

Close by Edward Foley

Download	HTML	Title
		<b>Notes of Democracy Index Conference (Cemenska)</b> Created: Wed, 10/03/2007 11:33 AM EDT

## NOTES OF DEMOCRACY INDEX CONFERENCE (Nate Cemenska)

### Introduction

Foley

Goal: what are 10-20 metrics that would compose democracy index

Five areas: Registration, casting, counting, post-voting procedures, general considerations. Try to find 3-5 measurements on average in each area that are promising. Don't worry about what weight to assign each measurement for now.

Ground rules:

1. Scope of conference: Election administration, not campaign finance, ballot access.
2. Goal is to measure how good of a job administrators are doing. We're not trying to measure overall health of political economy by allowing cultural or other factors to be mixed into analysis (e.g., voter turnout).
3. Index needs to be bipartisan or non-partisan.

Gerken

Hard to get traction on electoral reform. How can we actually get our ideas to become reality? Providing more information might help this goal. Hold election officials accountable. Performs a "shaming" function that might give incentive to do better. Might also help election administrators lobby for more resources. But what would be realistic, what can actually be measured? We cannot decide whether DI could actually work until we know what it would look like.

### Dan Esty

Three goals. Share why this is valuable idea. Talk about methodology. Lay out some of the issues and challenges.

Vision of environmental metrics came from academic interest in having information control rather than command and control legislation. Grew out of idea of World Economic Forum competitiveness ranking. Other concerns need to have indexes so that economy does not get all the attention. We can learn a lot from businesses, which know that the only thing that matters is what gets measured.

Project started about ten years ago. Helped signal what issues needed to be addressed worldwide. For instance, Mexico's ranking caused the country to be concerned. Countries complain that they are unique and face special problems, which is true, but in some ways irrelevant. Cannot stand in the way of having some benchmarks.

You don't have to get it all right the first time. Took 80 years to develop GDP system.

Sustainability index initially contained 70 metrics, but they decided it was unwieldy because it covered past, future, and too many issues. Current version is based only on past performance and has fewer metrics. Put it together with literature review and series of three expert panels. About 24 different issues identified, and then tried to find out whether data sets already existed. Some did, but not enough. So, were there proxies available? There were some. Spent a lot of time looking at reliability of data collection methodology and reliability of information collection. Graded different data sets for quality; if it gets a D grade, they don't use it in calculating the metrics. Because of poor data, can't measure wetlands protection, toxic exposures, recycling, and other areas. However, they have drawn attention to data gaps, which has a shaming effect in itself, as governments to some extent now feel that data should be provided.

Not looking at inputs (money spent), but look at outputs (pollutants in water). Historically too much attention spent on inputs, which didn't tell you how much quality you got out.

Current approach is proximity to target approach. Set targets by hierarchy of 1. international treaties; 2. 4. what major countries are doing; 5. scientific consensus. Target doesn't really matter. Distance to target varies depending on what target you pick, but it doesn't matter. People will be influenced by target no matter what it is, and it will show useful measurement.

Put 50% of weight on human environmental issues, 50% on the rest of ecosystem. This was not disputed that much, based on science. Subcategories within each of these categories. Use common factor analysis in some cases or otherwise simple GPA across categories. Now ranks countries 1 to 133. Ranking absolutely essential to getting countries to care or respond. Past efforts that did not rank were ignored. Still, have to be careful to qualify analysis to make it clear can't guarantee that no 1 is really higher than no 5.

Looking at peer groups by disaggregating is more useful than just looking at 1 to 133. Haiti isn't trying to compete against Sweden.

One question is what to do with countries for which not all cells can be filled. Currently, they just drop them out of the index. Used to try to use proxies.

They have found a lot of factors, economy, demographics, etc.

Transparency. Publish all underlying data, issues they could not cover, assumptions they're working with, statistical sensitivity analysis using outside experts and audits of how they constructed index. 70 countries have given input on how to improve index.

Interesting to look at drivers behind performance. You see variation around regression line where countries with similar characteristics are performing differently. Now are trying to fine tune the analysis to look at government structures, to see how they influence outcome.

## Challenges

1. Data sets will be a huge mess. Shocking that we're not better than this.
2. Now working on filling gaps and making methodologies more consistent.
3. Questions about how to aggregate, how to weight factors. They use unweighted grade point average, but do add each category up. Must add up category because it makes more of a splash.
4. Countries are angry, in denial, want to be removed from index. But anger is good because it means they're paying attention. Korea has spent a lot of money double-checking because of anger.
5. Some say simplicity is hallmark of good index. But they disagree, because you've got to capture the complexity.

6. Accessibility vs. statistical rigor. Increasingly, accessibility is better. More important for people to be able to understand methodology

#### Thad Hall-question

If you drop people out for not providing information, do you punish. Answer: They don't report either because they don't track or because they're ashamed. To the extent that you think lack of info is a bad, you can impute low score to them to give incentive to report. In general, countries are embarrassed about being left off the index because it makes you look incompetent or fudging the numbers. Esty also audits data and kicks bad numbers out.

How do you try to collect new data where there is not existing data. Working with officials, trying to coax, trying to construct data sets for future. In next report, half of data sets will be reconstructed in one form or another. There are people who will work with you to get more and better data.

#### Huefner-question

After you aggregate, how do you decide how to weight? Put all of variables on common scale, don't let extreme outliers in just one area depress a country's score that is otherwise good. Except for 50/50 human/environment, do not weight but use unweighted average. Encourage people do weight the categories themselves and use their own weighting system. In future, technology will allow users to assign weight to chart to get the data the way they want.

#### Fortier--question

Did you consider using some input data? Yale still uses some input data, but before you do that look for proxy of output (ex: using just carbon monoxide emissions as proxy for all greenhouse gas emissions). If an issue is important, use input data because it's a shame to drop out the issue altogether. Signal that you're using input data to encourage development of more output data. By the way, try to avoid survey data because it's unreliable-only good for keeping an issue on the table.

#### Cain-question

Have you experimented to see how assigning different weights would change rankings. Answer: Yes, we have. However, changing targets doesn't change rankings much. Changing 50/50 weight does change results a lot.

#### McDonald--question

Are countries trying to change score by withholding data? Yale doesn't believe there are systematic attempts to withhold, although there may be systematic fake data (ex Soviet Union). There are some isolated gaps that make you wonder whether they are strategic.

#### Kropf-- question

What was the problem with survey data? Answer: Example of regulatory rigor, how seriously laws are taken. Surveys on this question are unreliable and what it means that laws are taken seriously varies across countries because of culture, etc.

### **Voter registration**

#### Tokaji-brief summary of panel's papers

1. Alger's paper. Is it appropriate to use title Democracy Index for something that just measures election administration? Thinks its important to keep eye on the ball-what percentage of people are turning out

to vote, and are they representative? Dean looks at two measurements: Citizen outreach and registration (accessibility of website for college students, ex-felons). 2. To what extent do states allow users to verify registration status? Minnesota currently lags behind. Online capacity is not enough, need other ways of checking that need to be measured.

2. Mac Donald's paper. Timeliness of processing applications using all registration methods. Could require data reporting from election officials. Date mailed from date put into database. However, there are data collection issues here: 1. concerns about accuracy and completeness
3. McDonald. Database accuracy. Database serves number of functions, not just checking to see whether you're eligible, but also for communications to voters, jury duty, etc. Two big problems are deadwood and errors (esp. data entry). Consequences are profound (waste of money by sending info to wrong place) and eligible voters being inconvenienced or disenfranchised. But how to measure? First, you could just look at internal database. Second, you could do external analysis, compare to percentage of eligible voters, percentage of records where there have been purges, etc. Can also look at provisional ballots cast, although there are many explanations for this.
4. Tokaji-- Measuring provisional ballots rejected because voter is not registered. There is lots of variation across states. Cautions against interstate measures of provisional ballots. There are all sorts of reasons why voter might not be registered. Could be an error on part of about five different entities, including voter. This might limit meaningfulness of knowing # of ballots rejected. Second, even if we thought this was meaningful, how can we measure the number of times ballots rejected for this reason. Dan still thinks this information is valuable and worth collecting, but too problematic to invest a lot of importance in.

Two questions

2. How can we deal with data collection issues.

Alger

Don't look at turnout for index, it's just good for background. Look at how you're reaching disadvantaged groups. CarterBaker said register as many people as possible. This is why voter outreach should be measured. Look at outreach for subgroups. Minnesota doesn't have much voter outreach or budget for this. Limited resources are a huge problem. Good start could be how accessible is website. Not even slightest hint of elections on Ramsey County webpage, where St. Paul is. Also have to think about people who aren't as technologically sophisticated. Alger could not come up with good metric. Citizens have to call county auditor to find whether they're registered. Not good, because administrators are working 12-16 hours per day before election.

Mac Donald

If you are perceived as an auditor, you will get obstructed by officials and won't get good information. Instead, look at data provided by administrators and look for things that don't add up. South Carolina had reported huge number of HAVA voters?, but Karin found a mistake. Administrators are overwhelmed by data collection requests and the person who gets responsibility for this is not trained, low level, will make mistakes. Karin says because of quantitative problems, qualitative problems should be part of index. Right now not enough cooperation between and among administrators and academics. Suspicion and lack of transparency because administrators are afraid of getting in trouble.

