Framework for Open and Reproducible Research Training

FORRT
### FORRT sale’s pitch

<table>
<thead>
<tr>
<th>The enemy</th>
<th>Why now?</th>
<th>The promised land</th>
<th>Obstacles</th>
<th>How?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bad Science</td>
<td>Credibility Crisis</td>
<td>Scientific Nirvana</td>
<td>Research Training</td>
<td>FORRT</td>
</tr>
</tbody>
</table>

---

**FORRT**

---
The Enemy

Questionable Research Practices (QRPs)
Q-Measurement-Ps
P-Hacking
Data Hogging
HARKing
Publication bias
Nefarious Incentives
Underpowered studies
Why now?

- Reproducibility Crisis
- Questionable Research
- Misaligned Incentives
- Re-evaluation
Why Now? Initiatives!
The Promised Land

The Psychological Science Accelerator (Chartier, 2016-Forever)
Becoming Cumulative Science (Mischel, 2009)
Obstacles

Current teaching and mentoring practices
Obstacles [1]

Teaching subject-matters without:
- probabilistic uncertainty
- research design
- Samples (type & quality)
- Measurements
Obstacles [2]

Thesis Supervision
- Questionable Research Practices (QRPs)
- Q-Measurement-Ps
- P-Hacking
- Data Hogging
- HARKing
- Publication bias
- Nefarious Incentives
- Underpowered studies
HOW?

Framework for Open and Reproducible Research Training

FORRT
Question

If science is a **process** of knowledge production, then is science **education** best **expressed** as teaching students the **process** or as teaching them the **knowledge itself**?
Teaching students the **accumulated knowledge**, then we are **not** actually teaching them science. Rather, we are teaching them science’s **products**, and indeed we are **misleading** them by substituting the teaching of scientific facts, **as if it were the teaching of science itself**

(Marks, 2009, p. 22)
FORRT:

Science education inevitably entails learning about how scientist learn what they know.
FORRT

KEY IDEAS

Purpose of Higher-ED
Epistemic Uncertainty
Ethics
Social Justice
What is Purpose of Higher-Education?

Insofar higher education institutions aim to prepare students for an increasingly technological workforce whose advances affect everyone’s lives in ever shorter time spans, its training ought to provide the compulsory knowledge for practical adjudication of scientific output.
Inasmuch as the verity of quantitative scientific findings hinges on probabilistic uncertainty, research design, measurements/instruments, sampling methods & representativeness, we should communicate the facts of science relative to the process by which it was acquired.
The Ethics

Merging the teaching of substantive topics with open and reproducible means to abide by principled teaching. E.g., to omit educating students about the replication crisis results in a false sense of certainty, which can be thought as misleading, if not unethical.
Maximize every student’s likelihood of present and future engagement with resources, facilitate the acquisition of knowledge and bolster opportunities that would otherwise be inaccessible to disadvantaged individuals.
The **why** of **open** in **teaching/mentoring**

---

**Social Justice**[2]

In science, **wealthy and elite-educated** individuals are **over-represented**, as are **males, whites, and citizens of western industrialized rich countries**.
The *why* of open in *teaching/mentoring*

*Social Justice*[3]

The current model of scientific production and teaching practices *reproduces global inequalities*. And as science is built on the same foundations of society itself, *it inherits* many of its *systematic barriers* hindering the success of traditionally *marginalized groups*. 
The why of open in teaching/mentoring

____________

Social Justice[4]

To help mitigate the detrimental effects in the access, learning and production of scientific content, it is paramount to create the conditions for knowledge to become a public good, accessible to all members of society.
Summary

Teaching and Mentoring of reproducible and open research practices is the **clearest indicator** of the degree to which institutions and/or departments **embody principles of credible science.**
Consequences of FORRT

Future Consumers of science

Future scholars

Collaborative Crowdsourced Citizen
Framework for Open and Reproducible Research Training

FORRT
FORRT

Educational Nexus

- Curated
  - Reproducibility and replicability knowledge
  - Conceptual and statistical knowledge
  - Reproducible analyses
  - Preregistration
  - Open data and materials
  - Replication research

