# THE QUESTION OF QUESTIONABLE RESEARCH PRACTICES

Stephan Guttinger

Department of Social and Political Sciences, Philosophy and Anthropology

University of Exeter, UK

University of Exeter

egenis

# OVERVIEW

- Claim: Abandon the idea of "questionable research practices" (QRPs)
- Clarification: What should be abandonned is not the idea of *questioning* practice, but the idea that there is a class of questionable research practices
- Reason: The idea that there is a class of QRPs is empirically inadequate and potentially damaging to science

# WHY THINK ABOUT QRPs?

Idea of QRPs dominates contemporary debate about replication crisis and scientific integrity

# EXAMPLE: BEGLEY AND IOANNIDIS (2015)

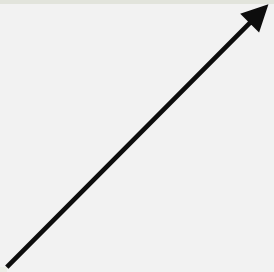
- Replicability crisis is not failure of scientific method but a "consequence of a system that is willing to overlook and ignore lack of scientific rigor and instead reward flashy results"
- "Although fraud is rare, use of questionable research practices seems to affect the majority of researchers"

# A QUESTION OF INCENTIVES

- Why do we have this problem? "We get what we incentivize" (Begley and Ioannidis 2015)
- Key incentive is prestige/credit
  - Prestige built through publishing new results
  - QRPs as shortcut to getting more/faster/higher-impact results

"Credit economy" → Need more output → Increased use of QRPs → Flawed scientific output

"Credit economy" → Need more output → Increased use of QRPs → Flawed scientific output



Lack of knowledge of "proper" method

[Note: In this one-dimensional view researchers are either greedy, dishonest, or stupid?
→ Change incentive structure, increase surveillance, and enforce more training]

# WHAT EXACTLY ARE QRPs?

NAS 1992 report on responsible science:

> "Questionable research practices are actions that violate traditional values of the research enterprise and that may be detrimental to the research process. […] Questionable research practices do not directly damage the integrity of the research process"

# EXAMPLES OF QRPs FROM NAS REPORT:

- Failing to store data for long enough
- Exploiting subordinates
- Not sharing unique research materials with other researchers
- Asking for co-authorship for specialised service that is not significantly related to results reported in paper
- …

# 2011-ISH: QRP CONNECTED TO CRISIS

- QRPs explicitly connected to idea of replication crisis (e.g., Simmons et al. 2011; John et al. 2012; Bakker et al. 2012)
- Used to "massage" or "torture" data
- Now seen as "detrimental" practices that directly damage actual research output (departure from NAS definition)
  - p-hacking
  - HARKing
  - lack of exp. controls
  - …

