Evidence on questionable research practices: The good, the bad, and the ugly

...and how effective are the interventions?

George C. Banks
University of North Carolina at Charlotte
Presentation for CARMA, November 2017
Overview of presentation

- What evidence exists regarding engagement in questionable research practices in the social sciences?
- How effective are interventions?
Evidence on Questionable Research Practices (QRPs)

- QRPs contribute to a growing concern regarding the credibility of research in the organizational sciences and related fields (Antonakis, 2017).
- Such practices can have harmful implications for evidence-based practice, theory development, and perceptions of the rigor of science (Grand et al., in press).
Popular Media Attention

When the Revolution Came for Amy Cuddy

As a young social psychologist, she played by the rules and won big: an influential study, a viral TED talk, a prestigious job at Harvard. Then, suddenly, the rules changed.

BY SUSAN DOMINUS OCT. 16, 2017

(Dominus, 2017; Photo illustration by Alec Soth)
Popular Media Attention

(Oliver et al., 2016)
Evidence on QRPs

- Examples of commonly discussed QRPs include:
  - Selective reporting of hypotheses
  - Excluding data post hoc
  - Hypothesizing after the results are known (HARKing)
  - Selective inclusion of control variables
- These practices can occur with or without intent to deceive, but exist out of normative assumptions around research.

(Bedian et al., 2010; John et al., 2012; O’Boyle et al., 2017)
Evidence on QRPs

- We conducted a systematic review to consider the evidence on QRPs.
- Using a triangulation approach (e.g., by reviewing data from observations, sensitivity analyses, and surveys), we sought to identify the good, the bad, and the ugly regarding evidence on QRPs.

(Banks, Rogelberg, Woznyj, Landis, & Rupp, 2016a)
Evidence on QRPs: Behavioral studies

- Of the 19 behavioral observation studies, 4 appeared to find little to no evidence of engagement in QRPs and the other 15 found more severe evidence.
- O’Boyle et al. (2017) illustrated that when dissertations became published articles, the ratio of supported to unsupported hypotheses more than doubled (0.82:1 vs 1.94:1).

(Banks et al., 2016a)
Evidence on QRPs: Sensitivity analyses

- No study in this area appeared to find little to no evidence of engagement in QRPs; Fourteen found more severe evidence.

- Many have found an overabundance of \( p \)-values immediately below the critical 0.05 threshold (Gerber & Malhotra, 2008a, b; Leggett et al., 2013; Masicampo & Lalande, 2012)

- O'Boyle et al. (2015) found that although sample size has remained relatively unchanged over a 20 year period (\( r = -0.007 \)), statistically significant moderated multiple regression (MMR) tests have risen from 42% (1995-99) to 52% (2000-04) to 60% (2005-09) to 72% (2010-14).

(Banks et al., 2016a)
Evidence on QRPs: Sensitivity analyses

(O’Boyle et al., 2015)
Evidence on QRPs: Self-report surveys

- Of the self-report survey studies, 1 appeared to find little to no evidence of engagement in QRPs and the other 16 found more severe evidence.
- Banks et al. (2016b) found that
  - Approximately 50% of researchers said that they selectively reported results and presented post hoc findings as if they had been determined a priori.
  - About a third of researchers surveyed reported engaging in post hoc exclusion of data and decisions to include/exclude control variables to turn nonsignificant results into significant ones.
  - The reporting of engagement in QRPs was not found to vary by academic rank.

(Banks et al., 2016a)
Evidence on QRPs: Observer report surveys

- Of the studies based on observer reports, 1 appeared to find little to no evidence of engagement in QRPs and the other 13 found more severe evidence.

- Bedeian et al. (2010) found that 92% percent of respondents reported having seen others present post hoc findings as those developed a priori and 78% saw others selectively report findings.

- In a meta-analysis of surveys asking about the behavior of colleagues, Fanelli (2009) found that 72% of respondents reported observing a variety of QRPs, such as data manipulation.
Interesting quotes

- “HARKING is a requirement of our journals, and the more prestigious the more it is emphasized. You MUST MUST MUST MUST invent some theory to explain your findings, and inductive research won't be published.”

