Open & Reproducible Science Symposium

Alice Trevail, Dan Padfield, Lotty Brand, Stephen Lang and MD Sharma

University of Exeter
Environment and Sustainability Institute

CELEBRATING TEN YEARS
#esi10

Exeter Data Analytics Hub

UKRI
Natural Environment Research Council

Event schedule

14:05 – **Prof. Mark Kelson: Reproducible and Open Science at Exeter**

14:50 - Tea break (20 mins)

15:10 – **Dr Erik Postma: Data fabrication in ecology, evolution and behaviour**

15:55 - Food refreshments (15 mins)

16:10 - **Open discussion with Erik, Dr Lotty Brand & Sofia Fernandes**

University of Exeter

Reproducible and Open Science at Exeter- perspectives and initiatives

What is reproducibility?

University of Exeter

		Data: Same	Data: Different
Code	Same	Reproducible	Replicable
Code	Different	Robust	Generalisable

Whitaker's matrix of reproducibility; 10 made available under the Creative Commons Attribution license (CC-BY 4.0).
https://www.researchgate.net/figure/Whitakers-matrix-of-reproducibility-Whitakers-matrix-of-reproducibility-10-made_fig1_354519195

Background

Brian Wansink

University of Exeter

Original blog 21/12/16

When she arrived, I gave her a data set of a self-funded, failed study which had null results (it was a one month study in an all-you-can-eat Italian restaurant buffet where we had charged some people ½ as much as others). I said, "This cost us a lot of time and our own money to collect. There's got to be something here we can salvage because it's a cool (rich & unique) data set." I had three ideas for potential Plan B, C, & D directions (since Plan A had failed). I told her what the analyses should be and what the tables should look like. I then asked her if she wanted to do them.

Every day she came back with puzzling new results, and every day we would scratch our heads, ask "Why," and come up with another way to reanalyze the data with yet another set of plausible hypotheses. "

University of Exeter

Emails February 2018

“I would like you to really dig into this to find a number of situations or people for which this relationship does hold - that is where the 1/2 price buffet did result in a difference.

Here’s some things to do. First look to see if there are any weird outliers (in terms of how much they ate). If there seems to be any reason they are different, pull them out but specially note why you did so, so that this can be described in the method.

Second, think of all the different ways you can cut the data and analyse subsets of it to see when this relationship holds males, females, lunch goers, dinner goers, people sitting alone, people eating with groups of 2, people eating in groups of 2+, people who order alcohol, people who order soft drinks, people who sit close to buffet, people who sit far away, and so on..

University of Exeter

Stapelgate

I'm getting a bad feeling in my stomach. I look around the vast hall of the station, trying to find a good spot. I turn and look and turn and look. I run to a different part of the hall. Turn and look and turn and look. People everywhere, pressing ahead, not looking around. It's the rush hour. This is completely different from how I'd imagined it. What an idiot! I can't find anywhere we could have put a row of five chairs. How had I set it up? How had we done the study? I walk over to the first track, and back to the last; to the left, to the right; past the supermarket and Burger King on one side of the massive hall, and past the newsstand and a café on the other. How is this possible? There isn't a good spot to do the research anywhere. There's nowhere you could put a row of five chairs anywhere along the side of the hall, against the wall or by a window. You could maybe fit them right in the middle of the hall, but that would attract too much attention. You can't expect people

University of Exeter

https://errorstatistics.files.wordpress.com/2014/12/fakingscience-20141214.pdf

Journal of Personality and Social Psychology
2011, Vol. 100, No. 3, 407–425

0022-3514/11/$12.00 DOI: 10.1037/a0021524

Feeling the Future: Experimental Evidence for Anomalous Retroactive Influences on Cognition and Affect

Daryl J. Bem
Cornell University

The term *psi* denotes anomalous processes of information or energy transfer that are currently unexplained in terms of known physical or biological mechanisms. Two variants of psi are *precognition* (conscious cognitive awareness) and *premonition* (affective apprehension) of a future event that could not otherwise be anticipated through any known inferential process. Precognition and premonition are themselves special cases of a more general phenomenon: the anomalous retroactive influence of some future event on an individual's current responses, whether those responses are conscious or nonconscious, cognitive or affective. This article reports 9 experiments, involving more than 1,000 participants, that test for retroactive influence by "time-reversing" well-established psychological effects so that the individual's responses are obtained before the putatively causal stimulus events occur. Data are presented for 4 time-reversed effects: precognitive approach to erotic stimuli and precognitive avoidance of negative stimuli; retroactive priming; retroactive habituation; and retroactive facilitation of recall. The mean effect size (*d*) in psi performance across all 9 experiments was 0.22, and all but one of the experiments yielded statistically significant results. The individual-difference variable of stimulus seeking, a component of extraversion, was significantly correlated with psi performance in 5 of the experiments, with participants who scored above the midpoint on a scale of stimulus seeking achieving a mean effect size of 0.43. Skepticism about psi, issues of replication, and theories of psi are also discussed.

