



# User Research of a Voting Machine: Preliminary Findings and Experiences

## Menno de Jong

University of Twente  
Faculty of Behavioral Sciences  
P.O. Box 217  
7500 AE Enschede  
The Netherlands  
m.d.t.dejong@utwente.nl

## Joris van Hoof

University of Twente  
Faculty of Behavioral Sciences  
P.O. Box 217  
7500 AE Enschede  
The Netherlands  
j.j.vanhoof@utwente.nl

## Jordy Gosselt

University of Twente  
Faculty of Behavioral Sciences  
P.O. Box 217  
7500 AE Enschede  
The Netherlands  
j.f.gosselt@utwente.nl

## Abstract

This paper describes a usability study of the Nedap voting machine in the Netherlands. On the day of the national elections, 566 voters participated in our study immediately after having cast their real vote. The research focused on the correspondence between voter intents and voting results, distinguishing between *usability* (correspondence between voter intents and voter input) and *machine reliability* (correspondence between voter input and machine output). For the sake of comparison, participants also cast their votes using a paper ballot.

The machine reliability appeared to be 100%, indicating that, within our sample, all votes that had been cast were correctly represented in the output of the voting machine. Regarding usability, 1.4% of the participants had cast the wrong vote using the voting machine. This percentage was similar to that of the paper ballot.

Practical implications as well as experiences with this type of usability testing are discussed.

## Keywords

Usability testing, voting systems, voting ballots, electronic voting

## Introduction

Although electronic voting systems have many obvious advantages compared to the traditional paper ballots, there is a vivid debate going on about the use of voting machines in elections. Most of these discussions focus on the trustworthiness of voting systems (Arbaugh, 2004; Di Franco, Petro, Shear & Vladimirov, 2004; Lauer, 2004; Simons, 2005). They either address the effects of potential software bugs (safety) or the vulnerability of voting machines to fraud attempts (security) (Pieters, 2006).

However, these are not the only important issues. For voters, casting a vote using a voting machine will be a far less transparent process than voting with a paper ballot. In a paper-and-pencil election, voters can easily verify their vote on the ballot, see how their ballot is cast in a box, and imagine how election officials will manually count all votes. In an election using voting machines, they must trust that basically the same things happen. 'Vendors and election officials are free to claim that elections have gone "smoothly," when there is, in fact, no evidence the votes counted had anything to do with the intent of the voters' (Dill, Schneier & Simons, 2003).

A survey by Herrnson et al. (2005), including a usability test, a field study, and an exit poll, underlines these concerns. Voters were predominantly satisfied about the various aspects of voting machines (81-93% positive), but there were also participants with negative experiences. For instance, 6% of the participants in the exit poll reacted negatively to the statement, "I am confident that my vote was actually recorded," and another 8% expressed at least some doubts in this respect. These voters leave the polling station without

knowing for certain whether or not they have actually participated in the democratic process of voting.

Usability is an important aspect of voting equipment, as well (Bederson et al., 2003; Norden, Creelan, Kimball & Quesenbery, 2006). Voting materials, whether they are paper ballots or voting machines, must enable voters to effectively and efficiently cast their vote. Earlier experiences, most notably the Florida 2000 elections, show that we cannot just rely on existing voting equipment. Roth (1998) evaluated the usability of several voting systems, and uncovered various important usability problems. Even though the voting equipment has evolved since then, a recent study by Conrad et al. (2006) showed that people still may make mistakes when they use a voting machine to cast their vote. They also came across substantial differences in usability between four voting machines. Interestingly, they found a significant (albeit weak) correlation between the usability of the voting system and voters' confidence that their vote was accurately counted.

In this paper, we will report on the design and results of a study into the quality of the Nedap voting machine, as used in the Dutch national elections (see Figure 1). We will not address security and safety issues, but instead focus on the question whether the voting machine helps users to effectively cast the vote of their choice. Two different aspects of this overall question are distinguished: the *usability* of the voting machine (the question whether the voting machine's input accurately represents the voting intentions) and *machine reliability* (the question whether the voting machine's output accurately represents the voters' input). To interpret the usability results, we asked

participants to cast their votes twice: both using the voting machine and using a paper ballot.



