Psychology as a Robust Science
Dr Amy Orben
Lent Term 2020

When: Lent Term 2020;
Wednesdays 2-4pm (Weeks 0-6; 15.1.2020 – 26.2.2020),
Mondays 2-4pm (Week 7; 2.3.2020)

Where: Lecture Theatre, Experimental Psychology Department, Downing Site

Summary: *Is psychology a robust science?* To answer such a question, this course will encourage you to think critically about how psychological research is conducted and how conclusions are drawn.

To enable you to truly understand how psychology functions as a science, however, this course will also need to discuss how psychologists are incentivised, how they publish and how their beliefs influence the inferences they make. By engaging with such issues, this course will probe and challenge the basic features and functions of our discipline. We will uncover multiple methodological, statistical and systematic issues that could impair the robustness of scientific claims we encounter every day. We will discuss the controversy around psychology and the replicability of its results, while learning about new initiatives that are currently reinventing the basic foundations of our field. The course will equip you with some of the basic tools necessary to conduct robust psychological research fit for the 21st century.

The course will be based on a mix of set readings, class discussions and lectures. Readings will include a diverse range of journal articles, reviews, editorials, blog posts, newspaper articles, commentaries, podcasts, videos, and tweets. No exams or papers will be set; but come along with a critical eye and a willingness to discuss some difficult and controversial issues.

Core readings

- Chris Chambers (2017). *The 7 deadly sins of psychology: A manifesto for reforming the culture of scientific practice*. Princeton, NJ: Princeton University Press.
- Zoltan Dienes (2008). *Understanding Psychology as a Science*. London, UK: Palgrave MacMillan.
- Sophia Crüwell, Johnny van Doorn, Alexander Etz, Matthew C. Makel, Hannah Moshontz, Jesse C. Niebaum, Amy Orben, Sam Parsons, and Michael Schulte-Mecklenbeck (2019). 'Seven Easy Steps to Open Science'. *Zeitschrift Für Psychologie* 227, no. 4: 237–48. https://doi.org/10.1027/2151-2604/a000387.

Structure

Below you will find the rough outline of each session of this course. Some sessions (e.g. the first and last) will abide by a different format.

1. Pre-session assignment – Before every session you will find an assignment in this syllabus which you should complete
2. Lecture – Introduction to the session's topic; overview of core papers and discussions; appraisal of initial arguments
3. Discussion – Short class discussion about key questions highlighted in the lecture
4. Tea/coffee break
5. Case Study – Presentations on set pieces of work (papers, commentaries, blog posts) by members of the class
6. Discussion – Longer class discussion about themes highlighted by this week's lecture and case studies

Overview of course structure

Week 1 – Introduction

Overview of the course contents and aims; the scientific process; Merton's norms; the replication crisis and how it began; defining Open Science

Week 2 – Replications

Replications and why they started a crisis; understanding replications; backlash and counter-backlash; replications in other disciplines; the Many Labs movement

Week 3 – Questionable Research Practices

Defining and discussing Questionable Research Practices and their influence on scientific inference in psychology; false-positive psychology and researcher degrees of freedom; multiverse approaches

Week 4 – Preregistration and Registered Reports

Pre-registration and Registered Reports; the split between exploratory and confirmatory hypothesis testing; arguments for and against the drive towards preregistration; scientific creativity; blinded analyses

Week 5 – Fraud, Errors and Scientific Self-Correction

Understanding what makes fraud different from errors; methods used to detect errors; whether scientific self-correction is myth or reality; scientific errors and debates about how to address them

Week 6 – The Statistics Wars (+ Power)

Error Rates; Null-Hypothesis Significance Testing; *p*-values; power; Bayesian methods (estimation and Bayes Factors); 'Redefine Significance' vs 'Justify your Alpha'

Week 7 – Journals, Publishing and Computational Reproducibility

Introduction to the for-profit publishing model; new forms of publishing; pre-prints; Open Access; peer review; Open Data and computational reproducibility

Week 8 – Psychology as a Robust Science

What makes robust inferences; triangulation and causality; theory and measurement; Psychological Science Accelerator; incentives and next steps

Week 1 – Introduction

Overview of the course contents and aims; the scientific process; Merton's norms; the replication crisis and how it began; defining Open Science

Assignment and readings
Before coming to this lecture, please complete the following three assignments:

1. Watch the following TED talk by Naomi Oreskes, *Why we should trust scientists*, May 2014 (https://www.ted.com/talks/naomi_oreskes_why_we_should_trust_scientists?language=en)
2. Listen to the following podcast by BBC Radio 4, *Analysis: The Replication Crisis*, November 2018 (https://www.bbc.co.uk/sounds/play/m00013p9)
3. Read the following magazine article: Engber, Daniel. 'Daryl Bem Proved ESP Is Real. Which Means Science Is Broken.' Slate Magazine, 7 June 2017. https://slate.com/health-and-science/2017/06/daryl-bem-proved-esp-is-real-showed-science-is-broken.html.

Core references
Merton, Robert K. 'The Normative Structure of Science', 1942. https://www.panarchy.org/merton/science.html. (or see this YouTube video for explanation: https://www.youtube.com/watch?v=00btFojQPiU&list=PLAKyhL4GNnqMVIdZDvSt3bqlHFJVRNOsF&index=2)

Feynman, Richard P. 'Cargo Cult Science'. presented at the Caltech 1974 Commencement Address, Caltech, 1974. http://calteches.library.caltech.edu/51/2/CargoCult.htm. (or listen to it: https://www.youtube.com/watch?v=yvfAtIJbatg)

Ravetz, Jerome R. Scientific Knowledge and Its Social Problems, 1971.