- “I routinely engage in selective reporting of significant hypotheses now, and all the co-authors I work with do. Unless researchers will be mandated to report (before collecting data) their study design, constructs included, and the a-priori hypotheses, such practices will remain in place.”

- “The practice of HARKing, while inappropriate when conducted under the guise of a prior hypothesizing is certainly inappropriate, this practice can also be a useful form of inductive theory development.”

Participant quotes from Banks et al. (2016b)
Summary of the QRP Research

- The vast majority of studies included in the current review identified evidence for the engagement in QRP (91% of studies)
  - This evidence was consistent regardless of the methodological approach
- The extent to which QRP are problematic varies by type and engagement frequency

(Banks et al., 2016a)
Summary of the QRP Research

- Some of the more common QRPs include HARKing and selectively reporting results with a preference for those findings that are statistically significant.
- Editors and reviewers may play a role in the prevalence of QRPs.
- Engagement in QRPs has not been shown to vary by academic rank.
- The vast majority of QRP research has focused primarily on practices that affect p-values.
- Statistical cutoffs can be problematic.

(Banks et al., 2016a)
Summary of the QRP Research

- Even if the data are overestimates, it is apparent that engagement is occurring at meaningful rates that could have an impact on our science (most likely the data provide underestimates)
  - Improvement to our research practices will have a positive effect
  - There isn’t one solution to this problem, but that shouldn’t discourage efforts to make improvements

(Banks et al., 2016a)
Open, high integrity, and reproducible science

How effective are interventions?
Do I have to wait for the journals to improve my research?

- Use the Open Science Framework
  - [https://osf.io/4znzp/wiki/home/](https://osf.io/4znzp/wiki/home/)
- Video tutorials
  - [https://cos.io/our-services/training-services/cos-training-tutorials/](https://cos.io/our-services/training-services/cos-training-tutorials/)
Pre-registration on Open Science Framework
Pre-registration

- For 10 exemplars see: [https://osf.io/e6auq/wiki/Example%20Preregistrations/](https://osf.io/e6auq/wiki/Example%20Preregistrations/)
- Tracking studies via Google Scholar that cite pre-registrations on the Open Science Framework we know of 242 papers
- Other non-clinical studies with preregistrations:
  - [https://www.socialscienceregistry.org/](https://www.socialscienceregistry.org/)
  - [http://ridie.3ieimpact.org/](http://ridie.3ieimpact.org/)
  - [http://egap.org/content/registration/](http://egap.org/content/registration/)
  - [http://egap.org/content/registration](http://egap.org/content/registration)
Can I get paid to pre-register?

- If you have a project that is entering the data collection phase, the Center for Open Science is giving away $1,000 to 1,000 research teams who preregister before they publish.
- Tutorial: https://cos.io/prereg/
- Our research question, in its simplest form, is “Is this policy working?”

(Toth et al., in progress)
Does pre-registration help improve our science?

- Results from an initial study indicate that pre-registrations sufficiently restricted opportunistic use of researcher degrees of freedom
  - Found no systematic differences between restrictions of the researcher Degrees of Freedom appearing in different phases of a research project.
- Pre-reg Challenge Registrations worked better than Standard Pre-Data Collection Registrations

(Veldkamp et al., in progress)
Pre-registration: Misconceptions

“Why would I elect to put my research program on hold for 6 months as I navigate pre-registering the report, revising IRB protocols, etc.? To be candid, avoiding a claim by some publication bias Nazi that I HARKed or p-hacked, simply isn’t a strong enough incentive for me to go through this exercise. Now, if ALL of the top journals required pre-registering, then that would be different.”

Participant quotes from Toth et al. (in progress)
Editors and reviewers in the review process

- **Halo Effect:** Drawing a general impression based on a single characteristic
- **Selective Perception:** A perceptual filtering process based on interests, background, and attitude
- **Confirmation Bias** Seeking out information that reaffirms our past choices and discounting information that contradicts past judgments
- **Hindsight Bias** Believing falsely that we could have predicted the outcome of an event after that outcome is already known

(Robbins & Judge, 2016)
Registered reports

- Registered reports: Peer review before results are known helps align scientific values and practices (Chamber et al., 2014).