McDonald

Information must be accurate, but he doesn't think there is good information. Not sure whether registration accuracy is where we want to go. Do we want to look at local level or state level? McDonald feels for administrator who have an impossible job. What is the benefit of shaming them? Instead, shame the state.

Connie McCormick

Accuracy is important. LA is very open about providing, but people's conclusions are often ignorant or statistically invalid. Lots of inaccurate data out there.

Olson

Can we measure data quality and have that be a part of the metric? Can we measure whether the infrastructure is in place to have good data collection.

Paul Gronke

Identifying errors in registration rolls could be somewhat easy. Academics could get credibility by doing good job on this (Fortier questions this, and McDonald said it would take about \$250,000 just to look at list). People in Oregon were proud of voter outreach (not registration, but encouraging to vote ballots). Gronke asked them whether outreach made a difference. But administrators use input metric and had no idea what output they got.

Alger

There's no allocation whatsoever for outreach.

Tolbert

Guy at Brown has been evaluating websites of 50 states. 20 different components (disability access, how many clicks to get where you want to go, etc). Are we able to make assumptions such as more registration is better? Can we assume that purges are good or bad? Can we say that being able to register close to election is better or worse? Environmental index has assumptions as well, but they are easier to make (greenhouse gases are bad).

McCormack

(couldn't hear). California has closed registration system.

Quinn

Virginia has closed system. Can figure out online whether you're registered, but other people can't look at this information.

Stein

Blaming people at local level might be counterproductive. There are some things that are objective (data entry). You can measure proxies like do we have technology in the registration system to minimize error, optical scan.

McDonald

There is some evidence to suggest data entry errors are going down with new statewide databases. More common-sense checks could be put in place. In Virginia, if there is a problem, officials will follow up by contacting voter.

Fortier

Can we break registration metric into simple yes/no questions (do they have a website?). Do they allow 1-800 to check registration? Do we know data entry is a big problem (not sure this is what he meant).

Thad Hall

Interoperability is an issue-need standard data. What about fragility? Florida just happened to break in one election. What makes system less likely to break?

Kropf

What are we measuring? Access, integrity, customer service. These are all different parts of registrations. Can't imagine punishing a local election official for making data entry errors. Instead, facilitating ability of people to confirm their own registrations are correct may be better. Many counties don't have research. Should it be part of index to measure whether it is part of DMV's job to make sure registrations get to correct office? Do we think it's okay to let third parties to be able to register voters (ACORN).

Alger

SoS tried to establish capacity to verify registration status, but it got show down by legislature. If this were part of metric, might make legislature more open.

MacDonald

Some county registrars don't even have email, are we shaming people who don't have control, don't have infrastructure. The state has only so much control over local people.

Fortier

Thinks we should concentrate on state level. Any state should have web access to find out whether you're registered.

Cain

We're leaping to how we use the data. Let's not think about shaming or other uses just yet. What do we do with bad registrations? Throw them out. In the past, goal has been to maximize registration. We should focus on people who tried to register. People who do not try to register should not be part of measurement of registration quality. Quickness of processing and transparency are important. And what about convenience, cost of registration?

Second, two causes of inaccuracy and problems within category of people who try. We're mostly talking about problems in administration, but we also have to talk about voter error from lack of voter education (and regular education), political stimulation, help from administrators. Is it really fair to hold an administrator to the same standards if they're dealing with problems of language barriers and lack of education.

McDonald

Got to look at locals because there the ones that have the primary data. Some data doesn't make it to the state level, and what does often isn't accurate at that point. Two ways to measure websites. Do you look at % of jurisdictions, or looking at % of registered voters in the state that are being served by websites.

Huefner

It would be easier to do metric if states had uniform registration systems. As a matter of administrative performance, we need to capture convenience. On data error problem, thinks deadwood is hard to find, and we can't tell whether they're still living there or are dead or moved. Grandfather data that we can't fix.

McDonald

Would like to know percentage of people on inactive roles that are actually voting. Do jurisdictions automatically purge the old data if you register somewhere else?

Quinn

Resources are limited and cost of developing output data is expensive. You need to look at the inputs that are being provided because the real challenge for administrators is not having enough resources. Also legal

and cultural limitations (like NVRA) stop from you from being able to clean up database in a lot of ways. Quinn tried to get \$ for voter education from the state, but it was opposed because that was locals' turf. You've got to look at how locals discourage state from taking on more power. It was a major battle just to get a database at all.

Accuracy can be measured by # of postcards that get returned. Ease of fixing errors should be considered. Are there systems in place to try to minimize voter error (for example, DMV clerk's computer system has ticker that says "Do you want to register to vote today?")

Stewart

Must think about what most important one or two. More registered is better than fewer. Accuracy is good. Not sure about which one is most important. Let's concentrate first order problems. Another question is, where is the data going to be coming from? Sustainability index is just a data aggregation. Probably want to improve state reporting, or we could try to do audits ourselves, but that would cost millions. Can we compare telemarketing data with voter registration data? What is the measure of the high order concepts? There are underlying hypotheses about whether fax machines improve registration, money improves it.

Chapin

Prefers found data.

McCormack

Experiment in California, some registrations didn't make it through and were not put in until canvassing (San Bernadino), but the people were able to vote and their provisionals were counted. Wants to know about EDR to see whether you can eliminate provisionals, which are expensive and compared to EDR create problems with putting registrations in at the last minute.

McDonald

Some states (N.C.) enter registrations immediately upon receipt, even when received after deadline (put date entered into database).

Foley

Thinks we have to decide values we think are top before we can start talking about metrics. Also need to decide whether we want to limit it to people who try. It's probably worth talking about how overly difficult registration process discourages people from even trying. Maybe state can be held accountable for whether people are registered, although they probably can't be held accountable for whether they actually vote.

Tolbert

We have to agree on unit of analysis-county, state or local? Maybe everybody should write down what the goals are for registration.

Gronke

Thinks that index should be designed to have categories for data that is available or at least susceptible of being collected. Certain things (people being turned away from the polls) just can't be collected at all in any systematic way.

Heather

Other people's memos

2. McCormack-votes not counted. Can we track this and use as measure of state performance. Lots of explanations for why absentee and provisionals aren't counted. Often voter error. Some are data problems. Others are things that the state could be doing better on a systemwide level. Widely variable practices about how states do provisionals and feedback information. California has problems because of "vote anywhere" rhetoric, which creates a "blip" that makes it look like the state is doing something wrong when it's not. Have to disaggregate data to see what reason is for ballot not being counted.
3. \_\_\_\_\_-- Is goal to maximize number of ballots cast? Transparency. No systemic bias. Can existing data sets get us there? No. Sample sizes are not large enough. Second, data is not tagged by state, but regionally. Need to improve collections to fix these problems. Do survey sampling of voters after election. What are driver's behind performance?
4. \_\_\_\_\_-- We want to decrease time and inconvenience of voting, but can we measure it? Should it be recommended as a best practice? How do we measure whether they work?

Tolbert

Stewart is right that we need measures that are not outputs of the election system themselves. Maybe do half of data based on individual experiences and evaluations. How do you find the data to measure these individuals? Might be able to use actual voter rolls. They have limited demographic data but you can cross reference it says neighborhood so you can cross reference it with other data to isolate particular demographics. CPS, current population survey, is not detailed enough. Could we update it to get better information for voting?

McCormack

There is great potential for misinterpreting database, such as having high volume of provisional voting and assuming that's due to a bad database. "Vote anywhere" has carried over to new elections. Connie would like to study who these people are. Promulgating best practices can improve things a lot (used updating provisional ballot envelope as example).

Stein

EDR, early voting have had marginal effects on turnout, just make it more convenient for people who are already voting. In Colorado, data found that voting centers created significant increased turnout among regular non-voters. States need to find ways to make an experience for the voter like Lowes and Wal-Mart create an experience for the shopper. These stores know the benefit of putting stores in logical locations where people will go, know the value of having a big parking lot, and know the value of having short lines. Poll workers should be just like the people at Wal-Mart. If the assumption is that we want more people to vote and for that to be more representative. Right now we make it really hard.

McCormack

Voting centers won't work in big jurisdictions except for early voting, where you use DRE. Trying to do optical scan ballots on demand but it's going to create all kinds of problems.

Stewart

CPS is only survey with large enough sample size although questions and screening are lousy. Hopefully it can get better. Could we get census bureau to pitch in? There are empirical questions about what best metrics will be. Different types of errors cause different types of problems. Some errors will not affect ranking, while others will. Whether survey technique is good or bad is empirical question. Esty might have had problems because they were asking people factual questions about things those people had no direct experience with. Might be able to do surveys of direct voter experience that is accurate and cost-effective.

Tolbert

CPS has 50K respondents, 1,000 per state, compared with most other surveys that are not statistically significant.

Gerken

Can we fix inconsistency in how states report data so ranking would be valid?

Stein

Problem with CPS is that they don't talk to non-voters.

Stewart

CPS has wildly different results from year to year, that makes it look unrealizable. Tolbert thinks its okay on the demographics.