Bem, Daryl J. 'Feeling the Future: Experimental Evidence for Anomalous Retroactive Influences on Cognition and Affect.' *Journal of Personality and Social Psychology* 100, no. 3 (2011): 407–25. https://doi.org/10.1037/a0021524.

Frank, M. What is Open Science Movement (twitter thread):
https://twitter.com/mcxfrank/status/1044254887075147776

Fecher, Benedikt, and Sascha Friesike. 'Open Science: One Term, Five Schools of Thought'. In *Opening Science: The Evolving Guide on How the Internet Is Changing Research, Collaboration and Scholarly Publishing*, edited by Sönke Bartling and Sascha Friesike, 17–47. Cham: Springer International Publishing, 2014. https://doi.org/10.1007/978-3-319-00026-8_2.

Just to mention…
If you are interested in further understanding philosophy of science in relation to psychology, I would recommend

- Zoltan Dienes (2008). *Understanding Psychology as a Science.* London, UK: Palgrave MacMillan. (Chapters 1 and 2)
- Paul Meehl's recorded 1989 lectures, *Philosophical Psychology* at the University of Minnesota. Find them here: http://meehl.umn.edu/talks/philosophical-psychology-1989
- To learn more about Early Career Researchers in this area, why not listen to this *ReproducibiliTea* podcast: https://soundcloud.com/reproducibilitea/episode-11-ivan-flis

Week 2 – Replications

Replications and why they started a crisis; understanding replications; backlash and counter-backlash; replications in other disciplines; the Many Labs movement

Assignment and readings
Before coming to this lecture, please do the following to get a better understanding of both the Open Science Collaboration and the Many Labs studies:

1. Listen to the following podcast by NPR Planet Money, *The experiment experiment [audio podcast]*, January 2019 (https://www.npr.org/sections/money/2016/01/15/463237871/episode-677-the-experiment-experiment)
2. Read the following accessible and quick piece about Many Labs 2: Owens, Brian. 'Replication Failures in Psychology Not Due to Differences in Study Populations'. *Nature*, 19 November 2018. https://doi.org/10.1038/d41586-018-07474-y.

Case study
There will be two 5-minute presentations about key replication attempts in psychology from members of the class [no more than 4 slides]:

1. Ego-depletion: Hagger, M. S., N. L. D. Chatzisarantis, H. Alberts, C. O. Anggono, C. Batailler, A. R. Birt, R. Brand, et al. 'A Multilab Preregistered Replication of the Ego-Depletion Effect': *Perspectives on Psychological Science*, 29 July 2016. https://doi.org/10.1177/1745691616652873. (also see this podcast about ego depletion: https://fourbeers.fireside.fm/31)
2. Neonatal imitation: Oostenbroek, Janine, Thomas Suddendorf, Mark Nielsen, Jonathan Redshaw, Siobhan Kennedy-Costantini, Jacqueline Davis, Sally Clark, and Virginia Slaughter. 'Comprehensive Longitudinal Study Challenges the Existence of Neonatal Imitation in Humans'. *Current Biology* 26, no. 10 (23 May 2016): 1334–38. https://doi.org/10.1016/j.cub.2016.03.047. (also see this article about the replication: https://theconversation.com/the-imitation-game-can-newborn-babies-mimic-their-parents-61732)

After a short discussion of these two high-profile non-replications, we will examine the debate about the value of direct replications by listening to two additional 5-minute presentations [no more than 2 slides] about the opinions voiced in the following commentaries from the same issue of *Perspectives in Psychological Science:*

1. Stroebe, Wolfgang, and Fritz Strack. 'The Alleged Crisis and the Illusion of Exact Replication'. *Perspectives on Psychological Science* 9, no. 1 (1 January 2014): 59–71. https://doi.org/10.1177/1745691613514450.
2. Simons, Daniel J. 'The Value of Direct Replication.' *Perspectives on Psychological Science* 9, no. 1 (2014): 76–80. https://doi.org/10.1177/1745691613514755.

For those interested, the following paper gives a good overview about the debate around replication, while the second one gives some suggestions about how to conduct replications in future:

Zwaan, Rolf Antonius, Alexander Etz, Richard E. Lucas, and Brent Donnellan. 'Making Replication Mainstream'. Preprint, 20 October 2017. https://doi.org/10.31234/osf.io/4tg9c.

Lucas, Rich. 'The Rules of Replication'. The Desk Reject (blog), 2 May 2017. /2017/05/the-rules-of-replication/.

Core references

Open Science Collaboration. 'Estimating the Reproducibility of Psychological Science'. *Science* 349, no. 6251 (28 August 2015). https://doi.org/10.1126/science.aac4716.

Gilbert, Daniel T., Gary King, Stephen Pettigrew, and Timothy D. Wilson. 'Comment on "Estimating the Reproducibility of Psychological Science"'. *Science* 351, no. 6277 (4 March 2016): 1037–1037. https://doi.org/10.1126/science.aad7243.

Anderson, Christopher J., Štěpán Bahník, Michael Barnett-Cowan, Frank A. Bosco, Jesse Chandler, Christopher R. Chartier, Felix Cheung, et al. 'Response to Comment on "Estimating the Reproducibility of Psychological Science"'. *Science* 351, no. 6277 (4 March 2016): 1037–1037. https://doi.org/10.1126/science.aad9163.

Bishop, Dorothy V. M. 'BishopBlog: Sowing Seeds of Doubt: How Gilbert et al's Critique of the Reproducibility Project Has Played out'. *BishopBlog* (blog), 27 May 2018. http://deevybee.blogspot.com/2018/05/sowing-seeds-of-doubt-how-gilbert-et.html.

