots are retained and can always be counted again – by hand if necessary – to confirm that the machine's count is accurate.

On the other hand, DREs – voting machines for which the voter directly manipulates a computer, usually with a touch screen – offer some advantages over optical scan systems. Whereas punch card and optical scan ballots are prone to ambiguity – dimpled and hanging chads for punch cards; stray marks and partially darkened ovals for optical scan – DREs provide election administrators with uniformly clean and unambiguous records. DREs can also include more mechanisms to catch common voter errors, such as overvotes and undervotes, as well as selecting a wrong candidate listed nearby on the ballot.[24] DREs can include interfaces to assist disabled voters, such as an audio interface for the blind or a sip-and-puff tube for individuals who cannot use their hands. Also, DREs can simplify the logistics involved in providing ballots in multiple formats or languages, which many jurisdictions must do in order to comply with the 1965 Voting Rights Act, and as is required for vote centers, the advantages of which we discuss above.

Today's DREs do not, however, leave a way for administrators to double check electronic tallies, as optical scan systems do. They can be fitted with printers designed to fulfill this purpose and, as the VVSG notes, *this is the only way to bring DREs into compliance with the VVSG using today's technology* (Introduction, Chapter 2.4.1). The software independence requirement can therefore be read, with the exception of the "innovations class" provision, which we discuss below, as a paper requirement.

It seems to us that, of the reasons offered by paper trail advocates, the most compelling reason to require DREs to produce a paper record is the argument that such records will guard against election tampering. However, we believe that paper records could encumber elections and obscure the legitimacy of results, and therefore should not be mandated.

Election fraud far predates the advent of mechanical and electronic balloting. Indeed America's storied history of vote tampering extends nearly two hundred years before lever machines came into popular use in the 1960s. One scholar finds strong evidence that the statewide vote to ratify the Massachusetts state constitution in 1780 fell far short of the required two thirds majority, a problem that the constitutional convention may have resolved by fabricating the necessary ballots.[25] Few will be surprised that this tradition continued with ballot boxes robbed at

---

[24] On the "proximity" effect, see R. Michael Alvarez, Melanie Goodrich, Thad E. Hall, D. Roderick Kiewiet, and Sarah M. Sled, "The Complexity of the California Recall Election," *PS: Political Science and Politics,* Vol. 34, No. 1 (Jan., 2004), pp. 23-26.

[25] Tracy Campbell, *Deliver the Vote: A History of Election Fraud, An American Political Tradition, 1742-2004* (New York: Carroll & Graf Publishers, 2006), p. 9.

gunpoint,[26] flagrantly stuffed,[27] and mysteriously "found" months after an election, [28] to name but a handful of myriad anecdotes.[29]

The common theme is that paper is an inherently vulnerable medium with which to conduct elections. More easily than electronic records, paper can easily be modified, forged, lost, or destroyed, casting doubt on election results. This trend has not been escaped in modern times, as is evinced by the controversial discovery, several days after an excruciatingly close Washington state gubernatorial election in 2004, of 150 misplaced ballots at a polling site.[30]

Of course, even if it is more difficult to secure and back-up paper records than electronic ones, perhaps they are still necessary as an audit mechanism with which to reinforce an electronic tally. "Two records are better than one," some are sure to argue. The pertinent question seems to be, "Does the addition of a paper trail to the voting system increase the security and accuracy of that system?" We answer this question in the negative. We fear that the addition of a paper trail may afflict election results with uncertainty that is different in nature, but no less severe, than the uncertainty that currently exists.

Consider one hypothetical way in which the addition of a paper trail could facilitate corruption. Suppose a Republican partisan sought to steal an election from a popular candidate he disliked. Whereas the subversion of a purely electronic vote count would presumably require both substantial expertise in computer science and knowledge of the voting system software, the addition of a paper trail opens up new avenues. Our thief could simply steal one or more boxes of ballots from a Democratic-leaning precinct, the sort of theft that has occurred numerous times in the past.[31] Alternatively, he might seek to replicate a number of counterfeit ballots and mix them in with the genuine ones. In many states, existing laws designate the paper ballots as the ballot of record in the event of any discrepancy between paper and electronic totals. One popular bill introduced by Representative Rush Holt (D-NJ) seeks to nationalize this policy.[32] In any case, even if the discrepancy were resolved in favor of the genuine winner, the thief would have succeeded in casting doubt on election's legitimacy.

The proposal to add paper trail to DREs calls to mind the metaphor of a fortress. What design would best provide for the fortress's safety: To have one gate heavily fortified with a full complement of soldiers, or to incorporate a second gate and divide the defenders between the two? It seems to us that just as the addition of a second gate lends no additional protection, so too might the creation of parallel vote tallies – especially those that are hastily conceived – increase the opportunities for a successful attack.

---

[26] *Ibid*., p. 123.

[27] *Ibid*., p. 176.