**Figure 1.** The Nedap voting machine

### **Dutch elections and the voting machine**

Before describing the design and results of our study, we will first provide some background information about the Dutch elections and the voting machine used. In the national elections in the Netherlands, voters are normally assumed to only cast one vote: they have to select one candidate from the candidate list. This candidate is a member of one of the participating political parties.

The Nedap voting machine presents voters with a list of candidates for all participating parties. The political parties can be found in the upper row; the names of the candidates in the columns below the names of each political party. The party's top candidate ('lijsttrekker') can be found in the number one row.

Voters have to select one candidate name and press that particular button. After that, their vote intention is displayed on a small LCD screen. If the vote that is displayed is correct, voters can cast their actual vote by pressing a red button. If the vote displayed is not the right one, voters can push a button to correct their vote and start anew.

### **Method and Process**

Selker, Rosenzweig & Pandolfo (2006) stress the importance of ecological validity for usability research of voting equipment. In our study, we tried to stay as close to a normal election as possible. The research took place on the day of the national elections for the Dutch parliament on November 22, 2006. The town hall of Enschede served as the research location. In total, more than 1800 voters visited the town hall that day to cast their votes. Of this group, 566 people participated in our study (31%). All participants were subject to the following protocol.

#### *Step 1: Casting the real vote*

The participants in our study were people who actually came to vote on November 22. Posters announcing the research were placed at the entrance of the town hall. The posters explained that voters could be asked to participate in a study. To avoid interference with real voting behavior, recruitment started when people left

the polling station. In the real elections, all voters had cast their vote using the Nedap voting machine.

*Step 2: Recruitment of participants*

Voters who left the polling station were randomly approached by a research assistant (recognizable as such by a university badge) and asked to participate in a study into the user-friendliness of voting equipment. They were told that the research would take a few minutes and that they would have a chance to win one of six € 50 gift certificates. Especially during peak hours at the polling station not all voters could be approached.

*Step 3: Research outline*

Voters who agreed to participate were guided to a separate room in the town hall where a polling station was simulated. They were given a brief outline of the research procedure. It was explained to them that they had to cast a vote once more, and that they had to do this using the voting machine and a paper ballot. The designs of the candidate lists on the voting machine and on the paper ballot were exactly the same as those used in the official elections. The candidate list used in our study, too, consisted of real candidate names and real political parties. We emphasized, however, that the votes in the study were 'fake' and that the research had nothing to do with the real elections.

*Step 4: Voting task*

Participants received a form with a participant number and three demographic questions (gender, age, and educational level) which they would have to fill out afterwards. A researcher instructed the participants that they had to cast their votes vote on a pre-specified candidate. The voting assignment consistently included:

(a) the name of the candidate, (b) the political party to which the candidate belonged, and (c) the candidate's number on the list of the political party. Translated from Dutch, the instruction was:

'Your goal is to vote for [first name plus family name candidate]. This person is number [#] on the list of [name of the political party].'

As participants were supposed to memorize their voting task, we repeated this instruction, and asked if their voting task was clear to them. This was also one of the reasons why we added the candidate's number on the list of their political party: the redundant information would probably help them remember the voting task. Furthermore, for some of the candidates, the number on the list is something which voters may know.

Because we wanted the study to cover all possible voting options, all participants were given different voting tasks. Participant 2, for instance, had to vote for the second person on the first list, participant 3 for the third person of the first list, and so forth. In this way, all participants were given a unique voting task, and all voting possibilities were included in the research. After the instructions, participants were guided to the place with the voting machine and the ballots.

*Step 5: Voting*

The participants were first either referred to the voting machine or to the voting ballot. The order of votes (machine vs. ballot) was systematically varied, based on participant number. The first 25 participants used the voting machine first and then the ballot, the next 25 participants started with the ballot and then used the voting machine, and so on.

To check whether the participants voted correctly on paper, all ballots were marked with participant numbers. To check whether the participants voted correctly using the voting machine, and to compare the input and output of the voting machine, a small video camera was installed, which recorded all participant actions with the voting machine. The camera was unobtrusively placed 3 feet above the display.

#### *Step 6: End of session*

When the participants had cast their votes using the voting machine and on paper, they were asked to fill out and hand in their short form with demographic questions. If the participants wanted to have chance to win one of the gift certificates they were asked to write their contact information on a separate list.

