



Research Transparency and Reproducibility Training (RT2)

Final Report

September 11-13, 2019

Washington, DC

A. Overview

The Berkeley Initiative for Transparency in the Social Sciences ([BITSS](#)), an initiative of the Center for Effective Global Action ([CEGA](#)), held a Research Transparency and Reproducibility Training ([RT2](#)) in partnership with the World Bank’s Development Impact Evaluation ([DIME](#)) group at the World Bank Headquarters in Washington, DC, September 11-13, 2019. This was the eighth training event of this kind organized by BITSS since 2014. RT2 DC was supported by the National Institutes of Health, the Hewlett Foundation, an anonymous donor, and revenue from tuition fees paid by nine of the 39 participants¹.

The [agenda for RT2 DC](#) included presentations² clustered around the following topics:

1. Threats to research credibility and reproducibility and their relation to the scientific ethos (*Mertonian Norms, Scientific Misconduct and Researcher Degrees of Freedom, and Transparency and the Research Cycle*);
2. Improved research design specification (*Pre-registration and Pre-analysis Plans, Power analysis with DeclareDesign*, and parallel hands-on sessions on *Pre-analysis Plans for Experimental or Observational Research*);
3. Ethical and open research (*Ethical Research*, and parallel hands-on sessions on *Data Management and De-Identification for R or Stata users*);
4. Tools and methods for research reproducibility and collaboration (parallel hands-on sessions on *version control with GitHub and the GitHub App or GitKraken*, parallel hands-on sessions on *Dynamic Documents for R or Stata users*, and *Transparency and Reproducibility with Administrative Data*);
5. Evidence synthesis, reproducibility, and interpretation (*Meta-analyses and Systematic Reviews, What do True Findings Look Like?, Replication, and Replicability and Reproducibility*); and

¹ BITSS invited limited submissions from applicants seeking support from BITSS, including tuition fee waivers, travel, and lodging. Submissions from eligible self-supported applicants were also considered on rolling basis.

² Presentation slides and other supporting materials from all sessions are available [here](#). Videos are available [here](#).



6. Open science innovations and initiatives (*Open Policy Analysis, Gigantum, Replication Markets, and BITSS Scholarly Communication Initiatives*).

As in previous RT2 events, BITSS prepared a [Participant Manual](#), which includes instructions for preparing for hands-on sessions and required software downloads, a Glossary with definitions of terms and concepts, a pre-event Reading List, bios of RT2 faculty and organizers, and a list with basic information about the participants. To further improve access to training materials for participants of this and future RT2 events, BITSS also created a publicly accessible [Zotero library](#) with open access versions of all required and suggested readings.

Beyond formal instruction on transparency and reproducibility, RT2 DC also provided opportunities for networking and collaboration, and served as a platform for BITSS to expand membership of the [Catalyst Program](#).

B. Faculty

BITSS selected a diverse group of faculty members for this RT2 event, striving for balance in terms of discipline, position, and gender. These include [Luiza Andrade](#) (World Bank, Economics), [Lorena Barba](#) (George Washington University, Engineering), [Graeme Blair](#) (UCLA, Political Science), [Fiona Burlig](#) (University of Chicago, Economics), [Benjamin Daniels](#) (World Bank, Economics), [Charlie Ebersole](#) (University of Virginia, Psychology), [Sean Grant](#) (University of Indiana, Public Health), [Fernando Hoces de la Guardia](#) (BITSS, Public Policy), [Maggie Jones](#) (US Census, Economics), [Edward Miguel](#) (BITSS/UC Berkeley, Economics), Michael Orevba (World Bank, Economics), [Joseph Simmons](#) (University of Pennsylvania, Psychology), [Soazic Elise Wang Sonne](#) (World Bank, Economics), and [Alexa Tullett](#) (University of Alabama, Psychology).

C. Participant Profile

BITSS received a total of 265 applications for RT2 DC, including 246 submitted with requests for financial support, and 19 submissions from self-supported applicants. Five additional participants were selected by our host partner DIME through an internal review process.

Accounting for all submissions, applicants represented a broad scope of social science disciplines, including economics (48%), public health, biostatistics, or epidemiology (16%), political science (13%), psychology (7%), sociology (4%), public policy (4%), and other social sciences (7%). The applicant pool was dominated by PhD students (37%), followed by research practitioners (e.g., impact evaluation specialists, staff scientists, policy analysts, etc.) (19%), Faculty (17%), Masters students (16%), Postdocs (6%), and individuals working other research-related roles (e.g., research managers, librarians, etc.) (5%). Applicants came from 52 different countries, with 76% of all submissions coming from applicants based outside the US. Table 1 summarizes the distribution of applicants.