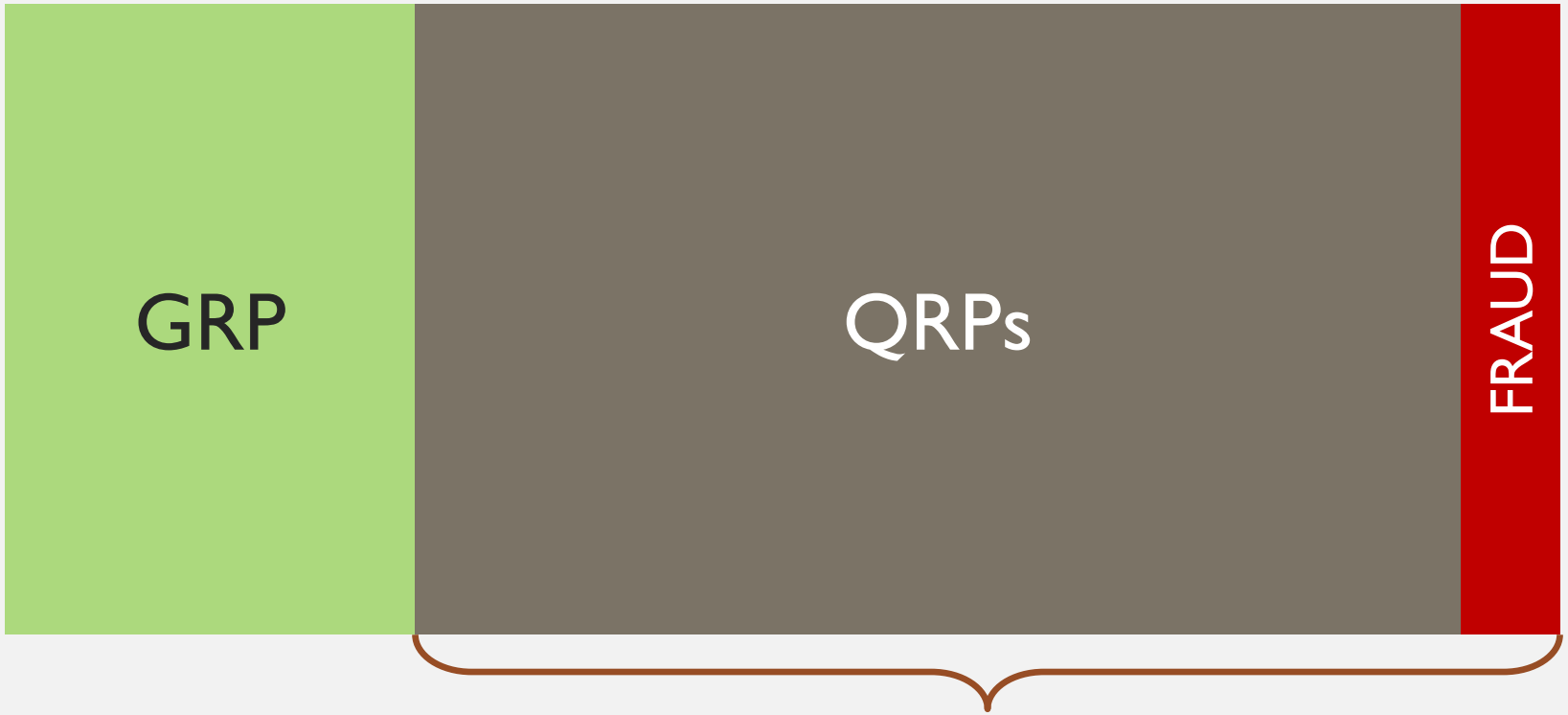
# 2011-ISH: QRP CONNECTED TO CRISIS

- Use of practices labelled as QRPs now seen as sign of "dishonest" science (e.g., Sarewitz 2016; Heesen 2020)