## Results

### *Sample*

Not all participants answered all demographic questions. Relatively many participants overlooked the question about gender (107), fewer participants did not answer the questions about age (23) and educational level (41). Of the participants who answered the question about gender, 56% were male and 44% female. The participants' age varied from 18 to 89 years old (mean age: 40). Compared to the Dutch general population, highly educated people were over-represented in the sample: 65% of the participants had a university or higher vocational education level, 7% had an intermediate level, and 28% had a lower level of education.

### *Usability of voting machine vs. ballot*

In 44 of all 1132 cases (4%) a deviation between voting task and actual vote occurred. However, the types of errors and the causes differ and are not all practically meaningful. Table 1 presents all deviations, along with a brief explanation.

**Table 1.** Voting problems and possible causes

	Machine	Ballot
Personal/situational mistakes	12	11
Practical research problems	4	3
Usability problems	9	6

The first two categories of problems can be seen as comments to the research design rather than as problems with the voting equipment.

Most of the deviations we found between the voting assignment and the actual votes must be attributed to *personal/situational mistakes*. Participants had trouble remembering their voting task and as a result voted for the wrong candidate. Two indicators for this type of errors were the consistency between the deviation on the voting ballot and the voting machine (participants made the same mistake twice) and the distance between the voting assignment and the actual vote (the mistake could not be caused by incorrect reading of the rows or columns).

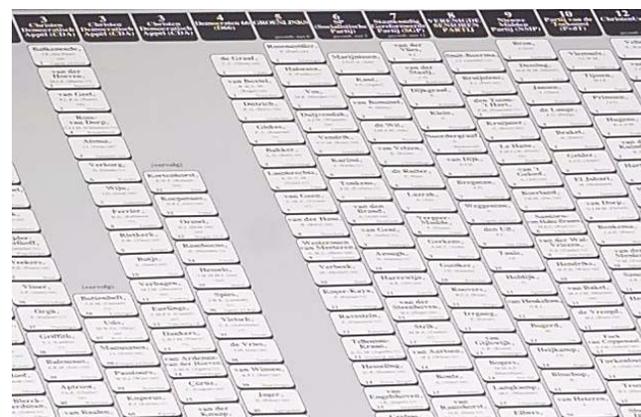
In five of the cases, participants appeared to have voted twice for the candidate and party of their own preference (like they had done somewhat earlier in the real elections): their two votes had nothing to do with the voting assignment. In three cases, participants voted for the first candidate of a party's list, and forgot to vote for a candidate lower on the list. In two cases,

participants seemed to vote for random candidates who had nothing to do with the voting assignment. One participant appeared to mix up the list number and the candidate number: instead of voting for candidate number 8 of list 1, this participant voted for candidate number 1 of list 8. In one case, a verbalization of the participant (on the camera recording) showed that the participant had forgotten the candidate number: on paper, he correctly voted for candidate number 73 of list 2; before using the voting machine, he said that he had to vote for candidate number 74 of the same list, and acted accordingly.

*Practical research problems* were the cause of another six deviations we found between voting assignments and actual votes. All these deviations occurred due to the sometimes hectic research context or to the artificial research situation of having to vote twice. The video recordings showed that one of the participants tried to cast a vote with the voting machine, which was not stored. This could happen because the voting machine was not released by one of the research assistants. Another deviation was caused by a difference in length between the candidate lists on the ballot and on the voting machine. When using the ballot, one participant noticed that his candidate was located in the first column, on the third row from below. He took the same strategy when using the voting machine, not noticing that the candidate was someone else. Another participant refused to vote both using the ballot and using the voting machine. Finally, one participant voted incorrectly twice, because he was misled by a wrong voting instruction by one of the research assistants.

The last (and most important) category of deviations consists of usability problems. Regarding the paper ballots, six deviations were likely to correspond with usability problems (1.1%). Regarding the voting machine, nine such deviations occurred (1.6%). The difference between these percentages was not statistically significant. Even though the number of usability problems appeared to be rather low in our sample, it is important to pay serious attention to them, since it is at the heart of the democratic process that each and every voter is able to cast the vote of his or her choice.

One of the participants, who had a severe physical handicap (spasticity), was not able to cast his vote with both types of voting equipment. With the paper ballot, he could not vote at all. With the voting machine, he only managed to cast a random vote. In the real elections, this voter had been helped by someone to cast his vote. The voting machine used was not specifically designed to be accessible.



