Longitudinal study of long term impact of birth trauma: Current best practice and future prospects in the era of Open Science

Carlo Schuengel

Plan

Open Science

Best practices research

Registration

Derivation chain

Effect size expectations

Overfitting issues

Code reproducibility

Strategy

CHAOS IN THE BRICKYARD

BRICKYARD by B.K. Forscher

3

https://massivesci.com/articles/chaos-in-the-brickyard-comic-matteo-farinella/

Richard Riley (R^2)
@Richard_D_Riley

The #academia career bootstrap:
1) Get existing data (call it 'big', ignore quality)
2) Form Qs based on variables recorded
3) Analyse until 'novel' findings
4) "Sensitivity analyses did not change results"
5) Publish
6) Inform comms team & media
7) Repeat 1-6 until tenure/chair

Tweet vertalen

10:33 a.m. · 28 feb. 2022 · Twitter Web App

2 Retweets   2 Geciteerde Tweets   15 Vind-ik-leuks

Maarten van Smeden
@MaartenvSmeden

Imagine how exciting every new analysis result would be if you didn't know about p-hacking, data dredging, confounding, colliders, measurement error, missing data, Simpson's paradox, table 2 fallacy, generalizability, transportability, model misspecification, robustness

Tweet vertalen

12:04 a.m. · 28 feb. 2022 · Twitter for iPhone

61 Retweets   5 Geciteerde Tweets   346 Vind-ik-leuks

Je antwoord tweeten   Beantwoorden

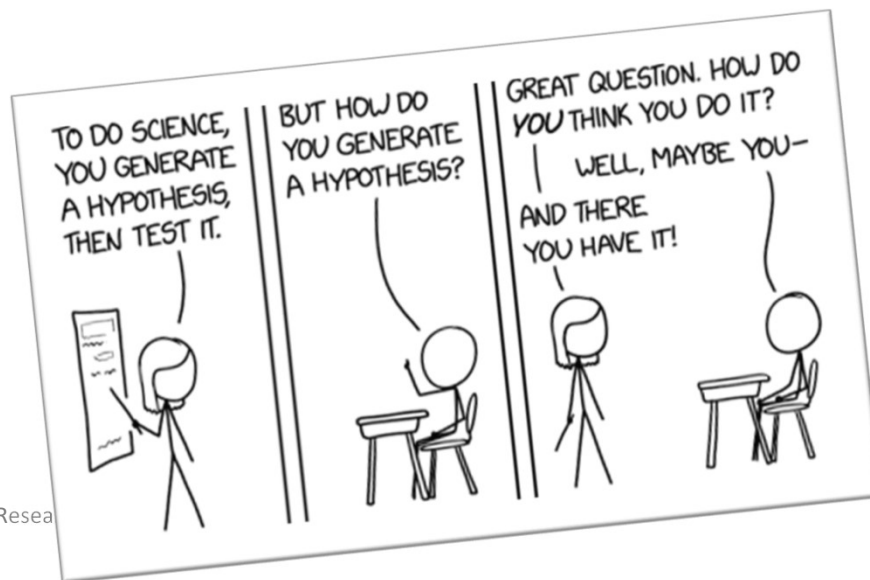**Maarten van Smeden** @MaartenvSmeden · 10 u
Als antwoord op @MaartenvSmeden
Type I and II errors, overfitting, sparse sample bias, winner's curse, non-collapsibility, ecological fallacy, competing risks, informative censoring, publication bias, spin, immortal time bias and anything about conditional probabilities
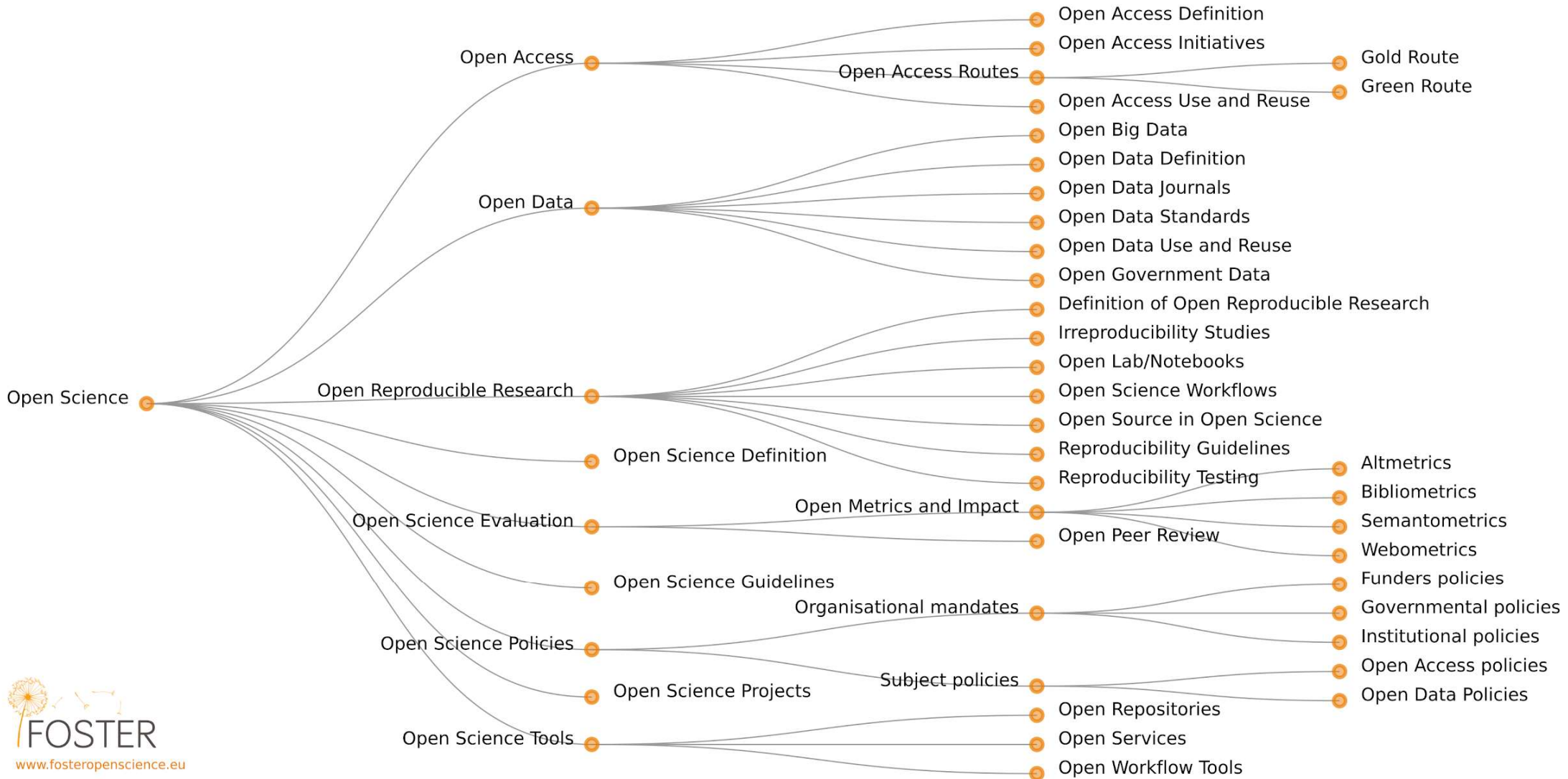
3   8   124