Kropf

Do we value making it really easy to vote, or is that voter's problem to do his job? She is afraid that we will use metrics that are systematically related to race or income, and build bias into the metric. If you used e-pollbook, you could reduce risk of giving wrong ballot in split precinct or voting center.

Stein

Thinks that election should be representative-that is the goal. This is not biased. The election is a survey process. When some people don't participate, that biases the survey. Voting centers increase turnout but not sure it's representative.

Fortier

AVBM hasn't increased turnout. How do we know voting centers will increase turnout.

Stein

Consolidated precincts in LA county was a trade off. Ease and speed increased, although some people were confused about where to go.

Fortier

States are developing more sophisticated ways of tracking polling places, but still not sophisticated.

Stein

They do polling place where lease is cheap.

Hall

Mode of voting made a big difference is satisfaction. Experience with your poll worker makes a big difference. Surveys have shown this is true. Confidence factors make it so that some people do not vote-they do not believe their vote will count.

Tolbert

Should subjective measure of voter confidence be part of survey? She thinks so, because it's so important.

Gronke

Oregon has little poverty or minorities so it's procedures might not work elsewhere. However, what would translate in other jurisdictions is "agility" (example of communications infrastructure).

Quinn

Virginia became concerned that DMV registrations weren't getting into the systems when there were basic mistakes (forgot to sign form). DMV could go back and track when person came in and find their application. Virginia went back to look at applications of people who claimed DMV didn't put their registration through; 88 out of 88 voters who complained about this didn't fill out voter registration part of form.

Cain

How easy was it for Connie to collect the info in her report? How easy would it be for other administrators? LA had 222 people who tried to vote who were not citizens.

McCormack

Non-citizen voting issue was addressed by legislature; it was just people who had achieved citizenship but this hadn't been finalized. Connie says it would not be hard to record why ballots aren't counted-someone is there making the decision, might as well record it at that time. Doesn't think you could get legislatures to support mandatory reporting.

Gronke

Problem with CPS is because it's all stick, no carrot. It tells you people didn't have a good experience, but doesn't say why.

Foley

Goals for casting ballots should be convenience, but what about training, # of poll workers (there are input measures but still might be good). Number of complaints generated by hotline-maybe give points for having a hotline. Do we need to separate problems from problems that result in disenfranchisement?

Stein

Must have reliable and sustainable database. Must form alliances with county officials, only they can provide the data. We will never get enough money on our own.

Fortier

You won't be able to get all states and areas to cooperate. How could this be done on a nationwide basis.

Quinn

Disagrees. Thinks that you would get a high level of cooperation. Go to NASED, EAC and local elections officials. But you must get input from a representative group of administrators.

Stewart

In order to get any accuracy, you've got to identify the representative voter or go into lots of different types of voters (early, absentee, etc). Modes of voting are so complex we might not be able to capture them and have to go for idea of representative voter or use proxy like residual vote rate. Some states stopped reporting turnout because they were afraid of being crucified. Other states have taken measures to be more accurate for fear of being crucified for being wrong.

McDonald

Got to take residual vote rate with a grain of salt, because there are people who doing "selective abstention" and it's not always the machine's problem. Abstention happens particularly in midterm elections. Most jurisdictions aren't going to report or simply don't have the ability to report data like LA County. If it never

gets programmed into your system, you can't report it. CPS wouldn't capture problems in particular areas within a state. 19% of people blame not voting on not being registered. .1% of people reported problems with voter idea, which means 100,000 people.

Foley

For future agenda, do we care about whether administration is partisan or non-partisan, including poll workers? And what about ballot design and other issues?

## Second Comparative Perspective

Lynn Olson

Think about what you're doing as a baseline. Even if information is incomplete or uncollected, keep reporting because other states will come on board with reporting.

Education report card has focused on states, not 50,000 school districts.

Started at looking at existing research, created national advisory group of experts online who could give advice about how to build metric.

Had debate about inputs vs. outputs, but in the beginning there was little output data and had to use inputs. Looked at state policy levers as well. Found there was not a lot of high-level information sharing going on across states.

Student achievement, resources, school climate, standards and accountability, teacher competence were the measurements. Must tie these values and assumptions into metric because public doesn't care about casting ballots per se.

In the beginning, newspaper naively thought that there was already good information was out there, but what they found was that most data did not exist or was extremely poor quality. Schools didn't even have basic information like graduation rates, how many kids had taken algebra. What little data there was was "data dumps" that weren't organized. Now use 50-state surveys every summer and vet answers against other documentation.

You can use less than perfect data and explain that it's imperfect, you can choose not to use bad data, or you can choose to use estimates or proxies. Newspaper actually chose to use some metrics where numbers were low or there were no numbers to make schools fix these problems.

Definitions matter. What is a provisional ballot? It depends on a state to state basis. For every question they provided definitions. Survey goes direct to state superintendent, but come in waves so that each section may be sent to whoever superintendent determines is appropriate.

When they send out survey, they also send how the state answered the year before, to allow them to just say nothing has changed.

Made sure to explain assumptions and values to administrators.

When they get survey cards back, they send answers back to state superintendent to make sure it checks out. Different part of bureaucracy don't necessarily talk to one another.

Every year the newspaper will do a special poll on an area that they wanted to draw attention to, even if there was no data collection.

Have subcontracted out for polls, other things.

Needed a definition of urbanicity to do report on urban education.

Local control states that don't have strong centralized bureaucracy will sometimes refuse to give data. Newspaper tells them it will still report on them and they won't get off the hook by not participating.

Locals do use metrics as leverage to go to their state representatives. Also used by advocacy groups.

Don't be afraid to make judgments. The reason why the newspaper is successful is because they do give grades and make judgments. It's just a bunch of numbers unless you rank the states. Originally decided to give letter grades for each area and not an overall grade, because felt that states could be good in one category but not others.

Avoided "horse race" indicators like SAT and ACT scores because of pushback. However, administrations wanted harder output data, and now rank performance on state achievement tests and transfer success index (support for preschool and college, etc.-13 benchmarks for performing at or above the national average). Some states do well with school but problems once out of school, whereas other places are different. D.C. has good economy but poor schools.

Masters degrees, number of years of experience, don't seem to make much of a difference in teacher quality.

Put your attention on outputs if you're not sure inputs are correlated with performance. If you are sure, then it's okay to use inputs.

In finance, they use an equity measure and McClunnen index (mean for state and what it would take to move everyone up to that level).

Biggest pushback comes from local control states that do poorly on state policy even though they may do okay overall. Other pushback comes from states who say they're doing everything they can but they can't get the locals to comply. She doesn't believe the arguments of those who argue they shouldn't be judged because every locality is different; nobody can know all the local influences but by providing some objective information at least there can be some comparison.

If you have subscores it can help show why a state did poorly in one area. This is why newspaper avoided overall ranking. Nevertheless, in other circumstances school rankings have made a difference (New Mexico example).

Mac Donald

Thinks you can do implementation research.

McDonald

What's your audience and how do you tailor index to audience?

Answer

Audience is the media. 1,000 newspaper stories on the day their report comes out. Researchers also use, but

not primary audience. Have different levels of detail that are adjusted depending on how much you want.

Foley

In voting, it's now always clear who deserves to be held accountable. Sometimes state legislatures or leaders, sometimes locals. Sometimes locals don't follow the law.

Olson

The outputs we use are graduation rates. No child left behind has caused states to report disaggregated information. In some states there are not enough minorities or poor to get significant data for certain communities, but report it anyway.

Gronke

Not sure voters care about how their state does on Democracy Index like parents care about how their schools are doing.

Olson

Parents don't read our report, except rankings that are published in their newspaper. Policymakers are the ones who actually read the report, although it does influence parents indirectly. The data is not that useful to parents because it's too much detail.

Gronke

Still doesn't think that voters care like parents.

Olson

People really did care in the 2006 election about voting problems.

Iowa is local control state with not much strong state policy. They are very well on achievement but it does have problems with minorities. By disaggregating the data, it shows how achievement is not the only important thing.

Huefner

Olson is going out and collecting her own data.

Olson

Yes, because there is little out there to aggregate, at least in the beginning. There might not even be enough to aggregate and give a fair picture.

Huefner

Education is similar to voting because there's not scientific standards for what is desirable, like with environmental science. How did you reach your standards, by excluding certain people from the room who disagreed with you? For instance, some would say that equity is not necessary as long as there is adequacy.

Olson

The reason why we have a metric is because we're not worried about getting 100% consensus before going forward. As long as you're clear about why you picked a particular target, you'll be okay.

## **Counting ballots**

Foley

Cameron's list: accessible/free expression of voter's actual intent/desire

Integrity/accuracy

Innovation/experimentation/reconsideration of methods/procedures

Clarity and transparency of processes/procedures

Something else

Should we be measuring accuracy or the counting processes themselves (processes)? Transparency?

Non-partisanship? Clarity and specificity of counting rules.

Methodological difficulties. When are you supposed to capture accuracy of the count? Do we do it by looking at certified counter or by audit? Do you have to review all ballots or is sample appropriate?

Foley's classification of counting errors: Tabulation errors, exclusion of valid ballots, inclusion of ineligible ballots, excess ballots.

