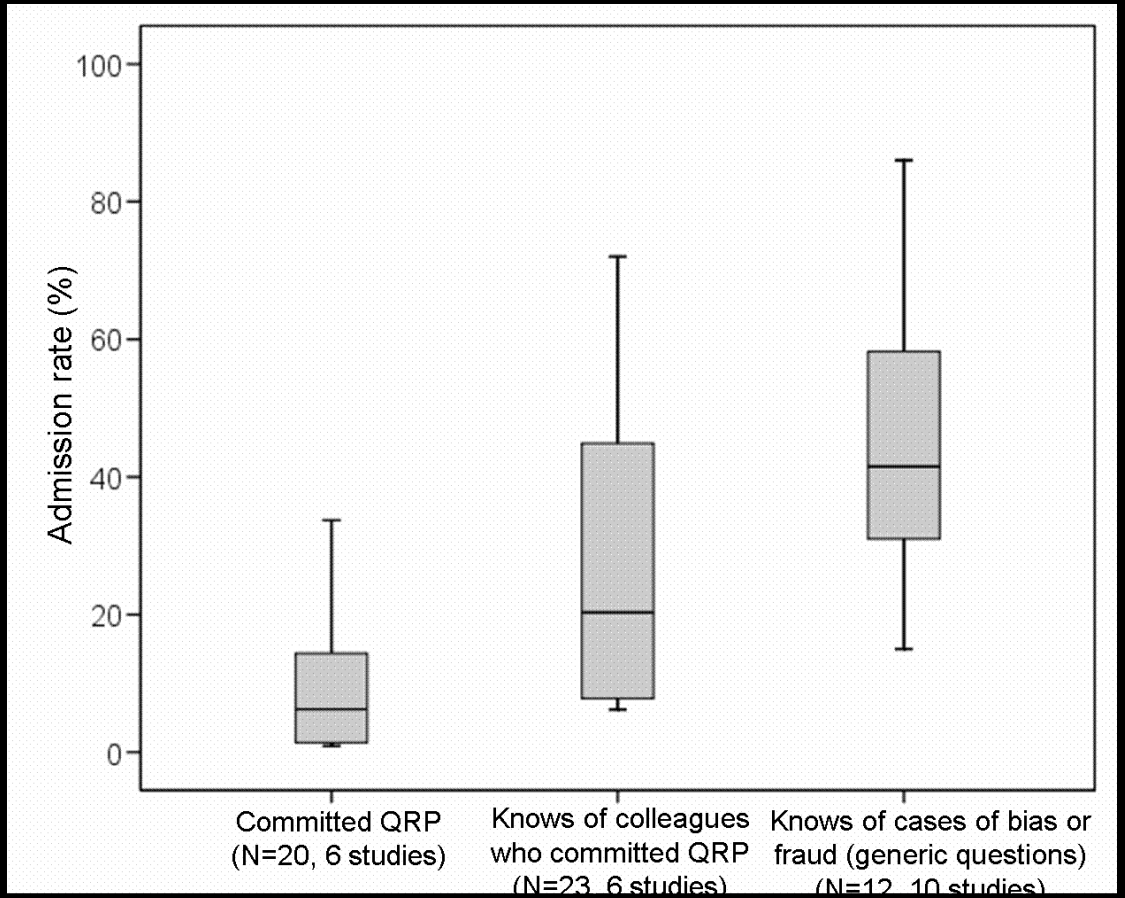
A classification and ranking of Questionable Research Practices reported in surveys

Daniele Fanelli, Alan Voodla, Siim Andres

Heriot-Watt University

University of Tartu

Too diverse to meta-analyse



(Fanelli 2009, PLoS ONE)

Bane or boon in disguise?

Bane or boon in disguise?

- Very commonly reported in surveys (e.g. John 2012)

Bane or boon in disguise?

- Very commonly reported in surveys (e.g. John 2012)
- Source of significant bias

Bane or boon in disguise?

- Very commonly reported in surveys (e.g. John 2012)
- Source of significant bias

- Not as common, when better defined (e.g. Fiedler and Schwarz 2015)

Bane or boon in disguise?

- Very commonly reported in surveys (e.g. John 2012)
- Source of significant bias

- Not as common, when better defined (e.g. Fiedler and Schwarz 2015)
- Might be neutral or beneficial, e.g.:
  - “file drawer” vs. “cluttered office” (de Winter and Happee 2013)
  - HARKing vs RHARKing (Rubin 2017)

Previous classification attempts

Table 1
A Taxonomy of Research Misconduct, Inappropriate and Questionable Conduct.
Source: An extensive review of the literature along with the authors' own experiences as journal editors.