Open Science practices

Go to menti.com or use the mentimeter app
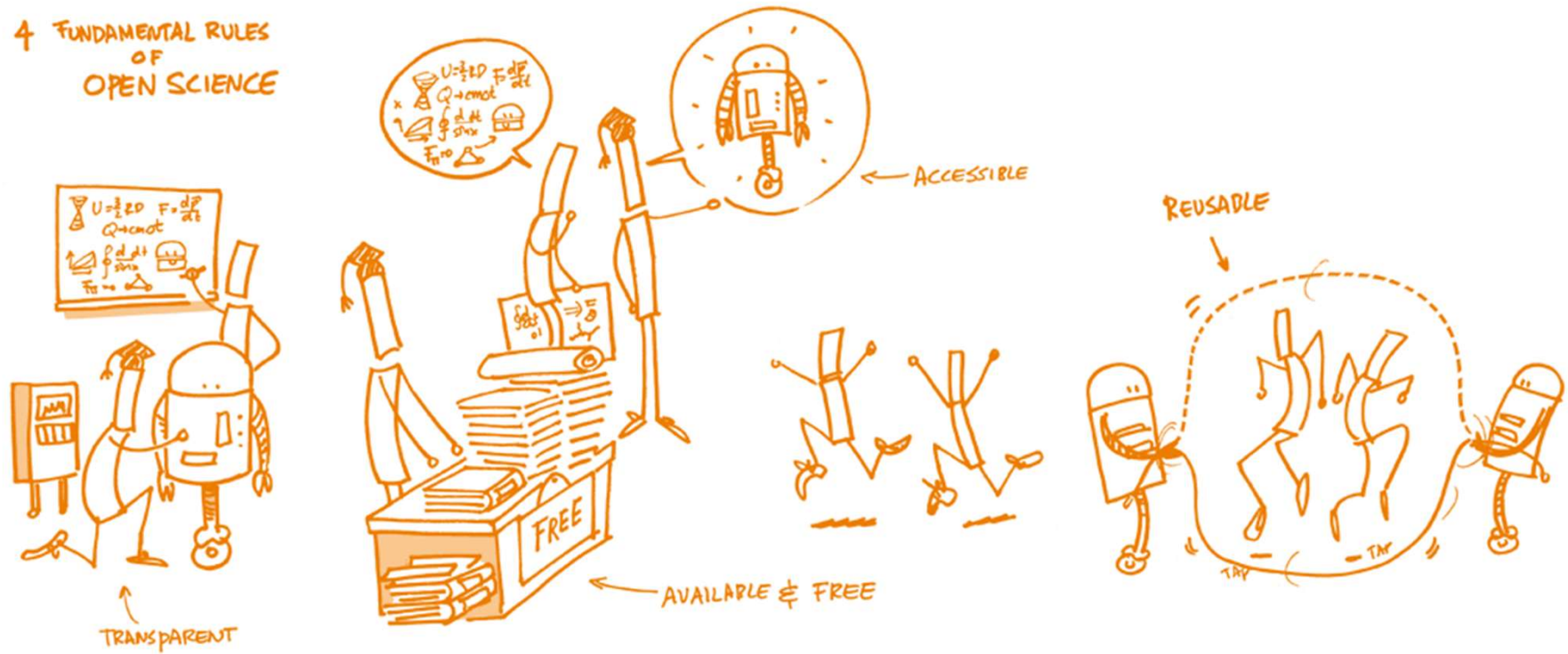
Open Science Taxonomy

Open Science

- Open Access
  - Open Access Definition
  - Open Access Initiatives
  - Open Access Routes
    - Gold Route
    - Green Route
  - Open Access Use and Reuse
- Open Data
  - Open Big Data
  - Open Data Definition
  - Open Data Journals
  - Open Data Standards
  - Open Data Use and Reuse
  - Open Government Data
- Open Reproducible Research
  - Definition of Open Reproducible Research
  - Irreproducibility Studies
  - Open Lab/Notebooks
  - Open Science Workflows
  - Open Source in Open Science
  - Reproducibility Guidelines
  - Reproducibility Testing
- Open Science Definition
- Open Science Evaluation
  - Open Metrics and Impact
    - Altmetrics
    - Bibliometrics
    - Semantometrics
    - Webometrics
  - Open Peer Review
- Open Science Guidelines
- Open Science Policies
  - Organisational mandates
    - Funders policies
    - Governmental policies
    - Institutional policies
  - Subject policies
    - Open Access policies
    - Open Data Policies
- Open Science Projects
- Open Science Tools
  - Open Repositories
  - Open Services
  - Open Workflow Tools

FOSTER
www.fosteropenscience.eu

Underlying schools of thought

Democratic
Pragmatic
Infrastructure
Public
Measurement

https://book.fosteropenscience.eu/

VU

4 FUNDAMENTAL RULES OF OPEN SCIENCE

TRANSPARENT

ACCESSIBLE

AVAILABLE & FREE

REUSABLE

https://book.fosteropenscience.eu/

Impact of open science

Go to menti.com or use the mentimeter app

Costs/challenges	Benefits/opportunities
Restrictions on flexibility: exploration vs planned hypothesis tests	Greater faith in research: registered reports guard against QRPs AND get published even with null-results
Time costs: archiving, documenting, quality control	New helpful systems: Github, Rmarkdown, Jupyter notebooks, OSF etc provide code, data, instructions, tools
No incentives: quantity still counts more than quality in mainstream science	Invest in future: mastering reproducible methods, collaboration increase market value

Allen C, Mehler DMA (2019) Open science challenges, benefits and tips in early career and beyond. PLOS Biology 17(5): e3000246. https://doi.org/10.1371/journal.pbio.3000246