**Figure 2.** Layout of the voting machine's candidate list

The rest of the usability problems, both with the ballot and with the voting machine, involved participants' misreadings of rows and columns in the lists of parties and candidates. Apparently, the layout did not sufficiently support the voting task (see Figures 2 and 3). The design did not prevent all voters from selecting the wrong row or column. This is something to pay serious attention to, since two adjacent columns may imply votes for very different political ideals.



**Figure 3.** Layout of the paper ballot's candidate list

One of the participants made the same reading error using the paper ballot and the voting machine. The other participants only made their reading error once: four of them with the paper ballot, and seven with the voting machine. The strongest signal for a usability problem is if participants make the mistake in their first voting attempt and after that vote correctly in their second attempt. In those cases, the influence of memory problems between the first and the second vote can be ruled out. Three of the four participants who made a mistake using the paper ballot did this in

their first attempt, and after that voted correctly using the voting machine. Only two of the seven participants who made a mistake using the voting machine did this in their first attempt.

#### *Machine reliability*

Using the video camera recordings, it was possible to compare the actual voting input with the output of the voting machine. Since the voting machine does not produce output that allows us to see which voter had cast which vote, we could only compare the totals of the voting input and the output of the voting machine with each other. There appeared to be a 100% correspondence between the votes that were cast by the participants and the election result produced by the voting machine. Based on these findings, the voting machine used can be characterized as entirely reliable within our sample. More research would be needed to generalize these findings to the entire population of Nedap voting machines. It must also be noted that we only tested the voting machine under normal circumstances; we did not include any safety or security threats in the research design.

#### **Discussion**

The user research reported in this article casts little doubt about the validity of the Dutch elections. Both with the ballot and with the voting machine, the vast majority of the participants appeared to be able to successfully complete their voting assignment. The results about the machine reliability are also positive, although it is important to stress that only one voting machine was included in the research.

In the literature on usability testing, there is a strong emphasis on the detection and diagnosis of possible

user problems. The usability study reported in this article put forward two concerns: (a) the ballot and the voting machine (at least this version of the voting machine) may not be suitable for voters with (severe) physical handicaps, and (b) the rows and columns structure used on the ballot and on the interface of the voting machine might benefit from a redesign.

Even though the numbers of problems detected in our usability test are far from spectacular, it remains highly useful to conduct this type of user research. First and foremost, it is a way of monitoring the validity of elections. Instead of relying on the voting equipment available, we empirically and rigorously checked its user-friendliness. Another advantage is that the results, if favorable, may be used to further the general public's trust in the voting equipment used. The results clearly show that people can effectively cast their vote using the ballot and the voting machine, and that the 'black box' of a voting machine indeed does what it is supposed to do: record all votes that were cast, and produce the correct election results.

Of course, it must be acknowledged that we are talking about a specific type of usability test in this article. Instead of focusing on detailed experiences of a very limited sample of potential users—for instance using think-aloud protocols—we drew a much larger sample, and only analyzed their observable behaviors. In terms of evaluation functions, we prioritized the verification function (getting an overall indication of the usability of the voting machine), and paid less attention to the troubleshooting function (getting an exhaustive list of potential user problems in need of revision) (cf. De Jong and Schellens, 1997). It is conceivable that a

think-aloud usability test would result in more fine-grained suggestions for improvement.

Despite the large sample size, our study was highly feasible. The polling station appeared to be an ideal location to recruit participants for a brief research session. People appeared to be very willing to participate. This may be due to the brevity of the research session, but also to the societal relevance of the research topic. Voters easily understand that it is important to evaluate and optimize voting equipment.

Our experiences raise some issues about the design of this type of usability test. The results show that most of the deviations we found were caused by the research design, and not by the voting equipment. The most problematic issue was that some of the participants seemed to forget the voting assignment that was given to them. For two reasons, however, we think it is important to work with voting assignments. First, it is the best guarantee that all voting options are included in a usability test. It seems plausible that voting on candidate number 1 of the first party will cause fewer usability challenges than voting on candidate number 23 of the twentieth party. Second, it is the only way of having an unambiguous check whether or not participants voted correctly. To solve the participants' memory problems, it may be an option to give them a reminder of the voting assignment on paper.

