

Recommendations for the protocol for Logic and Accuracy testing for instant runoff voting

Caleb Kleppner, Center for Voting and Democracy

Draft February 12, 2003: This is a very draft document that is intended for discussion purposes only by Dept. staff.

NOTE: This document contains information that the Center believes would be inappropriate to share with the voting equipment vendor. That information occurs in pages 3-5 and is marked “Do Not Disclose This Section To The Vendor” in the header and footer. The rest of the document would be quite reasonable to share with the vendor.

1. Introduction

L&A testing for the November 2002 election called for creating 2 test decks for each ballot type: an Election Test Deck and an Exception Test Deck. These test decks test every contest on the ballot, and every candidate receives an assigned number of votes, from 1 to 6. The basic routine is to run the Election Test Deck through every Eagle and to one the Exception Test Deck through one randomly chosen precinct per ballot type.

Since the totals from each test deck is known, and each candidate received the same number of votes in each precinct, the number of votes that each candidate received after totaling all test decks can be compared with the pre-determined totals to see if the equipment passed the L&A test.

In contrast to the current system, with instant runoff voting, there are two steps of an election: the storing by the equipment of ballot images consisting of each voter’s set of rankings and the instant runoff tabulation, which is performed on the aggregate set of rankings. The L&A test must test both portions: the storing of ranking and the instant runoff tabulation.

This paper lays out a methodology for testing both portions. It follows the current protocol as much as possible.

Note that in the November 2003, there will be 11 ballot types, one per Supervisor district due to ballot rotation, and there will probably be 2 cards. The 1st card will contain the mayor’s race on the front and the DA and sheriff’s races on the back. The other card will contain the local ballot measures. All of this L&A testing would only apply to the 1 card that contains the IRV races. The L&A testing for the ballot measure card would be the same as usual.

2. Testing the storing of rankings

The approach is the same:

- 1) Create an Election Test Deck and an Exception Test Deck for each ballot type that contains the same number of votes cast for each candidate (to ensure that candidates, and not ballot positions due to ballot rotation, receive the same votes).
- 2) Run the ballots through the Eagles with an Election Test Deck for each precinct and an Exception Test Deck for each ballot type.

- 3) Compare the output of the Eagles (data files containing each voter's ranking) with the pre-determined data file that corresponds to each test deck.

Since the output of the Eagle is no longer a vote total for each ballot position, we do not have to rely on patterns (1 vote for Candidate A, 2 votes for B, 3 for C, etc), since we can compare the actual output file with the exact data file that we manually create that corresponds to that test deck. Instead of comparing vote totals, we can literally overlay the output files from the test decks an manually created files that correspond to them to makes sure they are identical to the vote. This gives an even greater level of confidence than comparing vote totals, since we can compare every single vote.

One note: since the Eagles ideally will store ballot images (rankings) in a random order, to preserve voter anonymity, the votes in the Eagle's file and votes in the manual may be in different orders. The easy way to rectify this is to sort both files by alphabetical order before overlaying the files to compare.

For an idea of what the data files will look like, please visit:

<http://www.fairvote.org/sf/results.htm>

3. Testing the instant runoff tabulation

The testing of the instant runoff tabulation can use the exact same data files that resulted from testing the storing of rankings. You simply run those data files through the IRV tabulation software to get round-by-round vote totals for all candidates. Then you take the same data files, drop them into an Excel spreadsheet, and perform a manual tally of the instant runoff tabulation. If the round-by-round vote totals are identical, the tabulation passes the L&A test. If not, it doesn't.

4. Summary

L&A testing with instant runoff voting is quite similar with the current process: test decks are created and scanned by the voting equipment. Instead of comparing vote totals from the voting equipment, we compare the actual data files that contain all the votes (rankings) in the test deck. Then we conduct the instant runoff tabulation on the data files created from the test decks and compare the output to results manually generated in a spreadsheet.

Finally, the totals tape printed by each Eagle provides the L&A test of the ballot measures. It presumably also will list 1st, 2nd and 3rd choice totals for each candidate. These totals can be compared to the pre-determined totals from each deck (and using clever patterns, these totals could be 4-1st choice votes for each candidate, 3-2nd choices and 2 3rd choices). This provides another check of the data, including a guarantee that the 1st choices stored in the Data Pack correspond to the complete rankings stored in the PCM card.

[NOTE: The following sections contain information that should not be disclosed to the voting equipment vendor.]

5. Creation of test decks

Note that the number of possible valid rankings is quite large. For example, with only 3 candidates and a write-in option, there are 72 possible ways to cast a valid vote. With 15 candidates, there are over 4,000 valid combinations.

Number of valid vote combinations				
Candidates	1 ranking	2 rankings	3 rankings	Total
3	12	36	24	72
5	18	54	120	228
7	24	72	336	528
9	30	90	720	1,020
11	36	108	1,320	1,752
13	42	126	2,184	2,772
15	48	144	3,360	4,128

These numbers in the above table refer to the number of possible ways of filling out a valid ballot with the given number of candidates listed on the ballot along with one space for a write-in.