Camerer, Colin F., Anna Dreber, Felix Holzmeister, Teck-Hua Ho, Jürgen Huber, Magnus Johannesson, Michael Kirchler, et al. 'Evaluating the Replicability of Social Science Experiments in Nature and Science between 2010 and 2015'. *Nature Human Behaviour* 2, no. 9 (September 2018): 637–44. https://doi.org/10.1038/s41562-018-0399-z. (also see tweet here: https://twitter.com/siminevazire/status/1034800892317458432_)

Kaiser, Jocelyn. 'Plan to Replicate 50 High-Impact Cancer Papers Shrinks to Just 18'. *Science*, 31 July 2018. https://www.sciencemag.org/news/2018/07/plan-replicate-50-high-impact-cancer-papers-shrinks-just-18.

The 'Many Labs' studies

Klein, Richard A., Michelangelo Vianello, Fred Hasselman, Byron G. Adams, Jr Reginald B. Adams, Sinan Alper, Mark Aveyard, et al. 'Many Labs 2: Investigating Variation in Replicability Across Samples and Settings': *Advances in Methods and Practices in Psychological Science*, 24 December 2018. https://doi.org/10.1177/2515245918810225. (this is an accessible summary of the findings: https://www.nature.com/articles/d41586-018-07474-y)

Klein, Richard A., Kate A. Ratliff, Michelangelo Vianello, Reginald B. Adams, Štěpán Bahník, Michael J. Bernstein, Konrad Bocian, et al. 'Investigating Variation in Replicability'. *Social Psychology* 45, no. 3 (1 January 2014): 142–52. https://doi.org/10.1027/1864-9335/a000178.

Ebersole, Charles R., Olivia E. Atherton, Aimee L. Belanger, Hayley M. Skulborstad, Jill M. Allen, Jonathan B. Banks, Erica Baranski, et al. 'Many Labs 3: Evaluating Participant Pool Quality across the Academic Semester via Replication'. *Journal of Experimental Social Psychology, Special Issue: Confirmatory*, 67 (1 November 2016): 68–82. https://doi.org/10.1016/j.jesp.2015.10.012.

Klein, Richard Anthony, Corey L. Cook, Charles R. Ebersole, Christine Anne Vitiello, Brian A. Nosek, Christopher R. Chartier, Cody D Christopherson, et al. 'Many Labs 4: Failure to Replicate Mortality Salience Effect With and Without Original Author Involvement'. Preprint, 11 December 2019. https://doi.org/10.31234/osf.io/vef2c.

Yarkoni, Tal. 'What We Can and Can't Learn from the Many Labs Replication Project'. *[Citation Needed]* (blog), 27 December 2013. https://www.talyarkoni.org/blog/2013/12/27/what-we-can-and-cant-learn-from-the-many-labs-replication-project/.

Just to mention…
To find out more about current initiatives to look at the large-scale replicability of research, have a listen to a recent *Everything Hertz* podcast episode interviewing Fiona Fidler: https://everythinghertz.com/94

Week 3 – Questionable Research Practices

Defining and discussing Questionable Research Practices and their influence on scientific inference in psychology; false-positive psychology and researcher degrees of freedom; multiverse approaches

Assignment and readings
Before coming to this lecture, please do the following assignments:

1. Read the following paper: Simmons, Joseph P., Leif D. Nelson, and Uri Simonsohn. 'False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant'. *Psychological Science*, 17 October 2011. https://doi.org/10.1177/0956797611417632. (this YouTube video provides a really good explanation of their simulation: https://www.youtube.com/watch?v=uN3Q-s-CtTc)
2. Read the following article by Regina Nuzzo published in Nature, *How scientists fool themselves – and how they can stop,* October 2015 (https://www.nature.com/news/how-scientists-fool-themselves-and-how-they-can-stop-1.18517)
3. Play around with the p-hacking simulator: https://fivethirtyeight.com/features/science-isnt-broken/

Case study
We will listen to two 10-minute presentations about two papers that further developed ideas regarding flexibility in analytical choices [no more than 4 slides]:

1. Silberzahn, R., E. L. Uhlmann, D. P. Martin, P. Anselmi, F. Aust, E. Awtrey, Š. Bahník, et al. 'Many Analysts, One Data Set: Making Transparent How Variations in Analytic Choices Affect Results'. *Advances in Methods and Practices in Psychological Science* 1, no. 3 (2018): 337–56. https://doi.org/10.1177/2515245917747646.
2. Poldrack, Russell A., Chris I. Baker, Joke Durnez, Krzysztof J. Gorgolewski, Paul M. Matthews, Marcus R. Munafò, Thomas E. Nichols, Jean-Baptiste Poline, Edward Vul, and Tal Yarkoni. 'Scanning the Horizon: Towards Transparent and Reproducible Neuroimaging Research'. *Nature Reviews Neuroscience* 18, no. 2 (February 2017): 115–26. https://doi.org/10.1038/nrn.2016.167.

Core references
Simmons, Joseph P., Leif D. Nelson, and Uri Simonsohn. 'False-Positive Citations': *Perspectives on Psychological Science*, 29 March 2018. https://doi.org/10.1177/1745691617698146.

Gelman, A., & Loken, E. (2013). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time. Unpublished manuscript. http://www.stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf

Bishop, Dorothy Vera Margaret. 'The Psychology of Experimental Psychologists: Overcoming Cognitive Constraints to Improve Research'. Preprint, 12 July 2019. https://doi.org/10.31234/osf.io/hnbex.

John, Leslie K., George Loewenstein, and Drazen Prelec. 'Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling': *Psychological Science*, 16 April 2012. https://doi.org/10.1177/0956797611430953.

Botvinik-Nezer, Rotem, Felix Holzmeister, Colin F. Camerer, Anna Dreber, Juergen Huber, Magnus Johannesson, Michael Kirchler, et al. 'Variability in the Analysis of a Single Neuroimaging Dataset by Many Teams'. *BioRxiv*, 15 November 2019, 843193. https://doi.org/10.1101/843193.