- Overview of registered reports:
  - https://youtu.be/7MQ5XB1WUp8
Results-blind reviews

- Evaluation of editor, reviewer, and author perceptions of results blind reviews (Findley, Jensen, Malesky, & Pepinksy, 2016)
- Woznyj et al. (in progress): Participant quotes:
  - “Someone like Diederik Stapel would have still committed fraud, it’s not a solution to such behaviors.”
  - “Just as results can be massaged, research designs/methods sections can be too. So as long as the pressure stays high on publication quantity in particular journals, the temptation to make things look better than they are remains also high.”
  - “We might end up with large numbers of null findings being published, taking away space from other studies.”
Prospective meta-analysis

- **Retrospective** meta-analytic reviews are the primary way to cumulative scientific evidence in the social sciences
- A **prospective** meta-analytic (PMA) review involves planning in advance the hypotheses and research questions, designs, measures, and samples for a series of studies in order to triangulate on a phenomenon of interest
  - Triangulation using multiple study designs, settings, samples, and methods to examine a relationship of interest (Sackett & Larson, 1990)
Open data

- APA journals have authors sign a contract that data will be made available to peers for the purpose of verifying the findings.
- Yet, in one study, psychology researchers were approached for data published in papers less than 12 months earlier. Only 27% shared at least some of their data, the remainder failed to comply with the request (Wicherts et al., 2006).
What are the benefits of open data?

- Authors have indicated a greater willingness to share data if standards for citation and intellectual contributions are improved (Tenopir et al., 2011).
- Evidence indicates that openly sharing data leads to higher citations for their associated studies (Piwowar, Day, & Fridsma, 2007).
- Open data can reduce honest mistakes
- Journals that promote open science could have higher impact factors
Open data

- A mandatory open data policy was introduced at the journal *Cognition* for papers submitted after March 1st 2016. In brief, the policy requires that:
  - “All empirical papers must archive their data upon acceptance in order to be published unless the authors provide a compelling reason why they cannot.”

(Hardwicke et al., in progress)
Open data

- To the extent that we can draw causal inference about the policy - its clearly had a positive impact.
- Data availability goes up to about 80% in some of the more recently submitted studies and computational reproducibility looks reasonably healthy overall.
- Yet, only about 48% of articles in the post-policy period had understandable data sets (and for the subset we are doing reproducibility checks for, a substantial number have required author assistance).

(Hardwicke et al., in progress)
Badges seem silly. Do they work?

- Yes. Implementing these badges dramatically increases the rate of data sharing (Kidwell et al., 2016; Rowhani-Farid et al., 2017).
- View a list of journals and organizations that have adopted badges [here](#).
You’ve convinced me. What can I do to implement badges?

- Each journal may choose to award badges based on a simple author disclosure statement or through independent peer review.
- The badges are free to use with attribution of their source. Download the images and get started now: [https://osf.io/tvyxz/files/](https://osf.io/tvyxz/files/)
- The COS offers examples of how to display badges on publications. As long as the badge image and disclosure statement are presented, you can display them any way you like.
- For questions, see this [resources page](https://osf.io/tvyxz/files/) and email the COS at badges@cos.io
Conclusions

- Standards are changing for how we conduct research...
  - Editor Ethics 2.0 (https://editorethics.uncc.edu/editor-ethics-2-0-code/)
  - Committee on Publication Ethics or COPE (https://publicationethics.org/)
    - 3 of the 4 biggest scientific publishers are TOP signatories (Elsevier, Springer-Nature, and Wiley; Taylor and Francis no yet)
    - There are more than 5,000 journal signatories
  - American Psychological Association (APA) is now partnering with the Center for Open Science
Conclusions

- “Management is just starting to become aware of the social psych crisis, and people are largely unaware of the new guidelines and practices (including reviewers and editors). I hope that this changes—but until then “honest” researchers are punished in the review process.”
- “Many of these practices were normatively acceptable—for a long time. The changes in norms are wonderful, but they are quite new.”

Participant quotes from Banks et al. (2016b)
Conclusions

- Questions? Email: gbanks3@uncc.edu
References

References


