R³ISE Across the Disciplines: Questionable Research Practices - explained and applied

Gundula Bosch (PI), Sneh Sonaya, Ilinca Ciubotariu, Marsha Wills-Karp, Arturo Casadevall

R3 Center for Innovation in Science Education (R³ISE), Johns Hopkins University, Bloomberg School of Public Health

Advancing Innovations in Graduate Education
NSF-IGE Principal Investigators' Meeting
Hosted by the Council of Graduate Schools

Alexandria, VA
February 29 - March 1, 2024

Introduction

- Much has been published on the pedagogy of research ethics education. However, **many courses focus on contrasting flawless ethical practices with cases of misconduct or fraud**. The **poorly defined grey zone in between, the so-called "questionable research practices"** (NASEM 2017), are rarely taught formally. Yet, they play an important role for ethical decision making in science.

- Our project aims to fill that gap by producing and evaluating the **educational module "R³ISE across the disciplines"** (Research Integrity training supporting Rigorous, Responsible and Reproducible Practice in Science and Engineering). It **uses an error analysis approach and produces an openly accessible library of case scenarios**.

- The goal of this project comprises assessing the learning outcomes of the R³ISE module in a student population spanning the life and social sciences, as well as engineering and technology.

- We appraise students' **critical thinking** (conceptual understanding, application, and evaluative reasoning skills) **and communication skills** around analyzing scientific errors in general, and questionable research practices in particular.

R³ISE module template from "Anatomy of Scientific Error" course

Rigor
Reproducibility
Responsibility

Theory, History & Practice of Science

Science is **EVIDENCE** not opinion

Scientific Logic & Error Analysis

Practical Ethics

Social Justice in Scientific Practice

Communication

Methods & Innovation

Quantitative Reasoning & Data Science

Professional Development

CAREERS

ANATOMY OF SCIENTIFIC ERROR

- Concepts and applications **of errors in scientific practice**
- Ways to **reduce errors** in science
  - **Open science**
  - **Responsible science communication**
- **Serendipity, triumphs and disasters**

Questionable Research Practices (QRPs) in focus

POPULAR NAME OF QRP	DESCRIPTION
Cherry-picking	Selectively reporting studies / data that "worked"
Stopper	Stopping data collecting earlier than planned because one found the result that one had been looking for
Selective omission	Deciding whether to exclude data after looking at the impact of doing so on the results.
No surprise	Reporting an unexpected finding as having been predicted from the start
"Rounding off"	Reporting a p value of (e.g.) .054 as "significant"
Lack of transparency	Failing to report all of a study's conditions
HARKing	Hypothesizing After Results Known: Constructing or changing a hypothesis after the data have been collected and analyzed.
p-hacking	Data-dredging, significance-chasing, trying multiple things until you get the desired result

"Innocent" errors – Questionable research practices – Misconduct?