  → Idea of "QRP" becomes supercharged

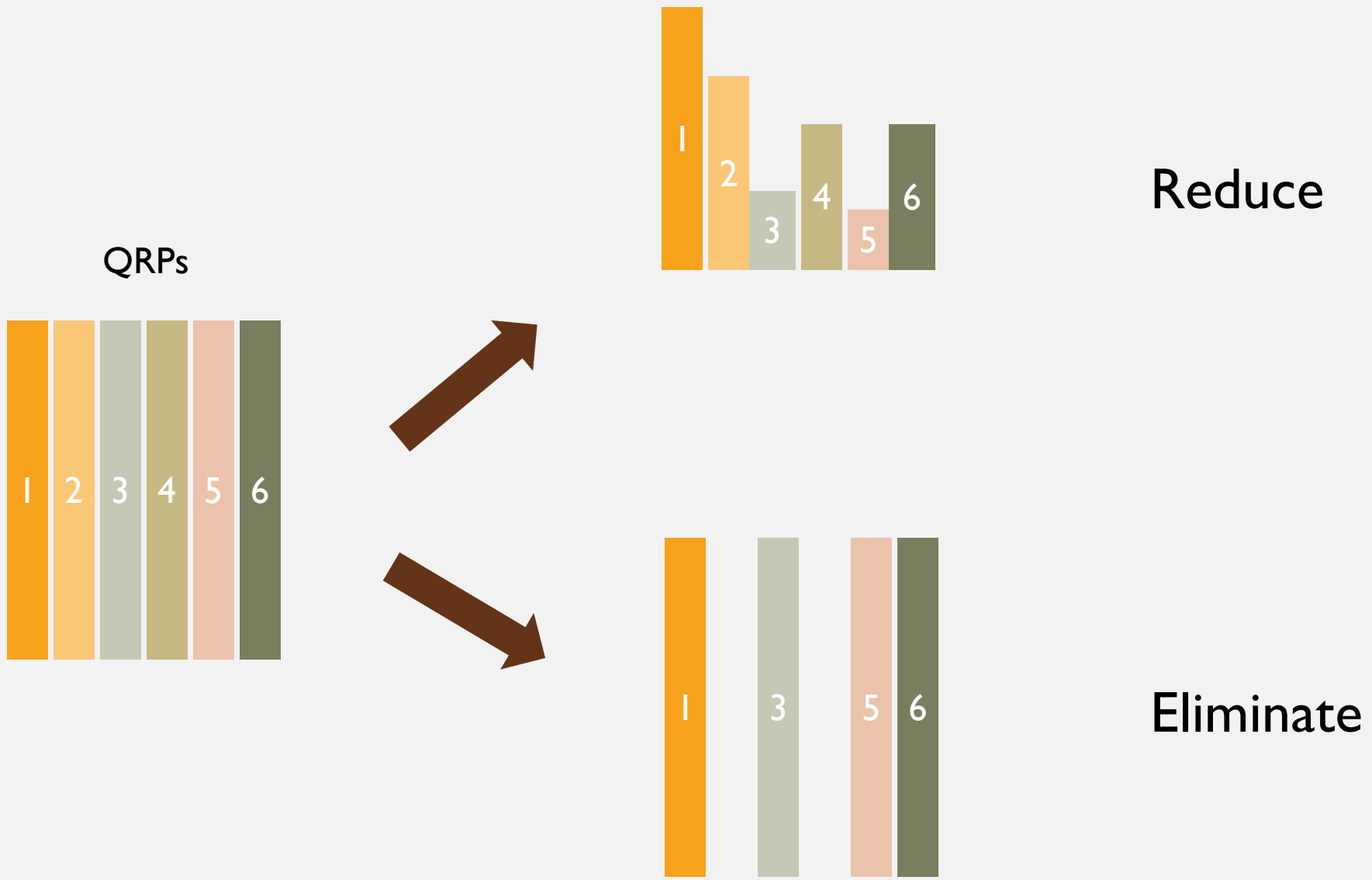
- Also becomes a tool to inform and guide debate about policy interventions
- Neat distinction promises neat solutions…

# SAVING SCIENCE




FFP: Falsification, fabrication and plagiarism

# SAVING SCIENCE


QRPs
1
2
3
4
5
6
1
2
3
4
5
6
Reduce
1
3
5
6
Eliminate

# WHAT SHOULD WE TARGET?

- Still no consensus on "the" list of QRPs within scientific community or among scholars on research integrity
  - Bouter et al. (2016): List of 60 QRPs
  - Ravn and Sørensen (2021): Interviews with 105 scientists results in list of 107 QRPs, then reduced to 34 by Ravn and Sørensen

# BUT DOES VAGUENESS MATTER?

- As long as we have *some* clear examples, surely we can live with vagueness/lack of neatness?
- Well-established examples:
  - Lack of controls (Begley 2[illegible] and Sørensen 2021)
  - P-hacking
  - HARKi[illegible]
  -

