Research Transparency and Reproducibility Training (RT2)
September 20-22, 2017
London, England

Welcome to the BITSS community! We are pleased to host you for our 3-day Research Transparency and Reproducibility Training (RT2), in London, England September 20-22, 2017. To ensure you get the most out of this three-day event, BITSS has prepared this **Participant Manual** with a Reading List and instructions for preparing for the hands-on sessions, including required software downloads.

Participants of RT2 will learn about a range of innovative practices and tools, including:

- *Pre-registration.* The registration of study designs in public repositories prior to data collection allows for better tracking of the universe of studies in a given domain, including studies with null results that are rarely published. This begins to tackle the "file-drawer problem" whereby only statistically significant findings are reported;
- *Pre-analysis Plans.* The design and use of a pre-analysis plan (PAP)—a step-by-step plan, written before data are accessed, describing hypotheses and strategies for analyzing data—can help protect against data mining and reduce researcher "degrees of freedom" in confirmatory research;
- *Meta-analysis.* Innovations in the design of meta-analyses—dealing with issues of bias, study sample size, and model selection—can improve the quality of inferences made from the analyses of pooled studies;
- *Data de-identification.* To facilitate open science, researchers must work toward public posting of the data and code needed to replicate findings of published studies. However, this requires understanding of and training on how to balance maximizing data's usability with protection of human subjects and data confidentiality by using methods for data de-identification; and
- *Tools for transparent workflows.* There are a plethora of software and online tools to facilitate transparent and reproducible workflows, such as the Open Science Framework (OSF), Git, R, and dynamic documents.

With this training, BITSS aims to *directly impact researchers' practices in favor of transparency and reproducibility*. We focus on topics such as pre-registration, pre-analysis plans, and version control so that you can apply these tools to your own work. BITSS hopes that RT2 events will have long-term, sustainable impacts on scientific norms and practices as learners and faculty like you continue to incorporate innovative tools and methods into curricula and coursework at your own institutions.

If you are interested in joining our community, please visit our website to learn more about the **BITSS Catalyst Program**. Please also visit our BITSS Preprints service for working papers on research transparency and reproducibility topics. We welcome submissions of working papers, pre/post prints by either posting directly from OSF project pages or sending to ucbitss@berkeley.edu.

Pre-Training Suggested Reading List

*This is a list of foundational and more recent literature related to social science research transparency and reproducibility challenges, as well as potential solutions and best practices. We suggest reading the **starred papers before RT2.*

Foundational literature

**Ioannidis JPA. 2005. "Why Most Published Research Findings Are False." *PLoS Med* 2(8): e124. doi:10.1371/journal.pmed.0020124. PMCID PMC1182327. Link.

Leamer, Edward. 1983. "Let's Take the Con Out of Econometrics." *American Economic Review*, 73(1): 31–43. Link.

Merton, Robert K. 1973 [1942]. "The Normative Structure of Science." in Merton, Robert K., *The Sociology of Science: Theoretical and Empirical Investigations.* Chicago: University of Chicago Press. ISBN 978-0-226-52091-9, OCLC 755754. Link.

**Miguel, E., C. Camerer, K. Casey, J. Cohen, K. M. Esterling, A. Gerber, R. Glennerster, et al. 2014. "Promoting Transparency in Social Science Research." *Science* 343 (6166): 30–31. doi:10.1126/science.1245317. Link.

Nosek, B. A., et al. 2015. "Promoting an open research culture: Author guidelines for journals could help to promote transparency, openness, and reproducibility." *Science* (New York, NY) 348.6242: 1422. PMCID PMC4550299. Link.

Open Science Collaboration. 2015. "Estimating the reproducibility of psychological science." *Science* 349, no. 6251: aac4716. PMID: 26315443. Link.

Rosenthal, Robert. 1979. "The file drawer problem and tolerance for null results." *Psychological Bulletin* 86.3: 638. Link.

Christensen, Garret, and Edward Miguel. 2017. "Transparency, Reproducibility, and the Credibility of Economics Research". BITSS PrePrints. Link.

Research Reproducibility

Begley, C. Glenn, and Lee M. Ellis. 2012. "Drug development: Raise standards for preclinical cancer research." *Nature* 483, no. 7391: 531-533. Link.

**Goodman, S. N., Fanelli, D., & Ioannidis, J. P. 2016. "What does research reproducibility mean?" *Science Translational Medicine*, Vol. 8. Ch. 341. Link.

P-curve

**Simonsohn, Uri, Leif D. Nelson, and Joseph P. Simmons. 2014: "P-curve: a key to the file-drawer." *Journal of Experimental Psychology: General* 143, no. 2: 534. Link.

