Estimating the prevalence and antecedents of questionable research practices in student theses in psychology from self-reports
What are QRPs?

False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant

Joseph P. Simmons¹, Leif D. Nelson², and Uri Simonsohn¹
¹The Wharton School, University of Pennsylvania, and ²Haas School of Business, University of California, Berkeley

Abstract
In this article, we accomplish two things. First, we show that despite empirical psychologists’ nominal endorsement of a low rate of false-positive findings (≤ .05), flexibility in data collection, analysis, and reporting dramatically increases actual false-positive rates. In many cases, a researcher is more likely to falsely find evidence that an effect exists than to correctly find evidence that it does not. We present computer simulations and a pair of actual experiments that demonstrate how unacceptably easy it is to accumulate (and report) statistically significant evidence for a false hypothesis. Second, we suggest a simple, low-cost, and straightforwardly effective disclosure-based solution to this problem. The solution involves six concrete requirements for authors and four guidelines for reviewers, all of which impose a minimal burden on the publication process.
Why QRPs?

Scientists employ QRPs because...

- Significant results get published and publishing is necessary for a career
  Shamoo & Resnick, 2003; Song et al., 2000; Tijdink et al., 2014

- They are socialized to do so by mentors and environments
  Swazey et al., 1993; Versteege, 2013

…but there is little research looking at proximal predictors for these processes!
Why QRPs?

Scientists employ QRPs because…

- Significant results get published and publishing is necessary for a career
  
  Shamoo & Resnick, 2003; Song et al., 2000; Tijdink et al., 2014

Possible proximal predictors:

- Expectation of reward from significant results
- Stress
Why QRPs?

Scientists employ QRPs because...

- They are socialized to do so by mentors and environments

Swazey et al., 1993; Versteeghe, 2013

Possible proximal predictor:

- Attitudes of mentors towards QRPs
- Belief in scientific value of significant results
- Motivation to do good work
If you haven’t read the original paper, or have forgotten it by now, let me refresh your memory: our original values describe the amount of carrots served to elementary school children (ages 8 to 11) at lunch in a control condition.

And, apparently, at least one of them is a Clydesdale horse.
Search results for “questionable research practices” in r/psychologystudents

Sorry, there were no post results for “questionable research practices”
QRPs in psychology

Students are an important group to investigate!

- Future researchers and leaders of the academic field

- Proper development of critical faculties for research findings as nonacademic professionals

- Purest expression of the field’s commitment to high scientific standards
What is going on with QRPs in the student world?
QRPs in psychology

• The final thesis is the most direct measure of a student’s ability to apply scientific principles

• Several possible structural similarities exist between predictors of QRP use in scientists and those in students
Why QRPs?

*Scientists* employ QRPs because...

- Significant results get published and publishing is necessary for a career  
  Shamoo & Resnick, 2003; Song et al., 2000; Tijdink et al., 2014

Possible proximal predictors:

- Expectation of reward from significant results
- Stress
Why QRPs?

*Students* may employ QRPs because of perceptions that…

- Significant results lead to better grades
- The advisor expects significant results
  - Significant results are more useful for the advisor!
- Stress
  - Trying to get a simple “story” from the data to ease writing the thesis
Why QRPs?

Scientists employ QRPs because...
- They are socialized to do so by mentors and environments
  
  Swazey et al., 1993; Versteeghe, 2013

Possible proximal predictor:
- Attitudes of mentors towards QRPs
- Belief in scientific value of significant results
- Motivation to do good work
Why QRPs?

*Students* may employ QRPs because of perceptions that...

- The advisor thinks of QRPs as good practice
- Significant results are a sign of good science

...or because of low motivation to do conscientious work.
Theses – research questions

- How prevalent are QRPs in student theses?
- What attitudes and circumstances lead students to engage in QRPs?
- What influence do advisor attitudes have on students’ QRP use?
- Do structural aspects of thesis work affect QRP use?
  - Students may have less control over design aspects of their work than over reporting or analysis
Hypotheses

H1:

QRP attitudes
Motivation
Stress

Reported QRP use

+
Hypotheses

**H2:**

- **Belief:** good science $\Rightarrow$ significant results
- **Perceived supervisor QRP attitudes**
- **Belief:** significant results $\Rightarrow$ better grades

QRP attitudes

QRP use
H3: All relationships should be stronger for reporting and analysis QRPs than for study design QRPs
Data collection

- Data from $N = 207$ students (varies by analysis due to nonresponses)
- QRPs taken from Fiedler & Schwarz (2016), Questions on:
  - implementation in thesis (yes/no)
  - own attitude
  - perceived supervisor attitude
  (attitude items scored 1 – problematic to 5 – sensible)
  - All questions had “don’t know” option
Survey contents

- Lists of QRP\textperiodcentered s taken from Fiedler \& Schwarz (2016), augmented with an item on HARKing and four distracter items (e.g. “performing a power analysis”)
  - Ich habe im Rahmen meiner Abschlussarbeit mindestens einmal… (ja/nein/weiß nicht)
  - [QRP] finde ich… (1 „problematisch“ – 5 „sinnvoll“ oder „weiß nicht“)
  - [QRP] findet mein Betreuer… (1 „problematisch“ – 5 „sinnvoll“ oder „weiß nicht“)
Self-reported prevalence