[28] *Ibid*., p. 218.

[29] For a current commentary on the prevalence and character of election fraud, see R. Michael Alvarez, Thad E. Hall, and Susan D. Hyde, *Election Fraud: Detecting and Deterring Electoral Manipulation* (Washington, D.C.: Brookings Institution Press, forthcoming in 2008).

[30] *Ibid*., p. 329.

[31] Campbell; Andrew Gumbel, *Steal This Vote* (New York: Nation Books, 2005).

[32] Voter Confidence and Increased Accessibility Act of 2007, HR 811, 110th Cong., 1st Sess. (2007).

A consideration of some of the exciting systems that would be *prohibited* by the software independence requirement further illustrates how it is an inferior course by which to secure elections. Punchscan, a system developed by David Chaum and researchers at the University of Maryland, Baltimore County, lets voters confirm not only that their votes were *cast* correctly – in other words, that one medium captured a correct vote at one instant in time – but also that their votes were in fact *counted* correctly; after the election, voters can use the Internet to view a partial image of their individual ballots as they were tabulated.[33] (Only part of the ballot is displayed, so that voters cannot prove for whom they voted, which could lead to vote buying.) Another system called Prime III, developed by researchers at Auburn University, would employ a separate electronic "witness" in each voting booth. The witness, which would operate independently of the DRE machine, could more efficiently double-check the DRE's tally of votes while safeguarding privacy and being more accessible to the disabled.[34] Although clearly very secure – far more secure than a DRE with a simple paper trail – neither of these systems satisfies the requirements of software independence because, although unlikely, they could theoretically be undetectably subverted by a computer hacker. It seems to us that the best way to ensure the long-term effectiveness of U.S. elections is to encourage systems like Punchscan, Prime III, improved forms of physical records, and others that have not even been imagined.

The VVSG does make one attempt to allow for new innovations. Chapter 2.7.2 of Part 1 describes a process by which a voting system manufacturer might earn certification for a system that does not use independent voter verified records – in a reasonable expectation, paper records. However, according to the current draft, *even systems seeking certification through this* "innovation class," *must still be software independent*, and so even with the innovation provision in mind, the VVSG offers no relief for systems like Punchscan and Prime III, which do not fit the definition of software independence. *It remains unclear to us what the purpose of the innovation class provision is, if not to offer some relief from the constraints of software independence.* The next draft of the VVSG should discard this provision or substantially reconsider its structure.

One possible alternative to the software independence provision would be a less restrictive requirement for *single agent* independence. This measure would preserve the "double check" principle, with redundant mechanisms that corroborate tallies, but leaves vendors the freedom to design double checks that rely on software. It should be noted, however, that developers might find even this standard to be restrictive. It would prohibit, for instance, the Punchscan system described above, as the system relies on a single cryptographic agent to decode votes.

Our beliefs that software independence would decrease election security and stifle innovation are the root of our concern about the proposition. However, we also call the reader's attention to a number of other factors to consider when deciding whether to fit DREs with a Voter Verified Paper Audit Trail (VVPAT):

[33] Stefan Popoveniuc and Ben Hosp, "An Introduction to Punchscan," George Washington University Computer Science Department, 15 October, 2006, available online at <http://punchscan.org/papers/popoveniuc_hosp_punchscan_introduction.pdf> (accessed 25 March, 2008).
[34] Philicity Williams, et al, "Prime III: Where Usable Security and Electronic Voting Meet," Auburn University Department of Computer Science and Software Engineering, available online at <http://www.usablesecurity.org/papers/primeIII.pdf> (accessed 25 March 2008).

- VVPATs greatly increase the logistical complexity of elections. VVPAT printers are prone to jamming – many fail on Election Day[35] – and frequently run out of their consumable supplies, ink and paper.[36] In addition, more voters ask for help when confronted with a VVPAT than when they use a paperless DRE.[37] As such, the average time per vote is increased, and election administrators using VVPATs should expect to purchase additional machines and hire additional poll workers in order to keep up with high turnout.

- At the same time, VVPATs do not necessarily increase voter accuracy. In several laboratory trials, voters made *more* mistakes when using a DRE with a VVPAT than a paperless DRE.[38] Numerous voters reported a mismatch between the votes recorded on the voting system and on the contemporaneous paper trail. It is almost certain that these perceived mismatches were due to misremembered votes or incorrect selections.[39] But most voters don't scrutinize the VVPAT at all.[40]

- The vast majority of VVPATs record ballots on continuous paper rolls, and so the cast ballots remain in the precise order in which they were cast, a characteristic that may present problems for voter privacy. While it would be difficult for a poll worker to match his or her neighbors with their votes in a large urban precinct, it would be much easier in a rural area where only a few dozen people arrive at the polling place on Election Day. In either location, without proper supervision and handling procedures, it would be comparatively easy to match the first and last voters to arrive, as their ballots would be easy to identify on the paper roll.