Simmons, Joseph P., Leif D. Nelson, and Uri Simonsohn. 2011. "False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant." *Psychological Science* 22, no. 11: 1359-1366. Link.

Researcher Degrees of Freedom

**Wicherts, Jelte M., Coosje LS Veldkamp, Hilde EM Augusteijn, Marjan Bakker, Robbie CM Van Aert, and Marcel ALM Van Assen. 2016. "Degrees of freedom in planning, running,

analyzing, and reporting psychological studies: A checklist to avoid p-hacking." *Frontiers in Psychology* 7. Link. If you cannot access the previous link, access a preprint here.

Pre-Registration and Pre-Analysis Plans

Casey, Katherine, Rachel Glennerster, and Edward Miguel. 2012. "Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan." *The Quarterly Journal of Economics* 127 (4): 1755–1812. Link.

Replication

Dafoe, Allan. 2014. "Science Deserves Better: The Imperative to Share Complete Replication Files." *PS: Political Science & Politics* 47 (1): 60–66. doi:10.1017/S104909651300173X. Link.

Hamermesh, Daniel S. 2007. "Viewpoint: Replication in Economics." *Canadian Journal of Economics/Revue Canadienne D'économique* 40 (3): 715–33. doi:10.1111/j.1365-2966.2007.00428.x. Link.

**Klein, Richard A., Kate A. Ratliff, Michelangelo Vianello, Reginald B. Adams Jr, Štěpán Bahník, Michael J. Bernstein, Konrad Bocian et al. 2014. "Investigating variation in replicability: A 'Many Labs' Project." *Social Psychology*. Link.

Data De-Identification

Goodman, Alyssa, et al. 2014. "Ten Simple Rules for the Care and Feeding of Scientific Data", *PLoS Computational Biology*, 10(4), e1003542. Link.

Kaiser, Karen. "Protecting respondent confidentiality in qualitative research." 2009. *Qualitative Health Research* 19, no. 11: 1632-1641. Link.

Playford, Christopher J., Vernon Gayle, Roxanne Connelly, and Alasdair JG Gray. 2016. "Administrative social science data: The challenge of reproducible research." *Big Data & Society* 3, no. 2: 2053951716684143. Link.

Responsible Data Forum. "The Handbook of the Modern Development Specialist." https://responsibledata.io/resources/handbook/ CC-BY-SA 4.0.

Sturdy, Jennifer, Stephanie Burch, Heather Hanson, and Jack Molyneaux. 2017. "Opening up Evaluation Microdata: Balancing Risks and Benefits of Research Transparency". BITSS Preprints. Link.

Zandbergen, Paul A. "Ensuring confidentiality of geocoded health data: assessing geographic masking strategies for individual-level data." 2014. *Advances in Medicine*.

Meta-analysis

**Borenstein, M., Hedges, L. V., Higgins, J. P. T. and Rothstein, H. R. 2007. "Fixed vs Random effects", in *Introduction to Meta-Analysis*, John Wiley & Sons, Ltd, Chichester, UK. Link.

Ioannidis, J.P.A., Fanelli, D., Dunne, D.D., Goodman, S.N. 2015. "Meta-research: Evaluation and Improvement of Research Methods and Practices." *PLoS Biol* 13(10): e1002264. doi:10.1371/journal.pbio.1002264. PMCID PMC4592065. Link.

Hsiang, Solomon M., Marshall Burke, and Edward Miguel. 2013. "Quantifying the Influence of Climate on Human Conflict." *Science* 341 (6151): 1235367. doi:10.1126/science.1235367. Link.

**Russo, Mark. 2007. "How to Review a Meta-Analysis." *Gastroenterol Hepatol* 3(8): 637–642. Link.

Transparency Reporting and Disclosure

Simera, et al. 2010. "Commentary: Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network." *BMC Medicine* 2010, Vol 8, Ch. 24. Link.

Power and Priors

**Benjamin, Daniel J, James Berger, Magnus Johannesson, Brian A Nosek, Eric-Jan Wagenmakers, Richard Berk, Kenneth Bollen, et al. 2017. "Redefine Statistical Significance". PsyArXiv. July 22. osf.io/preprints/psyarxiv/mky9j. Link.

Burlig, Fiona, Louis Preonas, and Matt Woerman. 2017. "Panel data and experimental design." *Energy Institute at Haas Working Paper #277*. Link. For a lighter read, please reference the blog post that tries to be less technical.