Publish and/or conduct next experiment
Publication bias

Generate and specify hypothesis
Failure to control for bias

HARKing

Design study
Low statistical power

Interpret results
P-hacking

Conduct study and collect data
Poor quality control

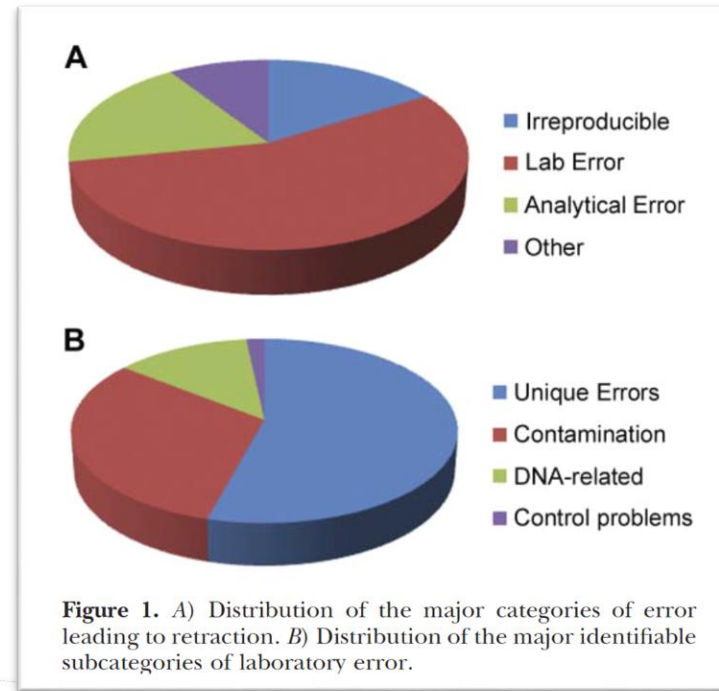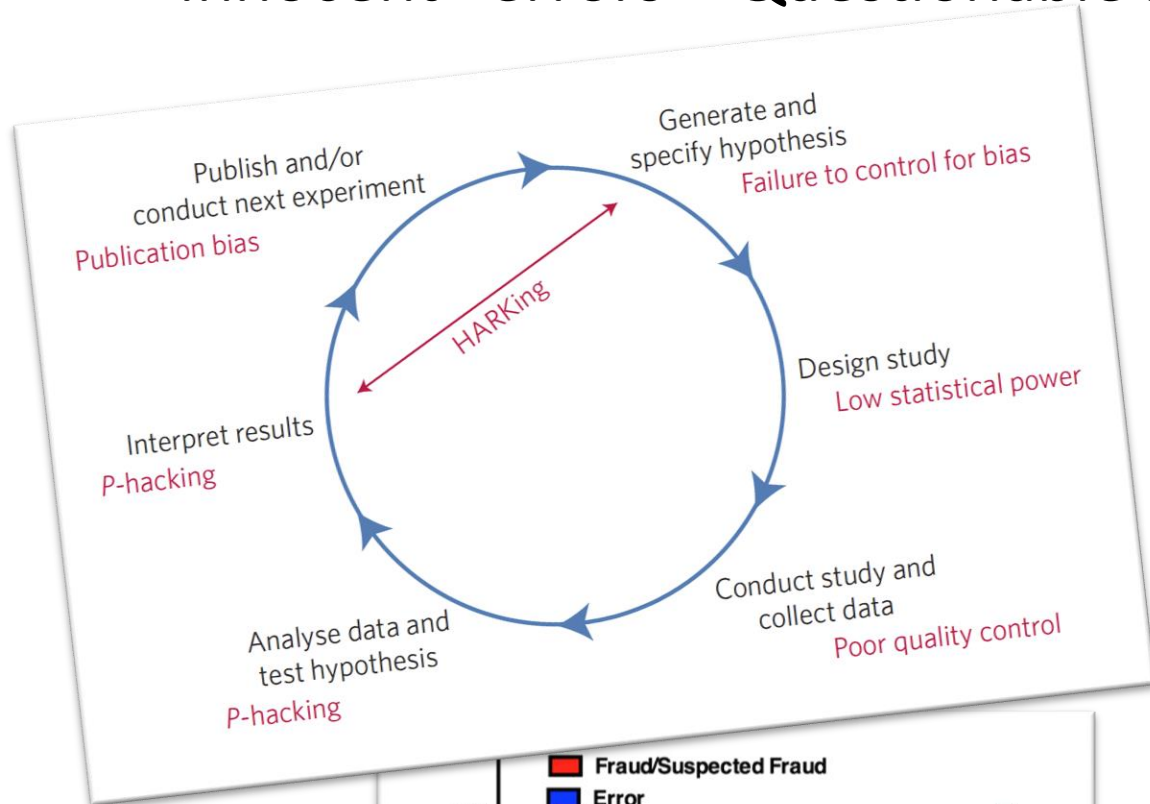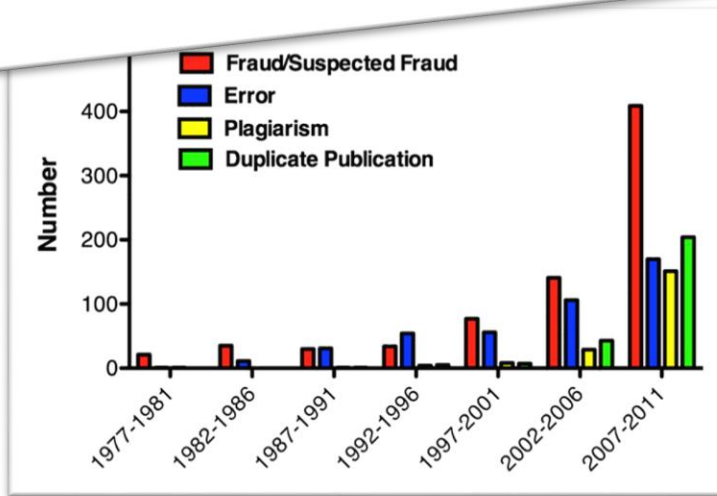Analyse data and test hypothesis
P-hacking

