



Standards and Scoring to Increase Transparency for Archived Public Opinion Data

Lynda Kellam, UNC Greensboro

presenting on behalf of

Kathleen Weldon, Roper Center for Public Opinion Research

https://openconf.org/IASSIST2019/modules/request.php?module=oc_program&action=summary.php&id=20





The Problem

Roper Center has used ***methodological and disclosure criteria*** to determine acceptability of acquisitions.

For methodology, guiding principle has been “the best of its time.” Over recent decades, that has meant ***probability-based polls*** and, if conducted by telephone, ***live interviewers***.

(exceptions: historical materials; list samples for special populations like high school principals, opinion leaders, or economists, usually when collected in conjunction with general population polls on the same topic).

Disclosure requirements have been aligned with AAPOR minimum disclosure requirements.

For decades this worked well. The best and most influential polls met both requirements: all major media polls, major non-profit survey research organizations, major academic polling organizations.

But things have been changing.....



Methodology: A Changing Environment



Rising surveys costs and ***falling response rates*** for telephone polls have increasingly led major institutions, like those above, to move to non-probability surveys for some or all of their polling.



Disclosure: Also a Changing Environment

In response to ***increasing heterogeneity in polling methods***, and in line with ***evolving expectations for data sharing and methodological disclosure*** in the social science, the polling industry moved to ***higher transparency standards***.



2006

Adopted new three-level Principles of Disclosure



2010

Launched Transparency Initiative, a membership-based program of organizations willing to commit to high levels of transparency



New emphasis: Peer-to-peer information sharing

AAPOR and NCPP had had disclosure standards in place for decades.

The focus of these standards: core information about polling that journalists, politicians, and the public need to know.

Literature on standards clearly reveals this intended audience.

Disclosure was supposed to help sort out “honest pollsters” from “bad actors”, increase public trust in polls – and circumvent the threat of government regulation

New standards broke out immediately available information (for public) from information to be made available on request (for other researchers).

New standards included ***information that would only be meaningful to sophisticated users.***

With these changes to the polling industry's core methodologies and disclosure standards, Roper Center's Board of Directors requested the Acquisitions and Transparency Committee review the possibility of changes to the acquisition policy of the Center.

In June of 2018, the Committee recommended a set of changes, which the Board approved:

- All methodologies would be accepted into a Recently Developed Methods collection
- This collection would be displayed as a distinct, separate collection from the Longstanding Methods collection.
- All Newly Developed Methods surveys would require a dataset for inclusion.
- A more robust set of Core disclosure elements, delineated in the Committee memo and aligned with new AAPOR standards, would be required for this collection and encouraged for Longstanding Methods.
- These Core disclosure elements, along with some Additional items, will be used to drive a transparency scoring mechanism on the website.
- The transparency score will be made visible on each study.
- A "face validity" test would also influence acquisitions, allowing staff to reject polls that could not make reasonable claims of representativeness.

The Charge to Staff: Work with the Committee to Figure out the Details

Scope of Scoring: How much of the collection will be affected? Historical material? Only current? What's the cutoff for "current"?

Disclosure elements: What information will be considered "qualifying" to score for each disclosure element?

Scoring: Will it be numeric/percentage/grade? Use language to describe levels? How will it be calculated?

Metadata: How do we capture the information required by the Committee?

Display: What should the scoring page look like? Should the score show on the results page, or only on the transparency page? What about the new acquisitions collection?

Once these issues have been addressed, staff will have to share the information with current data providers and get buy-in.

The Details: Scope, Defining Qualifying Responses for Disclosure, Scoring

Scope: For Longstanding Methods, transparency scoring will cover new surveys with field work starting after 1/1/2019. (Scoring mandatory for all Newly Developed Methods polls.)

Disclosure elements: After a review of polls from every current data provider, a number of questions about qualifying answers had to be addressed, including lack of complete question wording for omnibus polls, weighting benchmark sources that simply referenced “Census figures”, in-household respondent selection that was described simply as “random,” etc.

