Editorial

Progress in understanding survey data fabrication

Steve Koczela[a],* and Fritz Scheuren[b]
[a]*The MassINC Polling Group, Boston, MA, USA*
[b]*Editor-in-Chief Emeritus, Statistical Journal of the IAOS*

Writing in 1945, Leo Crespi warned "The prevalence of so-called 'cheating' by interviewers in the process of obtaining public opinion and market research data has become an increasingly grave concern to responsible opinion researchers" [10]. He could have said essentially the same today, though the cause for concern is more complex than in the past.

The forms of survey data fabrication we are facing today are more diverse, and in some cases more insidious than individual interviewers sitting on the proverbial curbstone and making up survey responses. Recent papers and presentations have described apparent cheating by field supervisors [7], wholesale fabrication of datasets by researchers [4], creation of final survey results without any apparent data collection [3,17], and the apparent mass duplication of all or part of survey responses most likely using a computer [9,16]. With the proliferation of types of fabrication, Crespi's charge from 70 years ago remains pertinent.

> "Any precise figures upon the incidence of fabrication are, in the nature of the case, difficult to obtain. But it is no secret that the prevalence and amount has been in many instances far from negligible, and it is widely agreed that the problem must be solved if the opinion research technique is to preserve its status as a reliable tool of inquiry".

Toward this end, a conference was held earlier this year at NORC's offices in Bethesda, MD, bringing together researchers working on survey data fabrication issues in various capacities. The NORC conference was the fifth in a series of events over the preceding 15 months, which have resulted in some of the papers in this volume and others appearing elsewhere. References, recordings, and websites for the papers and presentations from these sessions can be found in the appendix. The two authors organized each of the five events, in whole or in part.

1. Washington Statistical Society – "Curb-stoning, a Too Neglected and Very Embarrassing Survey Problem". December 2014, Washington DC.
2. New England Chapter of the American Association for Public Opinion Research (NEAAPOR) – "New Frontiers in Preventing, Detecting, and Remediating Fabrication in Survey Research". February 2015, Cambridge MA.
3. Washington Statistical Society – "Curb-stoning Part III". June 2015, Washington DC.
4. Total Survey Error 2015 conference panel, "Data Falsification: Who, Why and How to Detect It". September 2015, Baltimore, MD
5. NORC Conference, Bethesda MD, "New Approaches to Dealing With Survey Data Fabrication". February 2016, Bethesda MD.

The resulting papers, presentations, and conferences have moved the conversation forward in some important ways, both in terms of technical advancements, and in creating a community of like-minded professionals focused on the issue. Some of the papers in this volume describe the types of technical, organizational, and analytical improvements that allow for more effective approaches to data fabrication.

In this sense, the conferences and sessions have been enormously beneficial, since the literature on fabrication remains underdeveloped relative to the seriousness of the issue. Some instances of apparent data fabrication have never been covered in the literature or dissected for the technical information on fabrication that may be available. This is apparent when viewing the references section of this introduction, which can refer only to the media accounts of some high profile instances of data fabrication.

Some of the presentations from these recent sessions, have not been turned into written papers yet. The Department of State's Office of Opinion Research, for example, is implementing an impressive array of fabrication prevention and detection tools [7]. These are described in some detail in the video recordings of the NEAAPOR conference, but have not been written up for publication. A case study described by Parsons [15] describes a troubling case of apparent shortcutting in fielding a domestic face to face survey, but is only available in the NORC session recording. And Spagat's description of the difficulty in publishing findings on fabrication bears much further discussion [19].

This journal and others will hopefully publish papers from these sessions in future issues. For although recent progress is encouraging, there is more to be done, particularly in the following key areas.

1. A recognition of the occurrence of data fabrication. We know fabrication happens, and that it happens more in some places than others. It does not only happen to small research organizations with fewer resources, as was previously believed [12]. Recent instances involve the biggest and most names in the survey research business, academia, and the US Government. An acknowledgement by the research community that data fabrication is a common threat, particularly in remote and dangerous survey environments, would allow the community to be cooperative and proactive in preventing, identifying, and mitigating the effects of fabrication. Though the temptation is to see perfect data as the only acceptable outcome, this is clearly not a tenable position.