- Like optical scan ballots, VVPATs cannot treat blind voters in an identical manner as sighted voters. A blind voter would have to ask for assistance in order to read his or her VVPAT receipt.

These points are not to suggest that VVPAT systems are beyond reasonable implementation or that alternative approaches are without shortcomings, but rather to remind that there are problems with paper audit trails, as there are with all systems. Perhaps paper verification for DREs should be preserved as an *option* for places that are especially concerned about the specific kind of subversion that paper may prevent. *Mandating* a paper record would do little or nothing to improve election security, but it may do much to encumber election facility. Past experience tells us that the latter concern should take precedence.

---

[35] Paul S. Herrnson, et al., *Voting Technology,* chapter 6; Ted Selker, "Processes Can Improve Electronic Voting: A Case Study of an Election," Caltech/MIT Voting Technology Project, working paper (October, 2004), available online at <http://www.votingtechnologyproject.org/media/documents/vtp_wp17.pdf> (accessed 14 April, 2008).

[36] "DRE Analysis for May 2006 Primary..."

[37] Paul S. Herrnson, et al. *Voting Technology.*

[38] *Ibid.*, p. 134.

[39] *Ibid.*, p. 127.

[40] Ted Selker, "An Active Approach to Voter Verification," Caltech/MIT Voting Technology Project, working paper #28 (May, 2005), available online at <http://vote.caltech.edu/media/documents/wps/vtp_wp28.pdf> (accessed 18 January, 2008).

The VVSG takes an "anti-DRE" approach

It is worth recalling that the rush to adopt DRE modes of balloting was largely the result of federal initiatives and legislation in the aftermath of the 2000 election. The offer of federal funding prompted many election administrators to move from paper and punch card balloting systems to electronic voting machines. The adoption of DRE systems was made before the technology and operational procedures for these new balloting systems were sufficiently refined to avoid what are now seen as significant obstacles to their secure and efficient operation. In the whirlwind of bad press that has ensued, it is easy to forget some of the advantages that DRE voting can offer, such as an interactive environment; greatly increased accessibility for the disabled; clean, easy-to-tabulate ballots; error-checking mechanisms; and, we would argue, greater security.

In at least some ways, the VVSG takes an anti-DRE approach, holding electronic voting to a much higher security standard than paper records. The VVSG goes to great lengths to make its consideration of DRE vulnerabilities comprehensive, but does not note the substantial vulnerabilities that exist in alternative approaches. It is essential to understand that *all* approaches to voting are highly vulnerable without the more fundamental protections of chain-of-custody and monitoring; *any* voting system becomes highly vulnerable if left unattended, yet the VVSG does not scrutinize optical scan systems through the same lens.

Consider the standards that the VVSG applies to vulnerability testing. Standards, according to the usual definition, specify requirements for an object, based either on how the object is supposed to perform or regarding how it is to be designed. But the "open-ended" vulnerability testing required by the VVSG does not seem to fall into either category of standard, as they are by definition "open-ended." The intention of the "open-ended" testing priorities (because they really cannot be called standards in the typical use of the term) is to find security vulnerabilities, based on the expert opinion of the "open ended vulnerability testing team."

The provisions for vulnerability testing should clearly specify whether to use design or performance standards, and should be done to determine whether a voting system either meets or does not meet these standards. Otherwise, it will not be possible to judge whether or not a particular voting system has been tested adequately for vulnerabilities, because the current model relies mainly on the "experience" of the vulnerability testing team. As these teams have considerable discretion in the "open-ended" model, it will not be clear that any two voting systems have been tested to the same degree. For example, a vulnerability testing team in this "open-ended" model might concentrate on a certain set of tests for one voting system, and a different set of tests for another voting system; these inconsistencies will almost certainly arise for vulnerability testing teams in different testing laboratories.

A Time to wait

In May of 2007, Florida's Governor Charlie Crist (R) signed a bill requiring all election jurisdictions throughout the state to scrap touch-screen voting machines, and replace them with optical scan systems.[41] It is but one poignant example of waste,

---

[41] Fineout, "Voting Change Irks Supervisors."

considering that the systems being discarded were purchased since 2002 using $11.6 million federal tax dollars.[42] We strongly sympathize with the frustration of George Gilbert, Director of the Board of Elections of Guilford County, North Carolina, who remarked about his own state, "The problem for the future will not be technical obsolescence, but statutory obsolescence. I have about $3 million worth of equipment that still works, and I am trying to decide which landfill to take it to."[43]

Given time and sufficient liberties, American ingenuity will rise to the new challenges of democracy. But it is clear that much work has to be done before we fully understand the intricacies of the options before us. In the meantime, we are tired of the waste and headache that inevitably come with hasty reactions and partial solutions.