Additional questions are expected as we continue to score polls, particularly in the Recently Developed Methods collection.

Scoring:

Transparency Project scores are calculated as follows: $(10 \text{ points for providing a dataset} + 2 \text{ points for every other applicable core item} + 1 \text{ point for every applicable additional item}) / (\text{total possible points}) \times 10$ (Results rounded to .5). Sc

The resulting numeric score would be described with three possible categories:

9-10 Greatly Exceeds Requirements

8-8.5 Exceeds Requirements

6-7.5 Meets Requirements

Anything below 6 would not meet minimum requirements.

The Details: Metadata: Translating the Charge into Individual Elements, I

| Field | Acquisition Committee Memo Item | Field type |
|--|---|-----------------------|
| Survey sponsor | Survey sponsor, including all funding sources | Open text |
| Grant funding source | Survey sponsor, including all funding sources | Open text |
| Survey organization | Field work provider, if outsourced | Open text |
| Data collection dates | Interview dates | Date |
| Universe | The population of which the results are said to be representative, and the justification for this research claim, AND The universe from which the sample was drawn, and the proportion of that universe that had a nonzero chance of participation | Open text |
| Geographic coverage | | List |
| Justification for claims of representativeness | The population of which the results are said to be representative, and the justification for this research claim | Open text |
| Mode | Mode: RDD telephone, IVR; listed-sample telephone with live interviewers; listed-sample telephone via IVR; other telephone (describe); opt-in online panel; other online (e.g., river samples, mobile apps; hybrid or other (describe)) | List |
| Mode other: Description (filtered on previous) | | Open text |
| Sample size | Unweighted sample size | Numerical |
| Sampling procedure: Summary | Sampling method: Probability, non-probability or hybrid AND Mode: RDD telephone, IVR; listed-sample telephone with live interviewers; listed-sample telephone via IVR; other telephone (describe); opt-in online panel; other online (e.g., river samples, mobile apps; hybrid or other (describe)) | List |
| Sampling procedure: Respondent selection stage | Respondent selection procedure, or absence thereof | Controlled vocabulary |
| Sampling frame | Sample frame and a description of the universe from which the sample was drawn | Open text |
| Weight variable | Description of all sample weights and sources of weighting targets | Open text |
| Weighting benchmark source | | Open text |
| Variables used for weighting | | Open text |

The Details: Metadata: Translating the Charge into Individual Elements, I

| Field | Acquisition Committee Memo Item | Field type |
|--|--|------------|
| Response rate OR | Response rate calculated to AAPOR standards, or sample disposition data adequate for the calculation of AAPOR-standard response rates. When AAPOR-standard response rates cannot be calculated, completion or participation rates shall be provided using another method that is fully disclosed | Numerical |
| Disposition codes OR | | Checkbox |
| Completion or participation rate | | Numerical |
| Completion or participation rate details (filter on previous) | | Open text |
| Survey language(s) | Survey language(s) | List |
| Full question wording with all interviewer instructions, prompts and visual aids | Full survey questionnaire with all instructions, prompts, visual aids | Checkbox |
| External sample provider(s) | Sample provider(s), and, if multiple, the share of sample from each provider | Open text |
| Proportion of sample provided (filtered on previous) | | Numerical |
| Use of breakout routers or chains | Use of survey routers or chains | Checkbox |
| Breakoff rate | Breakoff rate (i.e., the percent of respondents who start the survey but do not finish it) | Numerical |
| Estimated size of the noncovered population | The universe from which the sample was drawn, and the proportion of that universe that had a nonzero chance of participation | Numerical |
| Use of incentives | Use of incentives | Yes/No |
| What incentive was provided (filter on previous) | | Open text |
| Quality control checks | Details of quality control checks (e.g., for logic, speeding, straightlining), including how they were performed and results of those checks, including percent of completed interviews excluded or dropped from the analysis | Open text |
| % of respondents removed due to checks | | |