Nieuwenhuis, Sander, Birte U. Forstmann, and Eric-Jan Wagenmakers. 'Erroneous Analyses of Interactions in Neuroscience: A Problem of Significance'. *Nature Neuroscience* 14, no. 9 (September 2011): 1105–7. https://doi.org/10.1038/nn.2886.

Orben, Amy, and Andrew K. Przybylski. 'The Association between Adolescent Well-Being and Digital Technology Use'. *Nature Human Behaviour* 3, no. 2 (February 2019): 173–82. https://doi.org/10.1038/s41562-018-0506-1. (see the tweet thread here: https://twitter.com/orbenamy/status/1084855999821959169?lang=en-gb)

Rohrer, Julia M. 'Run All the Models! Dealing With Data Analytic Flexibility'. *APS Observer* 31, no. 3 (28 February 2018). https://www.psychologicalscience.org/observer/run-all-the-models-dealing-with-data-analytic-flexibility.

Just to mention…
If you are interested in trying out some multiverse approaches yourself I would recommend having a look at the following package currently being developed by Phillip Masur and Michael Scharkow: https://twitter.com/MasurPhil/status/1208084330779750400

Furthermore here are some more papers to have a look at:

- Simonsohn, Uri, Joseph P. Simmons, and Leif D. Nelson. 'Specification Curve: Descriptive and Inferential Statistics on All Reasonable Specifications', 2015. https://doi.org/10.2139/ssrn.2694998.
- Steegen, Sara, Francis Tuerlinckx, Andrew Gelman, and Wolf Vanpaemel. 'Increasing Transparency Through a Multiverse Analysis'. *Perspectives on Psychological Science* 11, no. 5 (2016): 702–12. https://doi.org/10.1177/1745691616658637.

Week 4 – Preregistration and Registered Reports

Pre-registration and Registered Reports; the split between exploratory and confirmatory hypothesis testing; arguments for and against the drive towards preregistration; scientific creativity; blinded analyses

Assignments and readings
For this lecture please read the following two articles:

1. Wagenmakers, Eric-Jan, Ruud Wetzels, Denny Borsboom, Han L. J. van der Maas, and Rogier A. Kievit. 'An Agenda for Purely Confirmatory Research'. *Perspectives on Psychological Science* 7, no. 6 (1 November 2012): 632–38. https://doi.org/10.1177/1745691612463078.
2. Chambers, Chris. 'What's next for Registered Reports?' *Nature* 573, no. 7773 (September 2019): 187–89. https://doi.org/10.1038/d41586-019-02674-6.
3. Furthermore, if you are interested in learning about the experience of one of the first PhD students to write a Registered Report, I would suggest you listen to the *ReproducibiliTea* podcast episode interviewing Hannah Hobson: https://soundcloud.com/reproducibilitea/episode-18-hannah-hobson

Case study
There will be two short (max. 5 minute) presentations without slides. The first presenter will prepare by reading Sophia Scott's commentary about Registered Reports/preregistration; the second presenter will prepare by reading Susan Goldin-Meadow's APS presidential column about Registered Reports/preregistration and the comments left underneath it. *Please note that both authors' definitions of preregistration can be confusing and include some concepts of Registered Reports.*

1. Scott, Sophie. 'Pre-Registration Would Put Science in Chains'. *Times Higher Education (THE*), 25 July 2013. https://www.timeshighereducation.com/comment/opinion/pre-registration-would-put-science-in-chains/2005954.article.
2. Goldin-Meadow, Susan. 'Why Preregistration Makes Me Nervous'. *APS Observer* 29, no. 7 (31 August 2016). https://www.psychologicalscience.org/observer/why-preregistration-makes-me-nervous. [Please also read the comments left underneath this column]

Core references
Nosek, Brian A., Charles R. Ebersole, Alexander C. DeHaven, and David T. Mellor. 'The Preregistration Revolution'. *Proceedings of the National Academy of Sciences* 115, no. 11 (13 March 2018): 2600–2606. https://doi.org/10.1073/pnas.1708274114.

Warren, Matthew. 'First Analysis of "Pre-Registered" Studies Shows Sharp Rise in Null Findings'. *Nature*, 24 October 2018. https://doi.org/10.1038/d41586-018-07118-1.

Chambers, Christopher D. 'Registered Reports: A New Publishing Initiative at Cortex'. *Cortex* 49, no. 3 (March 2013): 609–10. https://doi.org/10.1016/j.cortex.2012.12.016.

MacCoun, Robert, and Saul Perlmutter. 'Blind Analysis: Hide Results to Seek the Truth'. *Nature News* 526, no. 7572 (8 October 2015): 187. https://doi.org/10.1038/526187a.

Just to mention…
Find a list of all journals offering Registered Reports here: https://cos.io/rr/

Also, this might be a good starting point if you are looking to write your own high-quality preregistration: Veer, Anna Elisabeth van 't, and Roger Giner-Sorolla. 'Pre-Registration in Social Psychology—A Discussion and Suggested Template'. *Journal of Experimental Social Psychology, Special Issue: Confirmatory*, 67 (1 November 2016): 2–12. https://doi.org/10.1016/j.jesp.2016.03.004.
If you are interested in preregistering an analysis of pre-existing data, have a look at these recent papers:

- Weston, Sara J., Stuart J. Ritchie, Julia M. Rohrer, and Andrew K. Przybylski. 'Recommendations for Increasing the Transparency of Analysis of Preexisting Data Sets'. *Advances in Methods and Practices in Psychological Science* 2, no. 3 (1 September 2019): 214–27. https://doi.org/10.1177/2515245919848684.
- Van den Akker, Olmo, Sara J Weston, Lorne Campbell, William J. Chopik, Rodica I. Damian, Pamela Davis-Kean, Andrew Nolan Hall, et al. 'Preregistration of Secondary Data Analysis: A Template and Tutorial'. Preprint, 20 November 2019. https://doi.org/10.31234/osf.io/hvfmr.