Addressing Voting System Usability

An essential topic

Although the lion's share of attention paid to election reform in recent years has focused on election security, the most serious shortcomings of recent elections have come from simple human error, whether it was the product of administrative or technological shortcomings. As we note above, the quagmires of Florida in 2000; Montgomery, Cuyahoga, and Sarasota Counties in 2006; and others, were all attributable to human mistakes. Standing political science research also suggests that a significant portion of the errors and failures in the outcome of elections originates with administrative behavior.[44]

It is easy for usability errors, such as a poorly arranged ballot, a voting machine with unclear instructions, or a poorly-calibrated touch screen to escape notice until after an election has been held. Yet because such errors are often replicated across a number of jurisdictions, they have the potential to affect many votes. As we saw in 2000, even a small ambiguity, such as which hole in a punch card ballot corresponds to which candidate, can lead to a large number of spoiled or improperly cast ballots. Additional errors may be the result of challenges faced by poll workers, including setting up, maintaining, and properly closing down voting equipment.

It is fitting, therefore, that the VVSG dedicates significant attention to voting system usability. Unfortunately, the pertinent section (Part 1, Chapter 3) and the literature on which it is based appear to be ill conceived. The VVSG's treatment of usability overlooks standard practices of social science and, as a result, arrives at usability measures that are almost entirely arbitrary. Unfortunately, we find this section to be beyond the remedy of marginal changes. A thorough rewrite is in order.

---

[42] "Report to Congress on State Governments' Expenditures of Help American Vote Act Funds," United States Election Assistance Commission, July 2007, available online at <http://www.eac.gov/clearinghouse/reports-on-state-expenditures-of-hava-funds-1/clearinghouse/docs/eac-report-to-congress-on-state-expenditures-of-hava-funds-july-2007/> (accessed 16 January 2008).
[43] Remarks delivered at "Research on the U.S. Voting System: Matching Needs with Knowledge," *American Association for the Advancement of Science*, Washington, D.C., November 27-28, 2006.
[44] Stephen Ansolabehere and Charles Stewart III, "Residual Votes Attributable to Technology," *Journal of Politics*, Vol. 67, No. 2 (May, 2005), pp. 365-389.

We begin by discussing some of the shortcomings of the VVSG's usability section, and then move on to suggestions for how it might be rewritten.

Deriving usability benchmarks

The heart of the VVSG's usability requirements relies on a separate document, "Usability Performance Benchmarks for the VVSG," which lays out the methodology through which the VVSG's benchmarks were derived.[45] For instance, the VVSG calls for 98% of a given machine's test users to be able to successfully complete a ballot, stating that this number allows "better systems to pass the test, while preventing certification of poorer systems" (Part 1, Chapter 3).

Unfortunately, the laboratory trials through which the VVSG's usability benchmarks are derived were poorly conceived. We begin with the method by which voting systems were selected for trial. It seems that the VVSG's usability benchmarks are entirely based on trials with fewer than 200 people and only four voting systems which, in order to protect manufacturer anonymity, are never named. Not only is this a disturbingly small window into the wide spectrum of available systems, but it appears that the four machines were used simply because they are the machines that manufacturers decided to make available. We are left wondering whether manufacturers had any incentive to skew the test results by providing especially high quality machines, especially low quality machines, or to opt out of the testing altogether, and are wholly unconvinced that these four machines are in any way representative of the market in general.

The next step in the TGDC's tests to develop usability benchmarks was to ask voting system manufacturers to design a ballot for a mock election. The manufacturers were given lists of fictitious candidates and asked to design a ballot so as to maximize usability. Here our concerns are twofold. First, by inviting system manufacturers to design the ballot, the TGDC introduces an unnecessary variable into the analysis. Usability scores will be influenced not only by system design, but also the designers' competence in ballot design. As such, it is more difficult to isolate the causes of any performance differences between any two machines. A better approach would be to have the *same* ballot design expert construct all the ballots, so that this variable is held constant across trials, or to randomly assign real election officials to design each ballot.[46]

Second, while we understand that the use of fictitious candidate and party names was intended to remove visceral reactions from behavior in a mock election, this step away from reality should be supplemented with other approaches. For instance, the TGDC should conduct some additional tests where subjects are asked to read a voting booklet with short descriptions of people (perhaps even real candidates), to write down their choices, and then to "go to the polls" and enter their votes. We also recommend asking people to include some information about the intensity they attach to prospective votes. Then researchers could compare the choices written in the voting booklet with the choices on the voting systems. The disconnect from

---

[45] "Usability Performance Benchmarks For the Voluntary Voting Systems Guidelines," prepared at the direction of the Human Factors and Privacy (HFP) Subcommittee of the Technical Guidelines Development Committee of the Election Assistance Commission, 17 August, 2007, available online at <http://vote.nist.gov/meeting-08172007/Usability-Benchmarks-081707.pdf> (accessed 22 January 2008).
[46] See, for example, Herrnson, et al., *Voting Technology*.

reality exhibited in the TGDC's narrow approach seems to ignore decades of political science and survey research on how to measure and think about voting choices.

