

# USING THE “LIST EXPERIMENT” TO IDENTIFY BIAS IN SURVEYS ON QUESTIONABLE RESEARCH PRACTICES

Michael Bang Petersen  
Department of Political Science  
Aarhus University, Denmark



DEPARTMENT OF POLITICAL SCIENCE  
AARHUS UNIVERSITY

6TH WORLD CONFERENCE ON RESEARCH INTEGRITY  
5 JUNE 2019

MICHAEL BANG PETERSEN  
PROFESSOR



# POTENTIAL BIAS IN SURVEYS ON QRPS

---

How prevalent are questionable research practices?

- "Any attempt to uncover the true base rates of QRPs is obfuscated by the very practices that make them questionable (i.e., lack of reporting or misreporting)." (Banks et al., 2016)

The field has turned towards surveys – but this has raised new concerns:

- "As with any unethical or socially stigmatized behavior, self-reported survey data are likely to underrepresent true prevalence." (John et al., 2012)
- "Considering that these surveys ask sensitive questions and have other limitations, it appears likely that this is a conservative estimate of the true prevalence of scientific misconduct." (Fanelli, 2009)

# APPROACHES TO BIAS CORRECTION

---

Traditional approaches to bias correction:

- Direct measurement of individual differences in social desirability (Marlowe & Crowne, 1977)
- Increasing incentives for truth-telling by tying donations to truthful reporting (John et al. 2012)

Yet, these does not eliminate respondent concerns about anonymity and repercussions (e.g., Tourangeau & Yan, 2007)

# POSSIBLE SOLUTION: THE LIST EXPERIMENT

---

## **Racial Attitudes and the "New South"**

James H. Kuklinski, Michael D. Cobb, Martin Gilens

*The Journal of Politics*, Volume 59, Issue 2 (May, 1997), 323-349.

A family of methods (Blair & Imai, 2012):

- Block total response method (Raghavarao and Federer, 1979)
- Item count technique (Miller 1984)
- Unmatched count technique (Dalton, Wimbush, and Daily 1994).

# THE LOGIC OF THE LIST EXPERIMENT

---

1. Respondents are randomized into two conditions
2. In each condition, respondents are presented with a list of items that could apply to their publications
3. Respondents are asked to report how many items (but not which ones) apply to their publications
4. In one condition, the list includes 4 items without any QRPs
5. In the second condition, the list includes the same items + 1 QRP item
6. The difference in reported items between the experimental conditions reveals prevalence of the QRP without any possibility of determining the conduct of any individual respondent

# THE LIMITS OF THE LIST EXPERIMENT

---

Social desirability bias takes many forms:

1. Other deception: Reluctant to admit problematic practices to others
  2. Self-deception: Won't admit problematic practices to themselves
- The List Experiment only addresses #1

Sensitive to sample size (Blair & Imai, 2012):

- Used in Fox et al. (2018) on small sample of researchers (referred to as "Unmatched count technique"): "...it is likely the relatively low number of participants in our UCT (n = 279) led this calculation to be overly sensitive to individual responses, and as such, we do not consider this estimate to be valid or accurate."

# THE DATA: THE "PRINT" PROJECT

---

The PRINT Project:

- Cross-field survey about questionable research practices
- Sample: All researchers (> PhD students) at all Danish universities (total of 8)
- Responses: 3402 "full" responses = 22% rate

Presented with 9-10 QRPs:

- 9 explicit questions, 1 list experiment (with or without QRP)
- Randomly drawn from total pool of 25 QRPs

Linked to publication record, so anonymity concerns might be especially salient

## Further questions on your research practices

We will now ask you about your research practices using a method to alleviate concerns about privacy and appearance that are common in connection with surveys – even completely anonymous surveys like this one.

Below is a list of five practices. Please tell us how many of them (not which ones) apply to at least one of your most recent sole or co-authored publications.

A number smaller than 5 will not reveal which of the specific practices apply to your publications.

Practices:

- Invested significant time satisfying reviewers who were critical of the publication's conclusions.
- Worked in a large collaborative team to complete the research reported in the publication.
- Completed all research related to the publication on my own.
- Self-nominated the publication for more than one academic award.
- In a publication, failing to disclose relevant personal, financial, political or intellectual conflicts of interests.

