Wisconsin Election Integrity
Promoting our right to self-government through fair & secure elections
WisconsinElectionIntegrity.org

Date: December 17, 2019

TO: Wisconsin Elections Commission and staff

FROM: Karen McKim, Coordinator, Wisconsin Election Integrity

SUBJECT: Risk-limiting audits in Wisconsin – Purpose and Possibilities

First, thank you, Chair Dean Knudson, for requesting this paper at the Commission's December 2 meeting. I'll do my best to give you a concise vision of how a risk-limiting audit might work in Wisconsin.

The first three pages contain a bare-bones description of a hypothetical, basic, risk-limiting audit (RLA) of a statewide election.

However, I've explained this often enough to know that a brief explanation always raises questions. So the last 10 pages contain more detail about: 1) the function and attributes of RLAs (*page 4*); 2) how RLAs' function differs from the random voting-machine audits intermittently performed in Wisconsin (*pages 5-6*), and 3) more detailed description of how a simple RLA of two statewide contests on the November 2018 ballot might have been done (*pages 6-10*). I'll end with a review of some of the issues that might complicate implementation of RLAs, which include one national debate related to computer-marked ballots.

— The bare-bones explanation—

Purpose

The one core purpose of a risk-limiting audit is to ***limit the risk that election officials will fail to detect and correct, before certification, any outcome-altering miscount*** that may have affected the preliminary results.

However, RLAs' biggest value is probably deterrence. When RLAs are routine, would-be hackers know that even if they make the voting machines identify the wrong winner, routine administrative procedures will promptly correct the results. Therefore, thieves are less likely to bother. With that, of course, comes stronger voter confidence and officials' improved ability to defend elections.

Three basic steps

You can think of an RLA like an exit poll. Instead of interviewing a random sample of voters, officials inspect a random sample of ballots. If done correctly, both exit polls and RLAs produce a statistically solid measure of who won the election. Every RLA contains three core activities:

1. Select a random sample of ballots;
2. Manually count the votes on those ballots;
3. Compare the sample results with the Election-Night results.
    - If the same winner got roughly the same percentage of votes in both, you've confirmed the outcome.
    - If the results don't match closely enough, randomly select and count some more ballots. Do that until you've confirmed either the Election-Night winner or a different winner with a 100% hand count. (Note: an RLA will lead to a 100% hand count only when Election-Night results picked the wrong winner, so you won't be irresponsible if you don't budget for it.)

Method

There is no one 'recipe' for an RLA. The following steps describe only one of several possibilities for an RLA in Wisconsin. I use the Governor and US Senate contests of November 2018 as illustration.

1. ***Municipal officials would collect and provide information for a "ballot manifest."*** A ballot manifest enables officials to select ballots randomly from anywhere in the state. Local officials would collect and quickly report information that will allow state staff to create a ballot manifest that could look something like this. *(For more discussion, see page 7.*)

Ballot Manifest, State of Wisconsin, November 6, 2018

County	Municipality	Reporting Unit	Bag number	Number of ballots	Sequence: First ballot	Sequence: Last ballot
Adams	Town of Adams	Wards 1-3	Bag #1	310	#1	#310
Adams	Town of Adams	Wards 1-3	Bag #2	297	#311	#607
	etc…					
Wood	City of Wisc Rapids	Wards 11-29	Bag #9	300	#2,688,579	#2,688,879

2. ***State officials determine the sample size***. State officials would need to choose an arbitrary "risk limit" (*More information on page 7*). State staff would use a statistical formula (already developed and tested) to calculate the number of ballots needed for the sample. Decisive results can be confirmed with surprisingly small samples, but closer races need larger samples. As the Washington Post explained in an October 2019 editorial, RLAs are "time-consuming only in the tightest elections, or when something actually has been tampered with. Of course, that's when it's most worth investing the time." *(More starting on page 8.)*

   If WEC had conducted RLAs for the Governor and US Senate contests in November 2018 with a risk limit of 10%, initial sample sizes would have been as follows:

Contest	Number of ballots	Winner's percentage	Number of ballots needed in initial sample
Governor	2,688,879	49.54 %	38,062
US Senate		55.36 %	405

   The 38,062-ballot sample for the Governor's race is more than enough to confirm the Senate race. So, both contests would have been audited using only the larger sample.