Also problematic is the selection of test subjects. At no point does the "Usability Performance Benchmarks" document describe how test participants were recruited, other than to say that they were from certain areas and met certain demographic characteristics. Were the subjects recruited through an advertising campaign, through a website, through the National Institute for Standards and Technology, or through some other means? It is important to provide this information to allay concerns that the recruitment procedure could lead to a biased subject pool.

More transparently problematic is the conscious decision to eliminate certain demographic groups from the test pool. Specifically, the TGDC's tests included only young (under 54) and college-educated individuals, the sorts of people *least* prone to be affected by usability issues because of their general familiarity with technology. To justify this decision, the "Usability Performance Benchmarks" document says

> These [demographic] requirements were selected in part because if the test could detect usability differences and all the expected errors with this population, it would detect differences and errors with older and less educated populations as well.[47]

In other words, it seems that the TGDC sees the constriction of the subject pool as a way to set a high bar for excellence; if problems emerge among a well-educated and young testing pool, then surely they would among a less-educated, older population. There might be some truth to this assertion, but it falls before a serious concern. Voting system research has shown that there is often an interaction between voting system type and individual characteristics, such as education, such that those with lower education experience problems when those with higher education do not. One could imagine, therefore, that this testing procedure would give a pass to two systems – and even rate them equally in terms of performance – but that one of the two systems would exhibit an interaction with education such that those with lower education would do very poorly with it. Hence, one of the voting systems would clearly be better than the other.

To evoke a metaphor, the VVSG is calibrating the bar for a team of Olympic pole vaulters. Of course the athletes can consistently succeed even at a challenging level, but the average individual will consistently fail at the same level. It seems obvious to us that the best way to design systems so as to be usable by the disabled, less educated, or other disadvantaged populations is to make these groups, if anything, the *focus*, rather than the exclusion of study.

Perhaps the most serious problem with the TGDC's derivation of usability benchmarks is the lack of an experimental control. In the "experiments" that were done to test usability, there is absolutely no indication that the human subjects who tested the machines were randomly assigned to different voting system types – a step that could have provided the justification for the statistical tests that are used to measure differences in performance. Instead, it seems that separate pools of test subjects were recruited to test each machine. While there was an effort to balance certain characteristics of these groups (e.g. the male/female ratio), there is no way to eliminate the possibility of *unobserved* preexisting differences between them. In

---

[47] "Usability Performance Benchmarks," p. 5.

contrast, random assignment can ensure that there are no systematic differences between the groups assigned to each type of machine. (We should note that perhaps random assignment was used, but if so, the absence of any mention is in itself a serious omission.) We understand the complexities of using random assignment in a testing procedure that would serially test systems, but there is no excuse for not using this method in the development of standards.

The VVSG would also be well served to consider additional metrics for usability. There are a number of meaningful metrics that can be used to assess voter confidence that go beyond what VVSG's employs. Two such metrics are 1) being able to vote independently without feeling the need for help and 2) being able to vote without actually receiving help from a poll worker. These metrics are extremely important from the standpoints of both voters and election officials. For voters, needing help is related to voter confidence and trust in the voting system and could suggest the voter is experiencing confusion or frustration with several aspects of the voting process. Beyond that, voters who need help take longer to vote than others and give up a measure of privacy when making their selections.[48]

In addition, the TGDC could put the results of its tests in a more meaningful context by including tests of systems that are "known quantities." We wish the "Usability" report had compared the results of the DREs and optical scan systems tested with punch-card systems, which almost everybody (perhaps even everybody) can agree are flawed. Moreover, some thought should be given to how to link one set of tests (using one subject pool) with another set of tests (using another subject pool). Linkages could be made by having at least some of the same machines tested twice. Some input from social science statisticians could significantly improve the results.

Finally, the reasoning through which the "Usability" document interprets test results is entirely arbitrary. The standard employed is to choose benchmarks "so that they are achievable by some systems, but not so low as to be trivially easy to meet with any of the current voting system implementations."[49] It is comforting that the document then expresses considerable willingness to revise the benchmarks, because the approach used is misguided. There is little sense in setting benchmarks relative to the performance of existing machines, especially through the use of a tiny sample of convenience, as the TGDC does. The pertinent question for policymakers is not, "How many systems will fail?" but rather, "Remembering values of inclusiveness and egality, what level of performance do we find acceptable?"[50] The TGDC should more fully and transparently consider the effect of its benchmarks on real life ballots and clearly state what level of performance – how few miscast ballots, etc. – it strives to achieve.

Incorporating additional approaches

Although informative to some degree, laboratory experiments, such as those conducted by the TGDC, offer an incomplete view voting system usability. This is because the transparently artificial environment of a test laboratory may affect test subjects' behavior. One alternative approach that we hope will be adopted with

---

[48] Paul Herrnson, et al., *Voting Technology*.

