

# 11

## Pippa Norris

Harvard University and University of Sydney  
www.electoralintegrityproject.com

For the last five years, the [Electoral Integrity Project](#), an independent research project based at Harvard and Sydney Universities, has conducted the Perceptions of Electoral Integrity global study. Following the November 8<sup>th</sup> 2016 elections, this method was applied to compare the electoral integrity of 50 U.S. States plus DC (PEI-US-2016), with the survey gathering responses from over 700 political scientists.

The results were published by the EIP team in two blog reports comparing electoral integrity across US states, with a longer piece published on 24 December in [Vox](#) and a shorter piece published on 27 December in the [Monkey Cage](#). The results were also highlighted in commentary by Andrew Reynolds published on 22<sup>nd</sup> December in *News and Observer* "[North Carolina is no longer classified as a democracy](#)." The release went viral in the media, spawning dozens of articles.

There has also been considerable interest among scholars, for example, since becoming available on 15<sup>th</sup> December, the PEI-US-2016 datasets have been downloaded via Dataverse over 1,100 times, which must be something of a record.

The PEI study raises some important general issues about the use and construction of expert indices, as well as the specific methods employed by the Electoral Integrity project. Andrew Gelman highlights several questions about the methods which are used. This note provides a brief response to both issues.

## Constructing expert indices

Indices and datasets derived from expert surveys have become increasingly common in comparative social science, in risk analysis by private sector organizations, in evaluation research, and among NGOs and policy makers (Meyer & Booker 1991; Cooley and Snyder 2015). This data collection technique has been applied to diverse research topics such as the series of studies on party and policy positioning (Laver and Hunt 1992; Huber and Inglehart 1995; Saiegh 2009; Laver, Benoit, and Sauger 2006; McElroy and Benoit, 2007), the power of prime ministers (O'Malley 2007), evaluations of electoral systems (Bowler, Farrell, and Pettitt 2005); policy constraints horizons (Warwick 2005); campaign communications (Lileker, Steta and Tencher 2015); human rights and democracy (Landman and Carvalho 2010), and the quality of public administration (Teorell, Dahlstrom and Dahlberg 2011). Expert surveys have been widely used in research on corruption - the Corruption Perceptions Index (Transparency International 2013; Global Integrity); measuring democracy since the 1900s -Varieties of Democracy (Coppedge et al. 2011)- and electoral integrity (Norris, 2014; Norris, 2015; Martinez I Coma and Van Ham, 2015). The World Bank Institute Good Governance indicators combine an extensive range of expert perceptual surveys drawn from the public and private sectors. Indeed, among the mainstream indicators of democracy, Freedom House's estimates of political rights and civil liberties, Polity IV's classification of autocracies and democracies, and the Economist Intelligence Unit's estimates of democracy are all, in different ways, dependent upon expert judgments.

Expert surveys seem especially useful for measuring complex concepts that require expert knowledge and evaluative judgments; and for measuring phenomena for which alternative sources of information are scarce (Schedler 2012). Yet, expert surveys are far from risk free and several scholars have pointed out their limitations (Budge, 2000; Mair 2001; Steenbergen and Marks, 2007). Moreover, in contrast to mass social surveys, a common methodology to construct such surveys is lacking, as well as agreed technical standards and codes of good practice. The most heated debate has focused on the pros and cons of methods used to evaluate the spatial positions of party policies, and about the use of governance indicators more generally.

Nevertheless, by contrast there has been remarkably little discussion about the challenges of validity, reliability, and legitimacy facing the construction of expert perceptual surveys. Yet it is critical to consider these issues given the lack of a clear conceptualization and sampling universe of ‘experts’, contrasting selection procedures and reliance upon domestic and international experts, variations in the number of respondents and publication of confidence intervals, and lack of consistent standards in levels of transparency and the provision of technical information.

Moreover, more research needs to be done on how to evaluate the consequences of expert and context heterogeneity on the validity of expert judgments (Martinez i Coma and van Ham 2015), for example by using item response models to test and correct for expert heterogeneity (Pemstein *et al.* 2015), and using techniques such as ‘anchoring vignettes’ (King & Wand 2007) or ‘bridge coders’ (V-Dem) to test and correct for context heterogeneity.

