Open Science tools

PSM2 UCL

Isabelle van der Vegt

19 Feb 2019

Today's tutorial

- Pre-registrations: why and how
- Power analysis: how-to
- Effect size conversions
- R Markdown how-to

Your final project

- Answer RQ(s) with dataset you are given
- Requirements: pre-registration, reproducible code

Pre-registrations

What is a good/interesting scientific result?

What is a good/interesting scientific result?

- “significant difference”
- “x predicts y”
- “$p < 0.05$”

The research process

- Generate hypotheses
- Design study & collect data
- Analyse data
- Interpret data

Compare

- Generate hypotheses → Introduction
- Design study & collect data → Method
- Analyse data → Results
- Interpret data → Discussion

The *real* research process

QRP 1: HARKing / post-hoc theorizing

- Design study & collect data
- Analyse data
- Generate hypotheses
- Interpet data → *Look, a "good" result!*

The *real* research process

QRP 2: significance chasing/p-hacking

- Generate hypotheses
- Design study & collect data
- Analyse data
- Collect **MORE** data
- Interpret data → *Nice, just as expected!*

The *real* research process

QRP 3: selective reporting

- Generate hypotheses
- Design study & collect data
- Analyse data
- Exclude condition/data not in line with results
- Interpret data → *Wow, an interesting result!*

The problem

- Everyone wants interesting results → QRPs

The solution

Pre-registrations

Specifying your hypotheses, study design, and analysis plan **BEFORE** collecting/analysing data.

How?

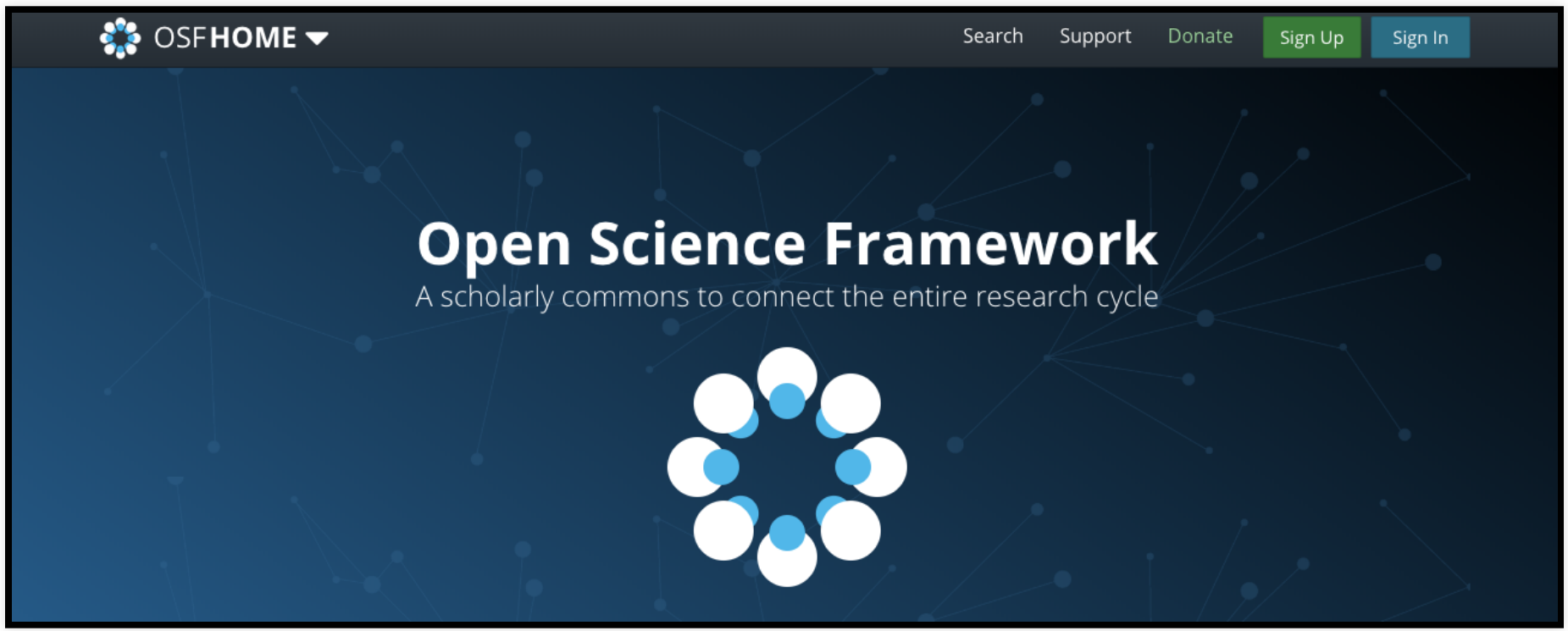
- Public, independent record
- Timestamps
- Explain deviations from pre-registered plans
- Distinguish confirmatory and exploratory results in paper

Open Science Framework

- Open a project
- Complete pre-registration form (free text or template)
- Upload additional files
- Register the project, this ‘freezes’ it
- Make public or share with specific people
- Anonymous links

Register

5 minutes: **osf.io**



OSF demo

Pre-registration practice in groups: 2-3

- Make a project titled ‘PSMII practice’
- Choose ‘registrations’ at the top
- Click ‘new registration’
- Choose the OSF pre-registration template

Pre-registration practice in groups: 2-3

You are interested in whether a lone-actor terrorist's ***ideology*** influences the ***number of hours spent on extremist forums*** and the ***number of ideological propaganda files*** they downloaded.

You are using an existing dataset of individuals who have been convicted for terrorism-related offences in the UK.

It includes information on 1) **the type of conviction**: attack, recruitment, operational support, 2) **ideology**: far-right, far-left, islamist, 3) **forum activity** (hours), 4) **number of propaganda files** found on computer, 5) **gender** of perpetrator. The sample size is 250.

Think about: analysis type, data exclusions & additional exploratory questions

Power

Power analysis

What is power?

Probability of rejecting H0 when it is *actually* false.

Example: 0.90 power = 90% chance of significant result when the effect is real. Also: 10% chance of "missing" the real effect.

What happens when power is low?

Decreased likelihood of true positive, increased likelihood of false negative.

See also: https://www.youtube.com/watch?v=7daQRvRO-NE

How to calculate power

- Post-hoc and a-priori
- G*Power Software: http://www.gpower.hhu.de/
- Also in R: see https://www.youtube.com/watch?v=ZEFSUm6JNQ0

G*Power demo

Power & ES practice

Your terrorist internet behavior study (N = 250) achieved an effect size of Cohen's *d* = 0.29 for far-left and far-right groups and forum activity. Using an effect size converter and G*Power, calculate the power you achieved. Use the statistical test you came up with in the previous exercise.

- Google "psychometrica effect size" and use the first result or https://www.psychometrica.de/effect_size.html
- G*Power download: http://www.gpower.hhu.de/

R Notebooks

R Notebooks

"In every project you have at least one other collaborator; future-you. You don't want future-you to curse past-you."
- Hadley Wickham

…You also don't want us to curse you for your code in your final project

R Notebooks

- Write text and integrate code
- Fully reproducible
- Different outputs: PDF, html, slides, etc.

R Notebook example

R Notebook practice

Generate an R notebook that contains the following elements:

- A short description of the terrorist internet activity study with a header, **bold** text, *italic* text, and a bullet point list
- A plot of the murder arrests versus urban population using the USArrests dataset (available in R). **Document your code!**
- Your favorite meme (as an image)
- Hint: Google R Markdown cheat sheet