<i>Discipline/Position</i>	PhD Student	Research Practitioner	Faculty	Current Masters student	Postdoc	Other Research-Related	<i>Total (Discipline)</i>
Economics	54	29	22	12	7	6	130
Public Health, Biostatistics, or Epidemiology	13	6	9	7	6	2	43
Political Science	16	4	6	7	1	1	35
Psychology	8	1	2	7	1	1	20
Other social science discipline	3	4	4	4	0	2	17
Sociology	3	2	2	2	1	1	11
Public Policy	2	5	0	4	0	0	11
Other (non-social science) disciplines	1	1	1	0	0	0	3
<i>Total (Position)</i>	100	52	46	43	16	13	270

Table 1: Distribution of RT2 DC applicants by position and discipline

Of the 271 applicants, 40 were selected and invited to participate, 39 of whom (21 self-identified women and 18 self-identified men) attended. Among those who attended, 13 were PhD students, 12 were research practitioners, five worked in research-related roles, four were faculty, three were postdocs, and two were Masters students. In terms of academic disciplines, 19 participants came from economics, eight from public health, biostatistics, or epidemiology, four from public policy, three from psychology and political science, respectively, and two from sociology.

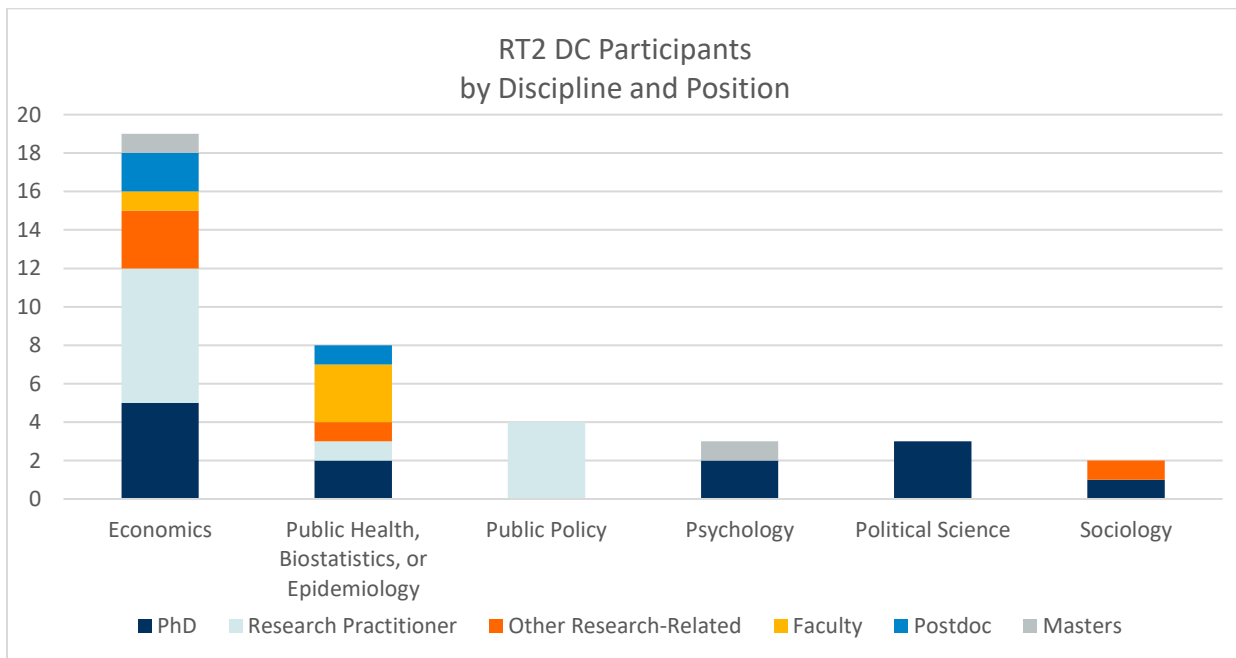


Figure 1: Distribution of RT2 DC Participants by Position and Discipline

D. Feedback and Lessons Learned



At the end of each day of the training, BITSS invited participants to submit feedback and assess each session in terms of *i)* the extent to which it improved their knowledge about each topic, *ii)* usefulness, *iii)* pace, and *iv)* the instructor’s preparedness on a scale from 1 to 5, with 5 being the highest rating. Daily evaluation forms had a 71% response rate on average. Participants provided additional comments in daily survey forms and during the final wrap-up session on Day Three. The following is a summary of this feedback.