Button, Katherine S., J.P.A. Ioannidis, C. Mokrysz, B. Nosek, J. Flint, E.S.J. Robinson, M. Munafo. "Power failure: why small sample size undermines the reliability of neuroscience." *Nature Reviews Neuroscience* 14.5 (2013): 365-376. Doi 10.1038/nrn3475 PMID: 23571845. Link.

Other Useful Resources

Manual of Best Practices

Christensen, Garret, and Courtney Soderberg. 2016. "Manual of best practices in transparent social science research." Berkeley, CA: University of California. Link.

Social Science Meta-Analysis and Research Transparency (SSMART) Projects

Findings from the Social Science Meta-Analysis and Research Transparency (SSMART) program are summarized here with direct links to their working papers.

BITSS Online Resource Library

We've compiled a wealth of tools and software, guidelines and templates, repositories, slide decks, and videos that you may find useful in making your research more transparent and reproducible. We also list a growing number of blogs, commentary, and podcasts discussing challenges and innovations in the evolving open science movement. Find the BITSS Resource Library here.

Stay tuned!

BITSS Faculty Director Ted Miguel and Project Scientist Garret Christensen are working on a forthcoming textbook on research transparency in the social sciences.

BITSS also plans to launch a second run of our Massive Open Online Course (MOOC) "Transparent and Open Social Science Research" on the FutureLearn platform in Winter 2017. Read more about the course here.

Pre-Training Actions

*BITSS requires all participants currently working on their own research to take the following actions **before September 20**. Participants who manage, fund, or publish research will be paired with other participants during hands-on sessions to observe the process.*

Please reference the Software section of our Resources page on the BITSS website for more information on this software.

1. Establish OSF Account

The Open Science Framework (OSF) allows you to store your research files and link together all your research across several platforms, such as Dropbox, Harvard's Dataverse, and GitHub. It version controls any files you upload and you can register a project to create a frozen time-stamped version with a persistent URL. So by writing a pre-analysis plan, you could prove to the world that your significant results aren't just a successful fishing expedition. Sign up for a free account here.

2. Review study pre-registration

In the Pre-Registration hands-on session, participants will be able to walk through how to register a study on the OSF. Please see **Appendix 1** for the information that should be prepared for a pre-registration on the OSF.

3. Install Git and Create Github.com Account

The date-and-initial version of keeping track of changes to your files doesn't really cut it when you're doing something complicated or you've got a lot of co-authors. If you want your work to be reproducible, use version control. It has a learning curve even for xkcd-type people, but it's worth it! (Read Gentzkow and Shapiro chapter 3 on why.) Software Carpentry and GitHub have great tutorials.

To get started, download the GitHub Desktop GUI app. Do not download from the top of the page under the most obvious "the new native" link, but rather, scroll to the bottom and download from the older "Not ready for Desktop Beta?" links. The difference is important! Note that this is only available for Windows and Mac users. Linux users can use the command line or pick one of the other GUIs listed here. If you are comfortable using the command line, we also recommend Windows users install Git Bash.

Next, create an account with GitHub.com. GitHub is a popular online storage platform for your repositories (folders/projects) that are version-controlled with Git.

4. Install software for Dynamic Documents

You can write your code and your paper in one place. This means you won't mess anything up copying and pasting, and you'll never have to wonder what code it was that produced which figure, where on earth you saved it, or whether the paper has the updated version.

In R, this can be done with R Markdown, which is built into R Studio - please download and install R and R Studio. When you open a new R Markdown file in R Studio, it starts with a really simple example, or you can learn more here.

In Stata, this can be done with the user-written command MarkDoc with the following commands:
ssc install markdoc
ssc install weaver
ssc install statax

The package may have been updated recently, so you might want to run "adoupdate" if you installed it a while ago. The syntax is explained in the built-in help file. For MarkDoc to work you also need to install Pandoc, a pretty cool Swiss-army knife that converts almost any markup file to almost any other, as well as wkhtmltopdf. If you install as above, these may be installed automatically, but you may have to click on a link that will show up inside Stata.

5: Install LaTeX

Microsoft Word is nice and easy for writing short papers, but when you start writing longer papers, or you want to include any equations or formatting it quickly becomes cumbersome. LaTeX is better for reproducibility since when you include your figures, you just refer to files, so there's no question of whether you remembered to update or not. LaTeX (download here) is also used by R Markdown when you make pdf's, so you have to at least have it installed in the background. *This is a large file, and you have to install the full version, so don't leave this until the last minute.* If you don't install this, you won't be able to make PDF's with the above dynamic documents software.

