



Berkeley Initiative for
Transparency in the Social Sciences

Research Transparency and Reproducibility Training (RT2)

2017 Report

June 7-9, 2017

University of California, Berkeley

OVERVIEW

The Berkeley Initiative for Transparency in the Social Sciences (BITSS), an initiative of the Center for Effective Global Action (CEGA), held its Research Transparency and Reproducibility Training (RT2) June 7-9, 2017. Formerly known as the Summer Institute, this RT2 was the fourth annual three-day training held at the University of California, Berkeley.¹ The event was sponsored by the National Institutes of Health (NIH), the John Templeton Foundation, the William and Flora Hewlett Foundation, and an anonymous donor.

In addition to learning about and discussing theoretical aspects of the reproducibility crisis, participants were introduced to a number of tools and methods to improve the transparency and reproducibility of their workflow, as well as the credibility of scientific research in general. Participants were introduced to the concepts of pre-analysis plans, replication, meta-analysis, and data de-identification. They were also given the opportunity for hands-on practice with tools such as version control using GitHub and Cmd Line Git, dynamic documents using Stata and R, and registration using the Open Science Framework (OSF). Finally, participants were introduced to a number of cutting-edge open science tools during a series of lightning talks, including pcpnl, Jupyter, TextThresher, and Docker.

RT2 also provided a number of opportunities for networking and served as platform for BITSS to identify and advise trainers for the [Catalyst program](#).

¹ All materials are available online at: <https://osf.io/tpd84/>



Participant Profile

BITSS received 107 applications for RT2 Berkeley this year, nearly double (1.78 times) the number of applications received for the 2016 Summer Institute. Applicants came from 26 disciplines and sub-disciplines, though Economics, Political Science, Psychology, and Health Sciences accounted for nearly three quarters (74.8%) of applicants' disciplines. Over a third (33.6%) of these applicants represented Economics. Nearly half (47.7%) of applicants were PhD students, while faculty and researchers each made up 18.7% of applicants. Other applicants included Master students (8.4%), Post-doctoral scholars (5.6%), others involved in research (3.9%), and librarians (1%).

Of the 107 applicants, 45 participants – 21 women and 24 men – were selected and invited to attend RT2. Of these 45 invited applicants, 40 accepted the invitation and were able to attend the training. BITSS also invited 7 African scholars from the Working Group for African Political Economy (WGAPe) meeting also organized by CEGA, as well as the BITSS Visiting Scholar. The total cohort of 48 participants included 29 PhD students, 2 Masters students, 9 research practitioners, 5 university faculty, 2 post-doctoral scholars, and 1 librarian. Attendees came from 19 countries, including 15 participants from 13 countries in the Global South.

Participants represented 12 distinct disciplines including economics, political science, psychology, public health, and others (see Figure 1). While only three participants at last year's Summer Institute were from a health-related discipline, this year, this increased to 7 participants from public health and medicine. This inclusion supports a larger interdisciplinary perspective and in-depth discussion of research transparency across many disciplines, as well as the inclusion of implementation or practice- based perspectives.