There have been recent debates about the value of preregistration both generally and in cognitive modelling specifically. These are three references for those interested in this:

- Szollosi, Aba, David Kellen, Danielle Navarro, Rich Shiffrin, Iris van Rooij, Trisha Van Zandt, and Chris Donkin. 'Is Preregistration Worthwhile?' Preprint, 31 October 2019. https://doi.org/10.31234/osf.io/x36pz.
- Wagenmakers, E. J. 'A Breakdown of "Preregistration Is Redundant, at Best"'. *Bayesian Spectacles* (blog), 5 November 2019. https://www.bayesianspectacles.org/a-breakdown-of-preregistration-is-redundant-at-best/.
- Quintana, Dan, and Heathers, James. 'Everything Hertz: 98: Episode Titles Are Redundant, at Best (with Sophia Crüwell)'. *Everything Hertz*. Accessed 5 January 2020. https://everythinghertz.com/98. (first half of the episode)

If you are interested in the debate around creativity and open science, these are two great papers to start with:

- Wagenmakers, Eric-Jan, Gilles Dutilh, and Alexandra Sarafoglou. 'The Creativity-Verification Cycle in Psychological Science: New Methods to Combat Old Idols'. *Perspectives on Psychological Science* 13, no. 4 (1 July 2018): 418–27. https://doi.org/10.1177/1745691618771357.
- Frankenhuis, Willem E., and Daniel Nettle. 'Open Science Is Liberating and Can Foster Creativity'. *Perspectives on Psychological Science* 13, no. 4 (1 July 2018): 439–47. https://doi.org/10.1177/1745691618767878.

Week 5 – Fraud, Errors and Scientific Self-Correction

Understanding what makes fraud different from errors; methods used to detect errors; whether scientific self-correction is myth or reality; scientific errors and debates about how to address them

Assignment and readings

I have set three newspaper and magazine articles for this week. The first gives you an insight into a 'pizza paper' scandal, while the other two will form the basis of our case studies this week. Please read them all carefully, especially the 2nd and 3rd piece, as they will form a crucial part of our discussion.

1. Lee, Stephanie. 'Sliced And Diced: The Inside Story Of How An Ivy League Food Scientist Turned Shoddy Data Into Viral Studies'. *BuzzFeed News*. Accessed 4 January 2020. https://www.buzzfeednews.com/article/stephaniemlee/brian-wansink-cornell-p-hacking.
2. Dominus, Susan. 'When the Revolution Came for Amy Cuddy'. *The New York Times*, 18 October 2017. https://www.nytimes.com/2017/10/18/magazine/when-the-revolution-came-for-amy-cuddy.html?rref=collection%2Fbyline%2Fsusan-dominus&action=click&contentCollection=undefined&region=stream&module=stream_unit&version=latest&contentPlacement=1&pgtype=collection.
3. Vazire, Simine. 'Criticizing a Scientist's Work Isn't Bullying. It's Science.' *Slate Magazine*, 24 October 2017. https://slate.com/technology/2017/10/criticizing-a-scientists-work-isnt-bullying.html.

Case studies

Our case studies will focus on the process of pointing out scientific errors and what counts as bullying. I set two articles which you all read before coming to this session. The first is an article about Amy Cuddy who was a proponent of 'power posing' a concept heavily criticised for scientific errors and low-quality science. The second blog post argues against the first, making the case that criticising others' work is not bullying. 1-2 students who have volunteered will prepare to lead and chair our discussion about error correction in science: reading the articles in detail, preparing questions, encouraging and leading the debate.

Core references

Quintana, Dan, and Heathers, James. '54: Cuckoo Science'. Everything Hertz. https://everythinghertz.com/54.

Quintana, Dan, and Heathers, James. '74: Seeing Double (with Elisabeth Bik)'. https://everythinghertz.com/74. (also see Elisabeth Bik's blog: https://scienceintegritydigest.com/2019/11/23/scanning-for-duplications/)

Brown, Nicholas J. L., and James A. J. Heathers. 'The GRIM Test: A Simple Technique Detects Numerous Anomalies in the Reporting of Results in Psychology'. *Social Psychological and Personality Science* 8, no. 4 (1 May 2017): 363–69. https://doi.org/10.1177/1948550616673876.

Heathers, James. 'Introducing SPRITE (and the Case of the Carthorse Child)'. *Hackernoon* (blog). Accessed 5 January 2020. https://hackernoon.com/introducing-sprite-and-the-case-of-the-carthorse-child-58683c2bfeb.

Nuijten, Michèle B., Chris H. J. Hartgerink, Marcel A. L. M. van Assen, Sacha Epskamp, and Jelte M. Wicherts. 'The Prevalence of Statistical Reporting Errors in Psychology (1985–2013)'. *Behavior Research Methods* 48, no. 4 (1 December 2016): 1205–26.

https://doi.org/10.3758/s13428-015-0664-2. (also see Statcheck the website: http://statcheck.io/)

Vazire, Simine. 'A Toast to the Error Detectors'. *Nature* 577, no. 7788 (30 December 2019): 9–9. https://doi.org/10.1038/d41586-019-03909-2.

Bishop, D. V. M. 'Fallibility in Science: Responding to Errors in the Work of Oneself and Others'. *Advances in Methods and Practices in Psychological Science* 1, no. 3 (1 September 2018): 432–38. https://doi.org/10.1177/2515245918776632.

Oransky, Ivan. 'Nobel Winner Retracts Paper from Science'. *Retraction Watch* (blog), 2 January 2020. https://retractionwatch.com/2020/01/02/nobel-winner-retracts-paper-from-science/.