3. ***State staff would identify the specific ballots to be in the sample***. Using methods that have been developed, tested, and used in other states, state staff would generate a list of 38,062 random numbers and then refer to the ballot manifest. For example, if one of the random numbers was 350, the ballot manifest would locate ballot #350 in Bag #2 in the Town of Adams.

4. ***Municipal clerks would send images of those paper ballots to Madison***. State staff would notify the municipal clerks to retrieve the selected ballots (the paper ballots, not the digital images created by the voting machines). State officials would have devised some method of transmitting the votes on those ballots to Madison, such as by fax. The clerks' selection and transmission of these ballots would need to be formally witnessed, but the physical ballots could remain secure in the custody of the municipal clerks. (*More on page 9*.)

5. ***State staff would manually count the votes in Madison, using the transmitted ballot images***.
6. ***State staff would calculate the risk that the real winner is a different one than identified in the preliminary results***. Using a statistical formula that has already been developed and tested, state staff would use the results of the manual count to calculate the risk that, given results from the sample, a full hand count of all the ballots would reveal a different winner. If that risk is the same as or less than the risk limit chosen in step 2, above, the correct winner is confirmed and the audit complete. If the risk is more than that, a larger sample of ballots is needed. (*Page 10*.) This continues until the risk limit is achieved.

Effort and cost of a statewide RLA

The hypothetical RLA I've described in this paper would not involve county clerks at all and would require less work from the municipal clerks than the s.7.08(6) voting machine audits, for which municipal clerks hand-counted 135,712 ballots in November 2018. That was more than 3.5 times the number of ballots they would have had to retrieve for a statewide RLA of the same election.

In other words, a state-level RLA could have confirmed the winners of the two highest-profile contests, more quickly and with less work for local officials than the s.7.08(6) voting-machine audits, which confirmed only individual voting machines rather than the overall results.

A statewide RLA could be less work than the s.7.08(6) voting-machine audits, cost less, and provide a definitive answer to the most important question: Did the voting machines identify the right winners?

States can choose to develop software to support the tasks described above, but all necessary software is available for free. Colorado incurred costs to replace all its voting equipment, but your staff and I agree that Wisconsin does not want to emulate the more complicated parts of Colorado's version of RLA, which necessitated new equipment.

To my knowledge, only Rhode Island has publicly reported the costs of its statewide RLA, a pilot effort. With a voting population roughly equal to the combined total of Milwaukee and Waukesha Counties, Rhode Island spent $12,705 on its first RLA. Of this amount, $3,900 was paid to a videographer hired to record material for a training video, and $2,279 went for food, drinks, and rental chairs and tables for lunch and dinner for participants. Other expenses were one-time purchases for the first audit, such as a document camera to project the ballots for a transparent, publicly observable manual count.

More detailed discussion follows.

A. What is a risk-limiting audit?

Election officials are in the unique position of having to depend on computers without any error-detection backup from anyone else. If a bank's ATM system develops a glitch, other banks' systems will start flagging errors even if managers at the first bank haven't yet noticed the problem. If the municipal treasurer mis-programs the computer that calculates property tax bills, thousands of property owners stand ready to 'help' the treasurer notice the problem. But if voting machines produce flawed vote totals, election officials are on their own to notice and correct it. They have no backup.

Another unique circumstance makes election officials' position even scarier: The bank manager and city treasurer can reverse errors even when they don't discover them for months. But government stability requires final election results to be truly final, so election officials must detect and correct any wrong results before they certify.

RLAs give a speedy answer to a narrow question:

Who won?

RLAs were developed to solve those two problems. They enable election officials to answer one critical question, and to answer it quickly. It is the only question they MUST answer correctly before certification: ***Did the preliminary election results identify the right winner?*** If the answer is "No," the speed of the RLA will enable them to correct the error before the certification deadline.