Figure 1. A) Distribution of the major categories of error leading to retraction. B) Distribution of the major identifiable subcategories of laboratory error.

A)
- Irreproducible
- Lab Error
- Analytical Error
- Other

B)
- Unique Errors
- Contamination
- DNA-related
- Control problems

- Fraud/Suspected Fraud
- Error
- Plagiarism
- Duplicate Publication

HYPOTHESIS MYOPIA	TEXAS SHARPSHOOTER	ASYMMETRIC ATTENTION	JUST-SO STORYTELLING
Looking one-sidedly for data to support, but not falsify a hypothesis	Overinter-preting patterns into data that aren't there	Rigorous re-testing of unexpected, but not of expected data	Storytelling to accommodate all data, or modify hypotheses

Our Goal

The <u>overall goal</u> of the project is to determine if the educational R^{3}ISE module (teaching common questionable research practices) is effective

- Upon implementation in an interdisciplinary course on Responsible Conduct of Research (RCR)

- By offering it to a greater variety of STEM disciplines than before

- By teaching a significantly larger learner population

- After transformation to a modular implementation format, usable in both on-site and online courses
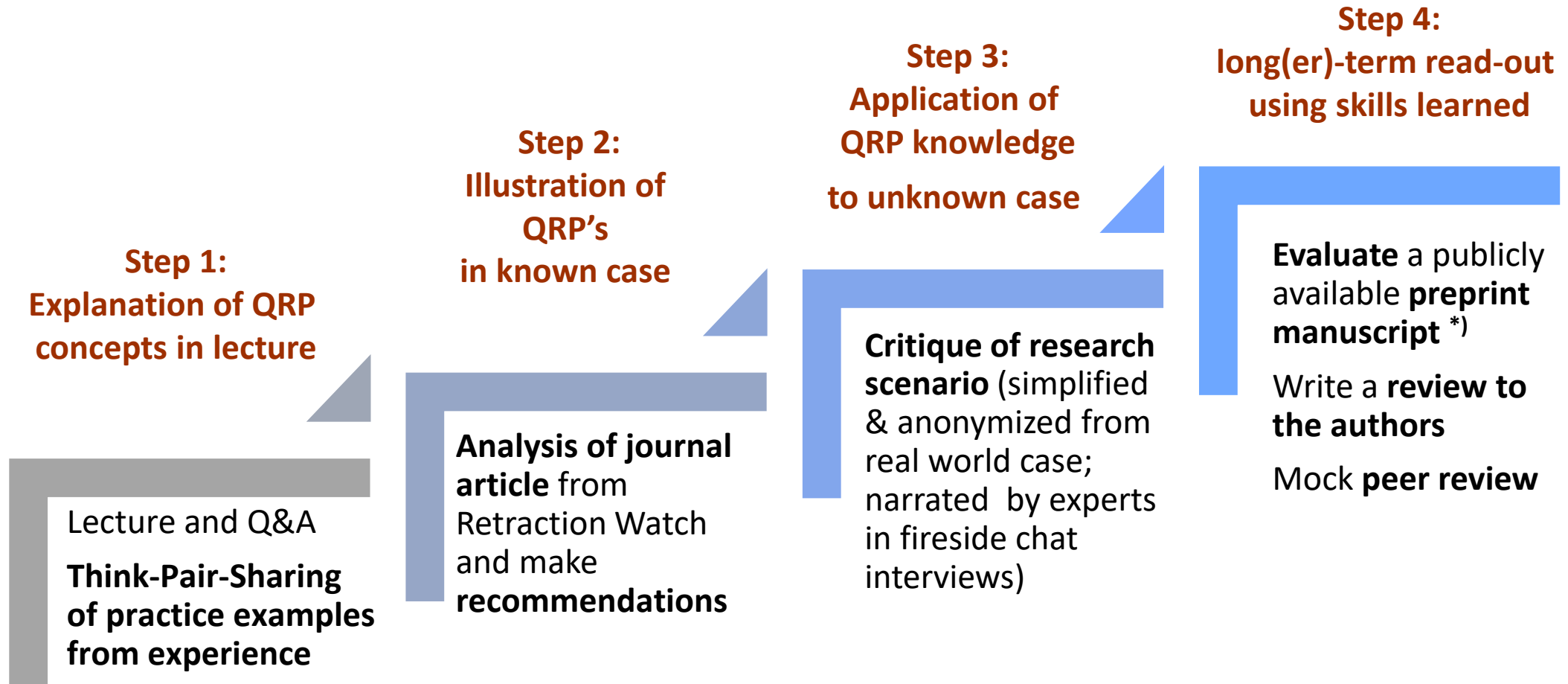(a flexible adaptation owing to the COVID-19 pandemic situation).

Our objectives

- We proposed to **assess how the learning outcomes of students** in the "Responsible Conduct of Research" course at Johns Hopkins University **who completed the R³ISE module**, will compare historically to students who did not, with respect to their:

  - **Conceptual understanding**
    => Measured by a concept inventory on basic ideas and frameworks regarding questionable research practices

  - **Application skills**
    => Assessed by checklist-evaluated case analyses of retracted research articles

  - **Short-term evaluative reasoning skills**