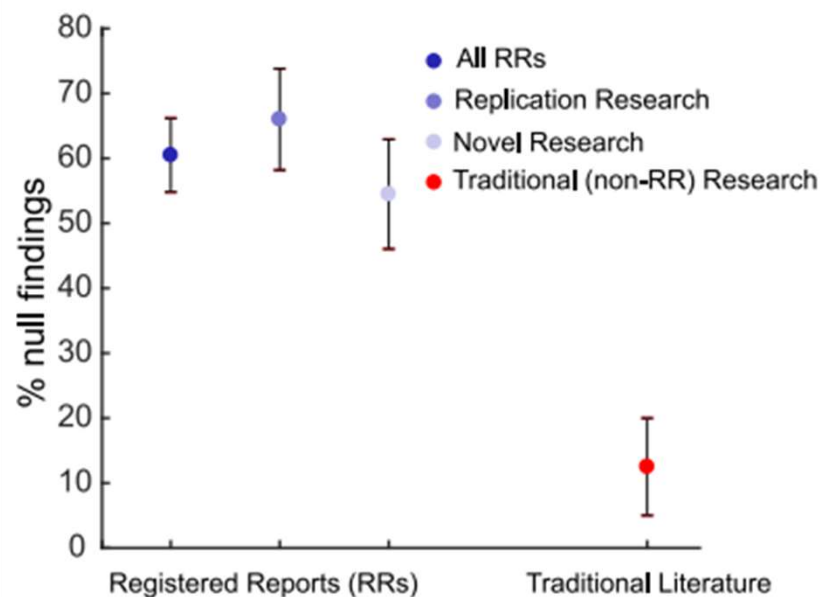| Costs/challenges	Benefits/opportunities

Percentage of null findings

Fig 1. Percentages of null findings among RRs and traditional (non-RR) literature [46,47], with their respective 95% confidence intervals. In total, we extracted $n = 153$ hypotheses from RRs that were declared as replication attempts and $n = 143$ hypotheses that were declared as original research. The bounds of the confidence intervals shown for traditional literature were based on estimates (5% and 20%, respectively) of null findings that have been previously reported for traditional literature [46,47]. Data is available on the Open Science Framework (https://osf.io/wy2ek/) and in S1 Data. RR, registered report.

https://doi.org/10.1371/journal.pbio.3000246.g001

Greater faith in research: registered reports guard against QRPs AND get published even with null-results

New helpful systems: Github, Rmarkdown, Jupyter notebooks, OSF etc provide code, data, instructions, tools

Invest in future: mastering reproducible methods, collaboration increase market value

Allen C, Mehler DMA (2019) Open science challenges, benefits and tips in early career and beyond. PLOS Biology 17(5): e3000246.
https://doi.org/10.1371/journal.pbio.3000246

11

11

Best practices research

Define and break down the process one seeks to improve

E.g.,
Recruitment
- Drawing attention
- Negotiate intermediairies
- Recruiting participants

