Trust BUT VERIFY: Research Integrity in the 21st CenturyTeachers

Scott W. Brown
University of Connecticut, USA

Research misconduct erodes the trust of the stake holders in the social science research enterprise. In order to return confidence and trust in social science research findings, we must emphasize research integrity through education and transparency. This essay discusses three fundamental components of research misconduct to raise awareness of examples of fabrication and falsification misconduct by several prominent social science researchers.

Trusting and Verifying

During the 1980's, when United States President Reagan and Soviet President Gorbachev were negotiating critical nuclear arms de-escalation, a critical time of tensions for geo-political powers, trust did not exist between the political leaders. A phrase was used by Reagan, traced back to a Russian proverb, became the turning point to successful negotiations – "Trust, but Verify", suggesting trust must be earned through verification, not just promises. As the verification processes were implemented, trust began to emerge, growing to a level that allowed for agreements and nuclear de-escalation for both sides. The verification process led to a level of earned trust, though limited, essential for progress.

The phrase, "Trust, but Verify", that was so critical to the progress of nuclear weapons agreements, also applies to the current state of the social science research enterprise and research integrity today.

As social scientists in the 21st century we are aware of the crucial role our research plays in the understanding of human behavior in society and the greater world. Our research work strives to provide insights into various aspects of human development and life, such as public health, education, culture, social interactions, and workforce preparation. However, for social science research to be judged valid, reliable and most of all – meaningful, our entire social science enterprise must be built upon a strong foundation of research integrity for both theory building and policy applications. This can only happen with ethical conduct, integrity, and transparency throughout the entire research process, from the proposal process to the dissemination of findings. If public policy-makers, the general public, and many in our own research communities do not trust the findings and implications of our research, then we must do a better job of demonstrating the values of scientific methods through transparency, verification, and replication. We must share our methodologies and procedures so that our work can be examined, and also share our analytical procedures and data sets for inspection by professionals in order to build trust for our collective work.

What Do We Mean by Research Misconduct?

For the purposes of this current discussion, research misconduct will be limited to two major categories of misconduct, fabrication and falsification, in proposing, performing, and/or reporting the results and implications of research. Fabrication refers to the researcher making-up data (creating fabricated data), altering analyses, the recording or reporting of data. Falsification is the manipulation of research materials, procedures or instrumentation, data manipulation that includes removing or modifying data to fit a specific outcome or results that are inconsistent with accepted procedures or described in the reporting process, such that the procedures are inaccurate and not transparent. It is important to note that in this essay I am deliberately differentiating research carelessness and/or incompetence, from intentional and deliberate misconduct actions.

There are many dimensions of research misconduct that are components of courses and professional development focused on research misconduct and integrity that are also important and should be included in a fuller discussion of the misconduct topic. These include issues of authorship rules, conflicts of interest, policies regarding human subjects, collaborations across institutions, data security, data sharing, plagiarism, and guidelines from professional organizations and associations for the ethical conduct of research, referred to as Responsible Conduct of Research, (RCR), (See Steneck, 2007).

Why Are We Discussing Misconduct?

Unfortunately, we all are aware of the existence of cases of research misconduct, whether through reports by professional organizations, funding agencies, professional publications, and the press. An internet search for "Research Misconduct" will yield pages of headlines about research misconduct across disciplines, including educational research. Some of these websites will reveal numerous reports of research misconduct allegations within specific funding agencies (e.g., The US National Science Foundation; The US Department of Health & Human Services), the majority of which were found to be in violation after they were funded, and the research conducted.

A recent essay in the Chronicle of Higher Education focused on the topic of research misconduct related to what was titled, statistical malfeasance and "cherry-picking" of data and data analytic findings (Jindra & Sakamoto, 2023). They present a strong case of some social scientists "undercontrolling, which can lead to omitted variable bias" and the danger that the researcher presents a conclusion that shows an effect without any certainty that the key independent variable studied is the cause for the findings, thereby misleading the reader and those researchers planning to build upon the stated phenomena. Undercontrolled research designs can be severely misleading and directly harm the research enterprise.