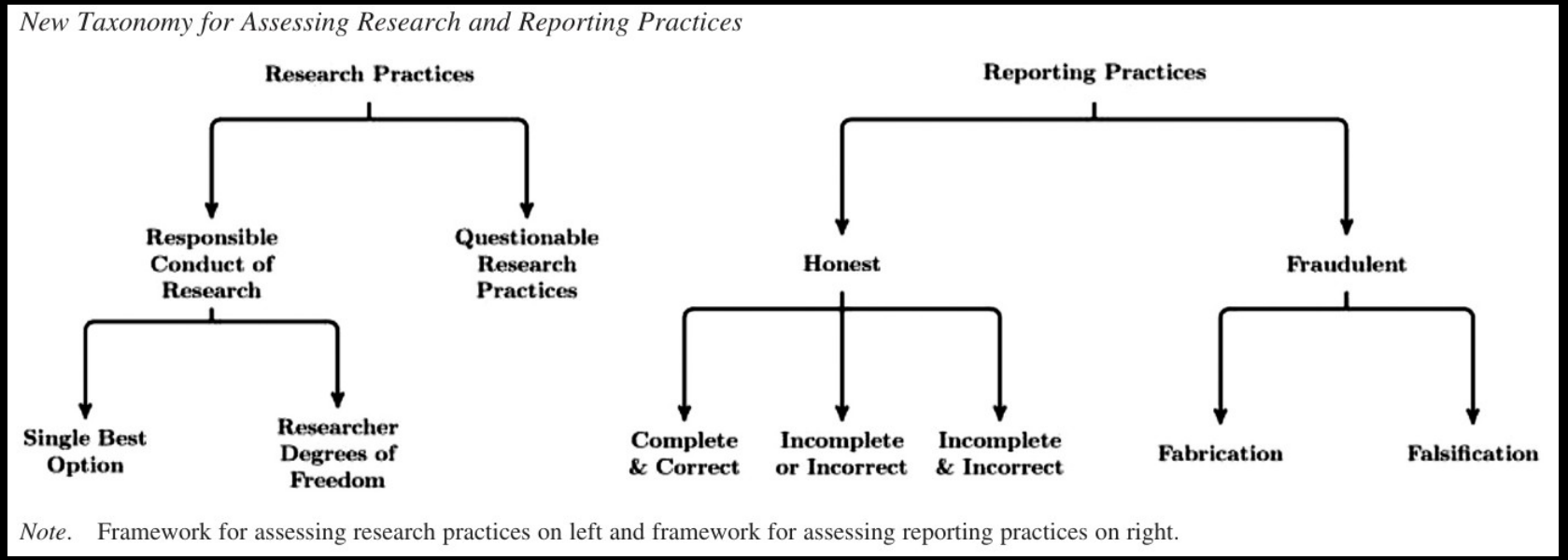
Nature of behaviour	Appropriate conduct	Questionable conduct	Inappropriate conduct	Blatant misconduct
	Widely accepted as good scientific practice	*Absence of clear rules but test of 'the reasonable reader'; perpetrator embarrassed/reluctant to be revealed*	*Rules exist although there may be some variation by field, country, institution and/or journal*	*Clearly defined and universally accepted rules*
Data mani-pulation	• Winsorization*	• HARKing	• Selective reporting • Omitted data	• Data fabrication (e.g. Stapel, Hunton) • Data falsification (e.g. Lichtenthaler, Smeesters)
Use of work by others	• Drawing from and building on work of others	• Short phrases lifted from others and not put in quotation marks	• Entire sentences reproduced without source or quotation marks • Failure to cite or acknowledge others	• Plagiarism of entire article, whole section(s) etc. (e.g. Gottinger, Antoniou) • Wilfully omitting an entire body of work (e.g. in a proposal)
Use of own work	• Making every effort to diffuse one's work • Avoiding excessive self-citation • Maximizing one's research output	• Hyping own work/excessive self-citation • Partial overlap with other papers by that author • Salami publishing	• Self-plagiarism (e.g. Frey, Lichtenthaler) • Redundant publication	• Using the same theory or data to arrive at different conclusions (just for the sake of publishing another paper)
Authorship	• Including as authors all who have made a substantial contribution	• Obligatory authorship (e.g. expectation that a PhD supervisor should be an author)	• Ghost authorship (e.g. Song) • Gift authorship • Gift colluding	• Failure to declare an interest (e.g. Yang & Tao)

Note: The above categories are not exhaustive; there are many other forms of misconduct (e.g. fake referees, citation cartels, journal impact factor (JIF) manipulation by editors) and the examples listed here are merely illustrative of the spectrum. Note also that some of the above examples may not fall neatly into a single 'box' but extend over two or more degrees of severity (e.g. use of a ghost author to improve the written English may not be considered as 'inappropriate' or even 'questionable').

* Winsorization is the assigning of lesser weight to an apparently spurious outlier or modifying its value so it is closer to other sample values (Dixon and Tukey, 1968; Ghosh and Vogt, 2012).

Hall and Martin 2019, Research Policy

Previous classification attempts



Manapat et al. 2022, Psychological Methods.

Our aim

- Re-classify definitions of QRP actually used in surveys
- Compare the frequency of different kinds of QRP across surveys

Our approach

Our approach

- Bottom-up
  - Based on QRPs being investigated
  - Based on the definitions given in research
- Information-centric
  - QRPs defined in terms of how information is manipulated
  - Information in/of any component of a study
    - Data, methods, analyses, authors, references…

Methods: Data collection

Methods: Data collection

Methods: Data collection

- Searched in the Web of Science for “Questionable Research Practices” OR “QRP” AND “survey”
  - Articles, reviews, editorials

Methods: Data collection

- Searched in the Web of Science for “Questionable Research Practices” OR “QRP” AND “survey”
  - Articles, reviews, editorials
- N=72 records, retrieved, inspected.