Define what constitutes a best practice for each element

E.g.,

- Eyeballs
- Low selection bias
- Control for self-selection bias

Identify exemplars through peer nomination and audit

E.g.,

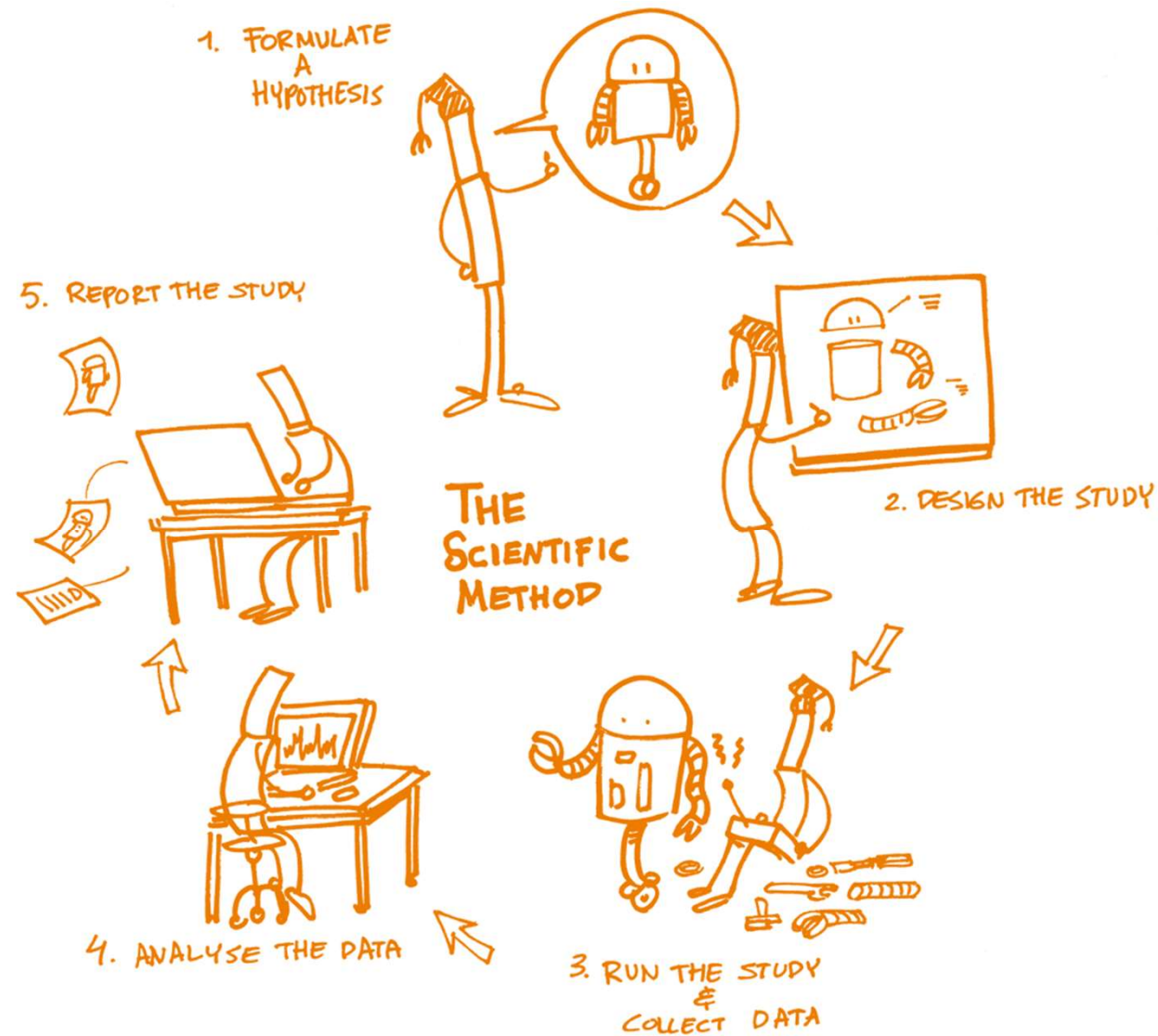- CAPTURE-group
- IMPACT Network
- MoBa

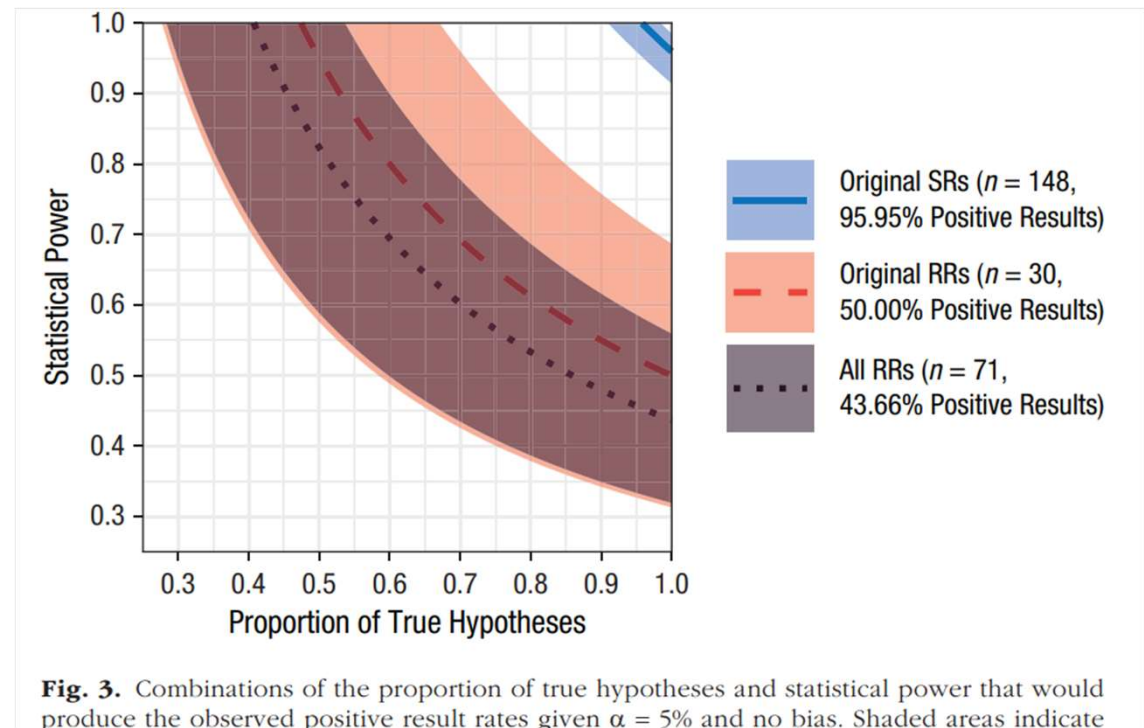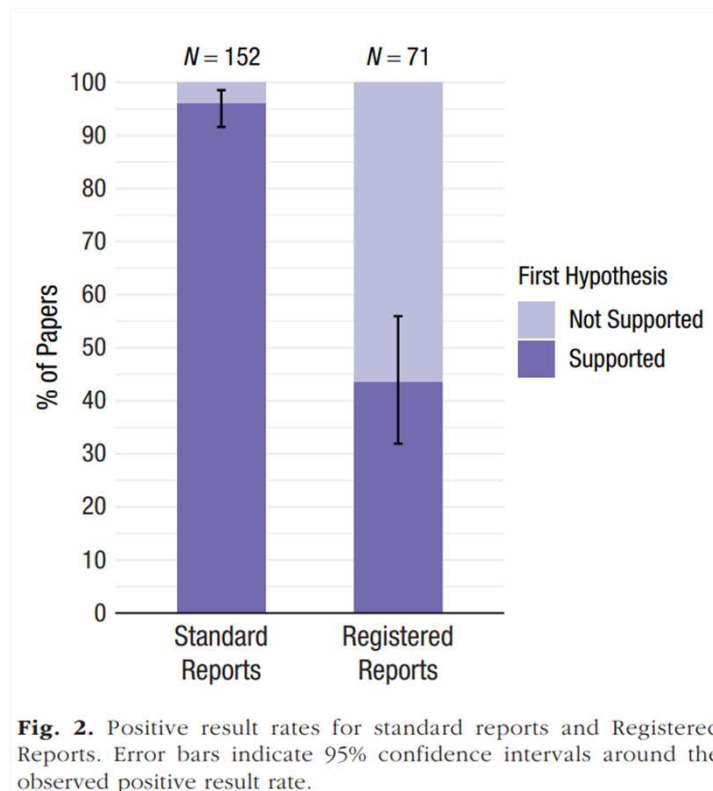### Combine methods used by exemplars into a best approach

E.g.,
Your project?

VU

Reducing wasteful research

- Reproducibility (of methods)

- Replicability (of results)

https://open-science-training-handbook.github.io/

Promoting replicability by increasing reproducibility: Study registration

Fig. 2. Positive result rates for standard reports and Registered Reports. Error bars indicate 95% confidence intervals around the observed positive result rate.

Fig. 3. Combinations of the proportion of true hypotheses and statistical power that would produce the observed positive result rates given α = 5% and no bias. Shaded areas indicate

Scheel, A. M., Schijen, M., & Lakens, D. (2021). An Excess of Positive Results: Comparing the Standard Psychology Literature With Registered Reports [Article]. Advances in Methods and Practices in Psychological Science, 4(2), 12, Article 25152459211007467. https://doi.org/10.1177/25152459211007467

14

Choices to make...