- Crowdsourced
  - As above
FORRT Assessment

6 core principles

- Reproducibility and replicability knowledge
- Conceptual and statistical knowledge
- Reproducible analyses
- Preregistration
- Open data and materials
- Replication research
<table>
<thead>
<tr>
<th>I. Reproducibility Crisis and Credibility Revolution</th>
<th>II. Conceptual and Statistical Knowledge</th>
<th>III. Reproducible analyses</th>
<th>IV. Open data and materials</th>
<th>V. Preregistration</th>
<th>VI. Replication research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attainment of a grounding in the motivations and theoretical underpinnings of reproducible and open research. Integration with field specific content (i.e., or grounded in the history of replicability)</td>
<td>Enacting this principle indicates that students attain a grounding in fundamental statistics, measurement, and its implications.</td>
<td>Reproducible analyses allow the checking of analytic pipelines and facilitate error correction. Enacting this principle requires students to move towards transparent and scripted analysis practices.</td>
<td>Enacting this principle indicates that students have attained a grounding in open data and materials in both; using and sharing.</td>
<td>Preregistration entails laying out a complete methodology and analysis before a study has been undertaken. This facilitates transparency and removes several potential QRPs.</td>
<td>Replication research takes a variety of forms, each with a different purpose and contribution. Reproducible science requires replication research.</td>
</tr>
</tbody>
</table>

i. Reproducibility crisis and credibility revolution.

ii. Exploratory and confirmatory analyses.

iii. Questionable research practices (its ‘theory’), and prevalence.

iv. Proposed improvement science initiatives on statistics, measurement, teaching, data sharing, code sharing, preregistration, replication.

v. Ongoing debates, (e.g. incentives for and against open science).

vi. Ethical considerations for improved practices.

i. The logic of null hypothesis testing, p-values, Type I and II errors (and when and why they might happen).

ii. Limitations and benefits of NHST, Bayesian and Likelihood approaches.

iii. Effect sizes, Statistical power, Confidence Intervals.


v. Questionable research (QRP)s & measurement practices (QMP)s.

vi. Understand the relationship between all of the above.

i. Strengths of reproducible pipelines.

ii. Scripted analyses compared with GUI.

iii. Data wrangling.

iv. Programming reproducible data analyses.

v. Open source and free software.

vi. Tools to check yourself and others; statcheck, GRIM, and SPRITE.

i. Knowledge of traditional publication models. Open access publishing, preprints.

ii. Reasons to share; for science, and for one’s own practices.

iii. Repositories; e.g. OSF, FigShare, GitHub.

iv. Accessing/sharing others data, code, and materials.

v. Ethical considerations.

vi. Examples and consequences of accessing un/open data.

i. Purpose of preregistration - distinguishing exploratory and confirmatory analyses, transparency measures.

ii. Preregistration and registered reports - strengths and differences.

iii. When can you preregister? Can you pre-register secondary data?

iv. Writing a preregistration.

v. Comparing a preregistration to a final study manuscript.

vi. Conducting a preregistered study.

i. Purposes of replication attempts - what is a ‘failed’ replication?

ii. Large scale replication attempts

iii. Distinguishing conceptual and direct replications

iv. Conducting replication studies; challenges, limitations, and comparisons with the original study

v. Registered Replication Reports

vi. The politics of replicating famous studies
FORRT
Assessment

**Breadth** describes how **widely** teaching is distributed

- **None.** Not yet enacted, minimal breadth, or no evidence
- **Some.** Opportunities for some.
- **Course requirement for all.**

**Depth** describes the **degree** to which students interact with the core:

- **None.** Not yet enacted, minimal depth, or no evidence
- **Knowledge.**
- **Practice.**
- **Application.**
## FORRT Assessment

<table>
<thead>
<tr>
<th>Depth:</th>
<th>None</th>
<th>Some</th>
<th>Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge</td>
<td>As part of an <strong>optional</strong> course/module, or workshop, or lecture given; attain knowledge of the importance of reproducible analyses, including data pipelines and scripting in R/Python/JASP/SPSS/etc.</td>
<td>As part of a <strong>required</strong> (assessed) course/module/dissertation project; attain knowledge of the importance of reproducible analyses, including data pipelines and scripting in R/Python/JASP/SPSS/etc.</td>
<td></td>
</tr>
<tr>
<td>Practice</td>
<td>As part of an <strong>optional</strong> course/module, or workshop, or lecture given; practice implementing reproducible data pipelines and analysis scripts with existing data.</td>
<td>As part of a <strong>required</strong> (assessed) course/module/dissertation project; practice implementing reproducible data pipelines and analysis scripts with existing data.</td>
<td></td>
</tr>
<tr>
<td>Application</td>
<td>As part of an <strong>optional</strong> course/module, or workshop, or lecture given; implement reproducible analytical pipelines in a research project.</td>
<td>As part of a <strong>required</strong> (assessed) course/module/dissertation project; implement reproducible analytical pipelines in a research project.</td>
<td></td>
</tr>
</tbody>
</table>
Where are we now?
FORRT needs you!

https://osf.io/bnh7p

https://osf.io/g29eu
Thank you