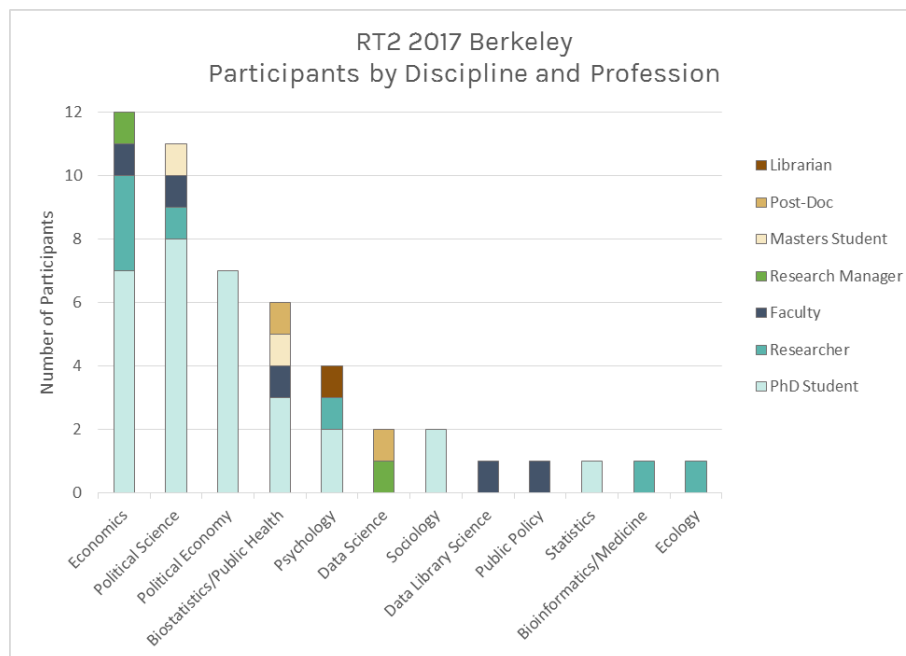


Figure 1: Distribution of BITSS 2017 RT2 Berkeley participants by discipline and profession. “Health Science” includes Biostatistics, Public Health, Bioinformatics, and Medicine. “Other Social Science” includes Public Policy and Statistics. “Data Science” includes both Data Science and Statistics.



Feedback and Lessons Learned

An evaluation survey was administered to participants after each day of RT2 to assess each session's usefulness, pace, and instructor preparedness. In addition to this ranking, a comments section enabled participants to provide additional feedback or suggestions regarding the sessions and the training as a whole. The following is a summary of the key findings from these survey responses.

The 3-day training consisted of 14 sessions. On a scale of 1 – 5 (5 being the most useful and 1 being least useful), participants rated overall curriculum at 4.64, which was an increase from last year's average rating of 4.43. They also rated overall faculty quality at 4.68, an increase from last year's average rating of 4.52.

The average session rating was 4.22, an increase from last year's average rating of 3.91. The five classes rated highest and found to be the most useful among participants were (1) Introduction and Mertonian Norms (4.65), (2) the Lightning Talks (4.58), (3) Improved Specification: Registration and PAPs (4.54), (4) Scientific Misconduct and Research Degrees of Freedom (4.47), and (5) Genoeconomics (4.41). Overall, participants felt the training improved their knowledge of both (i) problems facing the social sciences in terms of transparency and reproducibility and (ii) research transparency and reproducibility practices. With only one or two exceptions, a majority of the participants said the pace of each session was “just right.”

How would you rate RT2 in terms of:	Average Score (2017)	Average Score (2016)
Overall curriculum quality?	4.64	4.43
Overall faculty quality?	4.68	4.52
Improving your overall knowledge on the problems facing social science research transparency and reproducibility?	4.55	4.56
Improving your overall knowledge on research transparency and reproducibility best practices?	4.68	4.61

In response to suggestions from the 2016 Summer Institute participants, this year BITSS developed and distributed a Participant Manual before RT2. The manual included a brief overview of topics to be covered at the training, a suggested reading list, a list of actions to take before the training (software installation, OSF account registration, a final agenda, and lists of RT2 Faculty and Participants). All participants surveyed reviewed the manual before attending RT2 and, on average, gave the manual a rating of **4.16 out of 5 possible points**. Each participant also had an OSF account. Fifteen new accounts were registered using explicit instructions in the manual and the other surveyed participants reported that they already had an account. More information about software installation is below:



RT2 Materials and Software Installation	
Portion of participants who successfully installed the GitHub GUI or Bash before attending the GitHub session	93.8%
Portion of participants who successfully installed Version Control software before attending the GitHub session	80.8%

Based on feedback from the 2016 Summer Institute, BITSS focused on improving the gender balance of RT2 faculty this year. Of 15 faculty members, 6 (40%) were women (See figure 2). While this is an improvement on last year's gender ratio of 12%, BITSS will continue to strive for equity and diversity with both its faculty and participants.