Pre-registration

For descriptive, confirmatory, and secondary data analyses

Power:
A priori
Sensitivity
Tresholds
Informational value

	GitHub	AS PREDICTED	zenodo	OSF template	OSF open
★ Timestamp	✖	✅	✅	✅	✅
★ Indexed Registry	✖	✖	✅	✅	✅
★ Persistence	✖	✅	✅	✅	✅
Anonymity	✖	✅	✖	✅	✅
Additional Materials	✅	✖	✅	✅	✅
Sandbox	✅	✅	✅	✅	✅
Template	✖	✅	✖	✅	✖
Rich Formatting	✅	✖	✅	✖	✅
Flexibility	external	limited	external	limited	external
Collaboration	external	approval	external	sequential	external
Usability	skill	very easy	moderate	easy	moderate

VU

Registrations of birth trauma research

Hurdles

Go to menti.com or use the mentimeter app

Preregistration in longitudinal studies

Challenge

- Prior knowledge of data
- Research is not hypothesis-driven
- Ensure pre-registered analyses will be appropriate for data
- Lack of flexibility in data analysis

Motivation

- Provide transparency about potential biases from prior knowledge of data
- Provide transparency about the effects of analytic choices on results
- Pre-register research that does not have a hypothesis
- Incorporate exploratory research into a pre-registered, confirmatory approach
- Check data distribution and missingness before pre-registering
- Check characteristics of exposure data and covariates before pre-registering
- Plan for analytic issues without accessing any of the data in advance
- Apply new methods or conduct follow-up tests
- Prevent difficulties publishing null results arising from reduced analytic flexibility

Potential solution

- Declare prior access to data in pre-registration
- Conduct a multiverse analysis
- Pre-register research questions and conditions for interpretation
- Use a hold-out sample to delineate exploratory and confirmatory research
- Trial analyses on a blinded dataset
- Trial analyses on a dataset excluding outcome measures
- Pre-register a decision tree
- Transparently report unplanned analyses
- Conduct Registered Report; use methods to interpret non-significant results

VU

Reproducible derivation chain

Recommendations
- Formal modelling
- Machine-readable hypothesis tests
- Nonconfirmatory research
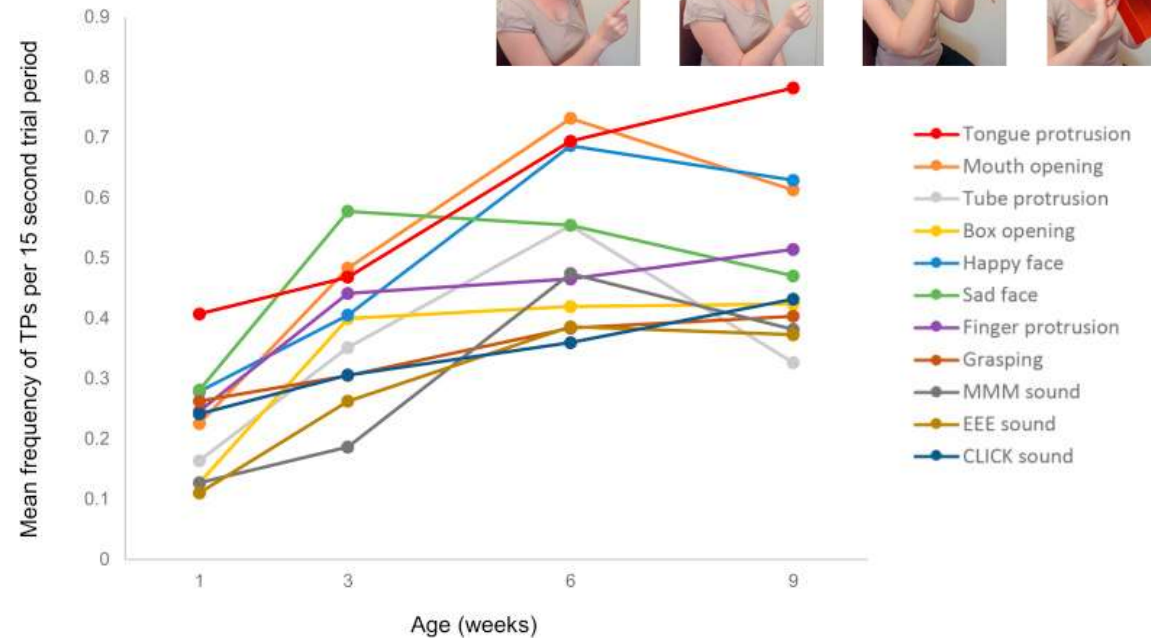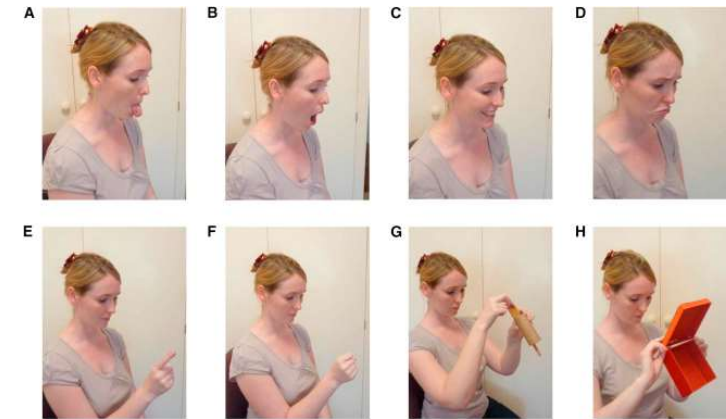
Oostenbroek, J., Suddendorf, T., Nielsen, M., Redshaw, J., Kennedy-Costantini, S., Davis, J., Clark, S., & Slaughter, V. (2016). Comprehensive Longitudinal Study Challenges the Existence of Neonatal Imitation in Humans. *Current Biology, 26(10), 1334-1338. https://doi.org/https://doi.org/10.1016/j.cub.2016.03.047*

Formal modeling of longitudinal effects

"Is the association between parents' birth experience and parents' symptoms of depression and anxiety at 2 years postpartum mediated by their relationship satisfaction at 14 months postpartum?"
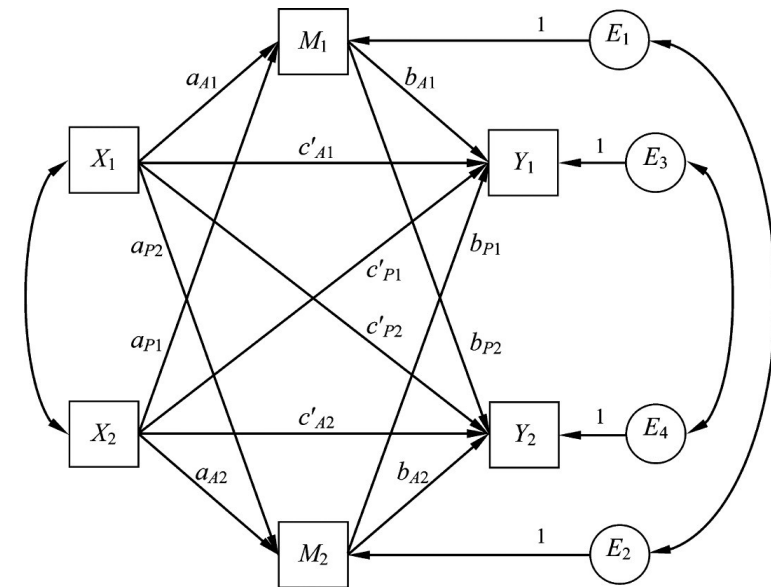
Ledermann, T., Macho, S., & Kenny, D. A. (2011). Assessing Mediation in Dyadic Data Using the Actor-Partner Interdependence Model. Structural Equation Modeling: A Multidisciplinary Journal, 18(4), 595-612. https://doi.org/10.1080/10705511.2011.607099

