

GAO REPORT NOT A CLEAN BILL OF HEALTH FOR VOTING MACHINES
Limited Scope Investigation Not Conclusive

Verified Voting Foundation (VVF) in consultation with computer science and voting systems experts evaluated the recently leaked draft report by the GAO for the Committee on House Administration's Task Force for the Contested Election in the 13th Congressional District of Florida, and concludes that the findings are not sufficient to exonerate the voting machines in determining what caused a massive undervote in the Florida District 13 contest of 2006.

The complexity of the computerized voting systems under investigation, and the poor quality of evidence trail they produce, undermines any investigation's capacity to determine conclusively what occurred in this election. (Had the voting system in use been a paper ballot system, it may not have failed at all. Examining the issues leading to an excessively high undervote rate may also have been significantly easier.)

Limitations on the scope of the investigation hinder further the ability to uncover whether bugs in the system, or user interaction issues, or both may have had an impact on the problems in the 2006 FL13 election. A significant number of factors raised by activists and areas of study recommended by other researchers have gone unexamined.

Further research is needed before conclusions can be drawn about whether the voting systems were at fault. VVF urges the Task Force to take into consideration the recommendations of other experts before drawing any conclusions, and to pursue those recommendations to examine the problem in a more exhaustive way.

SCOPE and FINDINGS INSUFFICIENT TO RULE OUT EQUIPMENT MALFUNCTION

In accordance with instructions given by the Task Force, the GAO conducted three tests: firmware verification, a limited ballot test, and a miscalibration test. They concluded the firmware matched that examined previously and that which was escrowed with the State, that the ballot test "worked" and that deliberate miscalibration did not mean test votes --though more difficult to cast-- could not be recorded accurately.

The testers also drew conclusions that other factors must have been at play in the anomalous undervote identified in Sarasota County's FL13 contest, including intentional undervoting and voters' inability to properly cast their ballots, possibly because of interaction problems (also undefined).

The investigation was limited in scope by agreement with the Congressional Task Force. The testing, however, was insufficiently ambitious to determine what *caused* an undervote rate many times higher than in any previous election in that contest, and many times higher than that experienced on other voting systems used in that same election.

Some types of tests could potentially have been included even within the limited parameters provided to the GAO testers, but apparently were not.

The draft report of the GAO study leaves most of the major questions unanswered. Many issues uncovered by activists have not been addressed (including a smoothing filter problem, comparatively

high iVotronic undervotes in other state races, low battery problems, and more),¹ and recommendations for further investigation from a study released many months ago have been ignored.²

The GAO testers **did not examine human factors and usability issues**, even though they point out in their conclusions that it is entirely possible "voters ... did not properly cast their ballot on the iVotronic DRE, potentially because of issues relating to interaction between voters and the ballot." Despite their conclusion, the tests conducted could not fairly include an examination of what it means for a voter to "not properly cast their ballot" nor an examination of what "issues relating to interaction between voters and the ballot" may have occurred.

To adequately support the hypothesis that the FL13 problem was due to ballot design would be to do user-testing (usability testing) of the various ballot designs. Such testing could produce important information about the capability of the system to capture voter intent adequately, and would have to be conducted before making any claims about those issues being potential culprits for the anomalous undervote.

As one iVotronic voting system expert, pointed out:

"The DRE issue that Sarasota brings home is that we have no way to audit the current generation of DREs to see if they are difficult to use. With mark-sense ballots, election workers see the frequency with which voters have trouble following the instructions. With current DREs, on the other hand, a recount cannot disclose anything other than the composite undervote figure. If you want to know more, you have to run a human factors experiment with paid subjects who may or may not behave like real voters."³

The GAO test checked for variance of software/firmware in some volume, and uncovered none, but the ballot testing portion of the review deployed only 224 ballots and 10 machines. In contrast, a volume test for functionality conducted by California's former Secretary of State Bruce McPherson examined 96 machines in a mock election to more closely simulate actual voter performance, and uncovered close to a 20% failure rate of the systems (failures including screen freezes, system crashes and printer jams).⁴

It is conceivable and even likely that, were there **bugs in the user interface** in the Sarasota machines, such as intermittent smoothing filter problems or other anomalies, these
(a) could have occurred regardless of whether the software/firmware were identical, and
(b) **would not have been uncovered without greater volume testing** with actual use of the interface, especially in conditions resembling election day use.