- Prevalence current study
- Prevalence Fiedler & Schwarz (2016)

- Deciding whether to exclude data after looking at the results
- Rounding off p values
- Claiming to have predicted an unexpected result
- Failing to report all relevant conditions
- Failing to report all relevant dependent measures
- Falsifying data
- Falsely claiming that results are unaffected by demographics
- Collecting more data in order to achieve significance
- Stopping data collection after achieving the desired result

Self-reported prevalence

<table>
<thead>
<tr>
<th>Behavior</th>
<th>Prevalence current study</th>
<th>Prevalence Fiedler &amp; Schwarz (2016)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deciding whether to exclude data after looking at the results</td>
<td>50%</td>
<td>5%</td>
</tr>
<tr>
<td>Rounding off p values</td>
<td>15%</td>
<td>5%</td>
</tr>
<tr>
<td>Claiming to have predicted an unexpected result</td>
<td>10%</td>
<td>5%</td>
</tr>
<tr>
<td>Failing to report all relevant conditions</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>Failing to report all relevant dependent measures</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>Falsifying data</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>Falsely claiming that results are unaffected by demographics</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>Collecting more data in order to achieve significance</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>Stopping data collection after achieving the desired result</td>
<td>5%</td>
<td>5%</td>
</tr>
</tbody>
</table>
Percentage of students reporting engaging in multiple QRPs

- 45% reported engaging in 0 QRPs
- 30% reported engaging in 1 QRP
- 10% reported engaging in 2 QRPs
- 5% reported engaging in 3 QRPs
- 3% reported engaging in 4 QRPs
- 2% reported engaging in 5 QRPs
- 2% reported engaging in 6 QRPs
- 1% reported engaging in 7 QRPs

Number of QRPs
Deciding whether to exclude data after looking at the results
Changing or formulating new hypotheses after analyzing the data
Rounding off p values
Claiming to have predicted an unexpected result
Failing to report all relevant conditions
Failing to report all relevant dependent measures
Falsifying data
Falsely claiming that results are unaffected by demographics
Collecting more data in order to achieve significance
Stopping data collection after achieving the desired result
Results

QRP attitudes

R&A: $B = .119 (.027)$***
S: $B = -.006 (.012)$ n.s.

Motivation to write thesis

R&A: $B = -.056 (.015)$***
S: $B = -.024 (.013)^*$

Stress due to thesis

R&A: $B = .001 (.016)$ n.s.
S: $B = .010 (.012)$ n.s.
Results

Perceived supervisor QRP attitudes

Good science leads to significant results

Supervisor rewards significant results

R&A: $B = .267 (.038)^{***}$
S: $B = .269 (.060)^{***}$

R&A: $B = .099 (.074)\text{ n.s.}$
S: $B = -.007 (.163)\text{ n.s.}$

R&A: $B = .056 (.064)\text{ n.s.}$
S: $B = .295 (.149)\text{ n.s.}$

R&A: $B = .107 (.036)^{**}$
S: $B = .002 (.15)\text{ n.s.}$

R&A: $B = .030 (.027)\text{ n.s.}$
S: $B = -.031 (.024)\text{ n.s.}$
Limitations

• Self-report study
  – Social desirability?
  – Fear of repercussions?
  – Participants competent to judge QRP use?

• Cross-sectional design
  – Causality issues: do attitudes predict QRP use or does QRP use bias attitudes? Do students infer supervisor attitudes from their own?
  – Participants able to judge supervisor attitudes accurately?
Conclusion

- Motivation appears to be a protective factor in student QRP use
- Student attitudes towards QRP predict their use
  - Only when they have influence, i.e. in reporting and analysis!
- Student’s perception of their supervisor attitudes shapes their own attitudes and affects their reported QRP use directly
- Beliefs about significance may affect student QRP use, but if so, their influence is small
Implications

• Teaching
  – Supervisors must be careful what attitudes they project to students!
  – Motivating students might encourage them to avoid QRP

• QRP use in academia
  – Students’ reported QRP use is comparable to that of career academics
  – Students do not endorse QRP use at the start of their career
  – Academic collaborators’ role (as data analysts, authors, etc.) can dictate in which QRP they engage
THANKS FOR THE ATTENTION!!!
Survey contents

• General motivation (3 items, $\alpha = .62$)
  „Ich finde das Arbeiten an meiner Abschlussarbeit spannend und interessant“

• Stress (3 items, $\alpha = .82$)
  „Die Arbeit an meiner Abschlussarbeit ist für mich belastend“
Survey contents

• Significant results important for grade (3 items, $\alpha = .81$)
  „Meine Note hängt davon ab, ob meine Ergebnisse signifikant sind“

• Good science produces significant results (3 items, $\alpha = .65$)
  „Ein Nulleffekt heißt für mich, dass ich keine gute Arbeit geleistet habe“