Although RLAs can be designed with extra steps to answer additional questions, basic RLAs answer only that one. They won't help to identify mis-handled mail-in ballots or the one voting machine where a mis-calibrated setting caused the machine to misread half the votes. An RLA would not have produced the audit evidence that enabled you to decertify the Optech Eagle. The RLAs' laser-like focus means that responsible election managers will still want to perform other kinds of audits between elections.

For example, when RLAs come to Wisconsin, the s.7.08(6) audits should still be done for the same purpose they are now—to assess the reliability of the various tabulation systems. But when you start doing RLAs, you will be able to allow municipalities to complete the s.7.08(6) audits at a more leisurely pace, because you don't ***need*** detailed information about voting-machine performance before you certify election results. But you do need to know you're certifying the right winners.

How are RLAs done?

There is no one definitive way to perform an RLA. For example, individual ballots can be sampled (least work) or entire wards can be sampled (more satisfying for voter confidence).

An election audit qualifies as an RLA if it: 1) objectively quantifies the risk that, if Election-Night results identified the wrong winner, the error will remain undetected during the canvass, and 2) continues counting votes until that risk is equal to or less than a "risk limit" chosen before the audit began.

Officials can choose whatever risk limit they want to tolerate. Consensus seems to me to be coalescing around a 10 percent risk limit. (*On page 7, I explain why that isn't anywhere near as lax as it seems*.)

Statistical formulas used in RLAs were developed more than ten years ago at the University of California-Berkeley; studied and endorsed by the American Statistical Association; and tested in pilot projects funded by the US Elections Assistance Commission in 2011.

B. Why are RLAs, but not random voting machine audits, considered to be security safeguards?

First, let's be realistic: Election-Night results will occasionally be miscounted. No one can eliminate human error, glitches, and random machine malfunctions. Far too many people have motive to manipulate our elections, and local officials have too little control over everyone (both within and outside Wisconsin) who has or can gain access to the software.

In that respect, election officials are no different than any other computer-dependent manager. Try as they might, they cannot protect the computers so completely and so reliably that they can trust them to be accurate every single business day without fail. That is why the National Institute of Standards and Technology (NIST), does not stop with "***identify** the risks"* and "***protect** the computers"* when it describes the five essential elements of an effective cybersecurity program. Managers must also take steps to ***detect*** any failures; ***respond*** so as to prevent or minimize damage; and ***recover***.

Credit: N. Hanacek, National Institute for Standards and Technology

For elections, we can understand ***detect*** to mean 'notice an incorrect outcome' and ***respond*** to mean 'determine and then certify the correct outcome.' 'Recover' is beyond the scope of this paper; for elections, that means to diagnose the cause of the miscount and fix it before it messes up another election.

Therefore, to function as a security measure, an election audit must:

1) be able to detect any outcome-altering miscounts (that is, miscounts that identify the wrong winner) and if such a miscount is detected,
2) identify the correct winner and replace the flawed preliminary vote totals with the correct totals in time to ensure that only the correct winner is certified.

That is why, when authorities such as the National Academies of Science make election-security recommendations, they mention no other type of election audit.[1]

To function as a security safeguard, election audits must be able to DETECT and CORRECT serious miscounts.

Random voting-machine audits, often called 'spot checks', may provide some useful information, but they ***do not reliably detect and correct flawed results***. Because they do not expand the sample when a miscount is detected, they cannot confirm the correct winner. Because they cannot confirm the correct

[1] "***States should mandate a specific type of audit known as a "risk-limiting" audit prior to the certification of election results.*** *By examining a statistically appropriate random sample of paper ballots, risk-limiting audits can determine with a high level of confidence whether a reported election outcome reflects a correct tabulation of the votes cast. Risk-limiting audits offer a high probability that any incorrect outcome can be detected, and they do so with statistical efficiency; a risk-limiting audit performed on an election with tens of millions of ballots may require examination by hand of as few as several hundred randomly selected paper ballots."*
National Academies of Sciences, Engineering, and Medicine. 2018. *Securing the Vote: Protecting American Democracy*. Washington, DC: The National Academies Press. https://doi.org/10.17226/25120.

winner, they cannot ensure that any outcome-altering miscounts will be corrected before certification. And because they cannot ensure the correct winner will be certified, they do not secure the election.

