**Day 2: Understanding challenges to reproducibility**

By David Souto

# Worksheet 4

## Challenges to reproducibility

What are challenges to reproducibility? (30')

- QRPs

- P-values misconceptions (dance of p-values & relation to power)

- Multiverse analysis

### Dance of p-values

Geoff Cumming pointed to a fundamental misunderstanding about p-values. While the language we use to report p-values suggests strength of evidence (e.g., “very significant”), a p-value tells little about how likely we are to see the same finding or the strength of the evidence. Cumming’s solution is to stop focusing on p-values but the precision of our measures, while others advocate improving p-value interpretation.

Use the shiny app to understand what results we can predict under different situations: [**https://shiny.ieis.tue.nl/d\_p\_power/**](https://shiny.ieis.tue.nl/d_p_power/)

### P-hacking challenge

We will be simulating data collection by using this Shiny app: <https://shinyapps.org/apps/p-hacker/>

We are going to start with entering the name of your favourite effect. A number of observations of XX is assumed. We are measuring 4 dependent variables to start with and the true effect is 0 (no effect).

Add participants by 5 and note save significant p-values.

### Multiverse analysis

A multiverse analysis is a great way of convincing yourself and others that analytical biases are not determining your conclusions. The example is based on this surprising claim that “female hurricanes” have more casualties! Is that true? In this exercise you will follow the file ex4\_3\_Multiverse.R

Level 2: There is a fully worked (advanced) example [here](https://cran.r-project.org/web/packages/multiverse/vignettes/visualising-multiverse.html).

**References**

**Shiny app:** [**https://shiny.ieis.tue.nl/d\_p\_power/**](https://shiny.ieis.tue.nl/d_p_power/)

**Lakens – Improving your statistical inference Misconceptions about p-values:** [**https://lakens.github.io/statistical\_inferences/01-pvalue.html#sec-lindley**](https://lakens.github.io/statistical_inferences/01-pvalue.html#sec-lindley)[**https://lakens.github.io/statistical\_inferences/**](https://lakens.github.io/statistical_inferences/)

The dance of p-values - [https://dx.doi.org/10.1111/j.1745-6924.2008.00079.x /](https://dx.doi.org/10.1111/j.1745-6924.2008.00079.x%20/) Video: <https://youtu.be/5OL1RqHrZQ8>

Definitions of reproducibility / replicability / generalizability (not universally agreed upon): [https://osf.io/cgdsn /](https://osf.io/cgdsn%20/) many analyst

On QRP: Nine circles of hell <https://journals.sagepub.com/doi/full/10.1177/1745691612459519?rss=1&ssource=mfr>

Simmons J., Nelson L. & Simonsohn U. (2011) False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allow Presenting Anything as SigniFicant. Psychological Science, 22(11), 1359-1366.

Examples across disciplines: Plenty examples across Medicine, Biology even Physics

Feinstein, A.R. (1988). Scientific standards in epidemiologic studies of the menace of daily life. Science, 242, 1257-1263.

Simonsohn, U. (2015). Small telescopes: Detectability and the

evaluation of replication results.Psychological Science,26,559-569

Belief in the law of small numbers - Kahneman & Tversky

What is replication? <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3000691>

Research misconduct and data fraud in clinical trials: prevalence

and causal factors <https://link.springer.com/article/10.1007/s10147-015-0887-3>