    => Rubric-assessed preprint evaluation capacities (course project)

  - **Long-term evaluative reasoning capacity**
    => Observation-guide measured communication and evaluative reasoning performance during a discussion panel at an on-site symposium or during journal club discussions

Our approach - pedagogy

Step 1:
Explanation of QRP concepts in lecture

Lecture and Q&A

Think-Pair-Sharing of practice examples from experience

Step 2:
Illustration of QRP's in known case

Analysis of journal article from Retraction Watch and make **recommendations**

Step 3:
Application of QRP knowledge to unknown case

Critique of research scenario (simplified & anonymized from real world case; narrated by experts in fireside chat interviews)

Step 4:
long(er)-term read-out using skills learned

Evaluate a publicly available **preprint manuscript** *)

Write a **review to the authors**

Mock **peer review**

Our approach – assessments - conceptualizing

Concept inventory

Step 1:
Explanation of QRP concepts

Lecture, peer instruction, Q&A

**Think-Pair-Sharing of practice examples from experience**

- Originally intended as individual pre-post quiz, it was found not meaningful in that modality due to different levels of pre-existing disciplinary experience (engineers and social sciences, versus most of life and natural sciences)

- Various course adjustments due to the Covid-19 pandemic made it necessary to **integrate the concept assessment into a peer instruction element** during class, and **administer the actual concept inventory test as an open book, online post-test** for self-assessment purposes

- The concept assessment is used is two different formats:

  - During the "think-pair share" phase in class, **student teams try their hands on formulating definitions for the QRP concepts** in a peer-instruction activity including practice examples and sharing those in their own words.