Seefeld, L., Handelzalts, J., Horesh, D., Horsch, A., Ayers, S., Dikmen-Yildiz, P., … Garthus-Niegel, S. (2022, March 30). Dyadic Analyses on the Longitudinal Association Between Parents' Birth Experience, Relationship Satisfaction, and Postpartum Depression and Anxiety. https://doi.org/10.17605/OSF.IO/2ZE9B

Realistic effect size expectations

Effect sizes often described using arbitrary standards (e.g., Cohen's benchmark)

Rather use well-understood benchmarks or concrete consequences

In psychology, small effect sizes (i.e., r = .10) may still be consequential in the long run—and more believable than very large effect sizes (i.e. r > .40)

Evaluating Effect Size in Psychological Research: Sense and Nonsense

David C. Funder and Daniel J. Ozer
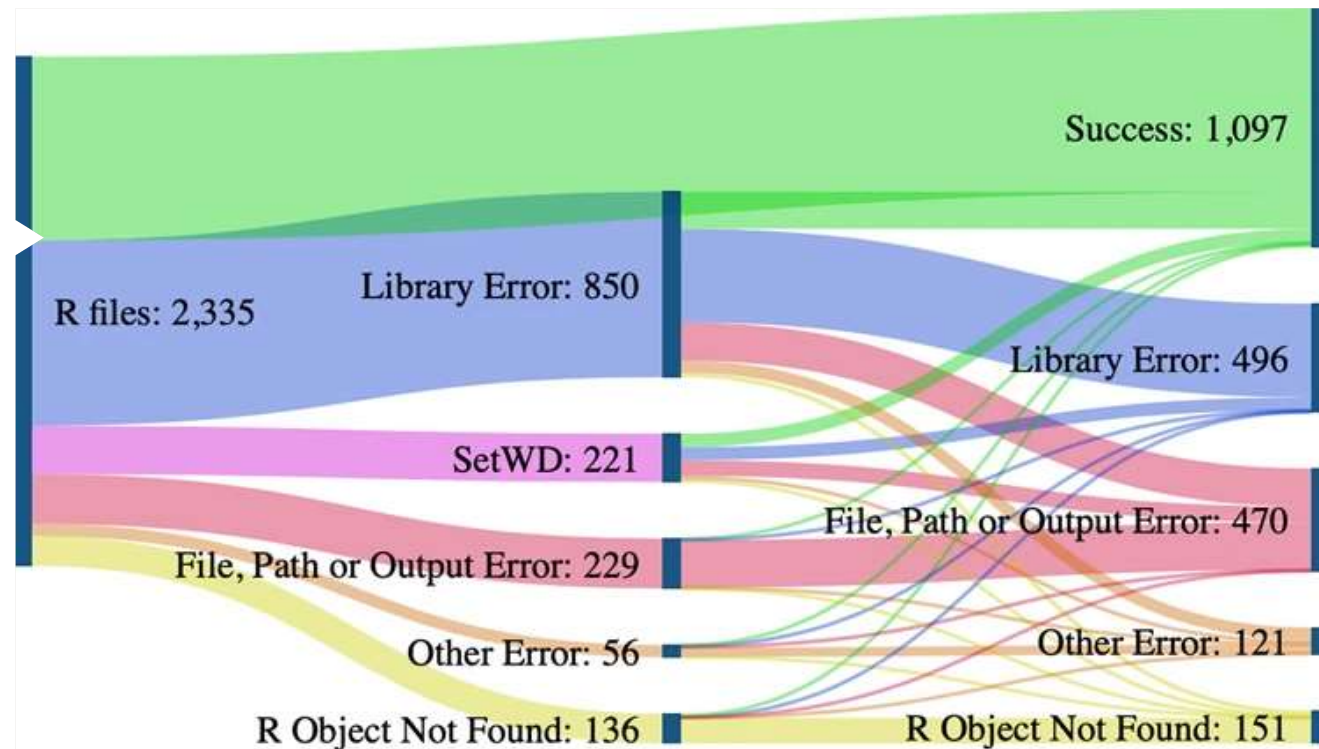Department of Psychology, University of California, Riverside

Meta-analyses of whole fields in psychology tend to find average e.s. r ~.20

VU

Code reproducibility

Recommendations
- Use basic good code practices (i.e., literate code)
- Data repository improvements
- Journal review and checklist