Rohrer, Julia Marie, Warren Tierney, Eric Luis Uhlmann, Lisa Marie DeBruine, Tom Heyman, Benedict C Jones, Stefan C. Schmukle, et al. 'Putting the Self in Self-Correction'. Preprint, 12 December 2018. https://doi.org/10.31234/osf.io/exmb2.

Ebersole, Charles R., Jordan R. Axt, and Brian A. Nosek. 'Scientists' Reputations Are Based on Getting It Right, Not Being Right'. *PLOS Biology* 14, no. 5 (12 May 2016): e1002460. https://doi.org/10.1371/journal.pbio.1002460.

Heathers, James (Twitter): https://twitter.com/jamesheathers/status/845696144999137280

Ioannidis, John P. A. 'Why Science Is Not Necessarily Self-Correcting': *Perspectives on Psychological Science*, 7 November 2012. https://doi.org/10.1177/1745691612464056.

Week 6 – The Statistics Wars (+ Power)

Error Rates; Null-Hypothesis Significance Testing; *p*-values; power; Bayesian methods (estimation and Bayes Factors); 'Redefine Significance' vs 'Justify your Alpha'

Assignment and readings
In this lecture we will be covering the difference between Null Hypothesis Significance Testing and Bayesian methodology. Furthermore we will be examining the importance of power. To prepare for this lecture please read the following papers:

1. Button, Katherine S., John P. A. Ioannidis, Claire Mokrysz, Brian A. Nosek, Jonathan Flint, Emma S. J. Robinson, and Marcus R. Munafò. 'Power Failure: Why Small Sample Size Undermines the Reliability of Neuroscience'. *Nature Reviews Neuroscience* 14, no. 5 (May 2013): 365–76. https://doi.org/10.1038/nrn3475.
2. Please have a look over the two papers we will discuss in the case study

Case study
In the case study we will be looking at the debate between 'Redefine Significance' and 'Justify Your Alpha'. Two groups of volunteers will act in a mock debate about 'This lecture theatre believes that significance levels should be adjusted to $p < 0.005$", having prepared at least three arguments to present for their 'side' [no slides] reading the following papers:

1. For the 'Redefine Significance' side: Benjamin, Daniel J., James O. Berger, Magnus Johannesson, Brian A. Nosek, E.-J. Wagenmakers, Richard Berk, Kenneth A. Bollen, et al. 'Redefine Statistical Significance'. *Nature Human Behaviour* 2, no. 1 (January 2018): 6–10. https://doi.org/10.1038/s41562-017-0189-z.
2. For the 'Justify your Alpha' side: Lakens, Daniel, Federico G. Adolfi, Casper J. Albers, Farid Anvari, Matthew A. J. Apps, Shlomo E. Argamon, Thom Baguley, et al. 'Justify Your Alpha'. *Nature Human Behaviour* 2, no. 3 (March 2018): 168–71. https://doi.org/10.1038/s41562-018-0311-x. (this is a reaction to Benjamin et al. 2018; also see Daniel Laken's tweet: https://twitter.com/lakens/status/1135469260686397440)

Note: This is also an interesting panel session and discussion between EJ Wagenmakers (pro-Bayes/redefine statistical significance), Daniel Lakens (pro-NHST/ justify your alpha) and Simine Vazire (pro-redefine statistical significance) held in 2017 at the BITSS Annual Meeting where they went head-to-head in the 'Redefine Statistical Significance' vs. 'Justify Your Alpha' discussion (~ 1 hour): https://youtu.be/4IgSVxkXMaM?t=2102

Core references
Nuzzo, Regina. 'Tips for Communicating Statistical Significance'. National Institutes of Health (NIH), 15 October 2018. https://www.nih.gov/about-nih/what-we-do/science-health-public-trust/perspectives/science-health-public-trust/tips-communicating-statistical-significance.

Fidler, Fiona, Felix Singleton Thorn, Ashley Barnett, Steven Kambouris, and Ariel Kruger. 'The Epistemic Importance of Establishing the Absence of an Effect'. *Advances in Methods and Practices in Psychological* Science 1, no. 2 (1 June 2018): 237–44.
https://doi.org/10.1177/2515245918770407.

Just to mention…
If you are interested in a better understanding of some core statistical issues in psychology and how you could perform better statistical inferences I highly recommend Daniel Laken's freely accessible Coursera course, *Improving your statistical inferences*:
https://www.coursera.org/learn/statistical-inferences#syllabus

Due to the success of this course Daniel Lakens has now recorded a second course, *Improving your statistical questions,* which you can access here: https://www.coursera.org/learn/improving-statistical-questions

If you are interested in Bayesian methods I would recommend the following:

1. Zoltan Dienes (2008). *Understanding Psychology as a Science.* London, UK: Palgrave MacMillan. (Last couple of chapters on statistical testing give a better and more detailed overview of the Bayesian ideas)
2. For an accessible but very rigorous introduction to Bayesian estimation with extremely helpful examples in *R* I would recommend Richard McElreath, *Statistical Rethinking: A Bayesian Course with R Examples* (2016, second edition coming out 2020). You can find a 19-part 2017 lecture series that accompanies his book on Youtube: https://www.youtube.com/playlist?list=PLDcUM9US4XdM9_N6XUUFrhghGJ4K25bFc
3. For an introduction into the scientific literature behind Bayesian statistics I would suggest Etz, Alexander, Quentin Frederik Gronau, Fabian Dablander, Peter Adriaan Edelsbrunner, and Beth Baribault. 'How to Become a Bayesian in Eight Easy Steps: An Annotated Reading List'. Preprint, 15 August 2016. https://doi.org/10.31234/osf.io/ph6sw.