Foley-question to team-do you think its important to think about measuring procedures and not just counting results

Chapin

He likes quantitative data, which is all that is out there right now anyway. This stuff can then be added into qualitative/input analysis later.

Quinn

You need both because good procedures result in accurate counts. In Virginia, local authorities double check precinct workers' count, then state officials looked at it again. Used example of recent recount where total only changed by about 30 votes.

Kropf

You can't just look at the outputs because they're not solely indicative of the inputs. There are problems with residual votes as a measure because they can be due to voter error.

Chapin

Yes, in the long run we should look at everything eventually. But in the short run we have to concentrate on the output data we now have. The kind of data that Connie provided is the exception to the rule. It's not necessarily important, for instance, whether it is important to have non-partisan administration or whether that affects accurately.

Quinn

I have been involved in two systems that were partisan and everybody thought that they were accurate. If you don't have partisans from both sides involved it gives your results less validity. Puerto Rico has 80% turnout despite short polling hours and ID requirements; their turnout is that they have "all the foxes guarding the henhouse."

Kropf

Research done on partisanship of election officials indicates there may be some effects. Research also shows professionalization of officials could help, more training. But has never seen a study that showed better training was correlated with accuracy.

Chapin

You have to have reliable outputs. Baseball analogy. All of the statistics in baseball are agreed upon. After you fill in the statistics, then you can start debating the meaning of those statistics like you do in baseball.

Hall

In Washington they remake 15% of all ballots, giving them a residual voting rate. This example shows how you have to know what the rules are to interpret the statistics. You have to disaggregate the data.

McCormick

Partisanship is not a problem, it's transparency. In some states the canvass period is too short to get results in and they just don't report the results that aren't counted before the deadline. This type of thing has to be considered.

Damschroeder

Most locals are happy to gather data if academics will tell us what information they want gathered ahead of time. Usually what happens is that we don't get the survey from the state of Ohio until the next year, which is worthless. It's a simple process. Usually just a matter of setting up fields in the database.

Gerken

Outputs/inputs. Problems happens when there's a lot of turnout. Do these output measures help measure fragility that will only lead in a breakdown once in a blue moon? Or do they not help measure that fragility?

Kropf

If you have good equipment, training, etc., it will reduce fragility.

Quinn

Florida had a great reputation until 2000. Outputs were good.

Chapin

I don't agree.

McCormack

Close races create failures. There will always be failures.

Gronke

Some outputs become inputs into other processes.

Stein

Doug is right that we don't have data on outputs. The problem is peeling back the layers of the residual vote. Could be voter mistake, machine mistake, voter choice. Is there any way of separating out the cause? Like Bender's audit procedures for checking drafts (checks).

Stewart

Residual vote is most useful in a comparative framework where you can hold a bunch of things constant. For instance, looking at jurisdiction across races, same races across jurisdictions, or same jurisdiction different machines. Difficult measure to use unless you've got that comparative perspective.

Kropf

If you break residual vote by absentee, early and election day.

Stewart

Discussion with Kropf about whether disaggregating residual vote data is necessary for ranking purposes in Democracy Index.

Gerken

There are trade-offs between voter confidence and new technology.

Foley

There is a trade-off between access and integrity, but that's okay because states will be ranked and they may find a way to reduce that trade off.

Cain

Absolute output is an accurate electoral decision, you can work your way back in the output chain and figure out what is absolutely necessary to have vote counted (registration, casting a ballot). Collect info on those first, then look at other variables that may or may not be affecting accuracy (absentee balloting, good poll worker training).

Gronke

Somebody needs to determine what is core to the index, but things that may or may not be related (election centers, residual votes), we don't need to fit all these ideas into the democracy index. Disaggregating absentee, early voting residual rates not critical for index.

Chapin

Two reasons why we have elections. 1. Make citizen feel like they have a voice. 2. Aggregate of all those votes needs to be accurate. Could these two things be a way to organize the index. Find a way to reward states that perform in these areas.

Foley

We could have an accurate system that is nevertheless perceived to be inaccurate. Should democracy index care about perception?

Chapin

He doesn't care whether people feel like process is inaccurate. Thinks what's important that system has the ability to prove that they are wrong, forever. Numbers are what's important, not perceptions.

Quinn

It's not fair unless you have clearly defined standards and rules going in (voter intent rules for counting votes).

Hall

Thinks confidence is important, but transparency may be more important than actual confidence. If you put the conditions for voter confidence, that's good.

Steward

Residual vote is calculated by valid votes who were counted/number of people who voted. It is always calculated with respect to a particular race. In Massachusetts its blank ballots/turnout. In other states its different, like having denominator be the number of people who ran a ballot through the machine.

Tolbert

Ballot initiatives at top of ballot can cause drop off for governor, President, etc. There is a problem with no universal definition of what is a residual vote.

Stewart

How much can recounts be used to measure whether the first recount was accurate? New Hampshire publishes data on every recount. Post-election audit data can be harvested, too.

McCormack

Agrees with Chapin that accuracy is important, not voter confidence. Voters are so colored by their emotions, they can't think clearly. In March of 2004 write-in ballots wouldn't count unless you filled in oval next to blank; lady lost election because not filled in. Law was amended so that it will not count except in recount if oval is not filled in. But recount will not happen unless election is close.

## **Recounts and Contests**

Huefner

Look at speed of counting provisional ballots and length of time to turn in additional information. Number of provisionals cast and why. Problems comparing across state because in some states if you're not in the poll book they tell you it won't count, in other states they don't.

Sometimes # of provisionals cast may be out of control of administrators (people not understanding rules vs. bad database, poor training).

Look at what # of provisionals are counted. Officials say they have little control. Can we measure the # that should have been counted but were not?

Recount issues: When should a recount happen? Ballots included in recount. How should it be conducted? Standards for recount. Transparency of recount. When would an audit be better than a memo.

Election contests: Hasn't gotten that much attention. Focus is typically on accuracy, but that's not the only important thing. Might still want to be thinking about access in how we treat provisional ballots. Might want to include finality and timeliness to outcome. Finality doesn't mean just ending, but having the result be respected.

Output measure for contest could be how many lawsuits are filed, but this doesn't say much. Other outputs: Might want to measure how long they take for resolution. Measure how often result is voided, changing outcome, how clear standards for relief are, neutrality of the forum.

Unresolved contests at the time somebody was expected to take office. Washington governor's race. Safeharbor date. Measure timeline for primary elections.

Question-how much weight should post-voting metrics have in DI? Recounts should satisfy, so if the action proceeds to contest, that is a type of failure.

Hall

Hall will not prescribe best practices for how recounts are performed. Audits vary a lot across some states-some only do DRE ballots. Others audit chain of custody.

Damschroder

We don't want to get too involved in details, we'll get lost. Ranking chain of custody doesn't necessarily help show whether outcome was fair. Chain of custody is a lot of minutia.

Foley

Is there a way to have chain of custody measurement be objective? Election contests might be a good proxy.

Damschroder

Evaluating frequency of recounts and contests becomes very subjective. There are some people who are unwilling to accept outcome no matter how objective that outcome is.

Hall

Number of recounts is going to depend on openness of process to obtain a recount.

Foley

There is some degree of inaccuracy in counting that cannot be eliminated. It doesn't matter that much unless result is changed. In those elections where the system wasn't able to distinguish a clear winner, that is a problem. You may think that recount is just result of lightning striking, then administrators can't be held accountable. However, if you think that administrators job is to be ready for the class 5 hurricane that occurs only once every 100 years, you can hold them accountable.

Gerken

There were a lot of lawsuits filed in Florida and Ohio in 2004. There were problems going on everywhere, so I'm nervous about how relevant this is. However, you can triangulate by talking to interest groups and other groups to find out how good chain of custody rules are, etc.

Quinn

A recount is a measure of how well your original count was, not a measure of how good your post-election processes are?

Fortier

Would it be fair to say that states that have recounts that end up the same as original count, would that say something?

Quinn

Yes, it's a good measure of accuracy of original count. For chain of custody, not sure how you could measure it. You get into issues of whether there is a law and also whether it is implemented. In Virginia, law was good but some clerks were keeping ballots in unsecure way. To have a good system you have to design it to withstand lightning striking. For lack of resources, most states haven't designed the process that way.

Fortier

This area seems to be harder to measure than other areas. Peer review might be one way, expert reviews might be another way (ex: rating database by having experts look at them). Otherwise probably not going to be able to capture.

Chapin

Check recounts and election contests altogether. They're heavily weighted against access and in favor of finality. Contests generally get thrown out. A more likely thing to look at are audits. Right now we are just talking about audit 2.0. Audits are a post-voting procedure analysis, and we don't need a post-voting procedure analysis of this procedure.

Huefner

Prevalence of recounts and contests could be a proxy.

Chapin

We run the risk of double-counting if we let recounts and election contests in.

Huefner

One way of measuring health is how well contests are handled. Do states prevent them from becoming politicized destabilizing events?

Fortier

There needs to be something about recounts and election contests because that's what voters think of.

Hall

Audits could be valuable, but right now laws are poorly structured or procedures are new.