## The Electoral Integrity Project

With these general points in mind, and to address the issues raised by Gelman more directly, what approach and techniques are used by the Electoral Integrity Project when constructing the Perceptions of Electoral Integrity index?

To start to gather new evidence, on 1<sup>st</sup> July 2012 the project launched the expert survey of Perceptions of Electoral Integrity. The design was developed in consultation with Professor Jorgen Elklit (Aarhus University) and Professor Andrew Reynolds (University of North Carolina, Chapel Hill).

## Global Coverage:

The PEI survey of electoral integrity focuses upon independent nation-states around the world which have held direct (popular) elections for the national parliament or presidential elections. The criteria for inclusion are listed below. The elections analyzed in the most recent release (PEI-4.5) cover the period from 1 July 2012 to 30 June 2016. In total, PEI 4.5 covers 213 elections in 153 nations.<sup>1</sup> The next release (PEI-5.0), expanding coverage to elections held during the last 6 months of 2016, will be in March 2017.

Criteria for inclusion in the survey	#	Definition and source
Total number of independent nation-states	194	Membership of the United Nations (plus Taiwan)
<b>Excluded categories</b>		
Micro-states	12	Population less than 100,000 in 2013, including Andorra, Antigua and Barbuda, Dominica, Liechtenstein, Marshall Islands, Monaco, Nauru, Palau, Saint Kitts and Nevis, San Marino, Seychelles, and Tuvalu.
Without de jure direct (popular) elections for the lower house of the national legislature	5	Brunei Darussalam, China, Qatar, UAE, and Saudi Arabia
State has constitutional provisions for direct (popular)	3	Eritrea, Somalia, and South Sudan

elections for the lower house of the national legislature, but none have been held since independence or within the last 30 years ( <i>de facto</i> )		
Sub-total of nation-states included in the survey	174	
Covered to date in the PEI 4.5 dataset (from mid-2012 to mid-2016)	153	87% of all the subtotal of nation-states

Because of the selection rules, elections contained in each cumulative release of the PEI survey can be treated as a representative cross-section of *all* national presidential and legislative elections around the world (with the exception of the exclusion of micro-states). The countries in PEI 4.5 are broadly similar in political and socio-economic characteristics to those countries holding national elections which are not yet covered in the survey, although being slightly larger in population size.

More recently the EIP project has also collaborated with local teams of scholars and conducted several sub-national surveys using similar methods but at the level of provinces, states or other sub-national units, including India, the US, Mexico, the UK, and Russia. The PEI uses the identical core 49 items across all sub-national studies, to maintain comparability, but also supplements the core with specific items most relevant to each particular context, such as violence and crime in Mexico. The PEI has now been conducted three times in the US, in 2012 (at national level), in 2014 (covering 20 states) and 2016 (covering all states). When merged, this will allow comparison over time as well as across states.

#### Respondents:

For each country or state, the project identifies around forty election experts, defined as a political scientist (or other social scientist in a related discipline) who had demonstrated knowledge of the electoral process in a particular country (such as through publications, membership of a relevant research group or network, or university employment). It should be noted that this is far more than is conventionally used in comparable expert-based surveys, like V-Dem. For the global PEI, the selection has sought a roughly 50:50 balance between international and domestic experts, the latter defined by location or citizenship. Experts are asked to complete an online survey. In total, 2,417 completed responses were received in the PEI-4.5 survey, representing just under one third of the experts that the project contacted (29%). For the PEI-US-2016, the survey received over 700 responses.

It should also be noted that PEI contacts experts one month after each election, when judgments are likely to be stable and memories fresh. By contrast, other expert-based surveys ask respondents for judgments far longer from the event, for example V-Dem asks their experts about electoral integrity in each country *for every year since 1900*, which we believe is not possible to assess with any degree of accuracy.

#### Concepts:

The idea of electoral integrity is defined by the project to refer to agreed international conventions and global norms, applying universally to all countries worldwide through the election cycle, including during the pre-election period, the campaign, on polling day, and its aftermath.<sup>2</sup>

What needs to be emphasized is that this new concept is far from equivalent to standard notions of liberal democracy. It remains difficult for scholars to break out of the familiar way of classifying regimes but instead the notion of electoral integrity is derived from international conventions and standards based on human rights. This provides a less tight theoretical concept but one which is both legitimate and authoritative for international programs of electoral assistance.