2. A cooperative approach to dealing with survey data fabrication. At the moment, the process, such as it is, often involves little more than discussion between the survey sponsor or field company and the person or group who discovered the apparent issue. Given the perceived stakes, these discussions often result in legal threats or even legal action of various types [3,5,17,20].

We can see the potential benefits when a cooperative approach is taken. Depending on the reaction of the researchers, the discussion of potential fabrication can bring significant benefits. Green's decision to allow his jointly authored *Science* paper with LaCour to be retracted cleaned up a corner of the scientific record that had been seriously muddied. Robbins reports that Percent-Match, their program developed following the discovery of duplication in some Arab Barometer surveys, has been used effectively by other organizations. Pew's choice to respond to Robbins and Kuriakose both at the NORC conference and now in the pages of this journal will undoubtedly improve our understanding of duplicate data [18]. And the State Department's acknowledgement of the apparent fabrication in their own surveys has led to more systematic prevention and detection protocols in their own work. They have shared this approach with other organizations at conferences such as those held by NEAAPOR and NORC [7].

3. Continued development of statistical processes for fabrication. Developments like *PercentMatch* and the detection method described by Mushtaq [14] are a good start. Winker's use of synthetic datasets to analyze fabrication patterns is impressive both in ambition and accomplishment. But the problem of fabrication is fundamentally one of co-evolution. The more detection and prevention methods evolve, the more fabricators may evolve to stay ahead. And to the extent we discover and confirm fabrication, we will never know whether we found it all, or caught only the weakest of the pack. With these truths in mind, more work is needed in developing and testing statistical methods of fabrication detection. This is made more difficult by the lack of training datasets, a problem prolonged by a general unwillingness to openly discuss data fabrication.

4. An evaluation of prevention and detection processes in hard to reach survey environments. Technical countermeasures during fielding are less useful in harder to survey areas, which also happen to be the areas where the incentive to fabricate data is the highest. Many of the recent advances in field quality control processes focus on areas where technical measures such as computer

audio recording, GPS, and other mechanisms can be used [6,13].

In remote and dangerous areas, where temptation to fabricate is the highest, technical countermeasures are often sparse [9]. And perversely, these are often the most closely watched international polls, since they often represent the hotspots of American interest and activity. Robbins and Kuriakose show a heavy skew in the presence of duplicate cases in non-OECD countries, potentially a troubling indicator.

These polls conducted in remote areas often have direct bearing on policy for the US and other countries. To get a sense of the impact of the polls, a brief review of the recently released Iraq Inquiry, the so-called Chilcot report, contains dozens of documents that refer, in most cases uncritically, to the impact and importance of polls.

5. Updated policies and procedures relative to fabrication from associations and oversight bodies. These recent sessions have shown our prior understanding of data fabrication to be outdated. Technical documents and policies may require updating. AAPOR's policy and technical documents that touch on fabrication, while helpful for the situation for which they were designed, leave many questions. They focus primarily on interviewer curbstoning and mostly look at situations with federal funded research under the purview of the Office of Research Integrity. With the evolution of the threat in recent years and the recent instances of fabrication elsewhere in the survey community, we suggest an update is warranted. Such an update would help to account for new forms of fabrication and begin to address the practical challenges of reporting, confirming and documenting instances of fabrication.

Despite much work to do, there is reason for optimism. We are encouraged by the growing community of researchers focused on this issue and the literature they are producing. We look forward to much more in the months and years ahead. We see progress in the continuous journey toward meeting Crespi's challenge, that "the problem must be solved if the opinion research technique is to preserve its status as a reliable tool of inquiry."

Appendix – Resources from panels and conferences

The authors have organized a number of recent conference presentations, panels, and workshops on data fabrication. To assist people or groups interested in learning more about this issue, we have complied a list of resources from the sessions. Below is a list of events from December, 2014 through February 2016 organized or facilitated by the authors.