Participant Manual

All respondents reported they had reviewed the Participant Manual before the training, and gave it an average rating of 4.17 in terms of improving their knowledge of research transparency and reproducibility. This is comparable to average ratings from RT2 Los Angeles (4.31), RT2 Amsterdam (4.24), RT2 London (4.13), and RT2 Berkeley (4.16).

Individual Sessions

All sessions were rated 4.24 on average. The three highest rated sessions were Version Control with GitKraken and the Github App (4.8), Mertonian Norms (4.7), and What do True Findings Look Like (4.61).

Day 1

RT2 opened with content on the motivation for the training and the open science movement in general, through sessions on [Mertonian norms](#) and [scientific misconduct and researcher degrees of freedom](#). Participants were then introduced to the conceptual framework of [the RT2 curriculum and the “RT2 Roadmap,”](#) a chart that provides an overview of how each session fits into the lifecycle of a typical research project. Faculty began each session by referencing the RT2 Roadmap to show where the topic of their presentation would be applicable. They also reflected on their perspective or “lens,” based on their discipline and research focus.

The second half of Day 1 was dedicated to sessions on improved design specification through pre-registration and pre-analysis plans (PAPs). This included an [introduction to pre-registration and PAPs](#), an [introduction to statistical power analysis using DeclareDesign](#), a software tool for describing and assessing research designs, and hands-on parallel sessions on [Observational PAPs](#) and [Experimental PAPs](#). This was the first RT2 event to feature sessions on power analysis and observational PAPs, both of which were introduced as a result of feedback from participants and faculty at previous RT2 events, and in order to further improve the relevance of RT2 for researchers across the social sciences.

Overall, sessions on Day 1 received the highest average rating (4.38) of all three days. Participants commended the logical and coherent flow of the agenda, and found the sessions a useful introduction to the main problems and solutions related to transparency and reproducibility. At the same time, a number of participants noted they were already largely familiar with the content of the motivational sessions and found them marginally useful in improving their knowledge about the topic. This can be interpreted as an



indicator of the growing awareness in the scientific community of problems related to research transparency and reproducibility. BITSS will consider condensing this content and/or replacing some of its parts with content on solutions for improved transparency, particularly in the form of practical exercises. To further improve the relevance of the RT2 curriculum on Day 1, participants also suggested we give a general orientation to issues surrounding interpreting statistical evidence (p-values and Bayes factors), a topic that was discussed on Day 3 of RT2 DC and previous RT2 events.

Day 2

The second day of the training started with an introduction to Ethical Research, including an overview of the ethical and legal standards for the protection of human subjects in the US, and a discussion of how data sharing and other open science practices fit in the review processes at Institutional Review Boards (IRB). Participants could then practice de-identifying mock datasets as part of parallel sessions in either R or Stata. Sixty-two percent of the survey respondents attended the R session, whereas 38% attended Stata. Those who attended the R session found it useful that the presenter had provided a script with all of the code in advance, as this minimized the need for troubleshooting during the session. Some participants would have appreciated receiving in advance detailed instructions on installing required software packages, as well as additional guidance in interpreting the code used in the exercise.

Participants who attended the Stata session found it clear and easy to follow. However, future sessions may benefit from improved understanding of and tailoring to participants' level of competency in the topic, as well as more context on the exercise. In order to improve the relevance and usefulness of this session, BITSS will explore the possibility of surveying the participants on their level of familiarity with R and Stata prior to the training, which could be used by instructors to accordingly adjust the level of instruction.

A large part of the remainder of Day 2 was dedicated to hands-on sessions on version control, where based on their prior experience, participants could attend one of two parallel sessions on version control with GitHub and either the GitHub App (for beginner users) or GitKraken (for advanced users). All participants were able to successfully install the required software in advance. Participants to both of these sessions were happy with the instructors' approach of hands-on instruction and found the tools relevant to their work.

For the final session of Day 2, participants could attend parallel sessions on Meta-Analyses and Systematic Reviews or Transparency and Reproducibility with Administrative Data. As part of the Meta-Analyses and Systematic Reviews session, participants learned about the advantages of evidence synthesis, various approaches to evidence synthesis, and were led through the process of conducting systematic reviews. During the session on Transparency and Reproducibility with Administrative Data, participants were introduced to the legal and administrative barriers to access administrative data in the context of the US, the implications of such barriers for transparency and reproducibility, and common solutions, such as the provision of syntax files, the use of synthetic data, and means to access parts of the US Census Bureau administrative data.



Day 3

Day 3 began with a presentation titled “What do True Findings Look Like?” that introduced threats to scientific validity (selective reporting, fraud, and errors) and their symptoms, as well as P-curve, a method for detecting publication bias. A presentation on Replicability and Reproducibility introduced participants to common definitions of the two concepts, and provided an overview of tools and best practices for reproducible social science. A second presentation on Replication guided participants through the process of producing conceptually and methodologically sound replications.