McCormack

Recounts and contests are more frequent than you think. Recounts teach us about voting systems. In Orange county, there was an election that was 7 votes apart; 70% of votes cast absentee optical scan, the rest DRE at polling places. During recount result so that other person now became winner by 7. Then election contest was filed. Judge looked at ballots and decided them one by one. Thinks audit procedure is important but isn't being used broadly enough.

Gronke

A word about transparency. You don't want to know how the sausage is made. Less transparent institutions are often the most trusted ones (Supreme Court). Transparent institutions like Congress are not trusted.

Huefner

Finality might be a goal we're working for, so we can't have a system where we go back and audit the intent of every voter. Have to accept trade-offs. When we have an outcome that is within the margin of error then it comes down to the flip of a coin. Ideally voters should be able to recognize that Florida 2000 was a statistical tie.

Cain

This is a necessary step in some races. Other processes are about accuracy, but there is also a fairness part of election contests that cannot be measured and is for lawyers.

Stewart

Wynn case, original loser got to set parameters about scope of recount. However, once tables were turned and original loser was on top, scope of recount could not be expanded and there was asymmetry.

Hall

In Washington, they were using 150 year old law that came from Northwest Territory. There were no standards, no provision of jurisdiction.

Chapin

No one agrees what a good recount or election contest process looks like. We'll never be able to agree on it.

Quinn

No, lawyers with experience in recounts and contests agree on how these processes should be conducted, at least within each state.

Chapin

Virginia is a state where you can't go back and look at ballots, have to use returns, but not everybody would agree that's a good idea. Some like the finality and orderliness, but others want to see ballots.

Huefner

Many recount and election contests rules are old and horribly designed. Are there some core areas of agreement that we can say these processes need to have certain characteristics.

Stein

If you're trying to evaluate whether post-election process is good, you have to know what you want the outcome to be. Replication of original count, voter confidence, finality, or just sufficient accuracy to avoid changing result.

Fortier

A lot of the problem with post-election procedures is they're too loose and leave too much up to the discretion of officials.

Huefner

One easy answer to our goal for outcome of post-election process is that we want it to give us a final result before the date of taking office.

Gerken

One question is not whether the state is following the absolute best procedure, just whether there are reasonable procedures. It might be feasible to "measure" whether the post-election processes meet some bare minimum standards.

Hall

In index, can we differentiate between administrators that are skillful as opposed to lucky. If we look at recounts and election contests, it will be hard to know whether they are being caused by lack of skill or lack of luck.

Rachel

What about inconsistency across jurisdictions?

Foley

Yes, maybe we can add consistency to our list of values and add it to our metric.

McCormack

Consistency is a big concern. Many recount laws are junk.

9/29/07

## General considerations

Chapin-brief review of others papers

1. Gerken-treating the voter as the unit of analysis.
2. Stewart-using surveys to do end run around barriers to gathering info from administrators.
3. Mac Donald-qualitative assessment of variation in the field.
4. Fortier-using expert peer analysis to evaluate jurisdictions.
5. Gronke-paper is about how to actually gather data and have it be valid for comparison.

Chapin

Who is audience for this index? Quality counts focuses on public sector accuracy. Environmental index focuses on private sector, which makes it more bottom line. Where do elections fall? Should we use voter as unit of analysis? How do you form ranking in a field that is utterly void of data collection?

Fortier

Thinks that we will need peer review because of big holes in data and things that aren't susceptible to measurement. Nevertheless, doesn't think that current administrators are the people to get opinions from. Instead, use ex-administrators, computer experts evaluating databases, voting machine code.

Chapin

If we did peer review, how could we be sure that we have a representative sample of evaluators?

MacDonald

How do you expect the evaluators to have a basis for their rankings.

Fortier

You just have them do it, they're smart and experienced, they know who is good.

Cain

Universities don't like U.S. News, it drives universities in directions that aren't good. Nevertheless, I think the DI is a good idea. I think we need to remember, keep it simple stupid. Think about what disaggregated data you want first. Then later figure out weighting, etc. Start with most disaggregated level. How can we get reporting requirements improved? We need to start collecting the data now, and we have to get it from election administrators.

Gronke

I like US News. People don't like it because it oversimplifies, but Gronke's response is that it's just one piece of information and other things can be considered, too. He also thinks the "gaming" of the system that is happening is good, not bad. It's improvement. Reputation is a good thing to measure education by because it's value and something people pay for. Use retired administrators to do fragility rankings, etc.

Gerken

Started thinking that the audience was voters, but after talking with Esty is that policymakers are the real audience who cares.

Stewart

Think about what are core measures with the net survey. What can we do about fact that most data cells are empty or unreliable? Or instead we can just concentrate on the core ones. Get these critical questions into

the EAC survey. Make the list of things the EAC is asking for shorter.

Quinn

You probably need and want both peer and expert review. As an administrator, Quinn knew what administrators to go to for good information. She knew what states had good databases, what states were accessible to the disabled. You begin see patterns and know who is good and bad. Experts might include people like political parties-do you think your state's elections are properly handled?

McCormack

Voter databases have to be evaluated. There's going to be database failure in '08 when a huge volume of registrations come in, they will freeze up etc. What percentage of data is getting inputted on average? In LA, about 1% historically due to voter error, but it went up to about 27% because of new database.

Hall

Peer review would help discourage practices like Georgia bell ringing to stop count.

Alger

Alger showed Gronke's paper to Minnesota elections chief and the guy didn't like it. He also said it would have taken one staffer two weeks to gather input data.

Chapin

We're going to have to develop our survey techniques so they are not perceived as clunky to election officials.

Alger

Thinks that election administrators are going to be uncomfortable criticizing the guy in the next state over.

Quinn

If you phrase the question right and get into subcategories ("Who would you go to to learn how to run a polling place?"), administrators would be allright with that.

Alger

Agrees that mid-level questions would be taken better.

Gronke

In EDS, after they collected it they sent it back to locals to double-check data was correct. The shaming element really worked then, and they kicked in and made sure the data was correct. They didn't want the wrong information in there.

Stein

Let's make a strategy for the next step. If there's some core pieces of information that if we don't have, there will be no index. Used Doug's three underlying concepts: 1. all eligible voters who want to should be registered; 2. all registered voters who want to should cast a ballot; 3. all ballots cast by registered voters should be counted. We have to build up trust and credibility with administrators to get them to give us information. Maybe we can set up a pilot that will collect a little information, get information that is useful to them, and build up some trust.

Fortier

Maybe we can start with two states.

Stein

Right now we're doing Texas and Colorado at Rice.

Chapin

Two alternatives: Ask for just some info from every state, or ask for all information from just a few states.

Kropf

Thinks audience is not only administrators, but also lawmakers and writers of administrative code. If this is audience, then metrics have to measure only those things that they can change. Can they change voter confidence? If not, don't measure it. However, more specific questions that you might ask in an exit poll might be useful, despite methodological problems. The question is, does this information give administrators an action plan.

Damschroder

There has been so much change recently that it has become difficult to know who is the best-databases, technology, administration. There aren't very many administrators who participate in national conferences, so many of them will be unknown even if they are good. Sometimes shaming policymakers doesn't work-they have their own policy agenda (voter ID) and they don't care how they look or the state's reputation.

Stewart

Election attorneys in each state might be able to rate administrators. Second, how credible will our index be? It seems like indices that are respected come from organization or publication with pre-existing credibility. Coming from group of law professors isn't enough.

McDonald

Virginia wants McDonald to do a pilot type program. Does anybody want to give him the money? One piece of data we haven't talked about is the cost of elections. It's been a very daunting task to collect this information due to decentralization. This is probably an important factor that would have a policy-changing effect. When we get raw data from administrators it is going to be a mess and will take months and months of cleanup. There will be blanks and also zeros-zero provisional ballots cast, etc. Faked data, for instance, number of pollworkers might be faked. There will be a huge data auditing problem.

Tolbert

Two-pronged attack-two states that do pilot project to get all the data we want. Second prong is looking at all 50 states with data that is already available. Can we break registration data to check registration rate for various disadvantaged subgroups.

McDonald

How are you going to get race information and information about eligible but unregistered voters.

MacDonald

Peer review won't work because it will hurt our index's credibility. Also, the people that people look up to aren't always the best. There will also be resistance to do this. This type of work takes special training because there's lots of pitfalls. We are losing a huge opportunity right now because the databases are being designed without the influence of the index. MacDonald tried to get names of fields that were in each database (such as email) but could not get this information because it's proprietary. Two out of four states did not keep email information, which is a big problem.

## Gronke

People don't choose their voting jurisdiction like they do their hospital, so individuals are not the audience. Legislators are the audience. Some of newer appointments at EAC are good and a lot of their failures are due to political pressures that are not their fault. Look at the NVRA report, compare registered voters that came off website with survey response-numbers changed a lot in some states.

## Foley

Five questions: 1. how many eligible citizens can't vote because of registration; 2. how many can't vote because of problems with casting a ballot; 3. what percentage of ballots that should have been counted were not; 4. what percentage of ballots are subject to procedural taints (such as chain of custody violations) where the consequence that they might not be counted; 5. how much local variation exists with respect to the previous four questions. We need to go out and figure out whether this data exists at any level. There seems to be some consensus of what the core measures should be.