1. Washington Statistical Society – "Curb-stoning, a Too Neglected and Very Embarrassing Survey Problem". December 2014, Washington DC.
   a. Website with presentation slides: http://washstat.org/seminars/2014/20141202_curbstoning.html.
   b. Partial audio recording: https://soundcloud.com/stevek-4/sets/curb-stoning-a-too-neglected-and-very-embarrassing-survey-problem.
   c. Presentations.
      i. Furlong, Cathy. Untitled presentation.
      ii. Kennickell, Arthur. "Curbstoning and Culture".
      iii. Koczela, Steven. "Think beyond the Curbstone: Taking Fabrication Detection and Prevention Beyond the Interviewer Level".
      iv. McCarthy, Jaki. Untitled presentation.
      v. Mushtaq, Ali. "Detection Techniques Applied".
   d. Published papers based on presentations.
      i. Kennickell, Arthur. "Curbstoning and culture", Statistical Journal of the IAOS 31, no. 2, pp. 237–240. May 2015.
      ii. Koczela, Steven, Cathy Furlong, Jaki McCarthy and Ali Mushtaq, "Curbstoning and beyond: Confronting data fabrication in survey research", Statistical Journal of the IAOS, vol. 31, no. 3, pp. 413–422, 2015.
2. New England Chapter of the American Association for Public Opinion Research (NEAAPOR) – "New Frontiers in Preventing, Detecting, and Remediating Fabrication in Survey Research". February 2015, Cambridge MA.
   a. Video recording.
      i. Part 1: https://www.youtube.com/watch?v=xRAMrCHSqgQ.
      ii. Part 2: https://www.youtube.com/watch?v=qsMjbj3GK-8.
   b. Website with slides and abstracts: http://neaapor.org/event-materials/.
   c. Presentations.
      i. Faranda, Regina. "The Cheater Problem Revisited: Lessons from Six Decades of State Department Polling".

ii. Robbins, Michael. "Preventing Data Falsification in Survey Research: Lessons from the Arab Barometer".

iii. Thissen, Rita. "Systems and Processes for Assuring Data Quality".

d. Published Papers based on presentations.

i. Robbins, Michael and Kuriakose, Noble. "Don't get duped: Fraud through duplication in public opinion surveys", Statistical Journal of the IAOS, current issue.

ii. Thissen, Rita and Myers, Susan. "Systems and processes for detecting interviewer falsification and assuring data collection quality", Statistical Journal of the IAOS, current issue.

3. Washington Statistical Society – "Curb-stoning Part III". June 2015, Washington DC.

a. Website with abstracts: http://washstat.org/ seminars/2015/20150609_curbstoning3.html.

b. Presentations.

i. Dajani, Aref and Rodrick J. Marquette, U.S. Census "Falsification Detection and Prevention at Census: New Initiatives".

ii. Faranda, Regina. "The Cheater Problem Revisited: Lessons from Six Decades of State Department Polling".

iii. Thissen, Rita M. "Technical Aids for Deterring and Detecting Falsification of Survey Data".

c. Published Papers based on presentations.

i. Thissen, Rita and Myers, Susan. "Systems and processes for detecting interviewer falsification and assuring data collection quality", Statistical Journal of the IAOS, current issue.

4. Total Survey Error 2015 conference panel, "Data Falsification: Who, Why and How to Detect It". September 2015, Baltimore, MD.

a. Presentations.

i. Furlong, Cathy. "Data Falsification: Who, Why and How to Detect It".

ii. Kennickel, Arthur. "Curb Stoning and Culture".

iii. Koczela, Steven. "Taking Fabrication Detection and Prevention Beyond the Interviewer Level".

iv. Kuriakose, Noble and Michael Robbins. "Detecting Data Falsification in Survey Research".

v. Olson, Timothy. "Understanding Core Reasons Why Falsification Occurs During Field Data Collection".

b. Published Papers based on presentations.

i. Kennickell, Arthur. "Curbstoning and culture", Statistical Journal of the IAOS 31, no. 2, pp. 237–240. May 2015.

ii. Koczela, Steven, Cathy Furlong, Jaki McCarthy and Ali Mushtaq, "Curbstoning and beyond: Confronting data fabrication in survey research", Statistical Journal of the IAOS, vol. 31, no. 3, pp. 413–422, 2015.

iii. Robbins, Michael and Kuriakose, Noble. "Don't get duped: Fraud through duplication in public opinion surveys", Statistical Journal of the IAOS, current issue.