Final included sample of surveys

StudyID	Country	Discipline	Career Level	Sample Size
1	Mixed	Medicine	Mixed	224
3	Mixed	Education	Researcher	1488
4	Mixed	Economics	Mixed	393
7	Mixed	Medicine	mixed	589
12	Canada	Psychology	PhD student	168
12	Canada	Psychology	MSc student	81
12	Canada	Psychology	UG student	171
16	USA	Psychology	Mixed	2155
17	USA	Linguistics	Mixed	322
19	Australia	Psychology	UG student	205
21	Mixed	Psychology	Researcher	257

StudyID	Country	Discipline	Career Level	Sample Size
23	Mixed	Communication	Researcher	872
25	USA	Psychology	Researcher	164
25	USA	Psychology	PhD student	110
26	Italy	Psychology	Researcher	208
27	Norway	Mixed	Researcher	7291
28	Mixed	Criminology	Researcher	1612
29	Mixed	Ecology	Researcher	807
30	Germany	Psychology	Researcher	1138
34	Croatia	Mixed	Researcher	237
37	Croatia	Medicine	UG student	220
41	Mixed	Psychology	Mixed	1166

Methods: classification

Methods: classification

- N=296 QRP definitions!

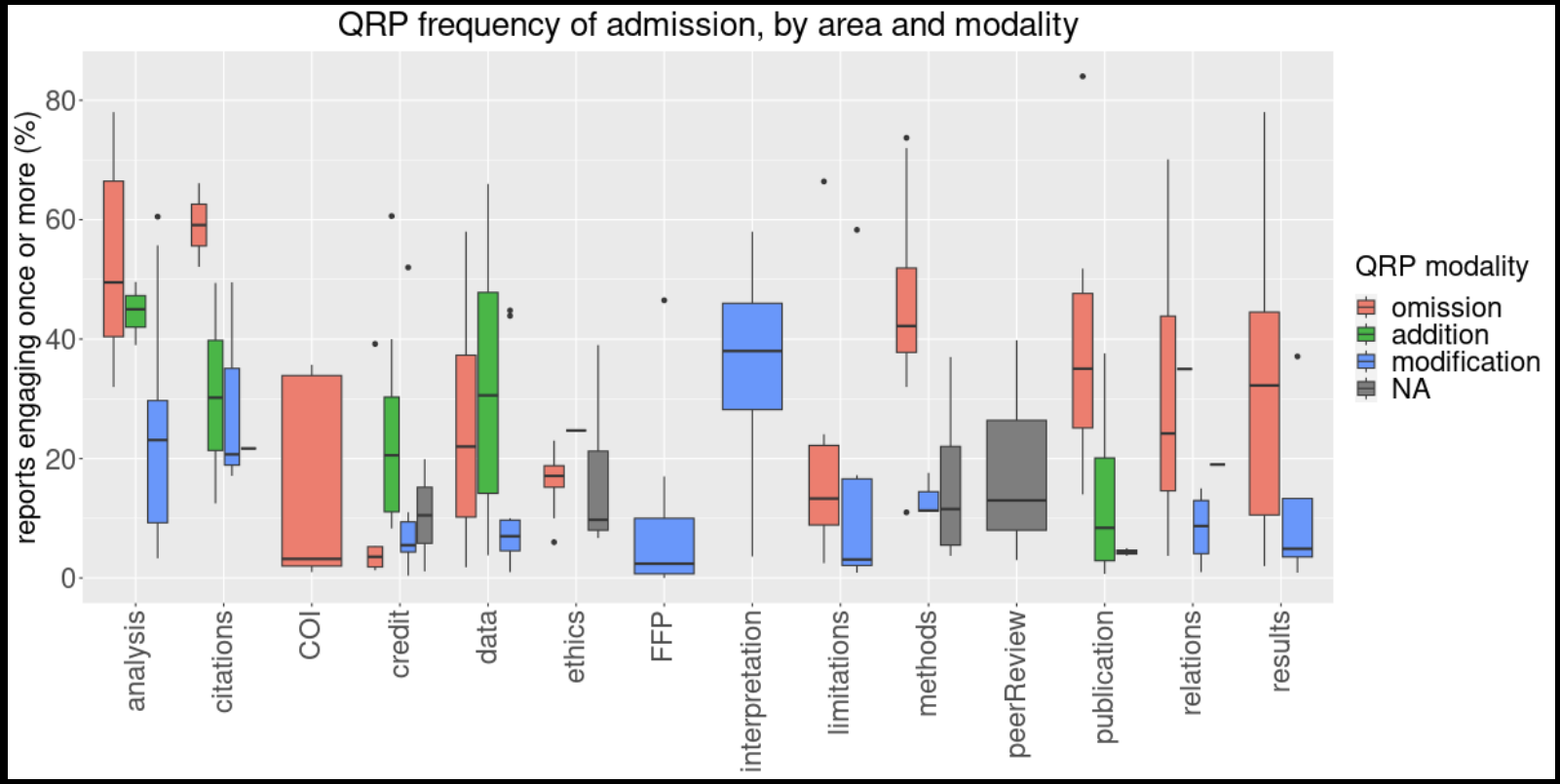
Methods: classification

- N=296 QRP definitions!
- Principled approach along 2 dimensions:
  - By “area” of research
  - By whether information was omitted, added, modified