In Wisconsin, neither voting-machine audits under s.7.08(6) nor counties' voluntary voting-machine audits reliably detects or corrects outcome-altering miscounts, for reasons including:

- County election audits cannot confirm winners in statewide contests.
- Neither the s.7.08(6) audits nor the voluntary county voting-machine audits have written standards and procedures requiring expansion if a miscounting machine is detected.
- The s.7.08(6) audit instructions have an additional feature related to interpreting voter intent that renders them unsuitable for documenting miscounts caused by human intervention.[2]

Only RLAs qualify as an election-security safeguard for an additional reason: Deterrence. When RLAs are routine, hackers will know that even if they defeat pre-election safeguards, they will be doing nothing more disruptive than allowing election officials to demonstrate that they will secure the final election results, nonetheless.

In contrast, spot checks without detect-and-correct capability can reveal miscounts without determining their extent or correcting them. This opens the door to controversy and chaos and may, counterproductively, make the jurisdiction more attractive to hackers who seek only to disrupt.

C. One possible, simple statewide RLA

Please don't consider the following description as anything more than a starting point for discussing how Wisconsin could verify statewide elections, a way of introducing the concepts and vocabulary. Every step described below could be done in a different, likely more complicated way. Many stakeholders and auditing experts, from both within and outside government, should have input into the design of Wisconsin's RLA practices.

In the following example, notice that I have used the term 'state staff' instead of 'WEC staff' when describing state-level tasks. I believe policymakers should not simply assume that WEC will perform the statewide RLAs, although there would be no problem with that. Election audits might be more effective in building voter confidence if responsibility for overseeing the elections is separated from the responsibility for auditing them. The Legislature could, quite sensibly, assign the responsibility to the Legislative Audit Bureau. No agency in this state is more respected or accomplished when it comes to both the theory and practicalities of governmental auditing. Other possibilities include a state constitutional officer, such as the Attorney General, or an administrative agency such as the Department of Administration.

To provide realism, this illustration uses actual contests from November 2018. The Governor's race is a good example of a narrowly decided contest, for which events demonstrably eroded voter confidence. Notice that a quick RLA confirming the statewide outcome might have helped to put the City of Milwaukee's delayed absentee results in perspective.

The specific type of RLA described here is a "ballot-polling audit," which is the simplest type. Describing the other types of RLAs is outside the scope of this paper.

[2] The reason why WEC instructs s.7.08(6) auditors to ignore voter intent is logical given the unique purpose for those audits, but the explanation is complicated and beyond the scope of this paper. Richard Rydecki can explain.

Step 1: Do the same basic things officials do for any post-election audit.

Certain precautions and preparatory steps are necessary but not unique to RLAs. Some of these, such as securing the paper ballots, are already familiar practice for Wisconsin clerks. As is done now with the s.7.08(6) voting machine audits, officials can select some contests before the election for audit and randomly select an additional race or two after the polls close.

Step 2: Choose a risk limit.

On page 4, an RLA was defined as one that "1) objectively quantifies the risk that, if Election-Night results identified the wrong winner, the error will remain undetected, and 2) continues counting votes until that risk is equal to or less than a "risk limit" chosen before the audit began."

States performing RLAs tend to set the risk limit at either 5% or 10%. That means if the Election-Night results identified the wrong winner, there is a 5% or 10% risk that the audit won't detect it. ***This does not mean any election officials are willing to tolerate a 10% risk of incorrect election results.***

To understand why, we need to look at two different risks:

1) The risk that an outcome-altering miscount will occur; and
2) The risk that, if one does, officials will fail to detect and correct it during the canvass.