Two sources that can help us understand the frequencies of research misconduct come from responses by scientists, faculty, and graduate students. First, in a 2005 report on a large-scale survey of early career and mid-career scientists conducted by Martinson, Anderson, and de Vries (2005) reported on the responses of 3,247 scientists indicating the specific actions they had conducted within the previous three years. While a very large sample for a study on research

misconduct, the authors noted a response rate of 52% and 43% for mid-career and early career samples, respectively. The authors also noted concerns of a potential non-response bias, where those scientists who were misbehaving "may have been less likely than others to respond to our survey, perhaps for fear of discovery and potential sanction" (p.737). The authors state that while the findings of the incidence for behaviors related to falsifying data, fabricating data, and plagiarism were in line with previous studies (1% to 2%) their findings also suggest, "that US scientists engage in a range of behaviors extending far beyond falsification, fabrication and plagiarism" (p.737). These additional behaviors included, "Withholding details of methodology or results in papers or proposals" (M=10.8%), "Using inadequate or inappropriate research designs" (M=13.5%), "Dropping observations or data points from analyses based on a gut feeling that they were inaccurate" M=15.3%), and "Inadequate record keeping related to research projects" (M=27.5%), and an alarming "33% of the respondents said they engaged in at least one of the top ten behaviors during the previous three years" (p. 738).

In a smaller study surveying education faculty and graduate students, reporting on their judgements to scenarios in research contexts (e.g., removing outlier data, modifying data sets, authorship, and lack of informed consent) (Artino & Brown, 2009). When asked if the participant had observed a behavior judged to be an ethics violation, 37% of the sample (N=120) indicated "yes" with similar rates for the groups of graduate students and faculty. The findings suggest that many of us have observed behaviors related to research misconduct. The next question to be asked should be, what did we do when we became aware of these behaviors? What action did we take? And if we took no action, why?

Cooking the Data

P-hacking, HARKing and Dry Labbing

In an article by Gary Smith titled "How Shoddy Data Becomes Sensational Research" (2023), we learn about three approaches to research misconduct related to data manipulation: P-hacking, HARKing and Dry Labbing. P-hacking is the manipulation of the data analytical procedures to conduct numerous statistical analyses in search of the Holy Grail of results, a probability value of p<0.05. Consider a researcher who has a bundle of independent and dependent variables that are dumped into the computer, a menu of various statistical procedures are checked and the enter button is pressed to run the analyses -- All the in search of a p<0.05. The belief is that if p<0.05 it is not found, the research project has no hope of being published in a reputable journal, and it is a complete failure and has been a waste of time and resources. But when there is no cost to running hundreds of analyses, with unimaginable speed, it can become a choice too easy to resist, for some. So, P-hacking is manipulating the p-value of research just to increase the chance of publication, and not for making significant contributions to the field of study and improving the human condition.

HARKing occurs when a project yields a finding with a low-p-value, drawing interest, but there does not seem to be an existing theory or model that can aid in the interpretation and implications of the finding. This can happen when out the sheer number of analyses conducted, there are a number of significant results, whether correlations, group differences, regression models, and/or factor loadings, which will result by chance alone (and not as a result of the

careful design and theory/model). In this case, the researcher engages in theory building that is often completely independent of models and theories. As an example, Smith (2023) describes a series of studies attempting to explain bitcoin prices by the National Bureau of Economic Research concluded that the statistically significant correlations between the cryptocurrencies dividends and certain market sectors and associated variables were not significant beyond pure chance, regardless of models created for their study. As researchers, we are always seeking to tie our results to models and existing literature so that we can advance a field. When that connection is not made our findings have little meaning.

Perhaps, the easiest and also most egregious behavior is described as Dry Labbing, a strategy based on "simply make up whatever data are needed to support the desired conclusion." One of the best-known examples of dry labbing is the case of Diederik Stapel, who was investigated in 2011 for having made up his data rather than collecting it (Enserink, 2012). Stapel, a social psychologist, retracted 58 papers and was removed from his position. In an interview with Stapel, discussing his behaviors over a period of years, Stapel reports, "I was not able to withstand the pressure to score points, to publish, to always have to be better." further stating, "I wanted too much, too fast" (Smith, 2023). This is a powerful example of someone responding to the academic pressures of "publish or perish" by breaking all the rules for advancement and recognition, without a concern for integrity or ethics.