[49] "Usability Performance Benchmarks," p. 21.

[50] Of course the answer to this question should also take account of groups that approach voting differently, such as individuals with disabilities.

greater frequency in the future is a field experiment, where voters are studied under real voting conditions.[51] There are several examples of very successful collaborations between university-based researchers and voting administrators, some of which have been funded by the National Science Foundation and the Carnegie Corporation of New York; others have been documented by the Pew Charitable Trusts.[52] One author of this report undertook such a collaboration just last year.[53] The difficulty of conducting a field experiment on a topic in election administration will vary by location, but some states have well-established laws enshrining municipal control over local elections, providing fertile ground for such experimentation.[54]

One means of obtaining a critical mass of field experiments on issues pertaining to voter satisfaction, confidence, and performance is to establish an industry-financed research fund managed by an independent third party. For example, industry representatives might endow a research fund to which qualified researchers seeking to study metrics related to the performance of various balloting devices would apply. Proposals would be peer reviewed for funding on relevant topics. All applicants would be required to have a "community-based" partner who agrees to allow researchers to assess polling places to observe, interview, and measure voting behavior. This approach, reminiscent of that used by the National Science Foundation's Digital Government Division, would assure that local election officials obtained answers to those issues and questions most critical to their needs.

It is discouraging and worrisome to see that the VVSG document places so much emphasis on the problems of computer hacking and voter verified ballots and so little on usability when the research on actual elections has found virtually no evidence of the former and abundant and serious evidence of the latter. It is especially discouraging when there are social science tools readily available that would make it possible to significantly improve the usability testing of voting systems. Unfortunately, the VVSG document and the "Usability Performance Benchmarks" do not use these methods.

The Proper Role of the VVSG

All things to all people

A landmark document like the VVSG fulfills many roles, and it is clear that one of the challenges its authors faced was designing a document that would have many different uses for many different people. The first paragraph states that the language was written so as to be usable not only by voting system manufacturers and test

---

[51] On this point, see Caltech MIT Voting Technology Project, "Voting: What is, What Could Be" (July, 2001), available online at
<http://www.vote.caltech.edu/media/documents/july01/July01_VTP_Voting_Report_Entire.pdf> (accessed 15 April, 2008). On the advantages of field experimentation in general, see James N. Druckman, et al., "The Growth and Development of Experimental Research in Political Science," *American Political Science Review*, Vol. 100, No. 4 (Nov., 2006), pp. 627-636.

[52] "Case Study: Election Partnerships," *Electionline.org Briefing*, June, 2007, available online at
<http://www.pewcenteronthestates.org/uploadedFiles/EB18.pdf> (accessed 8 February, 2008).

[53] Robert Stein and Greg Vonnahme, "Voting technology, election administration, and voter behavior," presented at the 2007 Midwest Political Science Association Conference, Chicago, IL, April 1-3.

[54] See, for example, Article XX, Section 6 of the Colorado State Constitution.

laboratories, but also "election officials, legislators, voting system procurement officials, various voting interest organizations and researchers, and the public at large" (Introduction, Chapter 1). Although this effort is laudable, it may be untenable. We are unconvinced that the nearly 600 page draft behemoth that currently exists is of much use to those who lack great familiarity with election reform issues; its language and requirements are far too complex for this purpose.

Rather than trying to be all things to all people, we recommend that the VVSG focus on the primary stakeholders, namely voting system manufacturers and test laboratories. In addition, the TGDC or the EAC might prepare shorter interpretive companion documents that could be more useful to individuals and groups that will be deterred by extreme detail. There is another concern, as well. In its current form, the document's impenetrability will likely lessen its effect. The VVSG is so formidable in its scope that it is easy for states to ignore its requirements which, after all, are voluntary. In some places, it may be difficult for a state to determine whether or not it is meeting a certain standard. Interpretive documents would serve to ameliorate this shortcoming.

Neglected topics

The VVSG attempts to focus solely on voting technology, considering issues of process and administration outside of its scope. (We have discussed above how this supposed agnosticism is, to an extent, illusory, since technology and process are inextricably linked.) It is disappointing that the federal government would conduct such a thorough review of voting technology – the VVSG truly is a monumental undertaking – without a similarly comprehensive effort to address election administration, which we consider to be a much more vulnerable component of US democracy. As such, even a much improved VVSG will be only a small step toward the basic goal of HAVA and the EAC – to provide for inclusive elections that accurately record and report all votes that are cast. The EAC should consider sponsoring a thorough scientific review of some of the following topics:

- **Ballot Design** – Although the VVSG does address some elements of ballot design (Part 1, Chapter 3), the EAC should go much farther. No mention is made, for example, of the problems associated with "banner effects," the difficulty that we believe is at fault for the recent election dispute in Sarasota County, Florida, discussed above. We hope that future efforts will give much more attention to the voluminous literature that exists for this important topic,[55] if not in the form of requirements, then at least in the form of "best practices."