5. NORC Conference, Bethesda MD, "New Approaches to Dealing With Survey Data Fabrication". February 2016, Bethesda MD.

a. Website: https://sites.google.com/site/ norcdatafabricationevent/home.

b. Video recording: http://norc.adobeconnect. com/p69urwdjs1f/.

c. Sessions and presenters (affiliations as of the conference date):

i. "Beyond the curbstone –new methods for detecting fabricated survey data".

1. Joe Murphy, RTI International.

2. Marsha Hasson, Westat.

3. Jennifer Parsons, UIC Survey Research Laboratory.

4. Sharan Sharma, University of Michigan.

5. Moderator: Steve Koczela, The MassINC Polling Group.

ii. "Fabrication in international surveys and difficult research environments"

1. Ken Gaalswyk, Abt SRBI.

2. Ali Mushtaq, Datafugue LLC.

3. Courtney Kennedy, Pew Research Center.

4. Michael Robbins, University of Michigan.

5. Moderator: Regina Faranda, Department of State.

iii. "Improving fabrication detection methods using synthetic data".

1. Peter Winker, University of Giessen.

2. Moderator: Ali Mushtaq, Datafugue LLC.

iv. "The role of organizational culture in preventing fabrication".

1. Jaki McCarthy, USDA National Agricultural Statistics Service.
2. Pat Moynihan, Department of State.
3. Charlene Peiffer, Bureau of Labor Statistics.
4. Moderator: Arthur Kennickell, Board of Governors of the Federal Reserve System.

v. "Rethinking data fabrication prevention, detection, and repair".

1. Ali Mushtaq, Datafugue LLC.
2. Cathy Furlong, Integrity Management Services, LLC.
3. Alan Zaslavsky, Harvard Medical School.
4. Moderator: Monica Dashen.

vi. "Ethical considerations for handling/ reporting fabricated data".

1. Courtney Kennedy, Pew Research Center.
2. Michael Spagat, Royal Holloway University.
3. Thomas Belin, UCLA School of Public Health.
4. Moderator: Brad Edwards, Westat.

d. Published Papers based on presentations.

i. Murphy, Joe, Paul Biemer, Chris Stringer, Rita Thissen, Orin Day, Patrick Hsieh. "Interviewer falsification: Current and best practices for prevention, detection, and mitigation," Statistical Journal of the IAOS, current issue.

ii. Robbins, Michael and Noble Kuriakose. "Don't get duped: Fraud through duplication in public opinion surveys", Statistical Journal of the IAOS, current issue.

iii. Simmons, Katie, Andrew Mercer, Steve Schwarzer and Courtney Kennedy "Evaluating a New Proposal for Detecting Data Falsification in Surveys," Statistical Journal of the IAOS, current issue.

iv. Spagat, Michael. "Comment on "Don't get duped: Fraud through duplication in public opinion surveys," Statistical Journal of the IAOS, current issue.

v. Winker, Peter. "Assuring the quality of survey data: Incentives, detection and documentation of deviant behavior," Statistical Journal of the IAOS, current issue.

Steve Koczela and
Fritz Scheuren
E-mail: skoczela@massincpolling.com

Acknowledgement

References

[1] AAPOR, Interviewer Falsification in Survey Research: Current Best Methods for Prevention, Detection and Repair of Its Effects, 2003, Accessed August 26, 2016 from https://www.aapor.org/AAPORKentico/AAPOR_Main/media/MainSite Files/falsification.pdf.