6: Install a Decent Text Editor

You need a good way to edit plain text. On a Mac, the simplest thing to do is use the built-in TextEdit, but you will need to change the default so plain text, not rich text (rtf) is the output format. On Windows, you can use Notepad if you like, but we suggest something a little more powerful, like Atom, or Notepad++, or Sublime Text. These have syntax highlighting, and add-on packages that can render markdown, and things like that.

Meet the RT2 Faculty!

Marjan Bakker (Scientific Misconduct and Researcher Degrees of Freedom)

Marjan Bakker is an assistant professor at Tilburg University in the methodology and statistics department. She teaches courses on regression analysis and multilevel analysis to psychology and research master students. Dr. Bakker is part of the Meta-Research Center at Tilburg University. Her research interests include scientific integrity, errors in statistics, power, publication bias, psychometrics, preregistration, and game theory. Dr. Bakker has published extensively on meta-research, including the well-known "The (mis)reporting of statistical results in psychology journals".

Sean Grant (Transparent Reporting and Disclosure)

Sean Grant is a behavioral and social scientist at the RAND Corporation and affiliated faculty at the Pardee RAND Graduate School. He is interested in advancing the overall transparency, openness, and rigor of intervention research for supporting evidence-based policy and practice. He principally evaluates interventions for substance use, though he is passionate about applying intervention research methods to topics spanning the behavioral, social, and health sciences. Current projects related to research transparency focus on reporting standards for social intervention research, and the feasibility and appropriateness of considering research transparency as part of the Institutional Review Board review process. Sean received a 2015 Leamer-Rosenthal Prize for Open Social Science and is a BITSS Catalyst.

Nicole Janz (Replication)

Nicole Janz is an Assistant Professor at the School of Politics and International Relations at the University of Nottingham. Her research interests include human rights, foreign direct investment, corruption, and slavery. In her current projects, she examines the effects of foreign direct investment on repression and labor rights; how expropriation hinders human rights; judicial delays, impunity and corruption; how human rights shaming influences FDI; and the IMF. Nicole is an ambassador at the Center for Open Science (COS) and a BITSS Catalyst. Before working at Nottingham, she taught statistics for social scientists at the Social Sciences Research Methods Centre, University of Cambridge. Nicole completed her PhD in Politics and International Studies at the Department of Politics and International Studies at Cambridge. Nicole is also a BITSS Catalyst.

Thomas Leeper (Dynamic Documents and Version Control)

Thomas J. Leeper is an Associate Professor in Political Behavior in the Department of Government at the London School of Economics and a BITSS Catalyst. His research on American and European public opinion uses survey and experimental methods to understand how citizens' political viewpoints reflect an interaction between the broader information environment (including the mass media and political elites) and individual-level attributes, namely citizens' expressed behaviors, psychological traits, social identities, and motivations. His work has been published in the American Political Science Review, American Journal of Political Science, Public Opinion Quarterly, Political Psychology, and

elsewhere. He has also published more than 25 packages for the R statistical programming language and regularly writes about open science and reproducible research issues.

Danae Roumis (Data Management and De-Identification)

Danae Roumis is a Program Director for Impact Evaluation at Social Impact, Inc. She serves as an evaluation specialist and technical advisor while also contributing to the management and growth of the Impact Evaluation division. She brings over ten years of experiencing designing and implementing research and evaluations, drawing on experience with a broad range of quantitative and qualitative evaluation methodologies. She is currently working on three urban water infrastructure impact evaluations for the Millennium Challenge Corporation (MCC) in Tanzania, Jordan, and Lesotho. She has also evaluated HIV/AIDS, TB, and WASH programming for USAID in Zimbabwe, Zambia, and Bangladesh. Prior to joining SI, Danae evaluated HIV/AIDS, maternal health, and malaria programs in Botswana, Kenya, Uganda, and Tanzania.

Arnaud Vaganay (Transparent Reporting and Disclosure)

Arnaud Vaganay is the founder and director of Meta-Lab, a consultancy that develops, implements and evaluates new tools to make research and teaching more cost-effective. He is also a visiting lecturer at the London School of Economics and Sciences Po. Arnaud is interested in defining what make 'good' research decisions. He also studies the economic, political, psychological, and philosophical factors driving these decisions. As a BITSS Catalyst, Arnaud has led transparency and reproducibility workshops in New Delhi, and will teach a short-course in open science at École Polytechnique Fédérale de Lausanne (EPFL) this summer.

Robbie Van Aert (Meta-Analysis)

Robbie van Aert is a PhD candidate in the Methodology and Statistics department at Tilburg University. Funded by the Netherlands Organization for Scientific Research (NWO), his research is about correcting for publication bias in meta-analyses. His other research interests include developing statistical methods for conducting meta-analyses (usually for the sake of convenience fully ignoring publication bias), detecting and correcting for questionable research practices such as p-hacking, and studying the reproducibility and replicability of science. Robbie also received a BITSS SSMART grant to study the extent of publication bias within psychology and medicine. He is also a BITSS Catalyst.