- **Voter Registration** – Voter registration systems are clearly critical for the conduct of elections. Section 221(e)(2)(A) of HAVA requires the TGDC to develop standards the statewide voter registration databases. Since the VVSG

---

[55] Frisina, et al.; David C. Kimball and Martha Kropf, "Dos and Don'ts of Ballot Design," *AEI-Brookings Election Reform Project Newsletter*, 3 October, 2007, available online at <http://www.electionreformproject.org>; David C. Kimball and Martha Kropf, "Ballot Design and Unrecorded Votes on Paper-Based Ballots," *Public Opinion Quarterly,* Vol. 69, No. 4 (Winter, 2005), pp. 508-529. Also see numerous studies conducted by the Center for American Politics and Citizenship at the University of Maryland, which one of the authors directs, at <http://www.capc.umd.edu/rpts/votingtech_par.html>.

does not address registration systems, it is essential that a separate effort do so in the near future.

- **Absentee Voting** – The VVSG primarily conceives of elections in a traditional fashion. That is to say that its focus is on elections conducted in polling places and on a uniform election day. By placing new restrictions on the way in which "traditional" elections are conducted, it may hasten the trend towards non-traditional voting, such as large-scale voting by absentee ballot. In 2004, 15% of the ballots counted nationally were absentee votes, with states like California and Arizona collecting more than 30% of their ballots through the mail.[56] Oregon conducts all-mail elections. The EAC and the TGDC should consider minimum standards for the usability and security of absentee ballots in tandem with new voting system standards. Absentee voting should not serve as a mechanism through which states sidestep the effort to make voting more friendly and secure.

- **Accessibility** – The VVSG provides accessibility requirements by developing standards not for all voting systems, but for what it calls an "Accessible voting station" (Part 1, Chapter 3). Thus, instead of working to develop voting systems that are universally accessible, the VVSG assumes that voters with disabilities may generally be relegated to use a voting system that differs from the voting system used by non-disabled voters. Such a system may in practice be expensive and complex to implement. The EAC and TGDC should develop universal accessibility standards that apply to all voting systems, and not just relegate accessibility standards to the "Accessible voting station." The EAC and TGDC should also include in testing of voting systems subjects who are vision impaired and blind, who have physical or cognitive disabilities, and for whom English is not their primary language.

- **Poll Worker Recruitment and Training** – The tradeoffs between different technologies may be influenced by how elections are administered and where they are held. There is evidence that the performance of different voting technologies is dependent on the quantity and quality of Election Day poll workers.[57] The number of adequately trained Election Day workers may have a direct impact on voter usability, satisfaction, and confidence with different voting technologies. Poll worker training should be an essential component of any recommendation to enhance our voting systems.

Conclusion

Since 2000, election conduct in the United States has undergone a revolution. The consistent theme of this report is that the revolution is still very much under way. It

---

[56] John C. Fortier, *Absentee and Early Voting: Trends, Promises, and Perils* (Washington, DC: AEI Press, 2004).

[57] R. Michael Alvarez and Thad E. Hall, "Controlling Democracy: The Principal-agent Problems in Election Administration," *Policy Studies Journal*, Vol. 34, No. 4 (Nov., 2006), pp. 491-510; Joseph Quin Monson, Ryan Claassen, and Kelly D. Patterson, "Confidence in Poll Workers: The Voter-Poll Worker Interaction," presented at the 2007 meeting of the American Political Science Association, Chicago IL, August 30th – September 2nd.

remains to be seen whether new technology and improved practices can live up to their promise to deliver results that are more reliable and worthy of public trust. Tempting though it may be to respond to the numerous problems and false starts that have occurred with strict regulation, we fear that such rigidity would substantially delay the arrival of a more efficacious election system. Now is the time for a slow, sober reflection on the changes that have been made thus far. Only with such a reflection will we have the understanding necessary to enact effective reform.

The fair conduct of elections is and always has been a vexing and controversial endeavor. This reality is unlikely to change dramatically in the near future, even as the conduct of elections continues to improve; it often takes years to realize the fruits of such improvements. Of course, this is hardly an excuse not to try. We look forward to playing an active role in this exciting process.

Members of the AEI-Brookings VVSG Task Force

R. Michael Alvarez, California Institute of Technology
Henry E. Brady, University of California, Berkeley
Paul S. Herrnson, University of Maryland
Ted Selker, Massachusetts Institute of Technology
Robert M. Stein, Rice University

Members *ex-officio*
John Fortier, American Enterprise Institute
Thomas Mann, The Brookings Institution
Norman J. Ornstein, American Enterprise Institute

Research Staff:

Molly Reynolds, The Brookings Institution
Timothy J. Ryan, AEI; *lead report architect*
Matthew Weil, AEI

Interns

Brian Haghighi
Jamie Hester

The report also benefited from the comments of Erik K. Antonsson, Professor of Mechanical Engineering at the California Institute of Technology.

