

Comment on “Don’t Get Duped:
Fraud through Duplication in Public Opinion Surveys”

This is an important paper.

Teachers know that many students will cheat on an exam or assignment if students believe they can get away with it. University workshops on how to prevent cheating are a normal part of the academic landscape and nobody thinks that University X has sunk to pathological levels of dishonesty because it hires PhD students to patrol classrooms during examinations and invests in plagiarism detection software. To the contrary, we would suspect universities without such policies of being plagued with rampant dishonesty and would expect the worst in any universities that use law suits to silence people who report cheating incidents.

Some humans eventually transition from being students to being survey researchers. At this point both producers and consumers of survey research seem to assume that these workers have outgrown the temptation to cheat. Serious strategies to prevent and detect data fabrication are considered unnecessary within this universe of presumed rectitude and the handful of researchers finding evidence of fabrication are generally ignored.¹

The Total Survey Error framework promises to consider “anything that can cause the information gathered in a research study to be of questionable or limited value” (Lavrakas, 2013) yet it ignores the possibility of data fabrication other than an occasional mention that such shenanigans can be viewed as a form of measurement error.² To be sure, some cutting edge survey research firms devote significant resources to deter and detect fabricated data (Thissen et al. (2015) although a veil of secrecy surrounds these operations:

“Many organizations and survey sponsors prefer that incidents of falsification be kept confidential;” (Thissen et al. (2015)

The side-lining of the fabrication issue runs deeper than neglect and silence. A few years ago I found evidence of fabrication in some public opinion surveys conducted in Iraq, wrote up my findings and shared them with the companies involved. They responded with a threatening letter from their legal team. This episode suggests that intimidation is also part of the landscape surrounding data fabrication in surveys.

¹ For example, Spagat (2010) provides evidence of fabrication that substantially inflated a survey estimate of the number of people killed in Iraq but the survey is frequently cited as if this evidence was never presented.

² For example, Biemer and Lyberg (2003, p. 41) mentions fabrication in passing in their book. Koczela et al. (2015) finds the literature on fabrication to be “fairly thin”.

Kuriakose and Robbins (KR) demonstrate that data fabrication in surveys is a central issue for survey research that must be brought out into the open and discussed forthrightly. Their analysis of 1,000+ mainstream public opinion surveys finds strong evidence of widespread data fabrication. In particular, they find at least 5% of the interviews to be fabricated in a quarter of the surveys conducted in non-OECD countries. Even 4% of the surveys in OECD countries reached this level of duplication. Moreover, the KR analysis considers just one form of fabrication so the full problem is probably worse than what they uncover.

KR search their mass of surveys for duplicated and nearly duplicated observations. There is a duplicate when two people give the exact same answer to all substantive questions about their opinions. Each survey contains dozens of questions while responses, and the way they are recorded, are known to be rather random. So one would expect that exact matches almost never occur. KR quantify this intuition by simulating 100,000 datasets with 1,000 observations of 100 binary variables. They never find more than one case with even 85+ out of the 100 variables matching. They also analyse many rounds from surveys known to have rigorous quality control and conclude that matches above the 85% level should be rare in real public opinion survey data.³

The obvious mechanism for producing matches is simple: cut-and-paste. The mechanism for near matches is cutting and pasting plus altering some entries. An advantage of this approach for fabricators is that individual interviews look real whereas interviews fabricated from scratch might exhibit strange and detectible patterns. Nevertheless, even full duplicates are readily detectable so their widespread presence suggests, at best, indifference among many survey managers.⁴ Near duplicates have been harder to find but KR contribute a new Stata program that does this work automatically.

In short, KR demonstrate that fabrication of survey data is a major issue that the survey world must confront as well as providing us with a useful tool to address the issue. At the same time we must not forget that duplication is just one form of data fabrication in surveys.

The background to the KR paper is important and illuminating. Steven Koczela informed researchers at Arab Barometer of duplicates in their data. Arab Barometer investigated and confirmed the problem. They then cleaned their existing data and successfully addressed the problem in their subsequent work. The Arab Barometer data are now substantially improved and they are now sharing their lessons learned with the broader survey research community. Arab Barometer's position is now analogous to that of a university that has addressed its cheating issue openly and honestly, reduced the problem and is disseminating its good practices. We should assume that survey organizations that remain in denial about

³ The surveys are the General Social Survey and the American National Elections Study.

⁴ Koczela et al. (2015) shows fabrication patterns in some surveys that must reach well above the interviewer level.

the fabrication issue have not reached the quality standards that Arab Barometer has at this point.

KR do not address the vital question of how conclusions from the many surveys covered by their study are affected by the fabrication problem. It is, of course, impossible to treat this question properly in a single journal article. The survey industry as a whole, not just the surveys covered in the KR article, must rise to this challenge. We must recognize that the fabrication issue is likely to be at least as serious for surveys with undisclosed data as it is for the more transparent ones that enabled the KR study.

Survey research is conducted by human beings and, unfortunately, many humans cheat and cut corners when it suits them to do so. Blinding ourselves to this reality is a recipe for disaster. We need to embrace reality.

References

P.P. Biemer and L.E. Lyberg, Introduction to Survey Quality. Wiley and Sons, Hoboken, New Jersey, 2003.

S. Koczela, C. Furlong, J. McCarthy and A. Mushtaq, Curbstoning and beyond: confronting data fabrication in survey research, *Statistical Journal of the IAOS* **31** (2015), 413-422.

N. Kuriakos and M. Robbins, Don't Get Duped: Fraud through Duplication in Public Opinion Surveys, *Statistical Journal of the IAOS*,....

P. Lavrakas, Presidential Address: Applying a total error perspective for improving research quality in the social, behavioral, and marketing sciences, *Public Opinion Quarterly* **77**(3) (2013), 31-850.

M. Spagat, Ethical and data integrity problems in the second *Lancet* survey of mortality in Iraq, *Defence and Peace Economics* **21**(1), 1-42.

R. M. Thissen and S. K. Myers, Systems and processes for detecting interviewer falsification and assuring data collection quality, *Statistical Journal of the IAOS*,....