Coosje Veldkamp (Improved Specification: Pre-registration and Pre-Analysis Plans)

Coosje Veldkamp recently completed her PhD dissertation at Tilburg University on the human fallibility of scientists and is currently working as a postdoctoral researcher in the department of Medical Humanities at VU University Medical Center. At Tilburg, she was part of the Meta-Research Center and her doctoral research focused on the effectiveness of methods aimed at reducing human error and bias in science, on trust in and among scientists, and on the psychology of the scientist in general. She co-chaired the committee that organized the International Symposium on Human Factors in Science in 2014, featuring highly influential scientists in the field of meta-research.

Eric-Jan Wagenmakers (Power and the Strength of Evidence)

Professor Eric-Jan Wagenmakers is a mathematical psychologist and a dedicated Bayesian. He works for the Psychological Methods unit at the University of Amsterdam and he is PI on the European Research Council grant "Bayes or Bust: Sensible Hypothesis Tests for Social Scientist", a grant that recently spawned the JASP open-source program for statistical analyses (www.jasp-stats.org). In 2016, he received a Leamer-Rosenthal Prize for Open Social Science for his design and leadership of the graduate-level course "Good Research Practices" at the University of Amsterdam. Dr. Wagenmakers also co-authored an influential paper on reproducibility and several replication studies, edited a special issue of Perspectives on Psychological Science on reproducibility in psychology, and received a SSMART grant from BITSS to develop a suite of meta-analytic techniques for Bayesian evidence synthesis.

Kaitlyn Werner (OSF in Detail + Hands-On Registration)

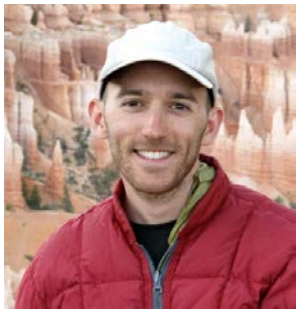
Kaitlyn is a PhD student in the Social, Personality, and Health Psychology program in the Department of Psychology at Carleton University in Ottawa, ON, Canada. Under the supervision of Dr. Marina Milyavskaya, her current research focuses on understanding the factors that facilitate successful goal pursuit. She also has a strong interest in advanced quantitative (e.g., multilevel SEM) and research methodology (e.g., experimental and longitudinal designs, transparency and better research practices). Outside of the lab, Kaitlyn is an Ambassador for the Center for Open Science, serves on Psi Chi's Research Advisory Committee, and is the student representative for CPA's Quantitative Methods Section.

Meet BITSS!

Edward Miguel (Faculty Director)

Edward "Ted" Miguel is the Oxfam Professor of Environmental and Resource Economics, as well as Co-Founder and Faculty Director of the Center for Effective Global Action (CEGA) and BITSS at UC Berkeley, where he has taught since 2000. He is also Co-Director of the Berkeley Opportunity Lab (O-Lab) and has served as Co-organizer of the Working Group in African Political Economy (WGAPE) since 2002. At BITSS, Ted leads research, supports partnership development, and is working on a forthcoming textbook on research transparency. His research focuses on African economic development and includes work on the economic causes and consequences of violence; the impacts of ethnic divisions on local collective action; interactions between health, education, environment, and productivity for the poor; and methods for transparency in social science research.

Garret Christensen (Project Scientist, Trainer – Dynamic Documents & Version Control)



Garret Christensen is a Project Scientist at BITSS and a Data Science Fellow with the Berkeley Institute for Data Science (BIDS) at UC Berkeley. Garret leads many of BITSS's trainings both in the U.S. and abroad, leads and conducts BITSS research, and is working on a forthcoming textbook on research transparency with Ted Miguel. He received his PhD in Economics from UC Berkeley in 2011 and has since conducted research for Innovations for Poverty Action (IPA) and Emory University in Kenya, and has taught economics at Swarthmore College. He is in interested in research

transparency and reproducibility and questions of causal inference in labor economics, particularly with regard to child health and education programs. Much of his current research focuses on the US Supplemental Nutrition Assistance Program (SNAP).