There are 7 ways to rank at least one candidate and cast a valid vote:

1. Pick 3 choices: 1st, 2nd and 3rd choices
2. Pick 2 choices: 1st and 2nd choices (but no 3rd)
3. Pick 2 choices: 1st and 3rd choices (but no 2nd)
4. Pick 2 choices: 2nd and 3rd choices (but no 1st)
5. Pick 1st choice only
6. Pick 2nd choice only
7. Pick 3rd choice only

6. Creation of Election Test Decks

The key to the Election Test Decks in the current system is using specified numbers of ballots that reveal patterns in the number of votes cast for each candidate: 1 vote for Candidate A, 2 for B, etc. This makes it very easy to compare the output of the voting equipment with the pre-determined totals.

With instant runoff voting, we don't need to use patterns of votes cast since we can easily compare the exact output for every vote to the pre-determined result. However, if the Dept. wishes to use patterns, it could use something like the following, with each row corresponding to a single ballot's first choice, second choice and third choice. A represents Candidate A, B represents Candidate B, and so forth. Let's imagine a race with candidates A through F plus a write-in, represented by G.

(A, B, C)

***** Do Not Disclose This Section To The Vendor *****

- (B, C, D)
- (C, D, E)
- (D, E, F)
- (E, F, G)
- (F, G, A)
- (G, A, B)

Note that it would not be necessary to create 1 ballot with the 1st ranking, 2 with the 2nd combination, and so on, since we can compare each individual ranking with the output. But if you want to have 1 ballot with (A, B, C), 2 with (B, C, D) and so on, that's fine.

Instead of using a pattern, such as the above pattern, it would also be possible to create valid votes randomly. It would be wise to test every candidate as a 1st, 2nd and 3rd choice in at least 1 combination.

The number of possible combinations suggest to me that the key is not to make test every possible combination but to test at least one combination that includes every possible ballot position using 3 rankings, and a sampling of 1- and 2- rankings.

Let's consider a race with 11 candidates. There are:

- 36 ways of ranking 1 candidate
- 108 ways of ranking 2 candidates
- 1,320 ways of ranking 3 candidates
- 1,752 ways total of casting a valid ballot

With a minimum of 12 ballots, you could cast ballots with 3 rankings that cover all candidates in all positions (in other words, candidate A would get votes as a 1st choice one time, as a 2nd choice another time and as a 3rd choice a third time).

It would take 36 ballots to cast ballots that cover all possible ballot positions with ballots that rank 2 candidates: 1st and 2nd, 1st and 3rd, and 2nd and 3rd, each of which requires 12 ballots.

Finally, to test all ballot positions with a single vote (1st, 2nd or 3rd), it would take another 36 ballots.

So of the 1,752 possible ways of casting a valid vote in a race with 11 candidates plus the write-in option, it's possible to cover all ballot positions and all types of rankings with 84 ballots (12+36+36). On the other hand, if you use 12 ballots to test all ballot positions in 3 rankings, you could cover all the ballot positions that have a 1st choice (1st only, 1st and 2nd and 1st and 3rd) and pick a sample of the 36 ways to hit all ballot positions (2nd only, 3rd only and 2nd and 3rd only). This would slightly reduce the test deck size. The cards in the test deck would look like this:

- 12 ballots with 3 rankings,
- 12 ballots with 1st and 2nd rankings

******* Do Not Disclose This Section To The Vendor *******

- 12 ballots with 1st and 3rd rankings
- 12 ballots with 1st only (1st choice are most important, so you want to hit all the options)
- 12 total with 2nd and 3rd, 2nd only or 3rd only (out of 36)
- 60 Total (instead of 84)

If you just only want to test valid votes with 3 rankings, you can do it with as few as 12 ballots.

The current Election Test Decks appear to have 1 or more undervotes in each race, 1 blank vote, and a complete overvote (every ballot position in every race). The test deck should probably have analogs to these votes as well.

7. Creation of Exception Test Decks

The current Exception Test Decks appear to include overvotes that cover each ballot position as well as an undervote for each race.

With instant runoff voting, there are additional ways to make errors. For example, in a race with 11 candidates plus write-in:

<u>Error vote</u>	<u>Number of cards required to test all ballot positions</u>
Overvotes for 1 st choice	12
Overvote for 2 nd choice	12
Overvote for 3 rd choice	12
Same candidate 1 st and 2 nd	12
Same candidate 1 st and 3 rd	12
Same candidate 2 nd and 3 rd	12

Here are sample decks of 12 cards that demonstrate overvotes for 1st choice and ballots with the same candidate ranked 1st and 3rd. In the following data set, the phrase, “A=B,” refers to ballot in which the voter overvoted for A and B for the given choice.

<u>Overvotes for 1st</u>	<u>Same candidate 1st and 3rd</u>
(A=B, C, D)	(A, B, A)
(B=C, D, E)	(B, C, B)
(C=D, E, F)	(C, D, C)
(D=E, F, G)	(D, E, D)
(E=F, G, H)	(E, F, E)
(F=G, H, I)	(F, G, F)
(G=H, I, J)	(G, H, G)
(H=I, J, K)	(H, I, H)
(I=J, K, L)	(I, J, I)
(J=K, L, A)	(J, K, J)
(K=L, A, B)	(K, L, K)
(L=A, B, C)	(L, A, L)