## Next steps

### Fortier

We reached consensus as using outputs as core or preference, with some other valuable measures to supplement. One question is, what can we do very quickly? Just to see what it would look like. We also thought maybe CPS could help us get data.

1. some sort of input on how state's reach out.
2. whether states let people check their registration
3. accuracy of the list/or use proxies by looking at data input processes
4. how many people try to register and don't; how many try to vote and can't; how many cast a vote that doesn't count
5. lots of overlap with casting and counting of ballots
  - a. convenience of voting-but not sure how much this should be weighted
  - b. provisional votes-but unclear what the meaning of various statistics might be
  - c. look at recounts as a check to see how accurate the first one was-measures precision as opposed to accuracy
  - d. is there a way to look at procedures in place in advance? Some thought it was too complicated and too variable across states

### Tasks

- find out what data exists
- find ways to improve the data
- do we want to break into working groups

### Caudell-Feagan

- the group that is here represents a lot more resources than just the people at this table
- PEW, Jeht
- indices are useful, but become much more useful if show data over a significant period of time
- let's be patient, we have a lot more clarity we need to develop, and we have to do it right so we're not perceived as partisan. Especially since it has electoral implications.
- better data, collection of data-may want to do ranking on this soon to help with further efforts
- in the next six months Pew is going to do some public opinion research on what would be most useful or desirable to people in an index

McDonald

Turnout and residual vote are already on my website. Also registration.

MacDonald

There's a lot of information about overseas and military voters. Might be able to do a little ranking based on this. How much technology do they have for this to make process easy.

Gronke

You can't compare states on turnout-it's apples and oranges. Otherwise you're penalizing states for being poor.

Tolbert

Sure, just compare them and then have a little discussion explaining how it's not perfect. You could also break information down by race.

McDonald

Gronke and McDonald have info about registration, turnout, mail voting and EDR, etc.

Stewart

Why should we be reluctant to punish states for being poor or disadvantaged. We can mitigate it by having equity measure and measures for improvement over time.

MacDonald

What about VRA compliance and disability access? Is there existing data on this?

Quinn

There's data about VRA suits on pretty much every state.

Foley

Let's amass the data and then decide whether it's useful for the rankings.

Hall

There is a data set on 50-state election codes.

Gronke

Who is going to update that?

Quinn

It doesn't need to be updated because there won't be any backsliding, it will at least be a baseline.

Gronke

No, there is backsliding, like ballot access in Oregon.

Foley

If people submit the information that is available we'll put in on our Democracy Index website.

Tokaji

There is an existing database of VRA s 2 cases maintained by university of Michigan law school (s 2 is nationwide). Second, it's important to measure turnout, but also representativeness in terms of income, education, age. This would help mitigate the Mississippi problem where states are penalized for being poor.

MacDonald

Let's put the data into a common format.

Fortier

How do we get to this core data?

Stein

Figure out what Secretaries of State are collecting data or at least have the technological capacity to collect it. Then set up pilot program in two of these states.

Foley

We have to figure out how our basic questions and values map into specific pieces of data.

Hall

There are international values (such as ACE project) for elections. They evaluate countries based on their ability to achieve these values.

Quinn

ACE is good for what people's practices are, but it doesn't rank or judge.

Chapin

Let's just get some data out there and not try to be perfect. At the same time, let's keep developing the theory and eventually they will come together.

Foley

I'm concerned about representativeness of election results. I think representativeness is important, but for public perception it will hurt us to include this on the index. People will say it's a tilted index by weighting a sub-community.

Gerken

No, when we talk about representativeness, we're not talking about punishing states for not being representative. Instead, we're talking about ways to not punish them for the greater problems that they have with certain communities.

Stewart

Thinks we can rank based on representativeness/race because of American history and history of disenfranchisement. Because it's a protected class, you can justify recording it, unlike the poor.

Quinn

That would only be fair if you compare turnout of minorities in one state with turnout in another. Not fair to compare turnout of minorities against whites.

McDonald

More blacks are registered in Alabama than whites, which is unrepresentative. Do we want to punish Alabama for this?

Quinn

If you look at who is getting sued and losing, it might be a proxy for this?

Foley

We can't have a measurement that is grading the government if the statistics are getting mixed up with the effects of people who don't try. Gronke said turnout was most important thing to measure for him, but Foley said at the beginning this is problematic.

Tolbert

Turnout is an output of the government because of voter ID, etc.

Cain

Project should start with one value, accuracy. Accurately translate preferences into choices. There's two types of errors: Errors made by the state, and errors made by voters. When we see errors made by voters, then we have to start looking at factors like convenience, simplicity of processes, and other things that election officials could actually do to help reduce voter error. Hard to say how far down that road we want to go. If you just try to get some good data on that one category, find out how much is administrative vs. voter error, then worry about the rest later, that will be good.

Tolbert

Utah, Wisconsin, Ohio, Louisiana are states with highest incidents of voters reporting trouble with registration.

Stein

Original EAC survey asked states to report how they reported their voting data.

Huefner

There are some probabilistic things that need to be in the core: Residual vote, because even though it's voter error there are things state can do to reduce error. Poll worker training is another thing (this is an input).

MacDonald

Registrars in California do not collect data on poll worker training. Even to get data just for California it would be a huge uphill battle. There was a battle to get a law passed to make registrars record what precinct absentee ballots came from.

HTML Translation provided by INSO Corporation's Outside In ® HTML Export software.

[Close](#) by Edward Foley

Download	HTML	Title
		<b>Cameron's List of Values [posted at conference]</b> Created: Wed, 10/03/2007 11:30 AM EDT Edited: Wed, 10/03/2007 12:38 PM EDT

Cameron's list:

Accessible/free expression of voter' actual intent/desire

Integrity/accuracy

Innovation/experimentation/reconsideration of  
methods/procedures

Clarity & transparency of processes/procedures

Fair, free, accurate & transparent



Legitimacy; belief in accuracy

Finality

Avoidance of fragility (robustness)

HTML Translation provided by INSO Corporation's Outside In ® HTML Export software.

[Close](#) by Edward Foley

Download	HTML	Title
		<p style="text-align: center;"><b>Bruce Cain's list</b> Created: Wed, 10/03/2007 12:31 PM EDT Edited: Wed, 10/03/2007 12:39 PM EDT</p>

Bruce Cain

In general, Connie's data is model

For each area:

- Registration
- Absentee/mail-in Ballot
- Vote Counted
- Provisional Votes

Collect this data:

- # Successful new reg/#attempts
- # failed reg due to:
  - Lateness
  - Admin. Error
  - Voter error
- # Successful Provisional Votes/# attempts
- # failed
  - Admin error
    - Voter not registered
  - Etc.

HTML Translation provided by INSO Corporation's Outside In ® HTML Export software.

[Close](#) by Edward Foley

Download	HTML	Title
		<b>Doug Chapin's list</b> Created: Wed, 10/03/2007 12:33 PM EDT

Doug Chapin


Number of eligible citizens who wish to register and do successfully

Number of eligible voters who wish to vote and do successfully

Number of eligible votes that are counted successfully

HTML Translation provided by INSO Corporation's Outside In ® HTML Export software.

[Close](#) by Edward Foley

Download	HTML	Title
		<b>Ned Foley's list</b> Created: Wed, 10/03/2007 11:33 AM EDT Edited: Wed, 10/03/2007 12:28 PM EDT

Ned's list for John:

1. What percentage of eligible citizens who tried to register and to cast a ballot that counts were unsuccessful in exercising the franchise because of obstacles encountered in registration procedures?
2. . . . because of obstacles encountered in the process of casting a ballot?
3. What percentage of ballots that should have been counted according to state law were not, and what percentage of counted ballots should not have been according to state law?
4. What percentage of ballots cast in each election on average are subject to procedural taints (such as chain-of-custody violations), with the consequence that their propriety is reasonably subject to question?
5. How much local variation exists with respect to (1), (2), (3) & (4)?
6. What percentage of a state's close elections, defined as margin of victory less that 0.1%, remain unresolved [with the runner-up refusing to concede defeat] by the date of taking office? (a measure of fragility)

HTML Translation provided by INSO Corporation's Outside In ® HTML Export software.


[Close](#) by Edward Foley

Download	HTML	Title
		<b>Heather Gerken's list</b> Created: Wed, 10/03/2007 12:19 PM EDT

- Heather Gerken (all stolen from Thad & Charles)
1. Residual vote on presidential races
  2. Turnout as a fraction of eligible voters
  3. Recount information/audit information
  4. Survey data on
    - a. length of lines,
    - b. trouble with registration/balloting,
    - c. poll worker experience,
    - d. how easy it was to find the polling place
  5. Failure to disclose data inundated by NVRA, UOCAUA, HAVA
  6. Conny's data on why absentee ballots & provisional ballots are discarded
  7. Expert review of laws/code/system

HTML Translation provided by INSO Corporation's Outside In ® HTML Export software.