Fernando Hoces de la Guardia (Postdoc, Trainer – Dynamic Documents & Version Control)

Fernando Hoces de la Guardia is a Postdoctoral researcher with the Berkeley Initiative for Transparency in the Social Sciences (BITSS). Fernando works on bridging research-to-policy gaps in regards to transparency and reproducibility, and supports BITSS trainings. He received his PhD in Policy Analysis from the Pardee RAND Graduate School where his research focused on increasing the transparency and reproducibility of policy analysis as a way to strengthen the connection between policy and evidence. Before RAND, he studied economics and conducted impact evaluations and economic analyses of various social policies. Fernando has also supported BITSS-led trainings in the past and will lead a series of Catalyst trainings in South America in October.

Kelsey Mulcahy (Program Manager)

Kelsey Mulcahy is BITSS's Manager, leading the development of BITSS programs, partnerships, and events. Previously, she served as the Sex Trafficking Policy Fellow at the Los Angeles County Commission on Human Relations, working to reduce the commercial sex trafficking of minors in LA County. She has also conducted quantitative evaluations of livelihoods-focused community driven development projects in South Asia. She holds a Masters in Public Policy from UCLA and a BA in Global and International Studies from UC Santa Barbara.

Katie Hoeberling (Senior Program Associate)

Katie Hoeberling is BITSS's Senior Program Associate, managing the BITSS MOOC and Catalyst program, as well as supporting grant management, communications, and event coordination. She has served as a Borlaug Fellow in Global Food Security studying savings-led microfinance and farmer-centered innovation initiatives in Cambodia. She has also supported an environmental impact assessment of the California almond industry and the revision of the Urban Forest Project Protocol for the California carbon market. She also interned at the Food Chain Workers Alliance and the Los Angeles Food Policy Council. Katie holds an MSc in International Agricultural Development from UC Davis and a BSc in Environmental Science from UCLA.

Jen Sturdy (Program Advisor)

Jennifer Sturdy splits her time between BITSS and the Millennium Challenge Corporation (MCC). At MCC, she spearheads several transparency initiatives, including the establishment of the MCC Evaluation Catalog and the MCC Disclosure Review Board for releasing de-identified, public use micro-data. Sturdy also instituted several internal protocols for strengthening the design and implementation of the MCC independent evaluation portfolio. Before MCC, she spent six years as a consultant for the World Bank, working on several large-scale impact evaluations in the health sector. She completed her MA in International and Development Economics at the University of San Francisco.

Research Transparency and Reproducibility Training (RT2)
September 20-22 | International Workplace| London, England

September 20	CRISIS OF REPRODUCIBILITY + EMERGING METHODS
09:00 am	**Introduction** Kelsey Mulcahy (BITSS)
09:30	**Transparency and the Research Cycle** Sean Grant (RAND Corporation)
10:30	*Coffee Break*
10:45	**Scientific Misconduct and Researcher Degrees of Freedom** Marjan Bakker (Tilburg University)
12:00 pm	*Lunch*
13:00	**Improved Specification: Pre-registration and Pre-Analysis Plans** Coosje Veldkamp (Tilburg University)
14:30	*Coffee Break*
15:00	**Replication** Nicole Janz (University of Nottingham)
16:30-17:30	**OSF in Detail + Study Registration: Hands-On** Kaitlyn Werner (Carleton University)
September 21	**EMERGING METHODS: PART II**
09:00 am	**Organizing Workflow and File Management** Kelsey Mulcahy (BITSS) and Arnaud Vaganay (Meta-Lab)
10:15	*Coffee Break*
10:30	**Data Management and De-Identification** Danae Roumis (Social Impact)
12:00 pm	*Lunch*
13:00	**Version Control w/ GitHub App and Command Line Git and GitHub I** Garret Christensen (BITSS, BIDS), Thomas Leeper (LSE), and Fernando Hoces de la Guardia (BITSS)
14:30	*Coffee Break*
15:00	**Version Control w/ GitHub App and Command Line Git and GitHub II** Garret Christensen (BITSS, BIDS), Thomas Leeper (LSE) , and Fernando Hoces de la Guardia (BITSS)
16:30 – 17:30	**Dynamic Documents using Stata and R: Hands-On** Garret Christensen (BITSS), Thomas Leeper (LSE) , and Fernando Hoces de la Guardia (BITSS)

September 22	EMERGING METHODS: PART III
09:00	**Lightning Talks:** Other open science tools and initiatives
10:30	*Coffee Break*
10:45	**Transparent Reporting and Disclosure** Arnaud Vaganay (Meta-Lab) and Sean Grant (RAND Corporation)
12:00 pm	*Lunch (Working Lunch for RT2 Faculty)*
13:00	**Meta-analysis** Robbie Van Aert (Tilburg University)
15:00	*Coffee Break*
15:30	**Power and the Strength of Evidence** Eric-Jan Wagenmakers (University of Amsterdam)
16:30	**Wrap-Up and Presentations** Garret Christensen (BITSS, BIDS) and Arnaud Vaganay (Meta-Lab)
17:00 –18:00	**Reception** To be held at Citadines Barbican

Materials for RT2 can be found on the OSF Project Page at https://osf.io/cdfh7/

Appendix I: OSF Pre-Registration

Prepared by Erica Baranski (UC Riverside)

Study Information

1. Title
    1.1. Provide the working title of your study. It may be the same title that you submit for publication of your final manuscript, but it is not a requirement.