**How many of the practices apply to at least one of your most recent publications?**

- 0
- 1
- 2
- 3
- 4
- 5
- 

Unable to answer

« Previous question

Save and proceed »



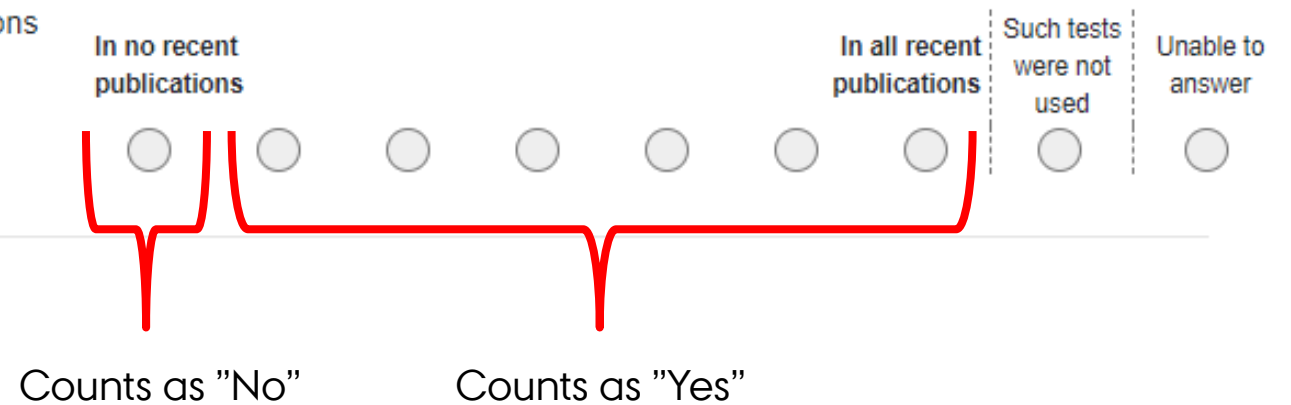
# COMPARISON WITH EXPLICIT QUESTION

Consider the following research practice:

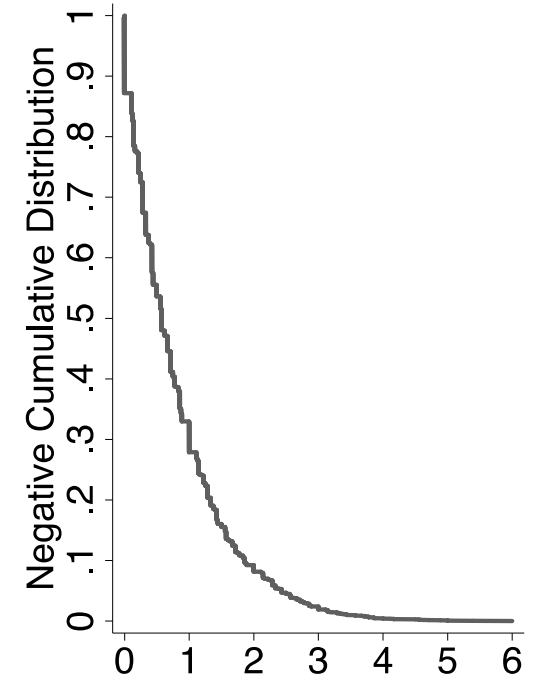
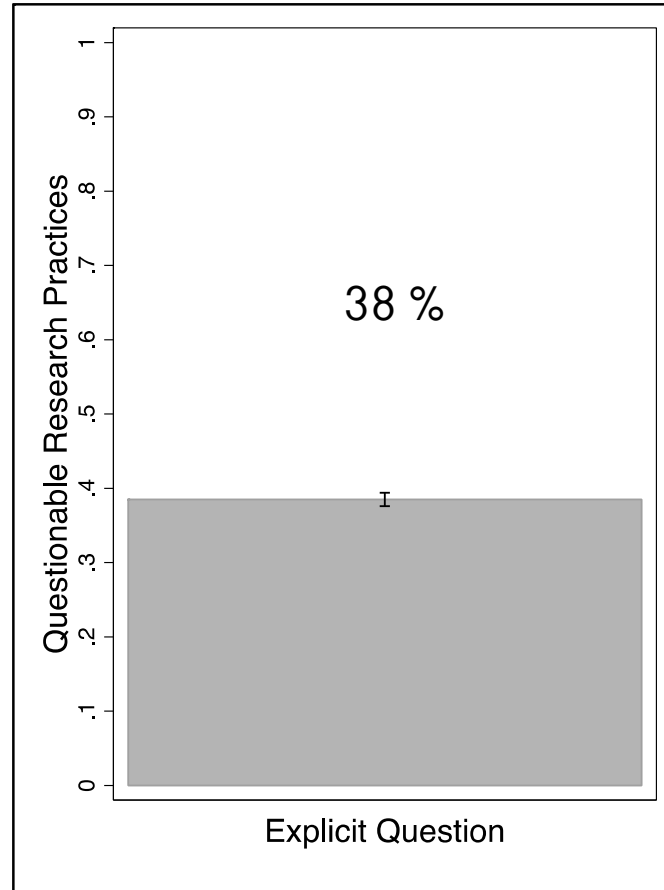
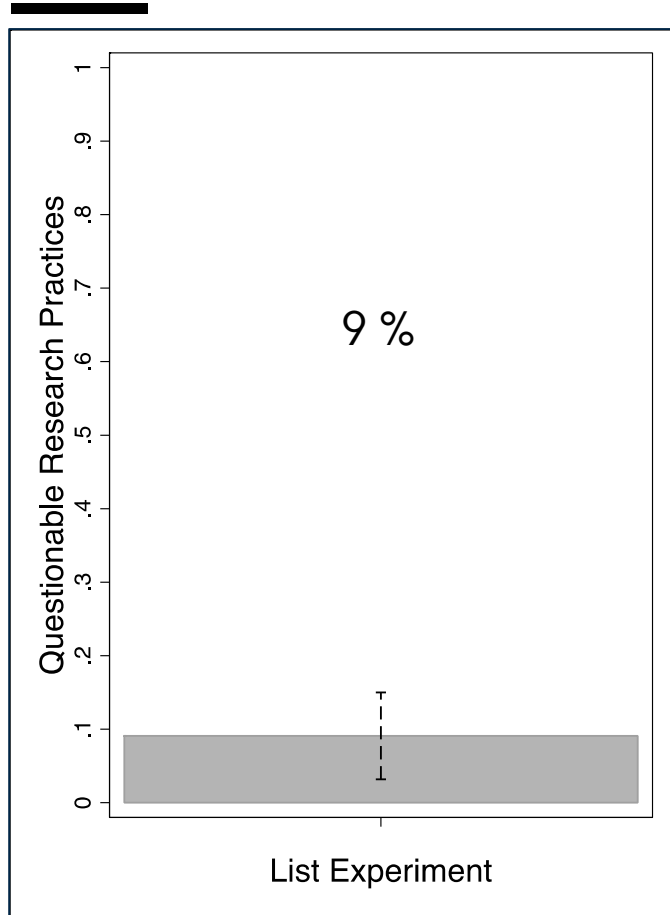
**Presenting statistically significant main findings, without distinguishing between their 'statistical significance' and their potential practical or theoretical importance.**