An audit's risk of failing to detect an outcome-altering miscount is 0% when there is no such miscount to detect.

The first risk is reduced mostly by pre-election security measures, while the second is reduced by RLAs. But RLAs help with the first risk, too. At the moment an RLA with a 10% risk limit begins, it has already reduced the risk that a miscount occurred—by putting hackers on notice that they face a 90% chance any hacked results will be detected and corrected.

An RLA's specific risk limit applies only to the second type of risk. To reduce that to zero, officials would have to perform huge hand counts after every election—at least 50% of the ballots in even the biggest landslides. If we're not going to reduce the risk of undetected miscounts to zero, we need to decide what level of risk we will accept. The 10% limit is arbitrary; officials can choose any limit they want. The lower the risk limit, the more work the audit will require. For example, if Election-Night results in a 3 million-ballot contest gave 52% of the votes to the winner, the initial sample sizes needed for ballot-polling RLAs with 10%, 5%, and 1% risk limits would be 2,902 ballots, 3,769 ballots, and 5,780 ballots respectively.

Step 3: Create a ballot manifest.

A "ballot manifest" is an index of ballots that enables election officials to locate specific ballots, so that they can be retrieved if chosen for the random sample.

States are choosing a wide variety of different methods to create ballot manifests. The following is only one possibility.

1) As poll workers place the ballots into bags on Election Night, they would need to count and record how many marked ballots they place in each bag. If they fill more than one bag, they would also need to assign a number to each bag.

   Currently, Wisconsin's poll workers are exceptionally reliable in verifying the total number of ballots cast in the polling place. Counting how many are sealed into each bag will be no extra

work in those polling places that can seal all ballots into one bag, and only a little extra work in other polling places.

2) The morning after the election, the municipal clerk would need to compile that information from all reporting units and create a municipal ballot manifest. For example, 607 ballots were cast in the Town of Adams, Adams County, in the November 2018 election. The town has one polling place serving three wards. On Election Night, poll workers might have sealed the ballots in two bags. That town's manifest would have looked something like this:

Ballot Manifest, Town of Adams, Adams County, November 6, 2018

Municipality	Reporting Unit	Bag number	Number of ballots
Town of Adams	Wards 1-3	Bag #1	310
Town of Adams	Wards 1-3	Bag #2	297

3) The municipal clerks would quickly report the information to the responsible state agency. That agency would compile the information into a statewide manifest and assign a unique number to every ballot in the state, as shown below.

Ballot Manifest, State of Wisconsin, November 6, 2018

County	Municipality	Reporting Unit	Bag number	Number of ballots	Sequence: First ballot	Sequence: Last ballot
Adams	Town of Adams	Wards 1-3	Bag #1	310	#1	#310
Adams	Town of Adams	Wards 1-3	Bag #2	297	#311	#607
etc…						
Wood	City of Wisc Rapids	Wards 11-29	Bag #9	300	#2,688,579	#2,688,879

An online reporting form would make the task easier for the municipal clerks, but legislation will probably be necessary to obtain their compliance. Because WEC staff can predict generally how many ballots to expect from each reporting unit, predictable problems can be avoided if such a form was set up to catch and force correction of large typos. Based on what we learned in the 2016 recount, it's a strong probability that at least a few municipalities will enter something like '44' when they meant to enter '484.'

Step 4: Calculate the initial sample size.

To calculate the needed sample size, state staff would also need to know the Election-Night vote totals for each candidate. So local officials would need to transmit that information, too, more quickly to the State than they now do.

State staff would then enter that information, plus the risk limit, into the RLA ***sample-size formula***. This formula has been endorsed by the American Statistical Association, tested, and used in other states. An easy-to-use application is available on the UC-Berkeley website.[3]

[3] https://www.stat.berkeley.edu/~stark/Vote/ballotPollTools.htm

As shown in the table on page 2, an RLA of the Governor's race would have needed a sample of 38,062 ballots. Only 405 ballots would have been needed to verify the US Senate race. Because the two races can be audited from the same ballots, only the larger sample would have been drawn.