The Details: Display Mockup

[← New Search](#)

iPOLL | [ROPER](#) | [Membership](#) | [Data Highlights](#) | [Data Archiving](#) | [Learn](#) | [About Us](#) | [Blog](#) | [Cornell University](#)

iPOLL Search Results

Selected Filters: Org: Pew Research Center ✕

pew research center Narrow Search

Organizations

Narrow Organizations

- ☒ Pew Research Center
- ☐ Princeton Survey Research Associates International
- ☐ ORC International
- ☐ Abt SRBI
- ☐ Social Science Research Solutions (SSRS)

[More](#)

Topics

Narrow Topics

- ☐ Information
- ☐ Economics
- ☐ Political Partisanship

| Questions (422) | Studies/Datasets (335) |
|--|---|
| View (per page): 50 ▾ Sort: Newest Study First ▾ | |
| <input type="checkbox"/> Select All | |
| <input type="checkbox"/> | <p>Pew Research Center Poll: 2017 Political Typology Survey (Roper #31I03004)</p> <p>Survey by Pew Research Center for the People & the Press. Conducted by Princeton Survey Research Associates International. Interview Date(s): 06/08/2017 - 07/09/2017 Country(ies): </p> <p> </p> <p>✓ TRANSPARENCY DETAILS</p> |
| <input type="checkbox"/> | <p>CNN/ORC Poll: 2017 Affordable Care Act/North Korea/Syria (Roper #31I02951)</p> <p>Survey by CNN. Conducted by ORC International. Interview Date(s): 04/22/2017 - 04/25/2017 Country(ies): </p> <p> </p> <p>✓ TRANSPARENCY DETAILS</p> |

| ✓ ROPER TRANSPARENCY PROJECT | | |
|--|-----------|----------------|
| Greatly Exceeds Requirements (9.5) | | |
| CORE | AVAILABLE | NOT APPLICABLE |
| Survey organization | ✓ | |
| External survey sponsor | ✓ | |
| Grant funding source | | ✓ |
| Data collection dates | ✓ | |
| Universe | ✓ | |
| Geographic coverage | ✓ | |
| Justification for claims of representativeness | ✓ | |
| Mode | ✓ | |
| Mode other: Description(filtered on previous) | | ✓ |
| Sample size | ✓ | |
| Sampling procedure: Summary | ✓ | |
| Sampling procedure:Respondent selection stage | ✓ | |
| Sampling frame | ✓ | |
| Weight Variable | ✓ | |
| Weighting benchmark source | ✓ | |
| Variables used for weighting | ✓ | |
| Response rate | ✓ | |
| Disposition codes | ✓ | |
| Response rate definition | ✓ | |
| Completion or participation rate | | ✓ |
| Completion or participation rate details(filter on previous) | | ✓ |
| Survey language(s) | ✓ | |
| Full question wording with all interview instructions, prompts and visual aids | ✓ | |
| ADDITIONAL | AVAILABLE | NOT APPLICABLE |
| External sample provider(s) | ✓ | |
| Proportion of sample provided (filtered on previous) | ✓ | |
| Use of breakout routers or chains | | ✓ |
| Breakoff rate | No | |
| Estimated size of noncovered population | No | |
| Use of incentives | | ✓ |
| What incentive was provided (filter on previous) | | ✓ |
| Quality control summary | Yes | |
| % respondents removed due to quality checks (filtered on above) | No | |

iPOLL Search Results

Selected Filters:

Country: United States



Collection 1
Longstanding Methods

Collection 2
Recently Developed Methods

Why Two Collections?

Recently developed methods:

This collection includes online nonprobability polls, mixed probability and non-probability methods polls, IVR polls, RBS sample polls, and other surveys utilizing recently developed survey-based public opinion research methods.

water



Organizations

- ☒ YouGov
- ☐ AP VoteCast

Topics

- ☐ Government
- ☐ Elections

Questions (1)

Studies/Datasets (1)

View (per page): 50

Sort: Newest Study First

Studies: 1-50 of 94 Page 1 of 2



Select All



Do you approve or disapprove of the way Donald Trump is handling his job as President?