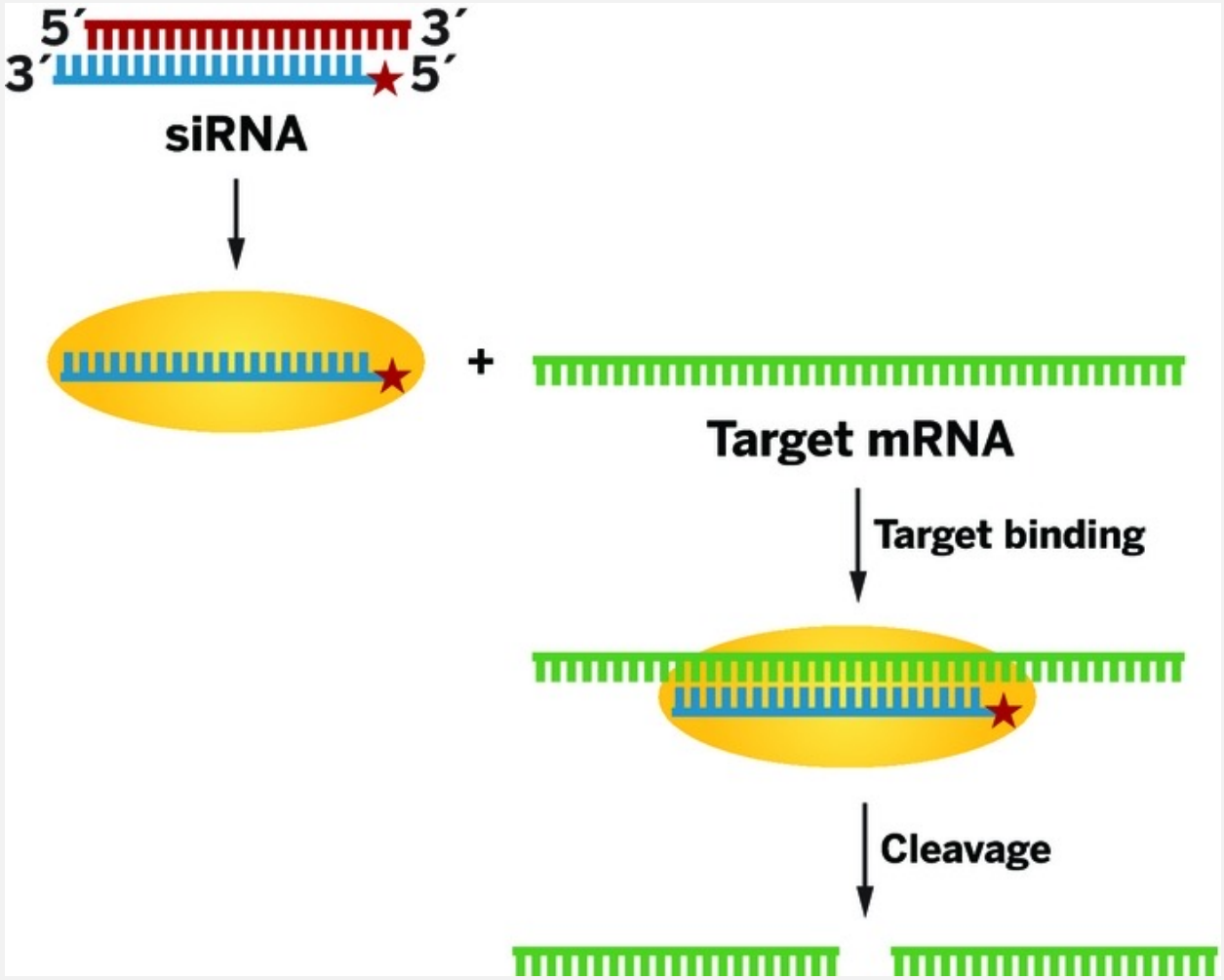
Eliminate/reduce these methodological "vices"?

# KEY QRP: LACK OF EXPERIMENTAL CONTROLS

- Enforce that only experiments with ”appropriate” controls are produced and published
  - Compulsory refresher courses on experimental design for all researchers? (Begley and Ioannidis 2015)
  - Monitor and reward investigator-compliance with peer-generated guidelines? (Begley and Ioannidis 2015)
  - Change in editorial practice, etc.

→ Enforce standardized landscape of controlling practices

RNA interference (RNAi)




https://cen.acs.org/articles/91/i46/Gene-Silencing-Design.html

siRNA targets specific mRNA for destruction → Depletes related protein from cell

Creates "mutant" cell → Study effects of absence

Negative control: same treatment but not targeting gene X (exclude artefacts of intervention per se)

Solution: "scrambled" siRNA. Same 21 nucleotides but in random sequence → Same chemical composition, but will not target X

All non-targeting siRNAs should have same effect (due to presence of reagents and manipulation)

But: Different non-targeting siRNAs show different effects on gene expression

Researchers started to leave out negative control as not "negative"

# NOT A QRP…?

- Enforcing use of standardized controls is problematic – some experiments (RNAi screens) could not be published
  - Lack of control is not QRP – *use of a control can be questionable*!
  - No fixed classification: What counts as QRP in one experiment is GRP in other context
- Contextual instability of classification within disciplines