A second remark pertains to the comparison of the voting machine and the paper ballot. The fact that all participants had just cast their vote in the real elections with the same Nedap voting machine may have worked to the advantage of the usability results of the voting machine. However, since the earlier voting session did

not include any feedback on the correctness of their votes, it seems less plausible that participants could have learned from mistakes they may have made in the real elections. More in general, however, the collection of usability data in the context of real elections raises questions about the relationship between the real task(s) and equipment and the task(s) and equipment used in the usability study.

A third problematic aspect was the sometimes hectic context of the research. From time to time, participants had to wait in line for their turn, and started talking with each other. The pressure led to a few errors by the research assistants. On the other hand, however, the non-laboratory setting closely resembled the real voting conditions in the Netherlands. The research was probably as ecologically valid as a usability test of voting equipment can get.

### Practitioner's Take away

- It is important to conduct user research of voting equipment. The research may focus on a fine-grained analysis of potential user problems or on an overall assessment of the usability of the voting equipment. It will be hard to serve both goals in one and the same study.
- User research can effectively be conducted in the vicinity of polling stations. This makes it easy to design an ecologically valid research context. People also appear to be willing to participate in such a study, provided that it does not take very long.
- It is important to clearly distinguish the test set-up from the real elections that take place in the same

building. This can be done by clearly announcing the research with posters, by visually stressing the researchers' affiliation, and by emphasizing the distinction in the instructions to the participants.

- The best way to be able to include all voting options in the usability study and to be able to verify the correctness of participants' votes is by using voting assignments. To avoid memory problems, it may be useful to support a voting assignment with a written note.

### Acknowledgments

The authors want to thank Nedap for the opportunity to use the voting machine in this study and for covering some of the research expenses. At the request of Nedap, the entire study took place in the presence of a notary. The authors also want to thank Jennifer Kleiss for her work as a research assistant.

### References

- Arbaugh, W.A. (2004) The real risk of digital voting? *Computer*, 37 (12), pp. 124-125.
- Conrad, F.G., et al. (2006) The usability of electronic voting systems: Results from a laboratory study. Paper presented at the Midwest Political Science Association, April 22, 2006, Chicago, IL.
- De Jong, M., Schellens, P.J. (1997) Reader-focused text evaluation: An overview of goals and methods. *Journal of Business and Technical Communication*, 11 (4), pp. 402-432.
- Di Franco, A., Petro, A. Shear, E., Vladimirov, V. (2004) Small vote manipulations can swing elections. *Communications of the ACM*, 47 (10), pp. 43-45.
- Dill, D.L., Schneier, B., Simons, B. (2003) Voting and technology: Who gets to count your vote? *Communications of the ACM*, 46 (8), pp. 29-31.

- Herrnson, P.S., et al. (2005) Early appraisals of electronic voting. *Social Science Computer Review*, 23 (3), pp. 274-292.
- Lauer, T.W. (2004). The risk of e-voting. *Electronic Journal of e-Government*, 2 (3), pp. 177-186.
- Norden, L., Creelan, J.M., Kimball, D., Quesenberg, W. (2006) *The machinery of democracy: Usability of voting systems*. New York: Brennan Center for Justice at NYU School of Law.
- Pieters, W. (2006) Acceptance of voting technology: Between confidence and trust. In: K. Stølen et al. (Eds.), *iTrust. Lecture Notes on Computer Science*, 3986, pp. 283-297.
- Roth, S.K. (1998). Disenfranchised by design: Voting systems and the election process. *Information Design Journal*, 9 (1), pp. 1-8.
- Selker, T., Rosenzweig, E., Pandolfo, A. (2006) A methodology for testing voting systems. *Journal of Usability Studies*, 2 (1), pp. 7-21. Simons, B. (2005) Computerized voting machines: A view from the trenches. In: S. de Capitani di Vimercati et al. (Eds.), *Esorics. Lecture Notes on Computer Science*, 3679, pp. 1-2.

## Author Bios



**Menno de Jong** is an Associate Professor of Communication Studies at the University of Twente (Enschede, the Netherlands). His research focuses on the methodology of applied communication research. He has published about usability testing, website evaluation and document evaluation.



**Joris van Hoof** is a lecturer/researcher in Communication Studies at the University of Twente (Enschede, the Netherlands). He has published about applied communication research, and is also working on a research project on adolescents and alcohol.



**Jordy Gosselt** is a lecturer/researcher in Communication Studies at the University of Twente (Enschede, the Netherlands). His research focuses on applied communication research, as well as on the issue of adolescents and alcohol.