If it seems counterintuitive that a statewide race can be audited with only 405 ballots, imagine a food scientist measuring the salt content in a soup. If he can determine the salt content in one tablespoon of the soup, it does not matter whether he draws that tablespoonful from a cup, a quart, a gallon, or a tanker-truckload of the soup—as long as the soup is well-mixed (or the ballots selected randomly.)

RLA sample sizes vary more based on the size of the victory margin than on the number of ballots.

Extremely close elections, such as the 2018 Governor's race (49.5% to 48.4%), will need larger samples, but that makes more intuitive sense. It is these contests that officials need to be most careful about, and it's the very close races that, when left unaudited, provoke the most resentment and suspicion.

Step 5: State staff Identify specific ballots for the sample and notify the municipal clerks.

Auditors and statisticians have several well-tested tools for generating random numbers in a way that observers can see is truly random. For the November 2018 election audit, state staff would have generated 38,062 random numbers.

They would then have referred to the ballot manifest to identify the location of ballots corresponding to those numbers. (A simple spreadsheet program could do that; publication of the ballot manifest and random-number list would make the process transparent.) For example, if the random-number generator had picked #350, the ballot manifest would have revealed that ballot was in Bag #2 in the Town of Adams.

State staff would then notify the municipal clerks with instructions for retrieving the ballots.

Step 6: Municipal clerks retrieve the selected ballots and transmit images to Madison.

To ensure objectivity and transparency, the municipal clerks' retrieval of the selected ballots will need to be officially witnessed and documented in some way.

Real ballots are not numbered, so Ballot #350 could be any in Bag #2 from the Town of Adams. This means that the clerk can randomly select any ballot from Bag #2 to be "Ballot #350." The instructions for random selection could be something like: "In the presence of observers, pull the ballots out of the bag, set them in a stack on the table, and roll a dice to determine how many times the stack will be cut like a deck of cards. Let an observer from each political party cut the stack that number of times, then cut the stack two more times, and select the ballot at the bottom of the last cut."

The municipal clerk would display the randomly selected ballot to the observers and transmit the ballot information to the state agency in a manner prescribed by the State. Possibilities include reporting the votes on a web form in the presence of observers, followed by faxing an image of the ballot to the state agency. The municipal clerk would then mark the selected ballot with red ink indicating it was the ballot selected for the audit; put it on the top of the stack of ballots; and reseal the bag.

Step 7: Conduct a manual count.

The precise method for manually counting the votes in an observable process would need to be worked out. Other states provide several different methods that could be considered for use here.

Step 8: Determine whether the risk limit has been reached.

When the manual count is complete, officials will enter the results into a second formula, the ***risk-limit formula***. By comparing the winner's proportion of votes in the sample with the winner's proportion of votes in the preliminary results, the formula calculates the risk that the preliminary results contain a still-undetected, outcome-altering miscount. If the same winner received a closely similar proportion of votes in both results, the risk will be below the chosen risk limit. In that case, the audit is complete, and the county canvasses can conclude their certification process as normal.

Important point:
RLAs stop only when a winner is confirmed.

However, if the winner's share of the sampled votes differs too much from the preliminary results, the formula will show the risk to be above the limit. In that case, auditors will select more ballots, count them, and run the risk-limit formula again. They will continue to do this until the correct winner is confirmed. This practice—called ***expanding the sample***—is the most critical feature of an RLA.

In practice, when Election-Night results identify the correct winner, nine out of 10 RLAs will confirm the winner with the first sample, and the tenth will likely need only one expansion. In the rare case when preliminary results identified the wrong winner, the auditors will be forced to keep expanding the sample until they have counted all the ballots. Sensible policies, though, would abort an RLA if expanded samples only increase the measured risk—a strong signal of miscounted results. Likely, state officials would determine it wise to stop the audit, declare a lack of confidence in the preliminary results, and order a full recount.