# ANOTHER EXAMPLE: FAR WESTERN BLOT

- Assay used to study protein-protein interactions
- Very powerful, easy/cheap to use
- But: Uses harsh conditions that destroys fold of proteins analysed
- Need to "re-fold" proteins before interaction tested → Difficult and error-prone process → Non-specific binding events because of unfolded proteins
- Assay long shunned by researchers and reviewers as questionable practice (high false positive rate)

# ANOTHER EXAMPLE: FAR WESTERN BLOT

- 1990/2000s: Discovery of proteins that don't need to be folded: "Intrinsically disordered proteins" (IDPs)
  - 30%-50% of all proteins
- Far Western Blot rehabilitated? – Unfolding/refolding problem is not relevant for proteins that never fold!
  - Assay works, at least for some proteins..?
- What counted as QRP has turned into potential GRP
- Diachronic instability of classification within discipline

# RAVN AND SØRENSEN (2021)

- Empirical finding: What scientists count as QRP in one discipline is seen as GRP in other discipline(s) (or even within one research area)
- "not only is there no unity in the creation and justification of science […], there is neither a unity in QRPs"

- Instability of classification within and across individual research areas

# BUT WHAT ABOUT THE "CLASSICS"?

P-hacking, HARKing etc. – surely these are *wrong*?

- Head et al. (2015); Ulrich and Miller (2020): p-hacking does limited/no damage to replicability (in some contexts)
- Rubin (2022): Premature to think HARKing is contributing to low replication rate
- Fanelli (2018): Presence of QRPs does not distort overall scientific output
- Bird (2018): High base rate of false hypotheses is key problem, not lack of quality in research

# BUT WHAT ABOUT THE "CLASSICS"?

~~P-hacking, HARKing etc. – surely these are *wrong*?~~

Cannot make *general* statement about the "goodness" or "badness" of these practices

# QUESTIONABLE QRPs…?

Even for those "QRPs" on which there is *some* consensus:

1. Diachronic instability: Former QRP becomes classified as GRP and vice versa (Far-Western blot example)
2. Synchronic instability: QRP can also be a GRP (and vice versa), within or across disciplines
3. Causal inertness: Presence of "established" QRPs does not have necessarily detrimental effect on results

# SUMMARY:

1. Hard to point to individual practices that are inside or outside of the class of QRPs

   *Label does not serve as a good guide to complex and shifting landscape of experimental practice (empirically inadequate)*

2. Label is used in universalising manner and has become supercharged ("the" detrimental practices)

   *Encourages blanket exclusions of certain practices and homogenised methodological landscape*

We have a label that is an inaccurate guide but which wields great normative power
→ Potentially damaging to science → Abandon talk of "QRPs"

# A WARNING FROM 30 YEARS AGO

NAS report 1992:

“Governmental or regulatory efforts to define "correct" research conduct or analytical practices can do fundamental harm to research activities if such efforts encourage orthodoxy and rigidity and inhibit novel or creative research practices” (p. 11).

# SUMMARY: WHERE NOW?

- Develop approach to "integrity" that focuses on local values and epistemic cultures (see e.g., Penders et al. 2009; Valkenburg et al. (2021); Ravn and Sørensen (2021)
  - Abandon universalising tendencies and look at local limitations (and potential) of different practices/cultures
  - Stop resorting to moral panic about "the" dishonest researcher and QRPs (see also Bird 2018)
  - Mobilise more pluralist approach to tackle issue of questionable research – empirical, conceptual, historical,…

# BROADER IDEAS ABOUT "QUESTIONABLE" PRACTICES

- Begley and Ioannidis (2015): Focus more on hypothesis-driven research (example of pharmacology)
- Sarewitz (2016): Reduce non-applied research and favour applied research instead, to keep scientists honest

→ Shift funding practices to control what researchers are doing

# IMPORTANCE OF EPISTEMIC DIVERSITY

- Plurality (of models and explanations) in biology required for effective science (Mitchell 2002)
- Variety of methods (e.g., exploratory vs hypothesis-driven) underlies major scientific advances (Steinle 1997; O'Malley et al. 2010; see also Penders et al. 2009)
- Drive for standardization and homogenization potentially dangerous for scientific research and its underlying epistemic diversity (Leonelli 2018; 2021; Malich and Rehmann-Sutter 2022)

# THANK YOU FOR YOUR ATTENTION!

References used in my talk: Please Email me

s.m.guttinger@exeter.ac.uk

University of Exeter

egenis