Survey by YouGov. Conducted by YouGov.com

Interview Dates: September 22 - 23, 2018

Country(ies):



TRANSPARENCY DETAILS

Final Project: Getting Buy-In from 33 Active Data Providers

Called each data provider – explained plan, showed mockups, shared a scored recent submission:

Meets requirements – 5 data providers

Exceeds requirements – 12 data providers

Greatly exceeds requirements – 16 data providers

Biggest questions:

Will we be able to see our scores before they publish? (Yes)

You aren't going to score older surveys, are you? (No)

Will I have to fill out forms every time? (No)

Comments:

"A real service to the field" – there's been a "feeling that the TI has plateaued" in terms of impact, this could "move things forward."

"Doesn't look too burdensome."

Looks manageable, but "I just hope it doesn't become something else to beat pollsters over the head with."

Shouldn't be called "disclosure" if it's something we can't provide because we don't calculate it.

"We want to beat Pew!"

Generally very positive. All were appreciative that they were getting a **long lead time** before implementation.

Consolidation of *fieldwork* (and therefore methodological reporting) in a few organizations made this easier. Most concern shown by small, academic polling operations that don't use the major survey houses.

Additional technical information has already been received as a result of conversations.

Returning to the question of how this project fits into the history of disclosure in polling...

The new requirements/encouraged best practices are far more thorough, and aimed at providing information to a sophisticated poll consumer who can understand the ramifications of each element.

Examples: sampling information, weighting information

| SAMPLING | Nedzi proposal | AAPOR 1967 | NCP (pre2006) | NCP (Current) | AAPOR (Current) | Roper (new) |
|---|-------------------|---------------|------------------|------------------|--------------------|----------------|
| Sampling method | X | | | Level 1 | Immediate | Core |
| Margin of sampling error | | X | | Level 1 | Immediate | |
| Whether these have been adjusted for design effect due to weighting, clustering, or other factors | | | | | Immediate | |
| Justification for claims of representativeness | | | | | | Core |
| Coverage of target population/Estimated size of the noncovered population | | | | Level 2 | Immediate | Additional |
| Sample design/sampling frame(s) | | | | | Immediate | Core |
| Name of the sample supplier, if sample/frame provided by third party | | | | | Immediate | Additional |
| Proportion of sample provided | | | | | | Additional |
| The methods used to recruit the panel or participants, if applicable | | | | | Immediate | |
| Respondent selection procedure (for example, within household), if any | | | | Level 2 | Immediate | Core |
| Description of any quotas or additional sample selection criteria during or post fielding | | | | | Immediate | |
| Maximum number of attempts to reach respondent | | | | Level 2 | | |
| Incentives | | | | | Within 30 days | |
| Other strategies to gain cooperation | | | | | Within 30 days | |
| Use of breakout routers or chains | | | | | Within 30 days | Additional |
| Details about other types of screening procedures | | | | | Within 30 days | |

| WEIGHTING | Nedzi proposal | AAPOR 1967 | NCPP (pre2006) | NCPP (Current) | AAPOR (Current) | Roper (new) |
|---|-------------------|---------------|-------------------|-------------------|--------------------|----------------|
| Description of weighting procedures (if any) used to generalize data to the full population | | | | Level 2 | Immediate | |
| Weighting benchmark source | | | | | Immediate | Core |
| Variables used to calculate weights | | | | | Immediate | Core |
| Identification of weighting variable in dataset | | | | | | Core |

Unresolved questions

Will Recently Developed Methods data providers be able or willing to offer such extensive information? The Center generally works with **sponsors**, not field organizations directly.

Will users find the Transparency Scoring useful – or simply confusing? What supporting materials will we need to build to ensure this is a **service** to users?

Will there be improvement in methodological disclosure for the Longstanding Methods collections? So far, the answer appears to be yes.

Will concerns about poor scores prevent new data providers from sending us older materials once the 1/1/2019 start date is long in the past?

Will this project ultimately be successful in promoting greater transparency in the field?

Have feedback, questions or concerns? Email kjw93@cornell.edu.