#### Measurement:

To measure this concept, the PEI survey questionnaire includes [49 items on electoral integrity](#) ranging over the whole electoral cycle. These items fall into eleven sequential sub-dimensions.

Most media attention in detecting fraud focuses upon the final stages of the voting process, such as the role of observers in preventing ballot-stuffing, vote-rigging and manipulated results. Drawing upon the notion of a 'menu of manipulation',<sup>3</sup> however, the concept of an electoral cycle suggests that failure in even one step in the sequence, or one link in the chain, can undermine electoral integrity.

Unlike many other summary indices, the results of the PEI survey can be broken down in far more granular detail to pinpoint specific weaknesses and strengths in each contest. For example, the data can be used to compare how elections rate across eleven stages of the electoral cycle, and across 49 indicators, such as in the processes of district gerrymandering, the opportunities that contests provide for women and minority candidates, the provision of equitable access to political finance, the fairness of electoral officials, and the occurrence of peaceful and violent protests after the announcement of the results, and so on. This is essential for the correct diagnosis of any problems – and thus identifying the appropriate reforms needed to strengthen integrity.

The electoral integrity items in the survey were recoded, where a higher score consistently represents a more positive evaluation. Missing data was estimated based on multiple imputation of chained equations in groups composing of the eleven sub-dimensions. The Perceptions of Electoral Integrity (PEI) Index is then an additive function of the 49 imputed variables, standardized to 100-points. Sub-indices of the eleven sub-dimensions in the electoral cycle are summations of the imputed individual variables.<sup>4</sup>

It could be suggested that the items should be weighted, for example by whether constitutional provisions or laws limit party competition. Nevertheless, legal restrictions are only one dimension of the procedures used to exclude or narrow party competition; in most electoral autocracies, today, multiple parties exist but there is no level playing field. Ruling parties limit opportunities for opposition forces through multiple mechanisms, whether blatant gerrymandering, intimidation and repression, patronage largess and corruption, or control over state media and public resources. The problems in Ethiopia, for example, differ sharply from those in Syria, Belarus, Haiti and Burundi, all countries with elections rated at the bottom by experts. Since different mechanisms are used in different states, each of these needs to be evaluated, and any single 'break in the chain' undermines integrity. Moreover, analyzing electoral integrity even in countries where there is no constitutional or legal right to organize political parties, or whether there remains very limited party competition, also provides an important benchmark to evaluate future developments in subsequent contests.

## Confidence intervals

When interpreting the results, it should be noted that modest differences in the PEI index are unlikely to be statistically significant at reasonable confidence intervals. It is more useful to focus on the range of indicators across the cycle and more substantial differences among elections or among countries. Confidence intervals are constructed at the 95 per cent interval for the summary PEI index, based on the number of experts who responded for each election and country.

## Validity and reliability tests:

The results have been tested for external validity (from sources of independent evidence), internal validity (consistency within the group of experts), and legitimacy (how far the results can be regarded as authoritative by stakeholders). The analysis, presented elsewhere, demonstrates substantial external validity when the PEI data is compared with many other expert datasets, as well as internal validity across the experts within the survey, and legitimacy as measured by levels of congruence between mass and expert opinions within each country.<sup>5</sup>

For *external* validity tests, the PEI Index was significantly correlated with other standard independent indicators contained in the 2016 version of the Quality of Government cross-national dataset. This includes the combined Freedom House/imputed Polity measure of democratization ( $R=.762^{**}$ ,  $N. 151$ ), and the Varieties of Democracy measure of electoral democracy (polyarchy) ( $R=.824^{**}$ ,  $N.140$ ).<sup>6</sup>

For internal validity purposes, several tests have been run with each release using OLS regression models to predict whether the PEI index varied significantly by several social and demographic characteristics of the experts, including sex, age, education, domestic and international institutional location, and familiarity with the election. In accordance with the findings from the previous versions, domestic experts and those reporting a higher level of familiarity with the election were significantly more positive in their evaluations, but other social characteristics were not significant predictors of evaluations.