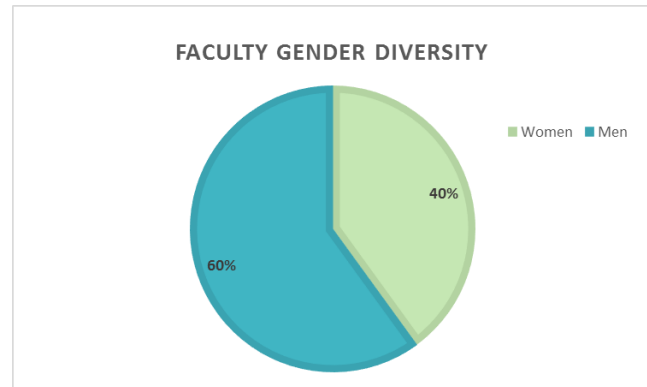


Figure 2: These numbers include individuals that may have presented more than once during the 3-day training.

Following positive feedback from 2016 Summer Institute participants regarding breakout sessions and hands-on practice, RT2 2017 included one breakout session in which participants could learn different software, depending on their experience and interest and one optional hands-on session. The breakout session focused on Version Control with either the Github app or Commandline Git was led by a BITSS Catalyst, serving as both a mechanism for instructing participants and developing the capacity of one of our supported trainers. The hands-on session, “OSF in Detail,” was optional and meant for participants who had little previous experience with the OSF. This session received an average rating of 4.13. Even with these breakout and optional hands-on sessions, when asked to “...describe any ways in which the workshop could be improved,” **the majority of participants requested deeper engagement with introduced tools and concepts that could be applied and integrated into a workflow.**

There were several comments related to the wide range of experience across participants. While many understood well that there is a reproducibility crisis in the social sciences and wanted more hands-on practice with tools, others felt a few sessions went into too much depth and detail for a newcomer. Some participants also suggested the use of case studies as a mechanism for better illustrating problems and solutions, especially when instructors’ and participants’ disciplines or levels of experience differ.

Overall, participants suggested more experiential learning strategies in addition to purely demonstrative teaching. In response to the 2016 Summer Institute participants’ suggestions,



BITSS worked with faculty to create a “roadmap” to better communicate the structure and flow of the workshop. While the roadmap was good first step, a number of participants mentioned the need for a clearer discussion during hands-on sessions of how the tool(s) might fit into a reproducible workflow. A few participants who serve as faculty at their own institutions also suggested that we provide training on how to teach open science practices and tools so that they could provide better instruction in classes they teach. An optional “Train-the trainer” session might be helpful, especially in the context of training would-be BITSS Catalysts.

Other common suggestions for improvement or use in future RT2 events include the following:

Suggestion	Number of Mentions
Address gap in participants’ experience with more breakout sessions, simplified material, more hands-on practice, or case studies	33
Offer more hands-on sessions or active learning, in general	23
Give conceptual overview for a tool before beginning hands-on portion	8
For denser sessions, give more time or hold earlier in the day when participants are more alert	7
Provide additional materials such as cheat sheets, glossary, list of guidelines, websites, and other useful resources	4

Long Term Impact and Broader Implications

Participants commented on the overall value of the workshop and how it informed their work. Their responses revealed (1) high overall satisfaction with the training, (2) desire for deeper and more active or applied learning strategies, and (3) the desire for BITSS and RT2 faculty to better communicate how tools fit cohesively into a broader reproducible workflow.

Similar to last year, many participants indicated they would like to further the transparency movement in their home institutions or engage in future collaborations with BITSS. In addition to expressing a desire to keep in touch with BITSS and other RT2 participants, a number of participants discussed wishes to join the Catalyst program to receive support for holding trainings at their own institutions.