Lastly, if you want to learn more about *R*, I would recommend Wickham and Grolemund (2017), *R for Data Science*, O'Reilly. The book is available online (in full!) on the author's website: https://r4ds.had.co.nz/

Furthermore you can look at some of the amazing *Software Carpentry* courses on R:
http://swcarpentry.github.io/r-novice-gapminder/
http://swcarpentry.github.io/r-novice-inflammation/

Week 7 – Journals, Publishing and Computational Reproducibility

Introduction to the for-profit publishing model; new forms of publishing; pre-prints; Open Access; peer review; Open Data and computational reproducibility

Assignment and readings

1. Buranyi, Stephen. 'Is the Staggeringly Profitable Business of Scientific Publishing Bad for Science? | Science | The Guardian'. *The Guardian*, 27 June 2017. https://www.theguardian.com/science/2017/jun/27/profitable-business-scientific-publishing-bad-for-science.
2. Tennant, Jonathan P., François Waldner, Damien C. Jacques, Paola Masuzzo, Lauren B. Collister, and Chris. H. J. Hartgerink. 'The Academic, Economic and Societal Impacts of Open Access: An Evidence-Based Review'. *F1000Research* 5 (21 September 2016): 632. https://doi.org/10.12688/f1000research.8460.3. (it is a long paper, so a scan through it is enough)

Case studies

We will all read the following paper, and six volunteers will present its core recommendation during the class [5 minutes each, no more than 2 slides]. Each of the volunteers will cover one of the following: data management, software, collaboration, project organization, keeping track of changes, manuscripts.

Wilson, Greg, Jennifer Bryan, Karen Cranston, Justin Kitzes, Lex Nederbragt, and Tracy K. Teal. 'Good Enough Practices in Scientific Computing'. *PLOS Computational Biology* 13, no. 6 (22 June 2017): e1005510. https://doi.org/10.1371/journal.pcbi.1005510.

Core references

Tennant, Jon. 'Elsevier Are Corrupting Open Science in Europe'. *The Guardian*, 29 June 2018, sec. Science. https://www.theguardian.com/science/political-science/2018/jun/29/elsevier-are-corrupting-open-science-in-europe.

Quintana, Dan, and Heathers, James. '93: Double-Blind Peer Review vs. Open Science'. Everything Hertz. Accessed 5 January 2020. https://everythinghertz.com/93.

Tennant, Jonathan P., Jonathan M. Dugan, Daniel Graziotin, Damien C. Jacques, François Waldner, Daniel Mietchen, Yehia Elkhatib, et al. 'A Multi-Disciplinary Perspective on Emergent and Future Innovations in Peer Review'. *F1000Research* 6 (29 November 2017): 1151. https://doi.org/10.12688/f1000research.12037.3.

Gowers, Timothy. 'The End of an Error?' *TLS*, 27 October 2017. https://www.the-tls.co.uk/articles/the-end-of-an-error-peer-review/.

Smith, Richard. 'Peer Review: A Flawed Process at the Heart of Science and Journals'. *Journal of the Royal Society of Medicine* 99, no. 4 (April 2006): 178–82.

Houtkoop, Bobby Lee, Chris Chambers, Malcolm Macleod, Dorothy V. M. Bishop, Thomas E. Nichols, and Eric-Jan Wagenmakers. 'Data Sharing in Psychology: A Survey on Barriers and Preconditions'. *Advances in Methods and Practices in Psychological Science* 1, no. 1 (1 March 2018): 70–85. https://doi.org/10.1177/2515245917751886.

This whole issue of AMPPS has some great discussions and tutorials about data sharing: https://journals.sagepub.com/toc/ampa/1/1

Morey, Richard D., Christopher D. Chambers, Peter J. Etchells, Christine R. Harris, Rink Hoekstra, Daniël Lakens, Stephan Lewandowsky, et al. 'The Peer Reviewers' Openness Initiative: Incentivizing Open Research Practices through Peer Review'. *Royal Society Open Science* 3, no. 1 (n.d.): 150547. https://doi.org/10.1098/rsos.150547.

Bishop, D. V. M. 'Open Research Practices: Unintended Consequences and Suggestions for Averting Them. (Commentary on the Peer Reviewers' Openness Initiative)'. *Royal Society Open Science* 3, no. 4 (n.d.): 160109. https://doi.org/10.1098/rsos.160109.

Stark, Philip B. 'Before Reproducibility Must Come Preproducibility'. *Nature* 557, no. 7707 (24 May 2018): 613–613. https://doi.org/10.1038/d41586-018-05256-0.

Just to mention…
If you are interested in learning how to use version control software I would look at this very good course from *Software Carpentry*: http://swcarpentry.github.io/git-novice/

Lastly, if you are thinking about preprinting your work, but want to make your manuscript look prettier (instead of using horrible APA formatting), have a look at Brenton Wiernik's great Word preprint templates that liken popular journals: https://osf.io/hsv6a/

If you want to make a website for yourself:
https://twitter.com/dsquintana/status/993410504570888192
https://debruine.github.io/tutorials/webpages.html

Week 8 – Psychology as a Robust Science

What makes robust inferences; triangulation and causality; theory and measurement; Psychological Science Accelerator; incentives and next steps

Assignment and readings
This week's assignments are a bit more extensive and challenging. If you struggle with any one of these, feel free to move on and we can discuss them in the lecture:

1. Yarkoni, Tal, and Jacob Westfall. 'Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning'. *Perspectives on Psychological Science* 12, no. 6 (1 November 2017): 1100–1122. https://doi.org/10.1177/1745691617693393.
2. Meehl, Paul E. 'Theory-Testing in Psychology and Physics: A Methodological Paradox'. *Philosophy of Science* 34, no. 2 (June 1967): 103–15. https://doi.org/10.1086/288135.
3. Smaldino, Paul E., and Richard McElreath. 'The Natural Selection of Bad Science'. *Royal Society Open Science* 3, no. 9 (n.d.): 160384. https://doi.org/10.1098/rsos.160384. (or read an accessible summary here: http://www.spsp.org/news-center/blog/bad-science-evolves)

Case studies
This week we will all take a couple of minutes to read the blog post below in the tea break. We will then discuss and examine its claims around incentives. This will lead us into a discussion about the current incentives in the field:

Yarkoni, Tal. 'No, It's Not The Incentives—It's You'. *[Citation Needed]* (blog), 2 October 2018. https://www.talyarkoni.org/blog/2018/10/02/no-its-not-the-incentives-its-you/.