The PEI-4 Codebook provides detailed description of all variables and imputation procedures. A copy can be downloaded from the project website [www.electoralintegrityproject.com](http://www.electoralintegrityproject.com)

## Datasets

The main PEI datasets are released on a bi-annual basis, as soon as they have been cleaned, so that they are available for secondary analysis by the community of users. The files are made available at country, election and expert levels along with the codebooks through the EIP Dataverse. They have been widely downloaded, for example the PEI-US-2016 attracted over 1,100 downloads in two weeks. A detailed report is also published bi-annually in an accessible format for practitioners and journalists. This transparency is important to allowing multiple tests beyond the capacity of the research team.

For data, go to:

<https://dataverse.harvard.edu/dataverse/>

## Conclusions

In short, the project has made considerable progress in developing the PEI methodology over the last five years and we are confident about the results. Nevertheless, there is always room for improvement, and, in particular, learning from comparisons across similar projects is very helpful to create a community or network. To this end, we organized the [workshops and panels](#) last year with V-Dem at IPSA in Posnan and APSA in Philly, bringing together representatives from major organizations generating political indices, including Freedom House, Polity IV, the Bertelsmann Institute, International IDEA, UNDP, the Manifesto Project and CHES. This dialogue can only benefit the process of generating reliable and valid indices, identifying best practices, as well as making the methodology more transparent.

This is only one part of the EIP project and we generally adopt mixed methods where we employ both elite and mass survey (WVS) data, as well as selected case studies, and qualitative methods.

The interest in PEI- US also demonstrates the need for scholars to think carefully about how social scientists can contribute evidence which is useful in the public debate about how to identify problems in electoral integrity and, then, what solutions might be most appropriate to overcome these. We seem to be heading into a fact-free zone where partisans assert that the world is flat but social science can still serve an important function in speaking truth to power, generating evidence of poor (and good) performance, and contributing towards the public sphere.

## References

- Benoit, Ken and Michael Laver. 2005. *Party Policy in Modern Democracies*. London: Routledge.
- Bowler, Shaun; David Farrel and Robin Pettitt. 2005. 'Expert opinion on electoral systems: So which electoral system is best?' *Journal of Elections, Public Opinion and Parties* 15(1): 3-19.
- Budge, Ian. (2000). 'Expert judgments of party policy positions: Uses and limitations in political research.' *European Journal of Political Research* 37(1): 103–113.
- Transparency International. 2013. *Corruption Perception Index*.  
[http://www.transparency.org/whatwedo/publication/cpi\\_2013](http://www.transparency.org/whatwedo/publication/cpi_2013)
- Huber, John and Inglehart, Ronald. 1995. 'Expert Interpretations of Party Space and Party Locations in 42 Societies', *Party Politics* 1:73-111.
- King, G. & Wand, J. (2007). Comparing incomparable survey responses: Evaluating and selecting anchoring vignettes. *Political Analysis* 15(1): 46–66.
- Landman, Todd and Edzia Carvalho. 2010. *Measuring Human Rights*. London: Routledge.
- Laver, Michael and Ben Hunt, B. (1992). *Party and Policy Competition*. London: Routledge.
- Laver, Michael, Kenneth Benoit, and Nicolas Sauger. 2006. 'Policy Competition in the 2002 French Legislative and Presidential Elections.' *European Journal of Political Research* 45: 667-697.
- Lilleker, Darren., Stetka, V. and Tenscher, J., 2015. Towards hypermedia campaigning? Perceptions of new media's importance for campaigning by party strategists in comparative perspective. *Information, Communication and Society*, 18 (7), 747-765.
- Mair, Peter. (2001). 'Searching for the position of political actors: A review of approaches and a critical evaluation of expert surveys.' In M. Laver (ed.), *Estimating the policy positions of political actors*. London: Routledge.
- Martinez i Coma, Ferran; Van Ham, Carolien . 2015. 'Can experts judge elections? Testing the validity of expert judgments for measuring election integrity.' *European Journal of Political Research* 54 (2): 305-325.
- McElroy, Gail and Kenneth Benoit. 2007. 'Party groups and policy positions in the European Parliament.' *Party Politics* 13:5-28. Meyer, M. & Booker, J. (1991). *Eliciting and analyzing expert judgment: A practical guide*. London: Academic Press.
- Norris, Pippa. 2014. *Why Electoral Integrity Matters*. NY: CUP.
- Norris, Pippa. 2015. *Why Elections Fail*. NY: CUP.
- Norris, Pippa. 2017. *Strengthening Electoral Integrity*. NY: CUP.
- Norris, Pippa, Richard W. Frank and Ferran Martinez i Coma. 2014. Eds. *Advancing Electoral Integrity*. New York: Oxford University Press.
- Norris, Pippa, Richard W. Frank and Ferran Martinez i Coma. 2015. Eds. *Contentious Elections*. NY: Routledge.
- Norris, Pippa and Andrea Abel van Es. 2016. *Checkbook Elections?* NY: OUP.
- Norris, Pippa, and Alessandro Nai. Eds. 2017. *Election Watchdogs*. NY: OUP.
- O'Malley, Eoin. 2007. The Power of Prime Ministers: Results of an Expert Survey' *International Political Science Review* 28(1):7-27.
- Pemstein, D.; Tzelgov, E.; Wang, Y. 2015. "Evaluating and Improving Item Response Theory Models for Cross-National Expert Surveys". *Varieties of Democracy Institute: Working Paper Series No. 1*.
- Saiegh, Sebastian. 2009. 'Recovering a basic space from elite surveys: Evidence from Latin America.' *Legislative Studies Quarterly* 34(1):117-145.
- Schedler, Andreas. 2012. 'Judgment and Measurement in Political Science' *Perspectives on Politics* 10(1):21-36.
- Steenbergen, Marco R. and Gary Marks. 2007. 'Evaluating expert judgments.' *European Journal of Political Research* 46: 347–366.
- Teorell, Jan. Carl Dahlström & Stefan Dahlberg. 2011. *The QoG Expert Survey Dataset*. University of Gothenburg: The Quality of Government Institute. <http://www.qog.pol.gu.se>