Participants could then attend one of two parallel hands-on sessions on Dynamic Documents using either R or Stata. Fifty-two percent of respondents attended the former, 48% attended the latter. Participants at both sessions found them useful and relevant to their work. To further improve the learning experience, particularly for beginners, participants suggested that instructors should prepare cheat sheets with basic code commands.

The final session of Day 3 was a series of four lightning talks that introduced participants to new tools and practices for research transparency and reproducibility, including [Open Policy Analysis](#), [Reproducible Workflows with Gigantum](#), [Replication Markets](#), and [BITSS Scholarly Communication Initiatives](#).

Participants rated Day 3 sessions 4.2 on average.

How would you rate RT2 in terms of:	Berkeley	London	Amsterdam	Los Angeles	Washington, DC
Overall faculty quality?	4.68	4.57	4.63	4.84	4.61
Improving your overall knowledge on the problems facing social science research transparency and reproducibility?	4.55	4.59	4.73	4.79	4.28
Improving your overall knowledge on research transparency and reproducibility best practices?	4.68	4.68	4.82	4.53	4.08

Overall Curriculum

On average, participants rated sessions at 4.09 both in terms of their usefulness and in terms of improving their knowledge of the topic. In terms of faculty preparedness and knowledge, participants rated sessions at 4.61 on average, which is on par with feedback from previous RT2 events, including Berkeley (4.68), London (4.57), Amsterdam (4.63), and Los Angeles (4.87).

Based on survey responses, this RT2 event however, was less effective in improving participants’ knowledge of the problems facing social science research transparency and reproducibility (4.28), as well as their confidence in their proficiency in transparency and reproducibility tools and practices they learned about RT2 (4.08). To further improve the effectiveness of RT2 in terms of helping participants master



reproducibility tools and practices, BITSS will consider allocating more time to hands-on, collaborative learning, as well as forms of continuous support, for example by leveraging the Catalyst Network. When asked for which topics discussed at RT2 participants would like more information or practice, the majority responded that they would find it useful to receive additional support with one of the hands-on sessions. For the sessions on version control, participants pointed out that they would have found it useful to receive more general guidance on reproducible workflows and collaboration, as many expected to find it challenging to onboard their collaborators.

Long-Term Impact and Future Considerations

BITSS drew a number of useful lessons for future training activities. Feedback from participants and faculty members has been extremely valuable in ensuring the relevance and usefulness of future RT2 events.

Many participants pointed out that, beyond its importance for skill development, **RT2 provides a unique opportunity for networking and collaboration** among proponents of research transparency. To further facilitate this, participants asked for more “structured” socialization during coffee and lunch breaks where participants could be paired or grouped with each other and faculty members to discuss how different RT2 topics fit into their ongoing projects. This could also be done through themed discussion tables during breaks, which could allow participants and faculty to exchange experiences around topics related to research transparency and reproducibility (e.g., practicing transparency as a journal referee, integrating transparency in course modules, etc.). BITSS will also consider allocating more time for questions at the end of presentations.

Hands-on sessions should accommodate different skill levels within each tool or practice. BITSS has partly addressed this need by providing parallel sessions with different, comparable tools. However, RT2 participants are increasingly intermediate or advanced users of certain tools who see RT2 as a place to troubleshoot and get practical advice from more advanced users. This may be addressed through additional pre-training introductory training materials that would ensure that all participants arrive to the training with at least basic familiarity. BITSS will also consider creating a “post-training participant manual” with practical advice and resources to help participants continue the learning process after the training.

RT2 (and the hands-on sessions in particular) has been effective in introducing participants to research transparency tools and practices. However, **researchers may need more sustainable support in more fully integrating them into their workflows.** Participants appreciate activities that incorporate learning-by-doing, and many pointed out that dedicating more time to hands-on sessions could help them further develop their proficiency in using and teaching open science tools and practices. For this purpose, BITSS will explore methods of instruction that help participants apply transparency tools and methods as they relate to their own work, producing tangible outputs using their own research projects (e.g., pre-register a study, de-identify a dataset, etc.). We will also explore the possibility of creating learning communities



that would connect RT2 participants with other members of the BITSS network, and allow them to get advice on practical implementation of transparency tools and practices.

In conclusion, BITSS was excited to learn that RT2 motivated participants to explore ways to further advance transparency and reproducibility in their own work. We also hope to empower them to change norms in their networks and at institutions through education and advocacy as [BITSS Catalysts](#).

We thank all of the RT2 faculty and participants for being part of this event and the broader open science movement!