Core references
Borsboom, Denny. 'Theoretical Amnesia'. *Open Science Collaboration Blog* (blog). Accessed 5 January 2020. http://osc.centerforopenscience.org/2013/11/20/theoretical-amnesia/.

Munafò, Marcus R., and George Davey Smith. 'Robust Research Needs Many Lines of Evidence'. *Nature* 553, no. 7689 (January 2018): 399–401. https://doi.org/10.1038/d41586-018-01023-3.

Flake, Jessica Kay, and Eiko I Fried. 'Measurement Schmeasurement: Questionable Measurement Practices and How to Avoid Them'. Preprint, 17 January 2019. https://doi.org/10.31234/osf.io/hs7wm.

Yarkoni, Tal. 'The Generalizability Crisis'. Preprint, 22 November 2019. https://doi.org/10.31234/osf.io/jqw35.

Moshontz, Hannah, Lorne Campbell, Charles R. Ebersole, Hans IJzerman, Heather L. Urry, Patrick S. Forscher, Jon E. Grahe, et al. 'The Psychological Science Accelerator: Advancing Psychology Through a Distributed Collaborative Network'. *Advances in Methods and Practices in Psychological Science* 1, no. 4 (1 December 2018): 501–15. https://doi.org/10.1177/2515245918797607.

Chartier, Chris, Randy McCarthy, and Heather Urry. 'The Psychological Science Accelerator'. *APS Observee* 31, no. 3 (28 February 2018). https://www.psychologicalscience.org/observer/the-psychological-science-accelerator.

Nosek, B. A., G. Alter, G. C. Banks, D. Borsboom, S. D. Bowman, S. J. Breckler, S. Buck, et al. 'Promoting an Open Research Culture'. *Science* 348, no. 6242 (26 June 2015): 1422–25. https://doi.org/10.1126/science.aab2374.

Spellman, Barbara A. 'A Short (Personal) Future History of Revolution 2.0'. *Perspectives on Psychological Science* 10, no. 6 (1 November 2015): 886–99. https://doi.org/10.1177/1745691615609918.

Pashler, Hal, and J. P. de Ruiter. 'Taking Responsibility for Our Field's Reputation'. *APS Observer* 30, no. 7 (31 August 2017). https://www.psychologicalscience.org/observer/taking-responsibility-for-our-fields-reputation.

Just to mention…
If you are interested in staying involved in the Open Science community or learning more skills here are some suggestions:

- Follow people and accounts on Twitter like APS, SIPS, PSA, Center for Open Science, Brian Nosek, Simine Vazire, Chris Chartier, Dorothy Bishop, Daniel Lakens, Anne Scheel, Katie Corker, Tal Yarkoni, EJ Wagenmakers, Marcus Munafo, Jessica Flake, Eiko Fried, Julia Rohrer, James Heathers etc.
- Do Daniel Laken's online courses or the Open Science MOOC (https://opensciencemooc.eu)
- Listen to Podcasts like Everything Hertz, ReproducibiliTea or Black Goat Pod
- Join the Society for the Improvement of Psychological Science (SIPS) and go to the SIPS Conference
- Keep your eyes peeled for events from the UK Reproducibility Network (UKRN)
- Have a look at ReproducibiliTea and join the Cambridge journal club: reproducibilitea.org
- Don't be worried about implementing everything at once, many people say Open Science is like a buffet where you can choose what is best for you and always go back to get more
- Email me if you have any questions, queries or ideas: aco35@cam.ac.uk

If you are interested in learning more about causality and causal claims, I would recommend the following books and papers:

- Rohrer, Julia M. 'Thinking Clearly About Correlations and Causation: Graphical Causal Models for Observational Data'. *Advances in Methods and Practices in Psychological Science* 1, no. 1 (1 March 2018): 27–42. https://doi.org/10.1177/2515245917745629.
- Pearl, Judea, and Dana Mackenzie. *The Book of Why: The New Science of Cause and Effect*. New York: Basic Books, 2018.
- Marinescu, Ioana E., Patrick N. Lawlor, and Konrad P. Kording. 'Quasi-Experimental Causality in Neuroscience and Behavioural Research'. *Nature Human Behaviour* 2, no. 12 (December 2018): 891–98. https://doi.org/10.1038/s41562-018-0466-5.
- Fiedler, Klaus, Malte Schott, and Thorsten Meiser. 'What Mediation Analysis Can (Not) Do'. *Journal of Experimental Social Psychology* 47, no. 6 (November 2011): 1231–36. https://doi.org/10.1016/j.jesp.2011.05.007.

Acknowledgements
I would like to thank Simine Vazire, Katie Corker, David Funder and Benjamin Le whose syllabi on the Open Science Framework both guided and inspired me (https://osf.io/vkhbt/). I also thank Marcus Munafo, Lisa DeBruine, Chris Chambers, Laura Fortunato, Dorothy Bishop, Jackie Thompson, Jon Tennant, Christina Riesenweber, Daniel Lakens, Julia Rohrer, Dominique Heinke, Kate Button, Verena Heise and Kirstie Whitaker whose slides made available on the OSF or Figshare helped influence and improve some of my lectures.