¹ Garber, Kitty and Pynchon, Susan, January 2008, "Sarasota's Vanished Votes: An Investigation Into the Cause of Uncounted Votes in the 2006 Congressional District 13 Race in Sarasota County, Florida"
<http://www.floridafairelections.org/CenterReports.htm#LostVotes1>

² Dill, David L. and Wallach, Dan S., "Stones Unturned: Gaps in the Investigation of Sarasota's Disputed Congressional Election", April 13, 2007 <http://www.cs.rice.edu/~dwallach/pub/sarasota07.pdf>

³ Jones, Doug; Email correspondence, February 7, 2008

⁴ <http://www.verifiedvotingfoundation.org/article.php?id=6257>

Further, a **combination of ballot design and a bug** in the machine could trigger a problem such as has been uncovered in other systems (e.g. the Diebold finger-dragging bug), but it is neither apparent from the report whether such combinations were looked for, nor likely that they could have been uncovered given the scope of the testing.

AREAS FOR FURTHER STUDY

Voter/Machine Interaction: Such testing should be *"directed by specific complaints about behavior of the voting machines, and observation and analysis of the tests should note any occurrences which might be related to vote selection or capture problems. Factors should duplicate typical machine usage in election conditions as closely as possible."* *"Tests should include touching the buttons on the screen for various lengths of time to understand the delays imposed by the smoothing filter, noting whether those delays are consistent with any variations in voter demographics, voter behavior, ballot design, voting machine calibration, or other related factors."*⁵

Internal Bug and Software Data Review: Testing software/firmware to ensure that what is on the machines corresponds to what was certified by the State will not necessarily identify any software bugs, nor whether such bugs, if any, contribute to the FL13 problem. From the Dill/Wallach April 2007 report: *"We likewise recommend that experts be given unrestricted access to ES&S's internal bugs databases and software repository, where they may find additional evidence that could lead to the discovery of what software bugs, if any, contributed to Sarasota's undervote rate."*

Separate Firmware Review for PEBs: Personalized electronic ballots, or "PEB" cards used to activate each voter's ballot on the DRE, contain firmware separate from the firmware for the DRE itself. The GAO report does not indicate whether such firmware was examined at all, nor compared with versions certified by the State, or more importantly whether it was what was used by Sarasota voters for the contest in question. Firmware on the PEB "card" which is inserted into the voting machine to activate it for each new voter could have an impact on performance of the system in various ways. Failure to examine this firmware which was in use in Sarasota County during the election for the questioned contest affects the validity of any conclusions drawn about the iVotronic system.

Possible Additional Ballot Style: Earlier reports by other researchers have recommended that the contents of the PEBs be examined.

Variances due to manufacturing quality: As investigative reporter Dan Rather and his team disclosed in recent reports, there is evidence that manufacturing quality may be quite poor. To the extent that there is variance between one device and another, this GAO test would not have uncovered it.

⁵ Dill, David L. and Wallach, Dan S., "Stones Unturned: Gaps in the Investigation of Sarasota's Disputed Congressional Election", April 13, 2007 <http://www.cs.rice.edu/~dwallach/pub/sarasota07.pdf>

DELIBERATE UNDERVOTING HYPOTHESIZED, ENTIRELY UNSUBSTANTIATED

The GAO summary continues to propose "deliberate undervoting" as one plausible explanation for the extraordinary undervote rate associated with the equipment in use in Sarasota County. No substantiation is offered, and no assessment of why the undervote rate on the iVotronic DRE – exclusively in this contest, in this county -- would be so dramatically higher than the undervote rate in this contest, either (a) on absentee ballots cast on paper within this county, or (b) undervote rates on systems in use in neighboring counties which form part of District 13.

Further, additional studies of undervote rates statewide have documented higher undervote rates on the Attorney General contest, in jurisdictions using this system as compared to jurisdictions using other systems.

The suggestion, without substantiation or review, that deliberate intent could account for the undervote anomaly that led to the formation of the Task Force and this inquiry is unwarranted and should be flatly disregarded.

CONCLUSION

The GAO report leaves us with almost as many questions as before, and the most important one – what caused the FL13 undervote problem – is left unanswered. The nature of the complex voting system in question, and the difficulty in auditing such a system, may mean such questions will remain indefinitely, but it is clear that more can and should be done to resolve the outstanding issues.

A variety of technological factors could have triggered an undervote rate five to ten times higher than that for previous contests for the same Congressional seat, but have yet to be examined by those with sufficient access to the technology to conduct such examinations.

Key concerns have been raised by other researchers but despite the resources brought to bear for this study, those questions remain unresolved. It is unclear from this report whether such prior reports were reviewed, either in developing the design for the testing conducted or subsequently.

Any argument, based on the limitations of the investigation, that bad ballot design was the problem instead of the software is a lot like putting bald tires on a Pinto. If a Pinto crashes because of bald tires, it does not indicate the rest of the Pintos on the road are safe.

The argument that voters decided not to choose in that contest – without first eliminating the possibilities that the technology made that decision for them – amounts to a disservice to American voters, and particularly those affected by still unresolved problems in this particular system.