Methods: classification

- N=296 QRP definitions!
- Principled approach along 2 dimensions:
  - By “area” of research
  - By whether information was omitted, added, modified
- “Iterative” process (i.e. trial and error)
  - Reaching balanced and homogeneous categories

Results: 2D classification



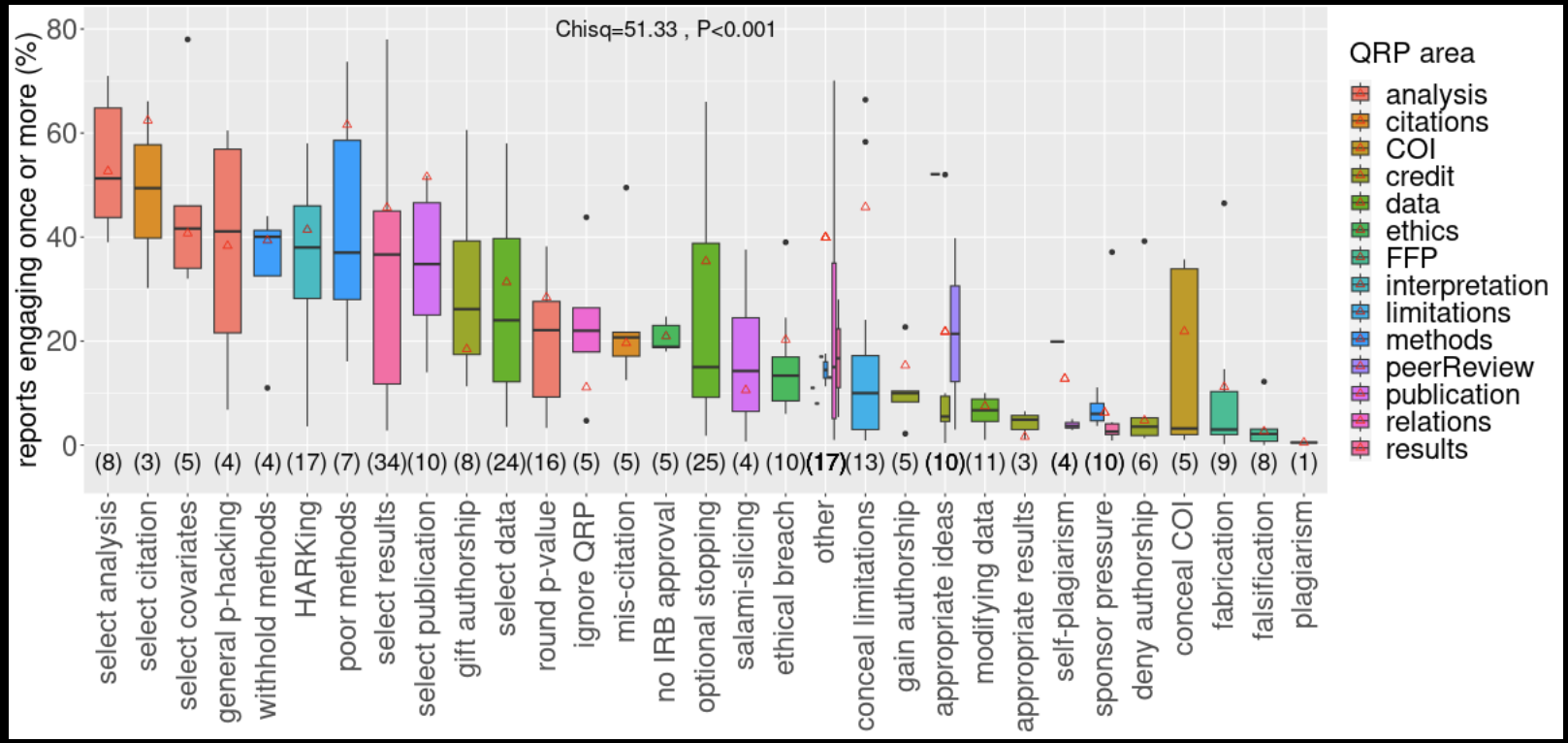
From 2D to 27 QRP categories

QRP area	QRP modality
Analysis	Addition
Analysis	Modification
Analysis	Modification
Analysis	Omission
Analysis	Omission