*For example, claiming a 'significant' finding solely based on the p-value and not the effect size. Notice, importance refers to clinical, economic, biological, psychological, sociological etc.*

To what extent is this practice used in your recent sole or co-authored publications reporting statistical significance tests?



# RESULTS: LOWER PREVALENCE IN LIST EXPERIMENT(!)



Proportion according to List Experiment equals proportion answering, on average, 2 or more on explicit questions.

# MODERATING FACTORS?

---

Estimates of QRP prevalences from the List Experiment were statistically unaffected by:

1. The main field of the researcher (e.g., social science, natural science)
2. The perceived severity of the QRP (...”damaging the trust in research findings...”)
3. The type of QRP (e.g., selective reporting, selective citing)

# CONCLUSIONS

---

Large-scale, cross-field survey of prevalence of QRPs using the "List Experiment" to reduce bias

Other deception is not a major source of bias in survey estimates of the prevalence of QRPs

- Across fields
- Across types of QRPs
- Across severity of QRPs

Survey estimates of prevalence of QRPS are unlikely to be strongly downward biased. Rather, these results raise concerns that explicit survey estimates are *upward* biased.

# ACKNOWLEDGEMENTS

---

Co-authors: Nick Allum, Jens Peter Andersen, Pernille Bak Pedersen, Asger Dalsgaard Pedersen, Nicole Foeger, Allan Rye Lyngs, Ana Marušić, Niels Mejlgaard, Philippe Mongeon, Mads P. Sørensen, Bobby Zachariae & Jesper W. Schneider

This work is supported by the PRINT project (Practices, Perceptions, and Patterns of Research Integrity) funded by the Danish Agency for Science and Higher Education (Ministry of Higher Education and Science) under grant No 6183-00001B.



**DEPARTMENT OF POLITICAL SCIENCE**  
AARHUS UNIVERSITY