Trisovic, A., Lau, M. K., Pasquier, T., & Crosas, M. (2022). A large-scale study on research code quality and execution. Scientific Data, 9(1), 60. https://doi.org/10.1038/s41597-022-01143-6
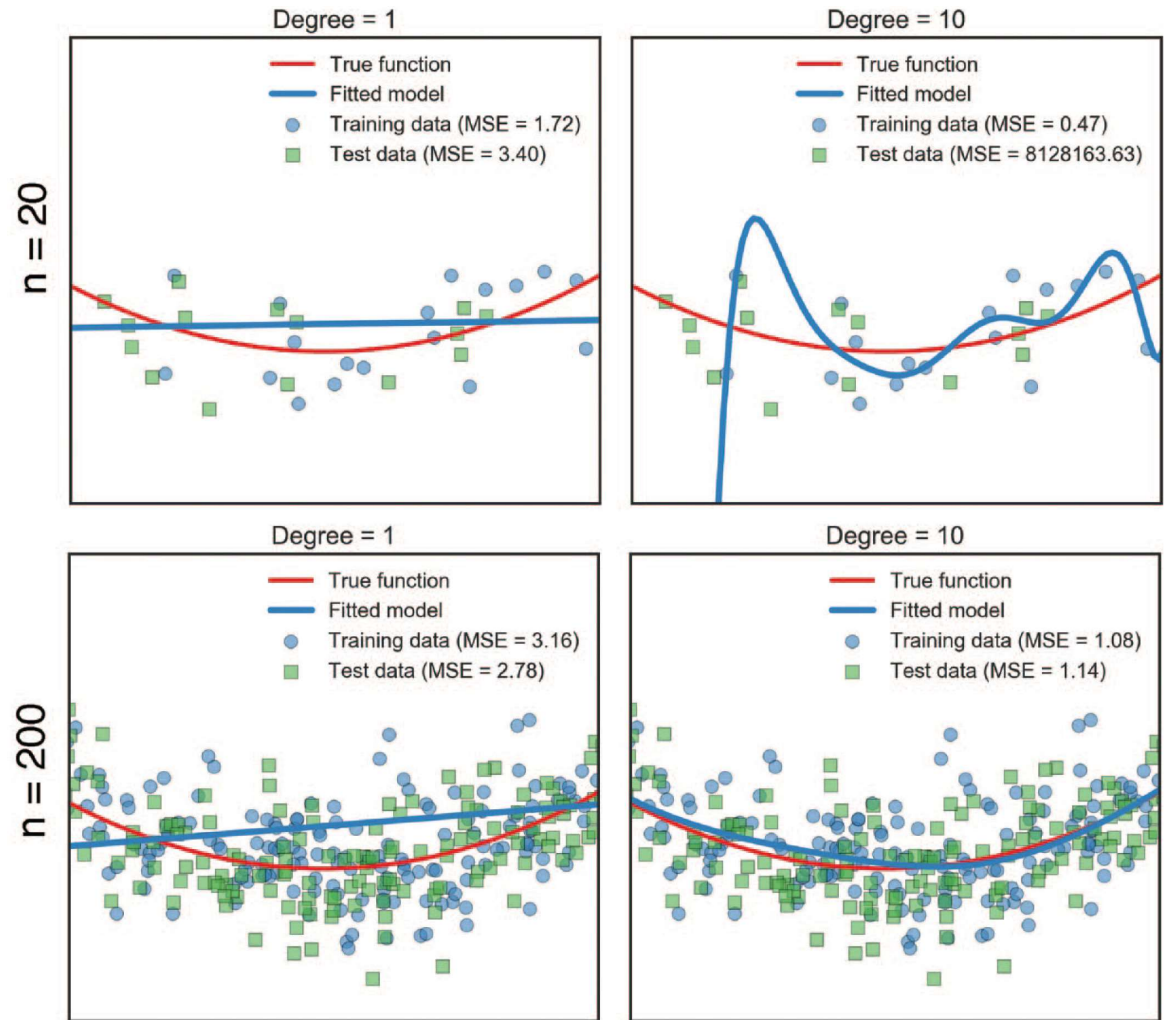
VU

To fit is to overfit

Esp. with more predictors, lower N, lower effect sizes

Procedural (p-hacking)

Cross-validation
Regularizing
(i.e., machine learning)

Also:
Sensitivity analyses,
multiverse analysis

Yarkoni, T., & Westfall, J. (2017). Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning. Perspectives on Psychological Science, 12(6), 1100-1122. https://doi.org/10.1177/1745691617693393

23

Artificial Neural Network for predicting PTSD after CS

"According to our findings, emergency cesarean section, pathology of gestation, preterm birth, the inclusion of neonate in NICU, absence of breastfeeding, psychiatric history, expectations from childbirth, and support from the partner are included in the set of important decision factors."
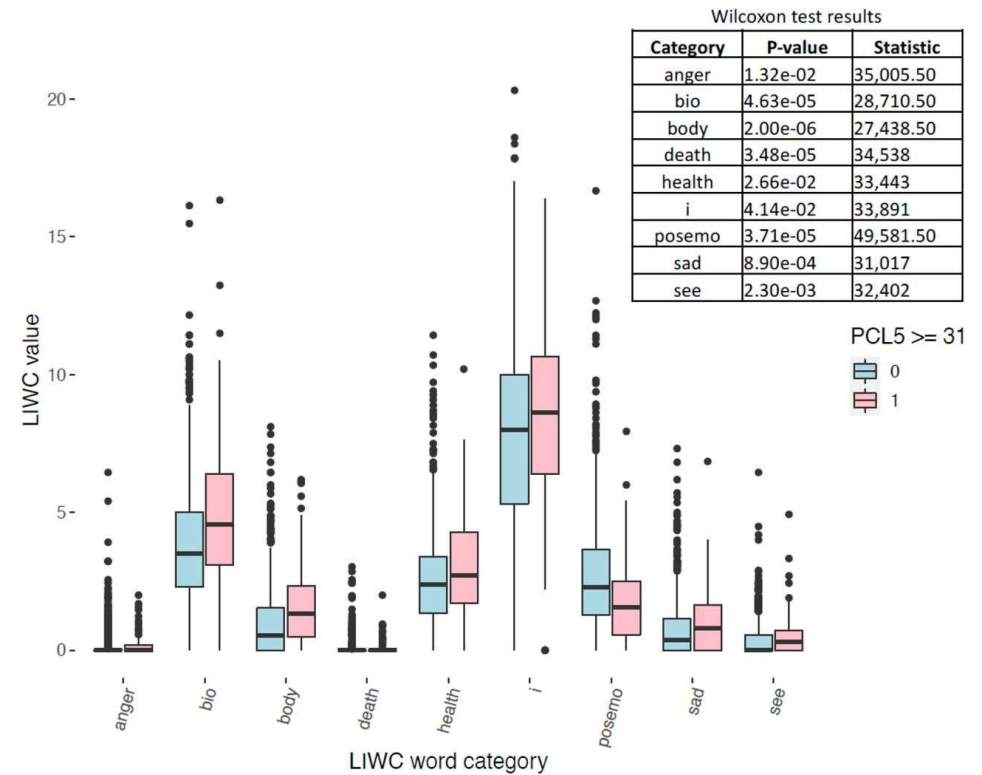