From 2D to 27 QRP categories

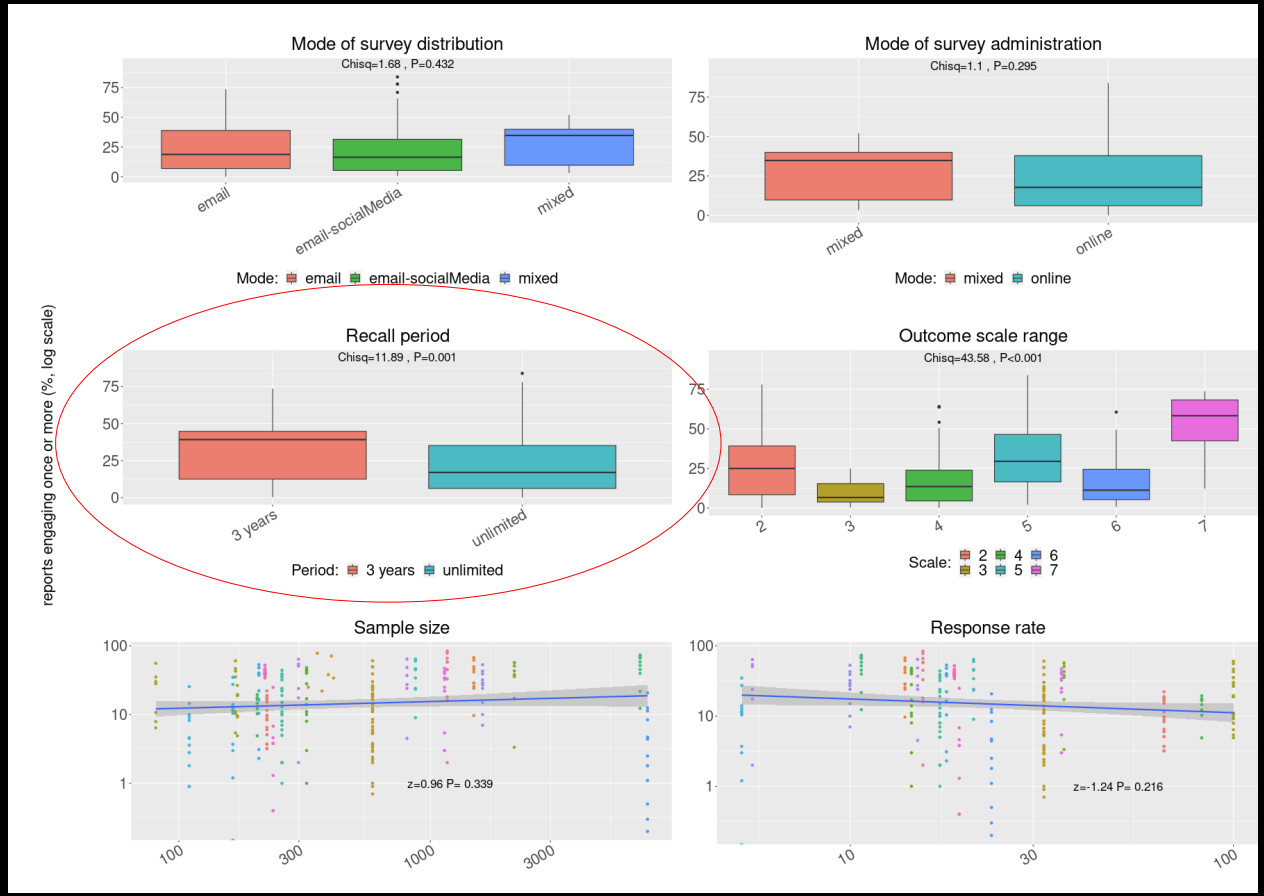
QRP area	QRP modality	QRP category	Example of survey phrasing	Reference
Analysis	Addition	Select analysis	Changing to another type of statistical analysis after the analysis initially chosen failed to reach statistical significance (e.G. P < .05) or some other desired statistical threshold	3
Analysis	Modification	General p-hacking	P-hacking	12
Analysis	Modification	Round p-value	Rounding off a p value or other quantity to meet a prespecified threshold (e.g., Reporting p = .054 as p = .05 or p = .013 as p = .01)	3
Analysis	Omission	Select analysis	Reporting a set of results as the complete set of analyses when other analyses were also conducted	3
Analysis	Omission	Select covariates	Not reporting covariates that failed to reach statistical significance (e.g., P < .05) or some other desired statistical threshold	3