Research Misconduct is Not Without Victims

Victims include students preparing for positions of social science researchers, who do not experience appropriate training in research integrity, or they are taught to ignore what they should know is misconduct. Further, students who read poorly conducted research in journals and online, may base their own research on a faulty research foundation. Funders may allocate funding resources to grant proposals that are based on false and unverified bases of research, not awarding proposals that may have been more conservative or of a smaller scope, but ethically and carefully designed. Policy-makers may be developing new innovative programs in education, public health care, or addressing global climate change, only to find that the research data supporting these programs has been fabricated, or falsified, thereby, wasting funds that could have been allocated to programs that are based on scientifically-based (and verified) data. And victims who are the doctoral students of those researchers who have been engaged in long-term misconduct, but who were completely unaware of the misconduct, who based their own research for dissertations and post-doctoral projects on the work of their advisor that has been based on fabrication and/or falsification of data, results, and conclusions. And these students and lab staff often are tainted with the unethical acts of their advisor/supervisor.

To better understand the depth and breadth of the damage done by research misconduct, I recommend the news descriptions of the alleged misconduct conducted by Diederik Stapel in social psychology (Enserink, 2012), Brian Wansink in food and consumer behavior (Lee, 2018), and Marc Hauser in cross-cultural cognition (Wade, 2010). I am not making judgments here about these researchers as people, only their alleged behaviors. I cite these three as examples of prominent researchers in their respective fields who held valued positions in their field for years,

demonstrating that if we, as social scientists, are to earn the trust of in our work, we must promote systems of verification replication and validation.

Recommendations

In order to establish, maintain and monitor research integrity I am proposing five actions to consider:

1. Education in ALL aspects of research integrity, RCR, for all students and faculty in higher education, professional associations, public and private laboratories, and government agencies engaged in intramural and external research, with a refresher professional development every three years, because research foci, statistical and qualitative analyses evolve with the development of new variable options (e.g., fitness trackers, professional survey responders, new statistical packages, and of course Artificial Intelligence).

2. Increased value placed on support for replication studies that will provide data confirming or disconfirming replications of results, instrumentation and/or procedures. and includes the publishing of replication studies.

3. Transparency of the research process, from proposal development to data collection, analyses, and final reporting and publication. This transparency must include all members of research teams in discussions of research integrity. These discussions must also include policies and procedures for reporting allegations of potential research misconduct.

4. The establishment of data repositories at institutions of higher education, professional associations, or government agencies, where the de-identified data sets of disseminated research projects can be securely stored.

5. Creation of independent oversight systems within institutions conducting and /or funding research, responsible for monitoring funded and unfunded research for compliance with research integrity.

If we do not take meaningful actions immediately, we will continue to have issues of trust with our stakeholders because we are not offering verification they can trust. We must promote the values of research integrity with clear and factual evidence that all our stakeholders can trust.

References

Artino, A.R., Jr. & Brown, S.W. (2009). Ethics in educational research: A comparative analysis of graduate student and faculty beliefs. *College Student Journal, 43*(2), 599-615.

Christensen Hughes, J., Eaton, S. (2022). Academic Misconduct in Higher Education: Beyond Student Cheating. In Eaton, S.E., Christensen Hughes, J. (Eds) *Academic Integrity in Canada. Ethics and Integrity in Educational Contexts, vol 1*. Springer, Cham. https://doi.org/10.1007/978-3-030-83255-1_4 .

Enserink, M. (2012, November 28). Final Report: Stapel Affair Points to Bigger Problems in Social Psychology. Science. https://www.science.org/content/article/final-report-stapel-affair-points-bigger-problems-social-psychology .