[Close](#) by Edward Foley

Download	HTML	Title
		<b>Paul Gronke's list</b> Created: Wed, 10/03/2007 12:35 PM EDT

Paul Gronke

1. Measures should not penalize administrators for issues out of their control.
2. To the degree possible, measures should be on a "per capita" (per turnout?) basis.
3. Need to separate quality of democratic procedures from quality of democratic electorate.

Measures

1. Turnout difference = Actual TO - predicted turnout (from good model)
2. Expert grade for administrative procedures (1.1.5) "Fragility" measure

HTML Translation provided by INSO Corporation's Outside In ® HTML Export software.

[Close](#) by Edward Foley



Download	HTML	Title
		<b>Bob Stein's list</b> Created: Wed, 10/03/2007 11:59 AM EDT

<p>Democracy Measure Indices</p> <p>Bob Stein</p> <p>Registration</p> <ol style="list-style-type: none"><li>1. Multiple entry points for registering i.e. web, NVRA</li><li>2. Electronic data entry and optical scanning for registration</li></ol> <p>Casting Ballots</p> <ol style="list-style-type: none"><li>1. Convenience of polling place i.e. access, parking</li><li>2. Voting places per capita weighted by population density</li><li>3. Voting system per voter (at precinct or polling place)</li><li>4. Poll workers per votes cast (at polling place)</li><li>5. EEEE standards for ballot design for DRE and paper ballot</li></ol> <p>Counting Ballots and post election audit</p> <ol style="list-style-type: none"><li>1. Residual Vote</li><li>2. Difference between election night vote and canvassed vote by candidate/race</li></ol>
--

HTML Translation provided by INSO Corporation's Outside In ® HTML Export software.

Close

by Edward Foley

Download	HTML	Title
		<b>Cameron Quinn's list of metrics for Index</b> Created: Wed, 10/03/2007 12:12 PM EDT

Cameron Quinn

On item 3: all eligible votes being counted properly [from list posted during conference]

1. Measure the clarity of each states rules for counting votes (not whether or not one agrees with the rules) using experts' review, and whether the rules are distributed to all polling places on election day.
2. While I'd like to see use of audits and recounts to determine proper counting where there are significant changes, I'm not sure every state will have audits or recounts annually. If the index would be done every 2 years after a federal election cycle, it is more likely that all states will have had at least some state or local recounts or audits during a 2 year time frame. The measure should be whether the results after the recount/audit are within a certain percentage (.5%, 1%?) of the certified state vote.

On items 1 & 2 [from list posted during conference, concerning disenfranchisement registration and ballot casting problems]

3. There needs to be a measure related to ineligible voters, but I'm blank on how to get there with what's currently available. On ineligible registrants, however, one could get the number of pieces of returned mail from a statewide database mailing and calculate a percentage.

Three add'l questions/measure for the index:

1. Does the state's laws on recounts and contests ensure that any recount/contest will ensure a winner in time for the electoral college deadline? If not, is there a process in place to ensure a state's electoral vote?

There will be few states that have addressed this. Frankly, there is no good excuse for a failure to have dealt with this since 2000-and are a measure of those states really trying to fix obvious problems prospectively.

2. Compliance with federal elections laws

Suits by DOJ against states amplified by private suits where legally allowed. Due to VRA reauthorization this data is available. States with higher suits they've lost should get lower rankings and in my informed view of these suites, there is a significant correlation between states that generally do a good job and fewer lost lawsuits.

Perhaps to be fair, given changes over time, these should weigh more heavily on more recent lawsuits.

3.If we truly value uniformity, then I'd also measure the states that elect *local* election officials. Those that do are significantly less likely to have uniform election rules or at least uniform enforcement of state rules.

Close

by Edward Foley

Download	HTML	Title
		<p style="text-align: center;"><b>Caroline Tolbert's list</b> Created: Wed, 10/03/2007 11:55 AM EDT</p>

## Caroline Tolbert

### Measures of a democracy performance index



1. Higher percent registered voters per voter eligible population (VEP) the better. %register/VEP
2. Higher percent of turnout of registered voters or VEP the better
3. More provisional ballots counted better than lower
4. Higher percent of absentee/mail voting better than lower.
5. Lower residual vote rates are better than higher error rates
6. Number of absentee/mail ballots NOT counted for following reasons: received after Election Day, returned by post, no signature on absentee ballot, void ballot, blank ballot, signature doesn't match, no ballot or 2 ballots in envelope, absentee ballot returned by unauthorized person.
7. Number of provisional ballots uncounted for same reason as #6
8. Higher percent of minority (black, Latino, Asian) voters registered to vote as percent of minority population.
9. Higher percent of minority turnout as percent of minority reg. Voters the better.
10. State allows voter to verify registration on the state government website.
11. Longer polling place hours are better than shorter hours.

### Survey Questions

1. How confident are you that your vote was accurately counted? (use 100 point scale)
2. Approximately how long did it take you to commute to your polling place and cast a vote? (Answer in increments of 10 minutes)
3. How much trust and confidence do you have in your state government? (0-100 point scale)
4. What percent of voters were asked to show an identification card on Election Day in ways contrary to state law? (combine/match survey responses with state laws)
5. Overall, how easy was it to vote, once you checked in and were in the voting booth?

HTML Translation provided by INSO Corporation's Outside In ® HTML Export software.

[Close](#) by Edward Foley

Download	HTML	Title
		<p style="text-align: center;"><b>Anonymous List # 1</b> Created: Wed, 10/03/2007 12:14 PM EDT Edited: Wed, 10/03/2007 12:16 PM EDT</p>

<p>Accuracy of VR list (deadwood, data entry error) Coverage of VR list (% of VEP) Residual vote rate (mode expansion) Uniform generally accepted standards     post election audits     training Public opinion research - (personal experience not confidence)</p>
--

HTML Translation provided by INSO Corporation's Outside In ® HTML Export software.

[Close](#) by Edward Foley

Download	HTML	Title
		<b>Anonymous list # 2</b> Created: Wed, 10/03/2007 12:25 PM EDT

Local level "infrastructure" (can be output data!) including number of full time employees per "eligible" voters, \$\$, other "jobs" of LEO (Local Election Official), VRA compliance cost of elections, funding structure

State level "performance"

Structure (i.e. availability of state - \_\_ [illegible word] "certification" program for LEOs; oversight of registration data collection/trouble-shooting of data systems (interactive ones); communication with LEOs on policy guidance and legal interpretation) transparency, voter outreach, meeting sec.203 VRA, disability access, cost of elections and funding structure, registration database


Voter access

Percent of registered voters; performance of registration database; collect data on precinct level: number of voters leaving polling place without voting; evaluation of provisional votes à which ones were just not counted and why (systematic patterns?); training of pollworkers (is every one trained prior to every election?); etc.; availability of second-language poll-workers, number of voters per polling place.

HTML Translation provided by INSO Corporation's Outside In ® HTML Export software.

Close

by Edward Foley

Download	HTML	Title
		<b>Methodology Statement (Michael McDonald)</b> Created: Wed, 10/17/2007 8:24 AM EDT

Attached is the Methodology Statement to accompany the data submitted by Michael McDonald, which has been posted separately on this page.

## Data Dictionary

**NOTE:** I embedded much of the data dictionary information into the accompanying spreadsheet to make it easier to program the database interface.

### Aggregate Turnout Rate Variables

#### Turnout Rates for Voting Eligible Population

These data are calculated by Michael McDonald. The numerator of the turnout rate calculation is what is commonly referred to as the "Vote for Highest Office," which in a presidential election year is the presidential vote and in a midterm election year is the highest vote of Governor or U.S. Senator, or lacking these offices, the combined vote for U.S. House (note that due to Louisiana's unique double-ballot system no statistics are available in 1982 for that state). A preferred numerator, the total number of ballots cast, is yet to be reported by all states and historical coverage is limited.

The denominator is those eligible to vote. This number is calculated by subtracting from an estimate of the voting-age population, estimates of non-citizens and ineligible felons (depending on state law). Note that estimates of permanently disfranchised felons, eligible voters within a state living overseas, and mentally incompetent (depending on state law) are not available and no adjustments are made for these people.

For more information about the methodology, see [http://elections.gmu.edu/voter\\_turnout.htm](http://elections.gmu.edu/voter_turnout.htm).

#### Percent Ineligible of the Voting-Age Population

The percentage of ineligible voters among the voting-age population demonstrates the growth in the ineligible population in recent elections. The percentage is calculated as the estimates of ineligible felons (depending on state law) and non-citizens as a percentage of the voting-age population estimate. Note that it does not include estimates of permanently disfranchised felons (depending on state law) and the mentally incompetent (depending on state law), and it does not add eligible overseas citizens to the voting-age population estimate since there are no reliable statistics for these populations.

#### Current Population Survey Variables

The Current Population Survey is a large-scale monthly survey conducted by the Census Bureau, primarily used to generate important labor statistics, such as the unemployment rate. In a November of a federal election year, a limited number of questions on voting and registration are added to the survey. The Current Population Survey's large sample size makes it one of the best sources of voting and registration data for states and among demographic subgroups.