Classification and ranking of QRPs



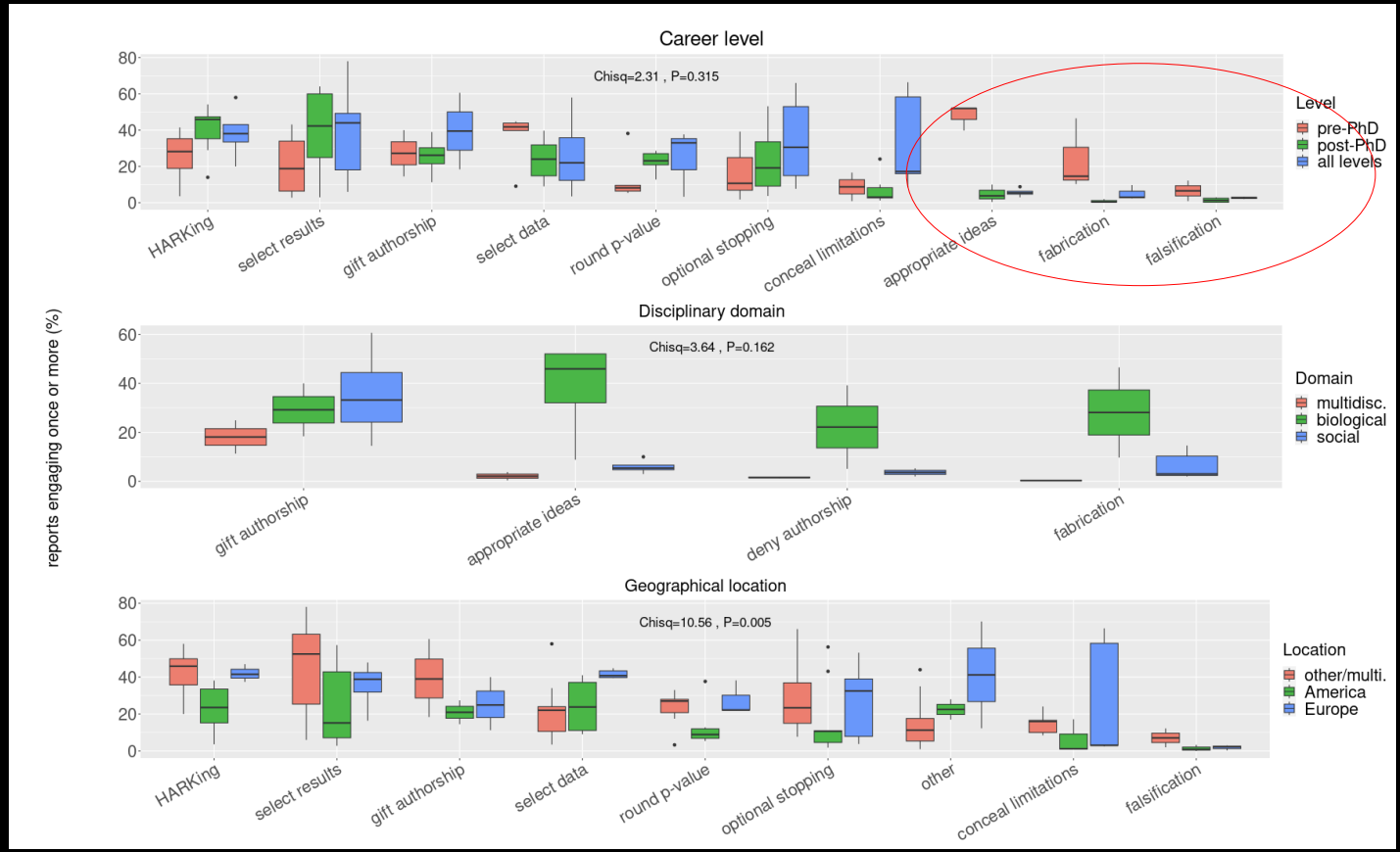
Comparison: survey characteristics

(restricted to QRPs with N>1 in each level)