In an effort to better understand the usefulness of and identify areas for improvement in our trainings, BITSS has begun implementing a formal survey strategy at BITSS and Catalyst trainings, including RT2. A pre-training survey was disseminated one month before the event and two more surveys will be sent to participants one month and six months after the event. We are excited to learn which practices have been fully adopted, as well as identify where BITSS can alter training strategies to facilitate higher uptake. We are especially excited to use what we’ve learned during this RT2 event and from the evaluations and surveys in designing our first [international RT2 event in London](#) this September!



Berkeley Initiative for
Transparency in the Social Sciences

BITSS is delighted to have had such an enthusiastic cohort of participants who were eager not only to improve the transparency and reproducibility of their own work, but also to pass on what they learned to their colleagues, peers, students, and wider community. The diversity of participants' disciplines, levels of experience, and roles in the research process helped drive insightful discussions. We are grateful that participants found the workshops useful and engaging! As we continue to keep in touch and evaluate the results from our forthcoming post-training surveys, we hope that participants apply the concepts and tools they learned and continue to develop their skills in performing transparent and reproducible research. We thank all those who attended for their participation and for being a part of the open science movement!



THANK YOU!



Appendix A: Agenda

JUNE 7 CRISIS OF REPRODUCIBILITY + EMERGING METHODS

- 09:00 am **Ice-Breaker and Introduction**
Edward Miguel (UC Berkeley, BITSS/CEGA Faculty Director)
-
- 10:00 **Transparency and the Research Cycle**
Sean Grant (RAND)
-
- 10:30 **Coffee Break**
-
- 10:45 **Scientific Misconduct and Researcher Degrees of Freedom**
Don Moore (UC Berkeley)
-
- 12:00 pm **Lunch**
-
- 01:00 **Replication**
Courtney Soderberg (Center for Open Science)
-
- 02:30 **Coffee Break**
-
- 03:00 **Meta-Analysis**
Eva Vivalt (Australian National University)
-
- 05:00 – 06:00 pm **OSF in Detail: Optional Session**
Erica Baranski (UC Riverside)
-

JUNE 8 EMERGING METHODS: PART II

- 09:00 am **Improved Specification: Pre-registration and Pre-Analysis Plans**
Edward Miguel (UC Berkeley, BITSS/CEGA Faculty Director)
-
- 10:30 **Coffee Break**
-
- 10:45 **Study Registration: Hands-On**
Erica Baranski (UC Riverside)
-
- 12:00 pm **Lunch**
-
- 1:00 **Transparency and Privacy: Data Management and De-Identification**
Danae Roumis (Social Impact)
-
- 02:30 **Coffee Break**
-



02:45 –

06:00 pm **The Magnes, 2121 Allston Way**
GitHub and Version Control
with GitHub App
Garret Christensen (BITSS, BIDS)

248 Giannini Hall, UC Berkeley

GitHub and Version Control
with Command Line Git

Harrison Dekker (University of Rhode
Island)

JUNE 9

EMERGING METHODS: PART III

09:00 am

Dynamic Documents using Stata and R: Hands-On
Garret Christensen (BITSS, BIDS)

10:30

Coffee Break

10:45

Transparent Reporting and Disclosure: Overview and Hands-On
Sean Grant (RAND) and Arnaud Vaganay (LSE)

12:00 pm

Lunch (*Working Lunch for Faculty*)

Lightning Talks: Other open science tools and initiatives
pcpanel

Fiona Burlig (UC Berkeley)

Text Thresher

01:15

Nick Adams (BIDS)

Jupyter

Alexandra Paxton (BIDS)

Docker

Chris Holdgraf (BIDS)

03:15

Coffee Break

03:30

Genoeconomics: A Primer Report and Progress
Dan Benjamin (USC)

04:30

Wrap-Up and Presentations
Garret Christensen (BITSS, BIDS) and Sean Grant (RAND)

05:00 –

06:00 pm

Reception