2. Authorship

3. Research Questions
    3.1. Please list each research question included in this study.

4. Hypotheses
    4.1. For each of the research questions listed in the previous section, provide one or multiple specific and testable hypotheses. Please state if the hypotheses are directional or non-directional. If directional, state the direction. A predicted effect is also appropriate here.

Sampling Plan

In this section we will ask you to describe how you plan to collect samples, as well as the number of samples you plan to collect and your rationale for this decision. Please keep in mind that the data described in this section should be the actual data used for analysis, so if you are using a subset of a larger dataset, please describe the subset that will actually be used in your study.

5. Existing data
    5.1. Preregistration is designed to make clear the distinction between confirmatory tests, specified prior to seeing the data, and exploratory analyses conducted after observing the data. Therefore, creating a research plan in which existing data will be used presents unique challenges. Please select the description that best describes your situation. Please do not hesitate to contact us if you have questions about how to answer this question (prereg@cos.io).
        5.1.1. Registration prior to creation of data: As of the date of submission of this research plan for preregistration, the data have not yet been collected, created, or realized.
        5.1.2. Registration prior to any human observation of the data: As of the date of submission, the data exist but have not yet been quantified, constructed, observed, or reported by anyone - including individuals that are not associated with the proposed study. Examples include museum specimens that have not been measured and data that have been collected by non-human collectors and are inaccessible.
        5.1.3. Registration prior to accessing the data: As of the date of submission, the data exist, but have not been accessed by you or your collaborators. Commonly, this includes data that has been collected by another researcher or institution.

5.1.4. Registration prior to analysis of the data: As of the date of submission, the data exist and you have accessed it, though no analysis has been conducted related to the research plan (including calculation of summary statistics). A common situation for this scenario when a large dataset exists that is used for many different studies over time, or when a data set is randomly split into a sample for exploratory analyses, and the other section of data is reserved for later confirmatory data analysis.

5.1.5. Registration following analysis of the data: As of the date of submission, you have accessed and analyzed some of the data relevant to the research plan. This includes preliminary analysis of variables, calculation of descriptive statistics, and observation of data distributions. Studies that fall into this category are ineligible for the Pre-Reg Challenge. Please contact us (prereg@cos.io) and we will be happy to help you.

6. Explanation of existing data

6.1. If you indicate that you will be using some data that already exist in this study, please describe the steps you have taken to assure that you are unaware of any patterns or summary statistics in the data. This may include an explanation of how access to the data has been limited, who has observed the data, or how you have avoided observing any analysis of the specific data you will use in your study. The purpose of this question is to assure that the line between confirmatory and exploratory analysis is clear.

7. Data collection procedures.

7.1. Please describe the process by which you will collect your data. If you are using human subjects, this should include the population from which you obtain subjects, recruitment efforts, payment for participation, how subjects will be selected for eligibility from the initial pool (e.g. inclusion and exclusion rules), and your study timeline. For studies that don't include human subjects, include information about how you will collect samples, duration of data gathering efforts, source or location of samples, or batch numbers you will use.

8. Sample size

8.1. Describe the sample size of your study. How many units will be analyzed in the study? This could be the number of people, birds, classrooms, plots, interactions, or countries included. If the units are not individuals, then describe the size requirements for each unit. If you are using a clustered or multilevel design, how many units are you collecting at each level of the analysis?

9. Sample size rationale

9.1. This could include a power analysis or an arbitrary constraint such as time, money, or personnel.

10. Stopping rule

10.1. If your data collection procedures do not give you full control over your exact sample size, specify how you will decide when to terminate your data collection.

Variables

In this section you can describe all variables (both manipulated and measured variables) that will later be used in your confirmatory analysis plan. In your analysis plan, you will have the opportunity to describe how each variable will be used. If you have variables that you are measuring for exploratory analyses, you are not required to list them, though you are permitted to do so.

11. Manipulated variables
    11.1. Describe all variables you plan to manipulate and the levels or treatment arms of each variable. For observational studies and meta-analyses, simply state that this is not applicable.