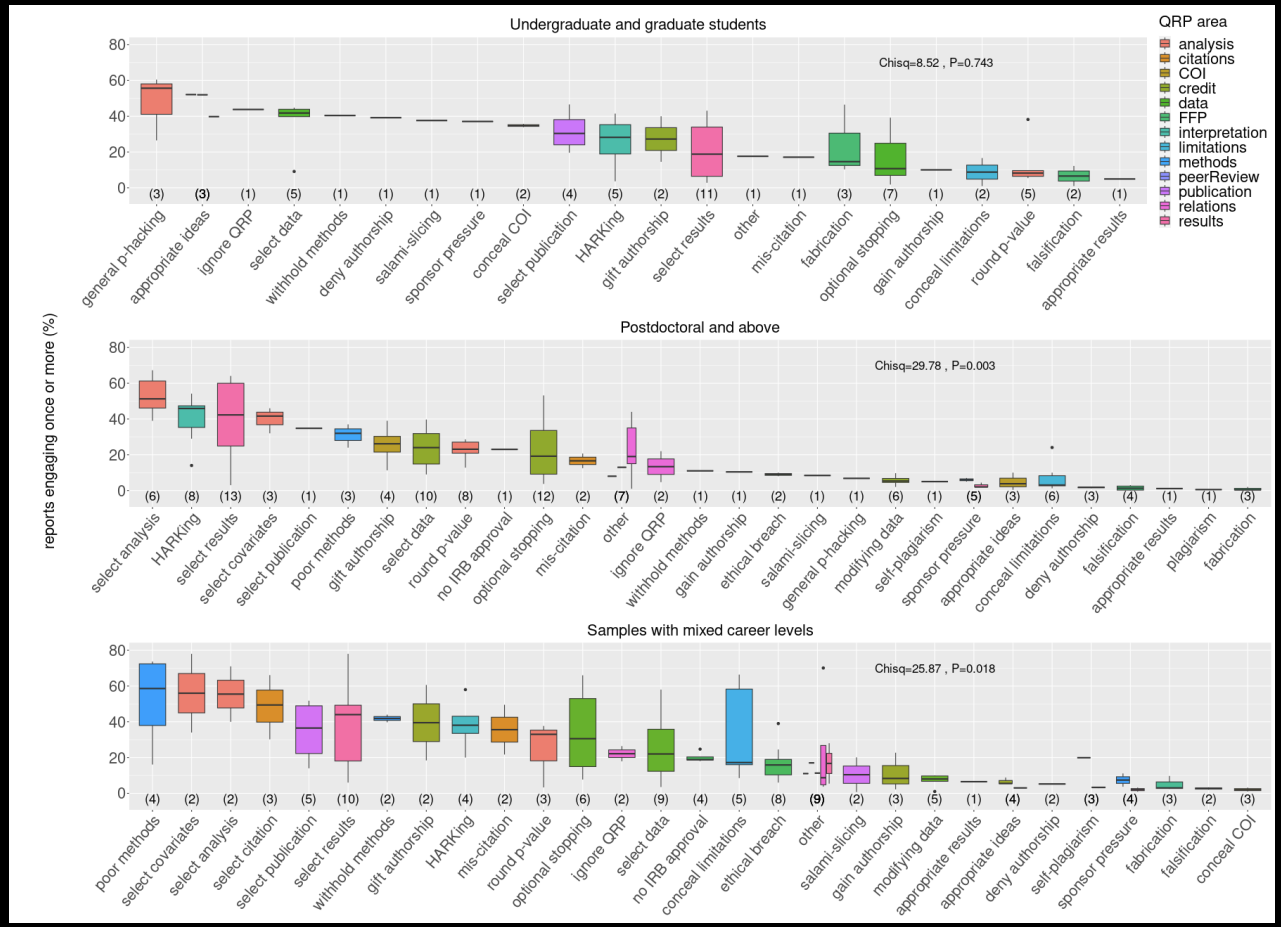


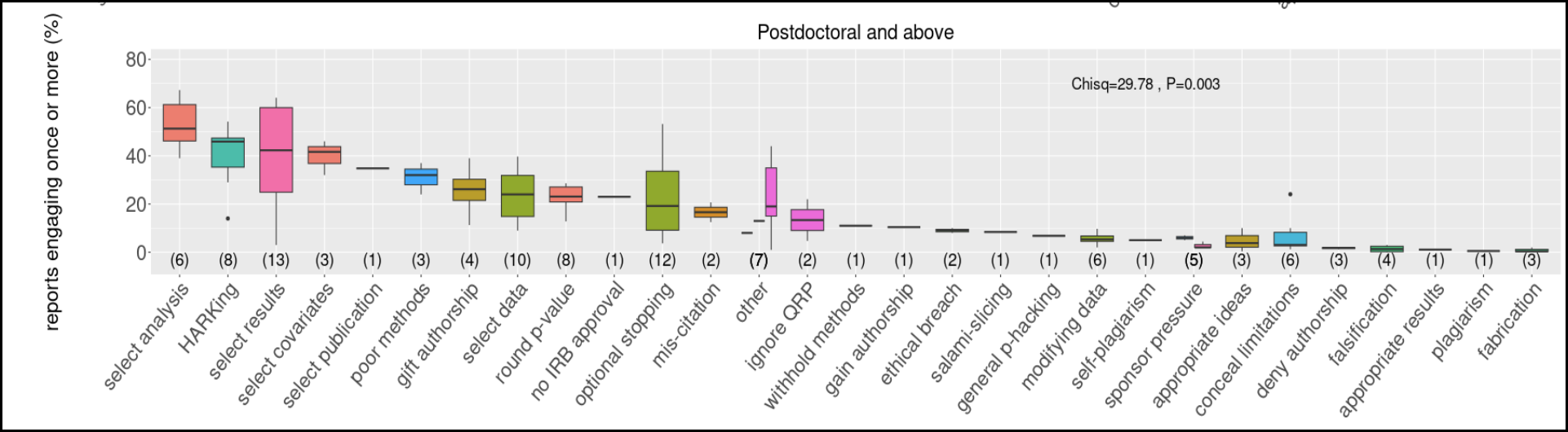
Comparison: demographics

(restricted to QRPs with N>1 in each level)