Warwick, Paul. 2005. 'Do Policy Horizons Structure the Formation of Parliamentary Governments?: The Evidence from an Expert Survey' *American Journal of Political Science* 49(2):373-387.

- 
- <sup>1</sup> In addition, in 2014 elections in Haiti, Lebanon, and Comoros were delayed or suspended. Those are thus not included in the dataset. The election in Thailand was held and later annulled. There were also elections in North Korea and Trinidad and Tobago, but with too few responses in these two cases meant that these are excluded from the dataset.
- <sup>2</sup> Pippa Norris. 2013. 'The new research agenda studying electoral integrity.' Special issue of *Electoral Studies* 32(4).
- <sup>3</sup> Andreas Schedler. 2002. 'The menu of manipulation.' *Journal of Democracy* 13(2): 36-50.
- <sup>4</sup> See the codebook for further information. [www.electoralintegrityproject.com](http://www.electoralintegrityproject.com)
- <sup>5</sup> Pippa Norris, Ferran Martínez i Coma and Richard Frank. 2013. 'Assessing the quality of elections.' *Journal of Democracy*. 24(4): 124-135; Pippa Norris, Richard W. Frank and Ferran Martínez i Coma. 2014. Eds. *Advancing Electoral Integrity*. New York: Oxford University Press; Ferran Martínez i Coma and Carolien Van Ham. 2015. 'Can Experts Judge Elections? Testing the Validity of Expert Judgments for Measuring Election Integrity'. *European Journal of Political Research* doi:10.1111/1475-6765.12084; Pippa Norris, Richard W. Frank and Ferran Martínez i Coma. 2014. 'Measuring Electoral Integrity around the World: A New Dataset' *PS: Political Science & Politics*, 47(4): 789-798.
- <sup>6</sup> Jan Teorell, Stefan Dahlberg, Sören Holmberg, Bo Rothstein, Felix Hartmann and Richard Svensson. January 2016. *The Quality of Government Standard Dataset*, version Jan16. University of Gothenburg: The Quality of Government Institute, <http://www.qog.pol.gu.se>.