Precision, Recall, Specificity and Accuracy in Testing Phase
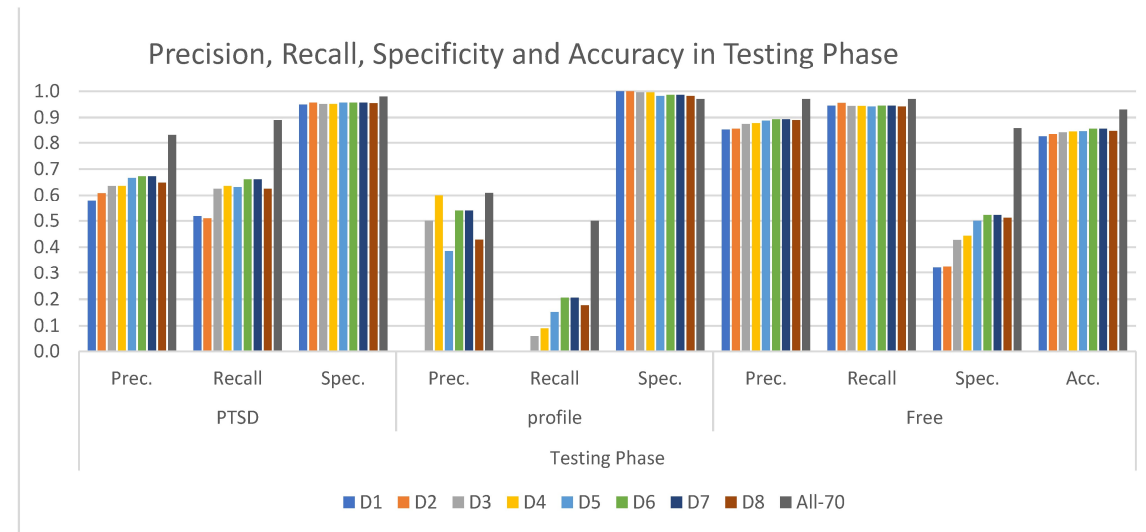
☹ Open code
☹ Data availability
✓ Cross-validation

Orovas, C., Orovou, E., Dagla, M., Daponte, A., Rigas, N., Ougiaroglou, S., Iatrakis, G., & Antoniou, E. (2022). Neural Networks for Early Diagnosis of Postpartum PTSD in Women after Cesarean Section. Applied Sciences, 12(15), 7492. https://www.mdpi.com/2076-3417/12/15/7492

VU

NLP predicting PTSD after delivery

"...personal childbirth narrative accounts generated in the early postpartum period and analyzed via advanced computational methods can detect with relatively high accuracy women who are likely to endorse CB-PTSD and those at low risk."

Wilcoxon test results

Category	P-value	Statistic
anger	1.32e-02	35,005.50
bio	4.63e-05	28,710.50
body	2.00e-06	27,438.50
death	3.48e-05	34,538
health	2.66e-02	33,443
i	4.14e-02	33,891
posemo	3.71e-05	49,581.50
sad	8.90e-04	31,017
see	2.30e-03	32,402

PCL5 >= 31
☐ 0
☐ 1

- ☹ Open code
- ☹ Data availability
- ✓ Cross-validation

Bartal, A., Jagodnik, K. M., Chan, S. J., Babu, M. S., & Dekel, S. (2022). Identifying Women with Post-Delivery Posttraumatic Stress Disorder using Natural Language Processing of Personal Childbirth Narratives. medRxiv, 2022.2008.2030.22279394. https://doi.org/10.1101/2022.08.30.22279394
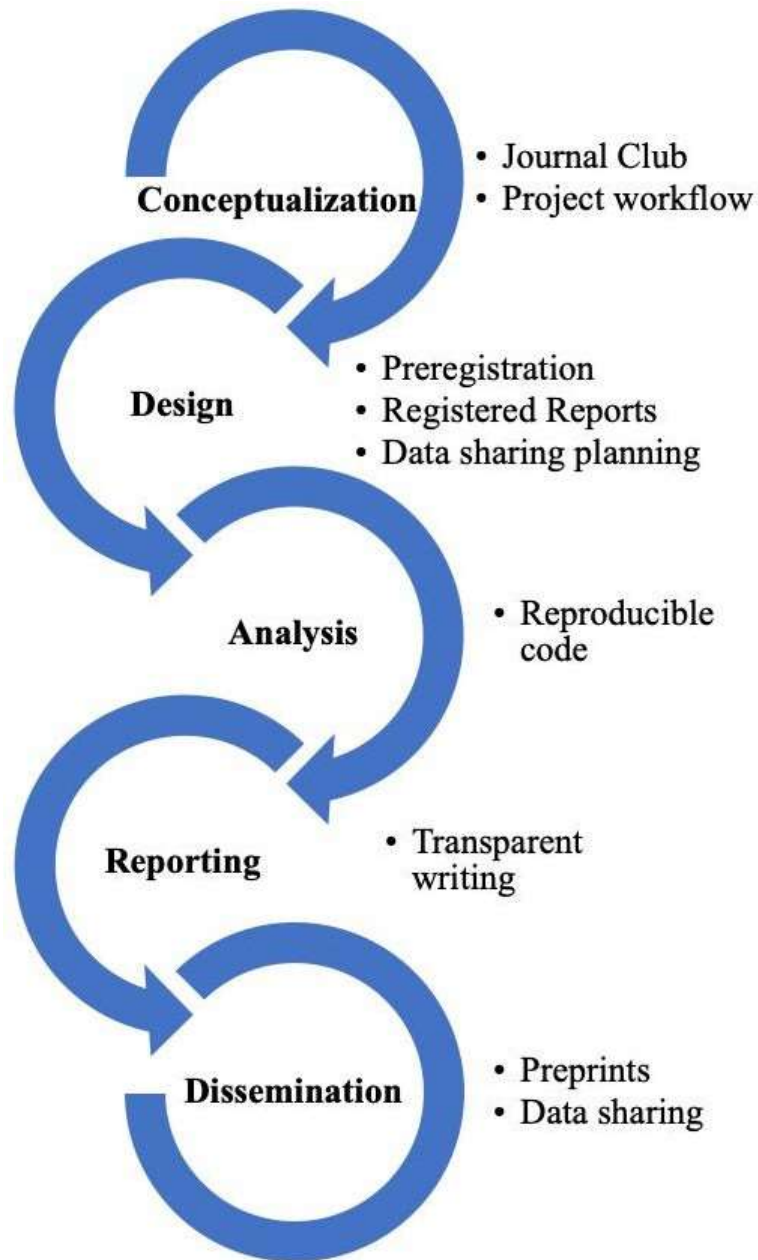
VU

Include stakeholders

Co-create

Align designs and operationalizations

Build biobanks

Collect rich perinatal data

Join consortia

VU

Using tools and resources that are openly accessible

Kathawalla, U.-K., Silverstein, P., & Syed, M. (2021). Easing Into Open Science: A Guide for Graduate Students and Their Advisors. *Collabra: Psychology, 7*(1). https://doi.org/10.1525/collabra.18684