



# Reply to Comment on: “The Methodology for Testing Voting Systems” by Whitney Quesenbery, John Cugini, Dana Chisnell, Bill Killam, and Ginny Redish

Ted Selker

Elizabeth Rosenzweig

Anna Pandolfo

We appreciate the authors (Quesenbery, Cugini, Chisnell, Killam, and Redish, in this issue) acknowledging the lack of research in the field of usability of voting systems. We hope that our early experiments guide people to push the work further, and to create experiments that are more efficient and are rich in useful data.

[http://vote.caltech.edu/media/documents/wps/vtp\\_wp24.pdf](http://vote.caltech.edu/media/documents/wps/vtp_wp24.pdf)

We all agree that in the area of voting technology, so much is at stake; voting interfaces are involved in deciding the fate and future direction of a society. Understanding the impact of voting interface in context then is important. Our studies were critical in demonstrating how voting simulation experiments in quasi-naturalistic settings added more complexity and difficulties, with a resulting downfall in the number of usable, valid data collected from the experimental subjects.

The complexity of training poll workers to work with standard and prototype voting machines increases setup and operational complexity. The number of extra steps added to the protocol generated certain confusion among the poll workers that assisted us during the

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Copyright 2006, UPA.

study. The similarity of the voting experience with the use of real candidates on the experimental ballots frustrated voters.

We have been studying the polling place problems for years as can be seen in our reference section. Polling place procedures must be followed in order for elections to run properly and the work in the field to date has only touched the surface of these problems. We hope that this work can point to a direction that will bear important research solutions for the futures.

We believe in the heuristic value of naturalistic studies and laboratory experiment, as well. Some important results from our past laboratory voting experiments (Cohen, 2005) had been validated and replicated in our quasi-naturalistic experiments.

The comparison shows that results of "lab style" tests compare in quality favorably to voting place simulation-style experiments without having so many problems. While the data is valuable from "lab style" tests, it does not uncover deeper process issues that occur with both the machines and the polling process

The message that testing in context can be complex and reduce the quality of data might be of interest to anyone designing usability experiments. We published this article precisely because of the difficulties we had with our extensive and carefully laid out voting experiments. Our previous publications had presented in detail problems in polling place operations from actual polling place observation. The low yield that we found in our *in situ* experiments was caused by difficulties with prototype equipment, poll worker blunders, video tape malfunctions, and protocol breeches. These problems occurred because renting polling places, transporting lots of things and training

actual poll workers add complexity that inevitably cost our experiments consistency and control. Managing the complexity in experiments that did not conflate learning about issues, polling place operations, and new technology yielded better data.

Our experiments uncovered some (in retrospect) obvious inherent risks in using real ballots with recognizable candidates; there is some value as well. We found participants who wanted to vote differently than instructed. Participants were in many cases unwilling to explicitly tell the experimenters that they were not following the experiment protocol and decided to vote according to their individual preferences.

The yield of usable data in the studies points to the difficulties that arise in polling places. In fact, the yield of data for all studies more than meets the minimum requirements for any usability test and we are satisfied with the data collected. Data problems occurred primarily in the first part of the day, which matches a number of patterns of trouble that actually occur in polling places. Since members of our research team have spent extensive time observing polling places in many municipalities, we are aware of the many problems that occur there. (See [http://www.votingtechnologyproject.org/media/documents/vtp\\_wp17.pdf](http://www.votingtechnologyproject.org/media/documents/vtp_wp17.pdf) and [http://electionupdates.caltech.edu/2006\\_11\\_05\\_archive.html](http://electionupdates.caltech.edu/2006_11_05_archive.html)) Simple experiments might yield the largest data set, but perhaps not the richest.

Our work at MIT is funded by the Carnegie and Knight foundations. The research is about creating and testing new ideas for ballots and voting machines. Since this is research and we are developing new technology, we use prototype machines. As part of the research, the testing is done on these prototypes.

Usability professionals commonly perform tests with low fidelity and high fidelity prototypes, often working around problems that occur as a result of using prototypes.

The article does cite as many important relevant experiments as we could fit in the space available. While we advised others on our ways of testing with real voting simulations, we have not seen others attempt to re-create the polling place experience in total as we did in these experiments.

A real election is a complicated process wrought with places for errors. US polling places have problems and by simulating them it creates a test process that is complex with many steps. It is useful and difficult to create a test that focuses on both the voting process and the machine itself, since it is during the use of the machine (during the voting process) that the errors and problems occur. Placing real and prototype machines in a real environment has the potential of yielding a rich data set.

More can be read about the NY study by at the voting technology project website:

<http://www.votingtechnologyproject.org/reports/chi-abstract-golerselker.pdf>

Thank you for providing us a venue to dialog on these important issues.