  - The exercise is followed by an interactive lecture component on the topic.

  - After class students complete an **individual term-matching quiz** on the QRP concepts (including confounding answer choices)

Our approach – assessments – recognizing

Step 2:
Illustration of QRP's in known case

Description of journal article from **Retraction Watch** *) and make **recommendations**

*) RetractionWatch.com
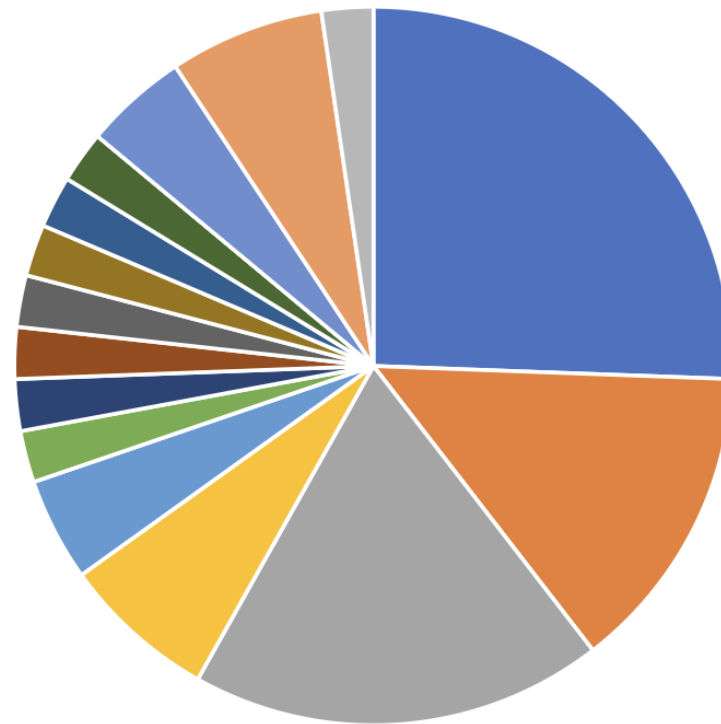
Frequency of reported QRP category

- Cherry-Picking
- Selective Omission
- Hypothesis Myopia
- Looking One-Sidedly for Data to Support or Falsify a Hypothesis
- Claiming Results Unaffected by Confounding Variables
- The Texas Sharpshooter
- Just-so storytelling
- P-Hacking
- Lack of Transparency
- Falsification/Fabrication of Data
- Reporting Unexpected Findings as Predicted
- Plagiarism
- Asymmetric Testing
- HARKing

Our approach – assessments – recognizing (c'd)

**Step 2:
Illustration of QRP's
in known case**

Description of journal article from Retraction Watch *) and make **recommendations**

Department / Program	QRP	Retracted / to be retracted article	Research Field
Applied Mathematics	Cherry-Picking	Criminal tendency detection from facial images and the gender bias effect. JoBD, journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0282-4	Machine Learning
Biomedical Engineering	Selective Omission	Safety and efficacy of favipiravir versus hydroxychloroquine in management of COVID-19: A randomized controlled trial. Nature, PMID: 33790308	Clinical Research
Biophysics	Cherry-Picking	Binding of DEAD-box helicase Dhh1 to the 5'-untranslated region of ASH1 mRNA represses localized translation of ASH1 in yeast cells; JBC, PMID: 28450395	Molecular Biology
Computer Science	Myopia	Timing Matters: When High-Performance Work Practices Enable New Venture Growth and Productivity JoM, 10.1177/0149206318763581	Artificial Intelligence
Chemistry	HARKing	Visfatin: A protein secreted by visceral fat that mimics the effects of insulin (PMID: 1560436, Science)	Biochemistry

Selected examples of retracted articles due to QRP's

Our approach – assessments – applying

Case scenario
(sample)

Step 3:
Application of QRP knowledge to unknown case

Critique of research scenario (simplified & anonymized from real world case; narrated by experts in fireside chat interviews)

Background: Autism affects an estimated 1 in 36 children in the United States today. This study looked at the timing of the first measles-mumps-rubella (MMR) vaccine given to infants and its hypothesized connection to autism.