All variables drawn from the Current Population are calculated using the base of *citizen-voting-age population* which is citizens age 18 and older. These calculations contrast with [Census Bureau reports](#) that for the exception of the most recent years calculate turnout rates for the *voting-age population* only. The distinction is important not simply because non-citizens cannot vote in federal elections, but also because the non-citizen population has grown considerably over time and their concentration varies among the states. The citizen-voting-age population is not the most accurate measure of the eligible population (see: **Turnout Rates for Voting Eligible Population**), but more precise measures of the eligible population cannot be obtained from the Current Population Survey (see: *Over-Report Bias* discussion for more information about the Current Population Sample).

Many of the Current Population Survey questions have changed over time, which can lead to subtle question wording error in addition to the more familiar sample error associated with surveys. All state level percentages are reported along with the standard error, a measure of the uncertainty as to where the true population percentage lies (see: **Note on Survey Error**).

### Citizenship

Prior to 1994, the Current Population Survey citizenship question was a simple "yes" or "no" question. In 1994, the Current Population Survey question on United States citizenship was changed to five categories that probed the method of naturalization for persons not born in the United States:

1. Native, Born in the United States
2. Native, Born in Puerto Rico or U.S. Outlying Area
3. Native, Born Abroad of American Parent or Parents
4. Foreign Born, U.S. Citizen by Naturalization
5. Foreign Born, Not a Citizen of the United States

Persons of voting-age who respond pre-1994 that they are not a citizen and post-1994 that they are "Foreign Born, Not a Citizen of the United States" are excluded from all turnout rate calculations.

### Voting

Following Census Bureau conventions, respondents who report that they voted are considered as voters. Non-voters are respondents who reported they "did not vote," "do not know," and "not reported."

### Registration

Following Census Bureau conventions, all respondents reporting voting are classified as registered to vote and respondents who report being registered if they "did not vote," "did not know," or "not reported." Non-registrants are persons who reported they "did not vote," "do not know," and "not reported" and who said they "did not register to vote," "do not know," or "not reported."

Note that North Dakota does not have voter registration, even though many respondents from North Dakota reported not being registered to vote.

### Age

We classify respondents into four age categories:

1. Age 18-29

2. Age 30-44
3. Age 45-59
4. Age 60+

### Gender

We classify respondents in response to their sex as being either "Male" or "Female."

### Race

The Current Population Survey has asked respondents' race by various questions over the years. We generally classify respondents into four categories:

1. Non-Hispanic White
2. Non-Hispanic Black
3. Hispanic
4. Non-Hispanic Other

For those unfamiliar with Census Bureau race classifications, it may seem odd that we separate respondents by their race and Hispanic origin. This is because the Census Bureau asks two separate questions, one for what people may generally think of as race and one for Hispanic ethnicity. Our race classifications therefore represent the collapse of these two questions into one mutually exclusive variable.

Prior to 1994, the Current Population Survey race question was limited to three categories, *White*, *Black*, (pre-1982, *Negro*) and *Other*. In 1994, the Current Population Survey race question was increased to five categories by adding response categories for *American Indian, Aleut, or Eskimo* and *Asian or Pacific Islander* to the existing categories. In 1996 it was reduced from five to four categories by dropping the *Other* category. In 2002, the categories of *Asian* and *Pacific Islander* became two separate categories and valid responses were expanded to permit persons to identify themselves with more than one race.

The Census Bureau asks respondents a separate question regarding their Hispanic origin, of which prior to 1994 several countries of origin are permitted as responses. In 1994, the question was changed to include a screening question that asked if a person identified themselves as Hispanic.

Prior to 2002, we classify *Non-Hispanic Whites* as respondents who report their race as *White* and do not report a Hispanic origin. We classify *Non-Hispanic Blacks* as respondents who report their race as *Black* and do not report a Hispanic origin. We classify *Hispanics* as respondents of any race reporting a Hispanic origin. While we would prefer to classify the remaining races with greater detail, the small survey population sizes within states would make the statistical error so large as to be meaningless, so we opt to classify *Non-Hispanic Other* as respondents providing any of the various other race categories and report no Hispanic origin.

For 2002 and after, where the Census Bureau permits multi-race responses, we classify *Non-Hispanic Whites* as respondents who reported *White* alone without any other combination and report no Hispanic origin. We classify *Non-Hispanic Black* as respondents who report *Black* alone and in any combination with other races and report no Hispanic origin. We classify *Hispanics* as respondents of any race reporting a Hispanic origin. All other respondents are classified as *Non-Hispanic Other*.

### Education

We classify respondents into four categories of educational attainment:

1. Less than a high school diploma
2. Received a high school diploma
3. Some college or a four-year college degree
4. Any post-year year degree schooling

Prior to 1992, the Current Population Survey measured education through two questions that asked highest grade a respondent attended and completed. In 1992, these two questions were combined into one question.

### Length of Residence

We classify persons by their length of residence whether or not they lived less than one year or more than one year at their current address. Respondents who reported living "less than one month," "1-6 months," or "7-11 months" were classified as living at their address for less than one year. Respondents reporting that they lived at their current address for "1-2 years," "3-4 years," or "5 years or longer" were classified as living at their address for more than one year. Respondents who "did not know," "refused," or gave "no response" are excluded from these two categories.

Length of residence was asked in the 1980 and the 1990-2006 Current Population Surveys.

### Early Voting

We classify persons as early voters if they reported casting a ballot before election day or if they voted absentee on the Current Population Survey. Unfortunately, an early vote or absentee vote question was not asked on the 1982, 1986, 1988, and 1990 Current Population Surveys. The question has also changed over time., From 1984, 1992, 1994, early voters are persons reporting their time of day voting as "voted absentee." From 1996-2002, early voters are persons reporting that they voted "in person before election day" or "voted by mail (absentee)." In 2004, early voters are persons that reported voting "before election day."

### Type of Early Voting

State early voting and absentee balloting laws are graciously provided by the Early Voting Information Center, <http://www.earlyvoting.net/>. States are classified into six categories:

- 1 = Traditional Absentee Balloting
- 2 = No Excuse Absentee Balloting
- 3 = No Excuse Absentee Balloting with Permanent Absentee Status
- 4 = Early In-Person Voting with Traditional Absentee Balloting
- 5 = Early In-Person Voting with No Excuse Absentee Balloting
- 6 = All Vote-by-Mail

**Note on Survey Error.** Caution is warranted when interpreting survey data as statistical and non-statistical methodology errors may affect results. A short description of some of the possible errors is presented here.

**Random Statistical Error.** Random statistical error arises from sampling methodology that randomly selects respondents to be surveyed from a population under study. This is the familiar margin of error often reported along with poll results. The percentages we report are our best guesses at the unknown true population percentages we would obtain if we accurately surveyed everyone. Our confidence that we observe a percentage close to the true population percentage is represented by confidence intervals, which

are related to the number of respondents with a population category. For small surveyed populations, confidence intervals can be quite large, reflecting that we are more uncertain the smaller the surveyed population.

Consistent with Census Bureau practices, we present 90% confidence intervals, which tell us that we expect the true population percentage is within the given range 90% of the time. This also means that 10% of the time the true population percentage is outside of the confidence interval. Care should be taken in interpreting any one number since 1 in 10 times the true population percentage will *not* fall in the confidence interval. When properly interpreting survey percentages, it is therefore wise to take into account the size of the confidence intervals, as smaller confidence intervals are related to more precise estimates, and to place a given percentage within the context of overall trends. A single number that deviates from a trend may simply be a consequence of randomly choosing a sample not reflective of the population.

Confidence intervals are calculated from standard errors, which are provided in the accompanying spreadsheets for each Current Population Survey percentage. To obtain the 90% confidence interval for a given percentage, add and subtract 1.8x times the standard error to the percentage.

**Non-Statistical Error.** Non-statistical survey error arises from many sources: the way in which questions are asked, the question wording, the personal interaction between the interviewer and the respondent, the way by which the survey is administered, and how the responses are weighted, among other factors.

*Over-Report Bias.* Another well-known source of election survey error is what is known as *over-report bias*, which is the tendency of surveys to report higher turnout rates than are observed using actual election returns. Over-report bias is observed by comparison of the national aggregate voter turnout rates and the Current Population Survey Turnout Rates.

[Insert Figure]

There is no consensus among public opinion researchers as to the primary source of over-report bias.

A natural conclusion is that some respondents are not reporting accurately whether by incorrectly providing the socially correct answer of being a good voting citizen to make themselves look good in the eyes of the interviewer or simply not remembering if they voted and providing the socially correct answer when asked. Interestingly, evidence suggests that people who fit the profile of a high propensity voter are more likely to feel pressured to give the socially desirable response.

While the foregoing is a compelling explanation, it is also possible that the sample universe is not reflective of the population being studied. One factor could be that persons who participate in a survey are more participatory in general and thus have a higher turnout rate. The Census Bureau spends considerable effort to reduce such error in the Current Population Survey since its primary use is to produce important labor force statistics, such as the unemployment rate. Reflective of this, the Current Population Survey has one of the lowest over-report biases among election surveys. However, the Current Population Survey sample is not reflective of the entire United States population as it is drawn from the residential (including military) non-institutional population, which for example does not include persons living in dorms, prisons, and high level nursing home facilities. While these differences are small, they may be sizable enough to produce the observed over-report bias on the Current Population Survey.