D. Unresolved issues

1. Does Wisconsin need a new law to have RLAs?

The biggest argument might be whether new legislation is needed. Given the current sluggish pace of new legislation, it may be unwise to wait for new laws if there is some way to conduct RLAs under current law. Technically, no new laws are necessary. No laws would be broken if the WCCA and WMCA worked collaboratively with the WEC to carry out the statewide RLA that I sketched out above. But sadly, it's unlikely enough clerks would cooperate. In reality, the Legislature will probably need to force local officials' cooperation.

Then the question becomes: Can WEC (with legislators' knowledge and consent) reinterpret s.7.08(6) and use that authority to order what is known as a "batch-comparison" RLA? Some thoughts:

- A "batch comparison" RLA would enable state staff to randomly select entire reporting units, rather than individual ballots, for the sample. However, each one of those reporting units would need to be audited with a 100% hand count, and the results would need to be transmitted to Madison more quickly and in more detail than the voting-machine audit results were sent in 2018. The amount of hurried work for the municipal clerks would be multiples of the work they would need to do for the ballot-polling audit I sketched out above.

- The size of the victory margin in the closest audited contest would determine whether a batch-comparison RLA would select more or fewer voting machines than were selected for the 2018 voting-machine audits. So, we cannot predict whether it would cost more or less.
- WEC would need to revise its instructions to tell auditors to read, not ignore, voter intent. If they were also instructed to record how many times this caused them to read a vote differently than the machine likely read it, the audit could serve both the purpose of an RLA and the purpose of the s.7.08(6) audits.
- WEC will need to think through how to handle expansion if the batch-comparison RLA fails to achieve the risk limit with the first sample. Without a specific plan and procedure for expansion, it could not fulfill the purpose of an RLA.

2. Too many computer-marked ballots make verifying outcomes impossible…but how many?

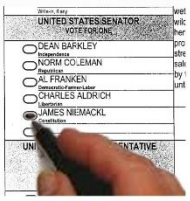
Nationally, the hottest debate about election audits involves the use of computer-marked ballots. Ballot-marking devices are necessary for voters with disabilities. However, if BMDs are used by too many voters, auditing election results becomes impossible. That is because although computer-marked ballots are *voter*-verifiable, they are not *auditor*-verifiable.

Credible audits comply with generally accepted audit standards, such as those promulgated by the US General Accounting Office. These standards require that auditors rely on "competent" evidence. That is, auditors must have a way to know the record they are looking at (e.g., the ballot) is an accurate record of the thing they are looking for (e.g., voter intent).

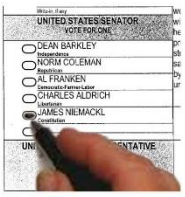
Hand-marked ballots are a textbook illustration of competent evidence of voter intent, as shown in this illustration. Auditors see the same candidates' names the voter saw, and they see physical evidence of the voter's choice: an inked-in oval.

In contrast, computer-marked ballots are a textbook example of insufficient audit evidence. Auditors cannot know what the voters saw when they touched the screen, nor can they know whether the voters reviewed the printed ballots.

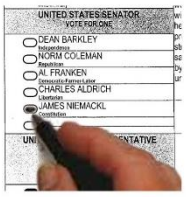
Election officials have been choosing to trust the ballot-marking computers and to assume voters catch and correct any Election-Day computer errors. But 'trusting' and 'assuming' are precisely what



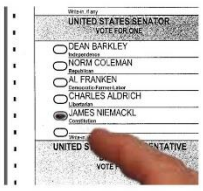
Voter touches ballot to indicate the intended vote.



Vote is recorded on paper.



Voter looks at ballot to verify that vote was recorded as intended.



This official sees what the voter saw, has proof the voter looked at the ballot.

Official looks at paper during a recount or audit, asks: "How do I know the voter's intent?"

This official cannot see the screen the voter saw or know if the voter read the printed ballot.

audits must avoid if they are to have credibility and value. If trust and assumptions are good enough, why bother with audits?