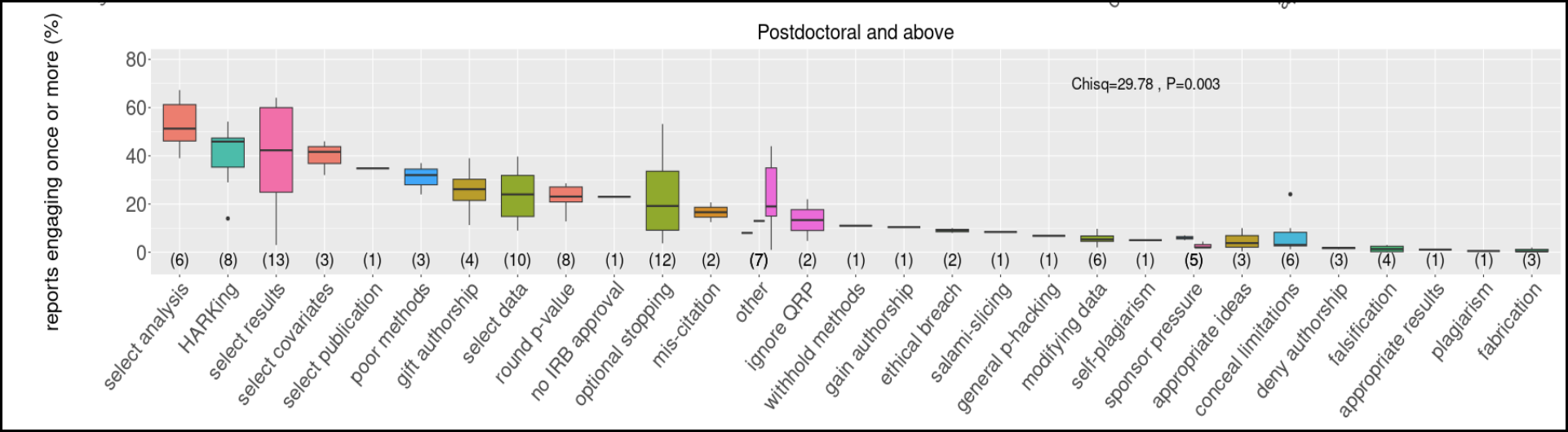


QRP ranking, by career level

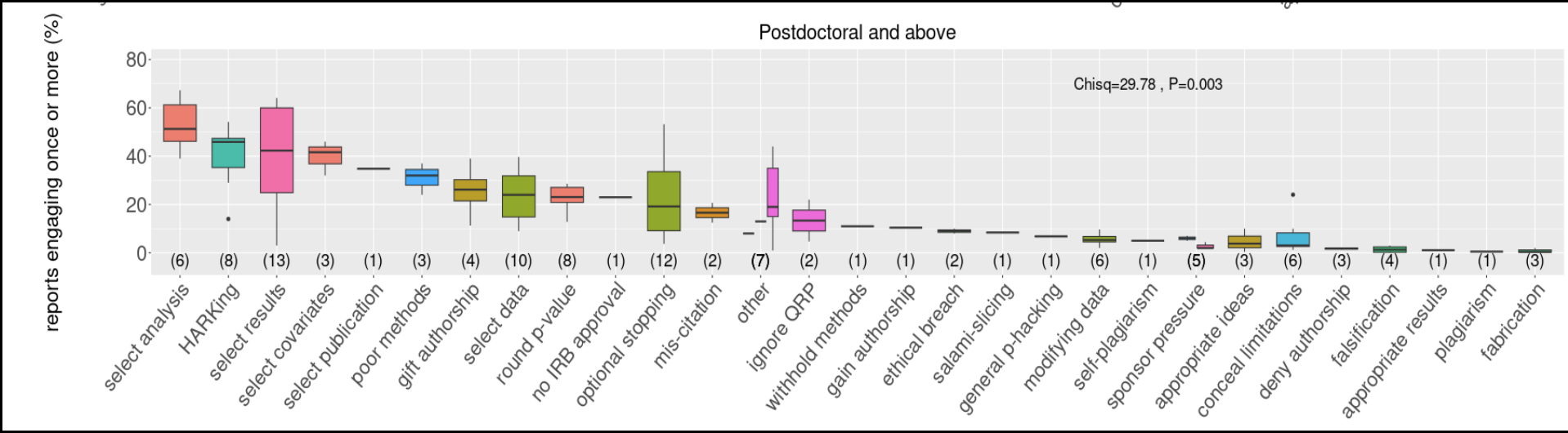



Postdoctoral and above
reports engaging once or more (%)
Chisq=29.78 , P=0.003
0
20
40
60
80
(6) (8) (13) (3) (1) (3) (4) (10) (8) (1) (12) (2) (7) (2) (1) (1) (2) (1) (1) (6) (1) (5) (3) (6) (3) (4) (1) (1) (3)
select analysis
HARKing
select results
select covariates
select publication
poor methods
gift authorship
select data
round p-value
no IRB approval
optional stopping
mis-citation
other
ignore QRP
withhold methods
gain authorship
ethical breach
salami-slicing
general p-hacking
modifying data
self-plagiarism
sponsor pressure
appropriate ideas
conceal limitations
deny authorship
falsification
appropriate results
plagiarism
fabrication

What explains the ranking?



What explains the ranking?



- Lies by omission
- Arbitrary thresholds
- Not always problematic

VS

- Lies by commission
- Clear demarcation
- Ethic./epistemically damaging

BEYOND

beyondbadapples.eu

Funded by
the European Union

Alan Voodla

Andero Usberg

Siim Andres

Join our lab!

TEMET

Theoretical and Empirical METascience lab

"temet nosce"

BEYOND

beyondbadapples.eu

Funded by
the European Union

Alan Voodla

Andero Usberg

Siim Andres

Join our lab!

TEMET

Theoretical and Empirical METascience lab

"temet nosce"

email@danielefanelli.com