Hamblin, J. (2018, September 24). A Credibility Crisis in Food Science. The fall of a prominent behavioral scientist tells of a system where research is judged not on merit, but on the attention it gets. *The Atlantic*. https://www.theatlantic.com/health/archive/2018/09/what-is-food-science/571105/

Jindra, M. & Sakamoto, A., (2023). When Ideology Drives Social Science: Statistical malfeasance and cherry-picking are rife. *The Chronicle of Higher Education*, March 6, 2023. https://www.chronicle.com/article/when-ideology-drives-social-science?sra=true&cid=gen_sign_in.

Martinson, B., Anderson, M. & de Vries, R. (2005). Scientists behaving badly. *Nature, 435*, 737–738 (2005). https://doi.org/10.1038/435737a

Smith, G. (2023, June 6). How Shoddy Data Becomes Sensational Research. *The Chronicle of Higher Education*. https://www.chronicle.com/article/how-shoddy-data-becomes-sensational-research.

Steneck, N. H. (2007). Introduction to the Responsible Conduct of Research.US Department of Health and Human Services. https://ori.hhs.gov/sites/default/files/rcrintro.pdf .

Wade, N. (2010, August 20). Harvard Finds Scientist Guilty of Misconduct. *The New York Times*. https://www.nytimes.com/2010/08/21/education/21harvard.html.

Appendix

Additional Resources

The US Department of Health & Human Services. Office of Research Integrity. https://ori.hhs.gov/infographics
Provides excellent resource for addressing research misconduct, training materials and a list of summaries of research integrity cases related to this agency by year.

American Psychological Association. Research Misconduct Resources.
https://www.apa.org/research/responsible/misconduct/

University of Connecticut Policy on Alleged Misconduct in Research. (2017).
https://policy.uconn.edu/2014/06/06/policy-on-alleged-misconduct-in-research/

US National Institutes of Health (NIH). (2019, November). A Guide to the Handining of Research Misconduct Allegations https://oir.nih.gov/system/files/media/file/2021-08/guide-handling_research_misconduct_allegations.pdf

US National Science Foundation. Regulation of Research. Resources for researchers and potential Whistle blowers.
https://www.nsf.gov/od/ogc/regulation.jsp

Further Readings on Allegations of Research Misconduct

Bartlett, T. (2019, September 11). 'I Want to Burn Things to the Ground' Are the foot soldiers behind psychology's replication crisis saving science --- or destroying it? The Chronicle of Higher Education. https://www.chronicle.com/article/i-want-to-burn-things-to-the-ground/ .

Donaldson, S. (2022, March 8). U. of North Carolina's Head of Research Cited for Research Misconduct. Chronicle of Higher Education. HTTPS://WWW.CHRONICLE.COM/ARTICLE/U-OF-NORTH-CAROLINAS-HEAD-OF-RESEARCH-CITED-FOR-RESEARCH-MISCONDUCT .

Gelman, A. (2023, June 27). How Academic Fraudsters Get Away With It. The Chronicle of Higher Education. https://www.chronicle.com/article/how-academic-fraudsters-get-away-with-it .

Lee, S.M. (2018, February 25). Here's How Cornell Scientist Brian Wansink Turned Shoddy Data Into Viral Studies About How We Eat. BuzzFeed News. https://www.buzzfeednews.com/article/stephaniemlee/brian-wansink-cornell-p-hacking .

Lee, S. M. (2023, June 16). A Weird Research-Misconduct Scandal About Dishonesty Just Got Weirder. The Chronicle of Higher Education. https://www.chronicle.com/article/a-weird-research-misconduct-scandal-about-dishonesty-just-got-weirder .

Mackintosh, N. J. (Ed.). (1995). Cyril Burt: Fraud or framed? Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198523369.001.0001

Scheiber, N. (2023, June 24). Harvard Scholar Who Studies Honesty Is Accused of Fabricating Findings. The New York Times. https://www.nytimes.com/2023/06/24/business/economy/francesca-gino-harvard-dishonesty.html

Sleek, S. (2022, December 29). On the Right Side of Being Wrong: the Emerging Culture of Research Transparency. APS Observer, January/February. https://www.psychologicalscience.org/observer/right-side-of-wrong