[2] AAPOR, Report on Interviewer Falsification, 2005, Accessed August 26, 2016 from http://www.aapor.org/Education-Resources/Reports/Report-to-AAPOR-Standards-Comm-on-Interviewer-Fals.aspx.

[3] M. Blumenthal, Daily Kos vs. Research 2000 Lawsuit Settled. May 27, 2011, Huffington Post. Accessed August 25, 2011 from http://www.huffingtonpost.com/2011/05/27/daily-kos-research-2000-lawsuit_n_867775.html.

[4] D. Brockman and K. Joshua, Irregularities in LaCour. Accessed August 26, 2016 from http://stanford.edu/~dbroock/broockman_kalla_aronow_lg_irregularities.pdf.

[5] Daily Research News Online. JIR Group Wins Respondent Data Falsification Case. Accessed August 26, 2016 from http://www.mrweb.com/drno/news22890.htm.

[6] A. Dajani and J. Rodrick, Marquette, U.S. Census Falsification Detection and Prevention at Census: New Initiatives. Paper presented at Washington Statistical Society – Curbstoning Part III. June 2015, Washington DC.

[7] R. Faranda, The Cheater Problem Revisited: Lessons from Six Decades of State Department Polling. Paper presented at New Frontiers in Preventing, Detecting, and Remediating Fabrication in Survey Research conference, NEAAPOR, Cambridge, MA, 2015.

[8] A. Kennickel, Curbstoning and culture, *Statistical Journal of the IAOS* **31**(2) (2015).

[9] S. Koczela, C Furlong, J. Mccarthy and A. Mushtaq, Curbstoning and beyond: Confronting data fabrication in survey research, *Statistical Journal of the IAOS* **31**(3) (2015), 413–422.

[10] L.P. Crespi, The cheater problem in polling, *Public Opinion Quarterly* **9**(4) (1945), 431–445.

[11] M. Lacour and G. Donald, When contact changes minds: An experiment on transmission of support for gay equality, *Science* **346**(6215) (12 December 2014), 1366–1369. (Retracted)

[12] P. Lavrakas, Encyclopedia of Survey Research Methods (2 volumes), SAGE, 2008.

[13] J. Murphy, B. Paul, S. Chris, T. Rita, D. Orin and H. Patrick, Interviewer falsification: Current and best practices for prevention, detection, and mitigation, *Statistical Journal of the IAOS*, current issue.

[14] A. Mushtaq, Detection techniques applied, Paper presented at Washington Statistical Society – Curb-stoning, a Too Neglected and Very Embarrassing Survey Problem, December 2014, Washington DC.

[15] J. Parsons and F. Isabel, Approaches for detecting fabricated survey data. Paper presented at New Approaches to Dealing With Survey Data Fabrication conference, NORC, Bethesda, MD, 2016.

[16] M. Robbins and K. Noble, Don't get duped: Fraud through duplication in public opinion surveys, *Statistical Journal of the IAOS*, *Statistical Journal of the IAOS*, current issue.

[17] Silver, Nate Comparison Study: Unusual Patterns in Strategic Vision Polling Data Remain Unexplained, September 26, 2009, FiveThirtyEight. Accessed August 25, 2016 from http://fivethirtyeight.com/features/comparison-study-unusual-patterns-in/.

[18] K. Simmons, M. Andrew, S. Steve and K. Courtney, Evaluating a new proposal for detecting data falsification in surveys, *Statistical Journal of the IAOS*, current issue.

[19] M. Spagat, Suspicious supervisors and suspect surveys, Stats.org, Accessed August 25, 2016 from http://www.stats.org/suspicious-supervisors-suspect-surveys/.

[20] M. Spagat, Comment on Don't get duped: Fraud through duplication in public opinion surveys, *Statistical Journal of the IAOS*, current issue.
bibitem21 P. Winker, Assuring the quality of survey data: Incentives, detection and documentation of deviant behavior, *Statistical Journal of the IAOS*, current issue.