Keywords: psi, parapsychology, ESP, precognition, retrocausation

University of Exeter

> Even so, I find myself in agreement with Tal Yarkoni's comment on his excellent blog: "It's important to note that none of these concerns is really terrible individually. Sure, it's bad to peek at your data but data peeking alone probably isn't going to produce nine different false positives. Nor is using one-tailed tests, or constructing measures on the fly, etc. But when you combine data-peeking, liberal thresholds, study recombination, flexible hypotheses, and selective measures, you have a perfect recipe for spurious results."

https://www.theguardian.com/science/2012/mar/15/precognition-studies-curse-failed-replications

Mundane problems

- Simple errors (top left cell of an excel spreadsheet)
- No time (precarity of contracts)
- Version control
- Timing of documentation
- No incentives

University of Exeter

Study retracted after posting data openly

> *A reader alerted me that the time stamps on the posted data for Studies 1 and 3 indicated that testing of subjects was blocked by condition, creating a confound between condition and date. This confound is particularly problematic because (a) the blocks were widely separated in time and (b) the primary dependent variable of interest in those studies was self-reported risk of flu, which might well vary substantially with date. After reviewing the data with the authors, we came to the consensus that the confound undermined interpretation of the results and warranted retraction of the article. As Editor of Psychological Science, I have decided to retract this article.*

> *This is a tough lesson for us. Open science does not mean people don't make mistakes; it simply means mistakes are more easily caught. We are very grateful for the reader who caught the mistake and the editor who coached us through how to handle a situation like this. We vow to be more thorough and careful in our future research. At the same time, this experience strengthens our commitment to open science as it shows that open science is working as intended.*

University of Exeter

Reproducibility team


Prof Mark Kelson


Dr Eilis Hannon

Prof Gavin Buckingham

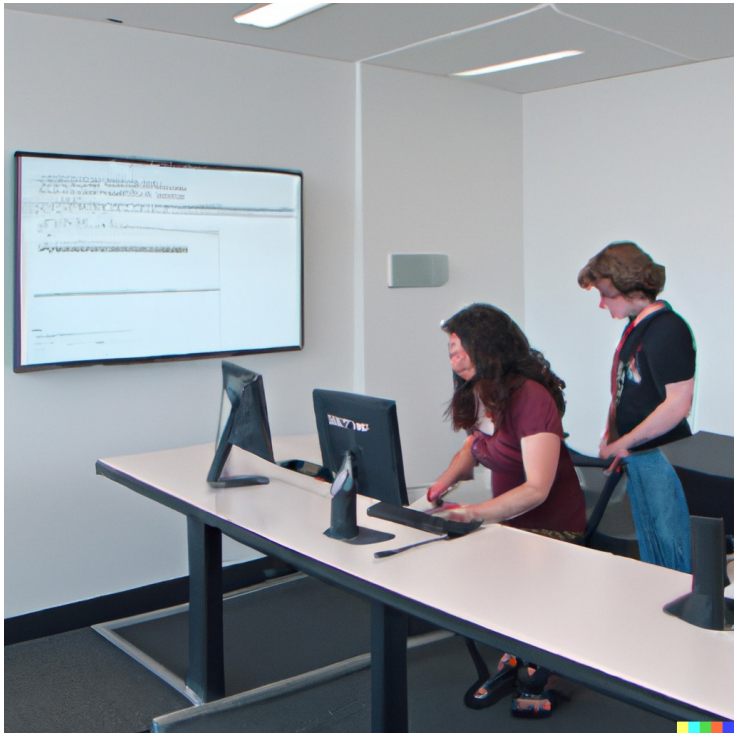

Dr Janice Ranson


Sofia Fernandes

University of Exeter

Coding for reproducible research

- https://uniexeterrse.github.io/workshop-homepage/workshops.html



University of Exeter

Summary of registrations

	All	Faculty of Health and Life Sciences	Faculty of Environment, Science and Economy	Faculty of Humanities, Arts and Social Sciences
All	639	284	312	43
PhD student	309	128	148	33
ECRs	236	111	116	9
Senior Academic Staff	65	36	28	1
Professional Services	14	6	8	0

	Penryn	St Lukes	Streatham	Other	Work Completely Remotely	Total
All	107	96	329	17	57	606
PhD student	50	47	167	5	24	293
ECRs	45	40	112	6	0	203
Senior Academic Staff	6	6	37	3	31	83
Professional Services	1	3	8	0	0	12

Summary of registrations

Workshop	Term	All	PhD student	ECRs	Senior Academics	Professional Services
Computational Thinking	Spring	73	42	26	2	0
Improve your R code	Autumn	56	34	20	2	0
Introduction to HPC	Autumn	66	34	28	4	0
Introduction to Python	Autumn	113	50	49	7	5
	Spring	77	44	21	9	3
Introduction to R	Spring	53	24	19	6	0
	Autumn	85	28	39	14	2
Introduction to Unix	Autumn	43	17	13	10	3
	Spring	44	29	8	6	0
Introduction to Version Control with Git and GitHub	Spring	30	7	13	5	2
Total		640	309	236	65	15

30 in person places; 15 online

Current delivery model

University of Exeter

Hybrid delivery:
30 in the room (Streatham) & 15 online via Teams

Develop and maintain a series of workshops on topics related to programming & reproducibility

Community of researchers & RSEs to facilitate the delivery

Delivery 2023/4

7 existing workshops to rerun + addition of intermediate & advanced sessions

Plan for 18 workshops of (1-3 sessions each)

Incentives

- Explicit statements on reproducibility and integrity issues in probation, promotion, and hiring documents, with specific sections/space to articulate these
- Build an internal network of academic ‘champions’ embedded within faculties who could champion (and provide a discipline-specific lens on) issues related to research reproducibility and integrity issues
- Open research awards

University of Exeter

Reproducibility and AI/machine learning

- Theme in the institute for Data Science and Artificial Intelligence (Janice Ranson)
- Roundtable event with the Royal Society on reproducibility in AI
- Working with the Tools, Practices and Systems group in the Alan Turing Institute

University of Exeter

Other things

- Scite pilot trial
- Train-the trainer GW 4 pilot scheme
  - Open Access
  - Data sharing
  - Open software / code
  - Preprints
  - Pre-registration

University of Exeter

Summary

- Training/support
- Incentives

Contact us: m.j.kelson@exeter.ac.uk

University of Exeter