Member biographies

R. Michael Alvarez is a Professor of Political Science at the *California Institute of Technology.* His areas of expertise include U.S. Presidential elections and American public opinion. Alvarez was selected by *Scientific American* magazine to be on the 2004 "Scientific American 50" for his outstanding scientific and technological contributions to help improve the U.S. voting system. He is Co-Director of the Caltech-MIT Voting Technology Project, which researches technological solutions to electoral problems, and has published scores of articles, reports and books on voting technology and election reform. His three books in this are *Point, Click and Vote: The Future of Internet Voting* (Brookings Institution Press, 2004); *Electronic Elections: The Perils and Promises of Digital Democracy* (Princeton University Press, 2008); and *Election Fraud: Detecting and Deterring Electoral Manipulation* (Brookings Institution Press, 2008).

Henry Brady is a Professor of Political Science and Public Policy at UC Berkeley. A former president of the American Political Science Association's Political Methodology Society, he is co-author of *Voice and Equality: Civic Voluntarism in American Politics* (Harvard University Press, 1995), co-editor of *Rethinking Social Inquiry* (Rowman and Littlefield, 2004), *Capturing Campaign Effects* (Michigan University Press, 2006), and the *Oxford Handbook of Political Methodology* (Oxford University Press, 2008), as well as the author of numerous articles on political participation and political methodology."

John C. Fortier is a Research Fellow at AEI. He is senior counselor to the AEI-Brookings Election Reform Project and executive director of the Continuity of Government Commission. He is a political scientist who has taught at the University of Pennsylvania, University of Delaware, Boston College, and Harvard University. Mr. Fortier has written numerous scholarly and popular articles, and his books include *Absentee and Early Voting: Trends, Promises, and Perils* (AEI Press, 2006), *After the People Vote: A Guide to the Electoral College* (third edition, AEI Press, 2004), and *Second-Term Blues: How George W. Bush Has Governed* (Brookings Institution Press, 2007). Mr. Fortier writes a column for *Politico* and is a frequent radio and television commentator on the presidency, Congress, and elections.

Paul S. Herrnson is Director of the Center for American Politics and Citizenship and a Professor in the Department of Government and Politics at the University of Maryland, College Park. He is the principal investigator of the National Science Foundation-funded "Assessment of Voting Technology and Ballot Design." He is co-author of *Voting Technology: The Not-So-Simple Act of Casting a Ballot* (Brookings Institution Press, 2008) as well as numerous books and articles.

Thomas E. Mann is the W. Averell Harriman Chair and Senior Fellow in Governance Studies at The Brookings Institution and co-director of the AEI-Brookings Election Reform Project. Mann is a fellow of the American Academy of Arts and Sciences and a member of the Council on Foreign Relations. He is a recipient of the American Political Science Association's Frank J. Goodnow and Charles E. Merriam Awards. His is author, most recently, of *The Broken Branch: How Congress Is Failing America and How to Get It Back on Track* (with Norman J. Ornstein; Oxford University Press, 2006).

Norman J. Ornstein is a Resident Scholar at AEI. He serves as senior counselor to the Continuity of Government Commission, co-directs the AEI-Brookings Election Reform Project, and is a member of the boards of the Public Broadcasting Service, the Campaign Legal Center, and the U.S. Capitol Historical Society. Mr. Ornstein was elected as a fellow of the American Academy of Arts and Sciences in 2004. His many books include *Second-Term Blues: How George W. Bush Has Governed* (edited with John C. Fortier; Brookings Institution Press, 2007) and *The Broken Branch: How Congress Is Failing America and How to Get It Back on Track* (with Thomas E. Mann; Oxford University Press, 2006).

Molly Reynolds is a Senior Research Assistant at the Brookings Institution. She is a 2006 graduate of Smith College.

Timothy J. Ryan is a Research Assistant at the American Enterprise Institute. He graduated from Tufts University in 2006 and will begin the doctoral program in American Politics at the University of Michigan in 2008.

Ted Selker is an Associate Professor of Context-Aware Computing at MIT and is co-Director of the Caltech/MIT Voting Technology Project. He has also served as a consulting professor at Stanford University. His area of expertise is the intersection of product design and human intention. He is the author of 18 patents and 20 papers in refereed journals and conference proceedings.

Robert Stein is the faculty director of the Center for Civic Engagement and the Lena Gohlman Fox Professor of Political Science at Rice University. His expertise includes urban politics and public policy. Dr. Stein's current research examines innovative election administration practices such as consolidated polling places and early voting. His articles on these topics have appeared in the *Journal of Politics*, the *Election Law Journal*, *Electoral Studies*, and other publications.

Matthew Weil was a Research Assistant at the American Enterprise Institute from 2006 until 2008. He is a 2006 graduate of the University of Pennsylvania.