We cannot abolish either accessible voting equipment or generally accepted audit standards. Effort will be needed to balance both until officials accomplish the impossible—that is, finding ways for auditors to verify that:

- The voters reviewed their ballots; noticed any incorrect or missing votes; and reported those errors to the poll workers; and
- Poll workers were able to tell the difference between voter error and machine malfunction, so that they were able to make good decisions about whether to help the voter re-vote or to turn off the machine.

The problem isn't yet solved, but neither is it yet catastrophic for audits in Wisconsin. The RLA formulas were developed and tested with the assumption that accessible equipment would be used by a small percentage of the voters. No one has yet answered the question of how many computer-marked ballots is too many. Formulas that expand the RLA sample size based on the proportion of unverifiable ballots could be developed but haven't been yet. Whether Wisconsin elections become unauditable depends on whether the use of ballot-marking devices grows faster than the use of the old DRE-VVPAT machines declines. [4]

In any event, the problem is almost certain to come to a head in jurisdictions where all voters cast computer-marked ballots. Lawsuits have already been filed in Pennsylvania and Georgia to force an end to overuse of ballot-marking devices.

3. Other issues

Just to give you a flavor, here is a list of some other issues. None are insurmountable, but when RLAs come to Wisconsin, they will need to be resolved.

- **Transparency** – Other government officials can establish credibility by turning to disinterested outsiders, like accounting firms, to verify that their work is honest and accurate. Election officials don't have that luxury because everyone has a stake in their work. So, election officials must do everything in a way that can be seen by anyone who cares to look. I skimmed over it in this short paper but maintaining transparency for each step in an RLA requires careful thought and planning.
- **Cross-jurisdiction races** – Selecting a random sample of ballots when a contest crosses municipal or county boundaries can be more complicated than the process I described above. US Congressional races are the best example.
- **Close local contests**—Because the size of the random sample varies more in response to the margin than to the number of ballots (*explained on page 9*), RLA samples for close local races can sometimes be not much different than a full hand-count. This forces local officials to confront the choices they are making regarding their willingness to trust versus their desire to secure the results, and regarding taking affirmative managerial responsibility for accuracy versus

[4] The math has not yet been done, but in Dane County where the county clerk reports that 30% of the ballots are now computer-marked, audits of *local* results are likely now impossible in all but landslide elections. But Madison's heavy use of computer-marked ballots may not yet be enough to prevent audits of *statewide* results. One solution would be for the City of Madison to provide its early off-site voters with blank-ballot-on-demand printers, so that the libraries could continue to offer all ballot styles without forcing all early voters to cast computer-marked ballots.

relying on candidates to demand recounts. Personally, I'm optimistic that as local officials get more experience with hand counts, they will have more confidence in their ability to do them efficiently and therefore more willing to check accuracy routinely, even if it means an occasional audit uses a 100% hand count.

E. Conclusion

In 2014, a municipal yes/no referendum in Stoughton, Wisconsin was electronically miscounted so badly the error was immediately obvious on Election Night without an audit. The problem was attributed to human error that made the ES&S DS200 tabulators look for votes in the wrong area of the ballot, coupled with botched pre-election voting machine tests by both Dane County and Stoughton.
In response, no change was suggested or made to local canvass procedures, state guidance, or state law.

In 2016, a municipal yes/no referendum in North Kingstown, Rhode Island was electronically miscounted so badly the error was immediately obvious on Election Night without an audit. The problem was attributed to human error that made the ES&S DS200 tabulators look for votes in the wrong area of the ballot, coupled with botched pre-election voting machine tests by both the vendor and North Kingstown.
In response, Common Cause-RI, the Rhode Island Board of Elections, Secretary of State Nellie Gorbea, and the Rhode Island Town and City Clerks Association came together to promote state legislation to make RLAs a routine part of that state's canvass procedures. The law passed, and from now on, Rhode Island will routinely secure its election results from electronic miscounts.

I will not believe that Rhode Island's civic community cares more than Wisconsin's about protecting our right to self-government through fair and accurate elections. I hope you agree, and that you exercise the leadership to start Wisconsin on the path we should have taken in 2014.