12. Measured variables
    12.1. Describe each variable that you will measure. This will include outcome measures, as well as any predictors or covariates that you will measure. You do not need to include any variables that you plan on collecting if they are not going to be included in the confirmatory analyses of this study.

13. Indices
    13.1. If any measurements are going to be combined into an index (or even a mean), what measures will you use and how will they be combined? Include either a formula or a precise description of your method. If you are using a more complicated statistical method to combine measures (e.g. a factor analysis), you can note that here but describe the exact method in the analysis plan section.

Design Plan

In this section, you will be asked to describe the overall design of your study. Remember that this research plan is designed to register a single study, so if you have multiple experimental designs, please complete a separate preregistration.

14. Study type
    14.1. Experiment - A researcher randomly assigns treatments to study subjects, this includes field or lab experiments. This is also known as an intervention experiment and includes randomized controlled trials.
    14.2. Observational Study - Data is collected from study subjects that are not randomly assigned to a treatment. This includes surveys, natural experiments, and regression discontinuity designs.
    14.3. Meta-Analysis - A systematic review of published studies.
    14.4. Other - please explain.

15. Blinding

15.1. Blinding describes who is aware of the experimental manipulations within a study. Mark all that apply.
  15.1.1. No blinding is involved in this study.
  15.1.2. For studies that involve human subjects, they will not know the treatment group to which they have been assigned.
  15.1.3. Personnel who interact directly with the study subjects (either human or non-human subjects) will not be aware of the assigned treatments.
  15.1.4. Personnel who analyze the data collected from the study are not aware of the treatment applied to any given group.

16. Study design
  16.1. Describe your study design. Examples include two-group, factorial, randomized block, and repeated measures. Is it a between (unpaired), within-subject (paired), or mixed design? Describe any counterbalancing required. Typical study designs for observation studies include cohort, cross sectional, and case-control studies.

17. Randomization
  17.1. If you are doing a randomized study, how will you randomize, and at what level?

Analysis Plan

You may describe one or more confirmatory analysis in this preregistration. Please remember that all analyses specified below must be reported in the final article, and any additional analyses must be noted as exploratory or hypothesis generating.

A confirmatory analysis plan must state up front which variables are predictors (independent) and which are the outcomes (dependent), otherwise it is an exploratory analysis. You are allowed to describe any exploratory work here, but a clear confirmatory analysis is required.

18. Statistical models
  18.1. What statistical model will you use to test each hypothesis? Please include the type of model (e.g. ANOVA, multiple regression, SEM, etc) and the specification of the model (this includes each variable that will be included as predictors, outcomes, or covariates). Please specify any interactions that will be tested and remember that any test not included here must be noted as an exploratory test in your final article.

19. Transformations
  19.1. If you plan on transforming, centering, recoding the data, or will require a coding scheme for categorical variables, please describe that process.

20. Follow-up analyses
  20.1. If not specified previously, will you be conducting any confirmatory analyses to follow up on effects in your statistical model, such as subgroup analyses, pairwise or complex contrasts, or follow-up tests

from interactions. Remember that any analyses not specified in this research plan must be noted as exploratory.

21. Inference criteria
    21.1. What criteria will you use to make inferences? Please describe the information you will use (e.g. p-values, Bayes factors, specific model fit indices), as well as cut-off criterion, where appropriate. Will you be using one or two tailed tests for each of your analyses? If you are comparing multiple conditions or testing multiple hypotheses, will you account for this?

22. Data exclusion
    22.1. How will you determine what data or samples, if any, to exclude from your analyses? How will outliers be handled?

23. Missing data
    23.1. How will you deal with incomplete or missing data?

24. Exploratory analysis (optional)
    24.1. If you plan to explore your data set to look for unexpected differences or relationships, you may describe those tests here. An exploratory test is any test where a prediction is not made up front, or there are multiple possible tests that you are going to use. A statistically significant finding in an exploratory test is a great way to form a new confirmatory hypothesis, which could be registered at a later time.

Script (Optional)

The purpose of a fully commented analysis script is to unambiguously provide the responses to all of the questions raised in the analysis section. This step is not common, but we encourage you to try to create an analysis script, refine it using a modeled dataset, and use it in place of your written analysis plan.

25. Analysis scripts (Optional)
    25.1. (Optional) Upload an analysis script with clear comments. This optional step is helpful in order to create a process that is completely transparent and increase the likelihood that your analysis can be replicated. We recommend that you run the code on a simulated dataset in order to check that it will run without errors.

Other

26. Other
    26.1. If there is any additional information that you feel needs to be included in your preregistration, please enter it here.