Methodology: Researchers reanalyzed data originally used in a 2004 study by the CDC. They focused on children born in low-income families from 1986 to 1993 in metropolitan Atlanta out of an overall study population of children born 1986 to 2000. The study checked if there was a link between the age at which kids got their first MMR vaccine and being diagnosed with autism.

Results: Results suggested that particularly among African American boys who received the first MMR vaccine before 24 months of age, there was a higher chance of autism. The study concluded that there is a potential link between early MMR vaccination and autism diagnosis in African American boys.

Measles-mumps-rubella vaccination timing and autism among young African American boys: a reanalysis of CDC data

Demographic Group	Autism Diagnosed	p-value
African American Boys	Yes	0.045
Caucasian Boys	No	0.784
African American Girls	No	0.091
Caucasian Girls	No	0.423
Hispanic Boys	No	0.321
Hispanic Girls	No	0.289

Our approach – assessments – applying (c'd)

Case scenario
(sample)

Step 3:
Application of
QRP knowledge
to unknown case

Critique of research scenario (simplified & anonymized from real world case; narrated by experts in fireside chat interviews)

Questions:

What is/are the questionable research practice(s) (QRP) that can be identified in this abstract?

- o **Stopping** - ending the analysis before the entire database is covered
- o **P-Hacking** – significance chasing
- o **Cherry Picking** – selective focus on one particular group, potentially based on pre-existing bias
  - ▪ => **Hypothesis myopia** is also accepted as answer here
- o **Lack of transparency** (missing information about sample sizes, error margins, etc)
- o **Overstating** / inappropriate generalization

Our approach – assessments – narratives

Critique of research scenario (simplified & anonymized from real world case; narrated by experts in fireside chat interviews)

Part I:
Fireside Chat with
Dr. Elizabeth Stuart

Ilinca Ciubotariu, Ella Jacobs, and Georgia Stavrakis interviewing:

Elizabeth Stuart, PhD
Associate Dean for Education, Professor, JHSPH
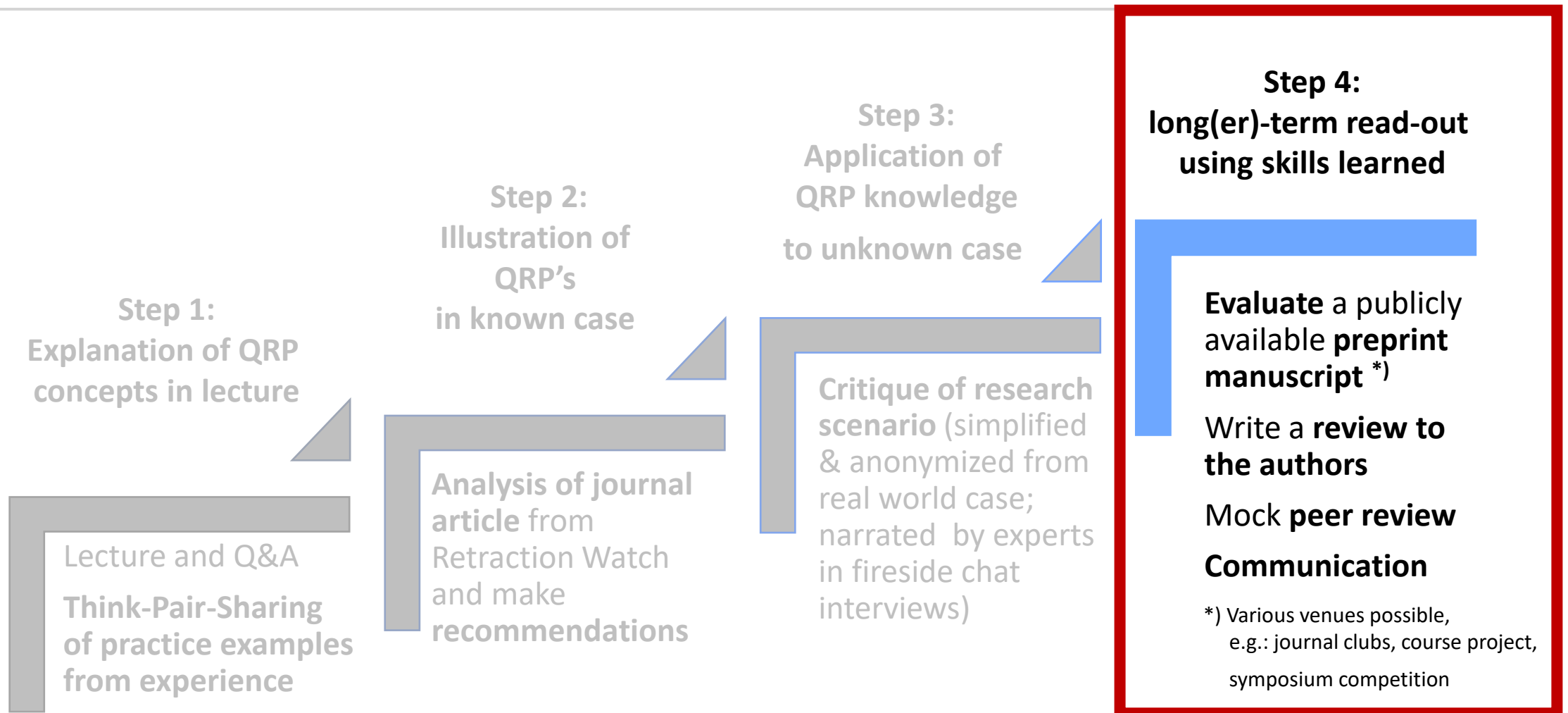Mental Health, Biostatistics, and Health Policy and Management

Part II:
Fireside Chat with
Dr. John McGready

Ilinca Ciubotariu, Ella Jacobs, and Georgia Stavrakis interviewing:

John McGready, PhD
Senior Scientist, JHSPH
Department of Biostatistics

Our approach – Longer term read outs

Step 1:
Explanation of QRP concepts in lecture

Lecture and Q&A

Think-Pair-Sharing of practice examples from experience

Step 2:
Illustration of QRP's in known case

Analysis of journal article from Retraction Watch and make **recommendations**

Step 3:
Application of QRP knowledge to unknown case

Critique of research scenario (simplified & anonymized from real world case; narrated by experts in fireside chat interviews)

Step 4:
long(er)-term read-out using skills learned

Evaluate a publicly available **preprint manuscript** *)

Write a **review to the authors**

Mock **peer review**

Communication

*) Various venues possible, e.g.: journal clubs, course project, symposium competition

References, resources, and acknowledgements

- The National Academies of Sciences, Engineering, and Medicine (NASEM) 2017. Fostering Integrity in Research. Washington, DC: The National Academies Press.

- The Retraction Watch Database. New York: The Center for Scientific Integrity. 2018. ISSN: 2692-465X.
Available from: http://retractiondatabase.org/

- Ciubotariu II, Bosch G. Improving research integrity: a framework for responsible science communication.
BMC Res Notes. 2022 May 15;15(1):177. doi: 10.1186/s13104-022-06065-5. PMID: 35570294

- GitHub repository: https://github.com/JHU-R3ISE

- Online Ethics Center for Engineering and Science (2024).
https://onlineethics.org/