



Transparency & Reproducibility Methods for Social Science Research

2016 Summer Institute Report

June 8-10, 2016

University of California, Berkeley

OVERVIEW

The Berkeley Initiative for Transparency in the Social Sciences (BITSS), an initiative within the Center for Effective Global Action (CEGA), held its third Summer Institute June 8-10, 2016 at the University of California, Berkeley.¹ The event was sponsored by the Alfred P. Sloan Foundation, the John Templeton Foundation, and an anonymous donor and coordinated with the Interuniversity Consortium for Political and Social Research (ICPSR).

In addition to learning about the many theoretical aspects of research transparency, participants were introduced to a number of tools and methods to increase the transparency and reproducibility of their workflow - including conducting a replication, version control using GitHub and Cmd Line Git, registration using OSF and how to de-identify data for public release.

Participant Profile

Of 60 applicants, 27 participants, 12 male and 15 female, were selected and invited to attend the Summer Institute. This cohort included 16 PhD students, 5 research practitioners and 6 professors/postdoctoral scholars. Attendees came from eight countries, including five developing countries. **Participants came from disciplines including economics, international development, psychology, public policy, and more. While over half of participants at last year's Institute came from economics, this year BITSS succeeded in diversifying the attendance list to**

¹ All materials are available online at: <https://osf.io/qh2nr/>.



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

include a number of new disciplines. This inclusion supports a larger interdisciplinary perspective and in-depth discussion of research transparency across many disciplines.

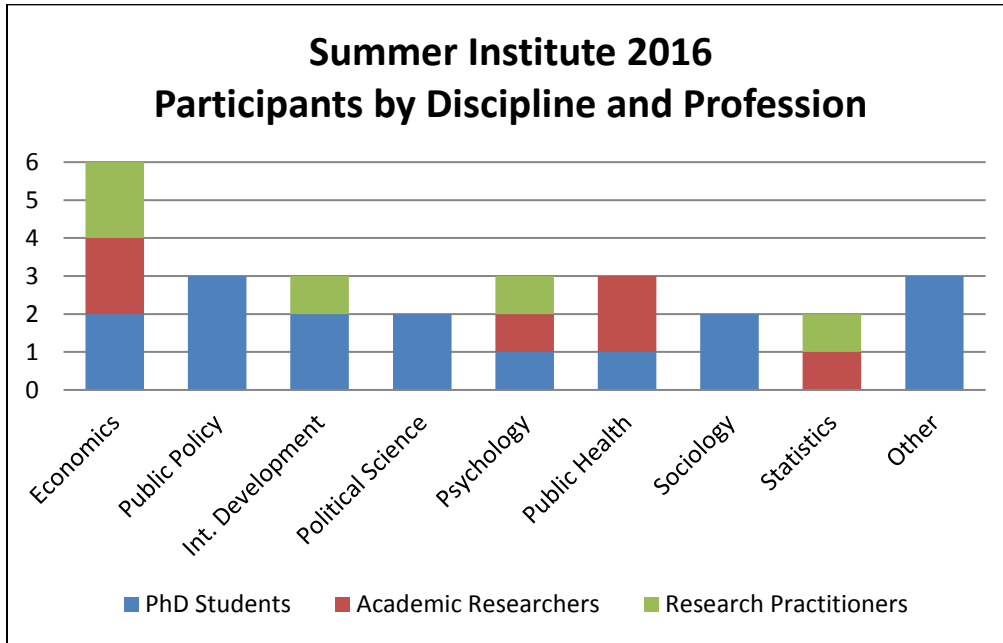


Figure 1: Distribution of BITSS 2016 Institute participants by discipline and profession.
The “Other” category includes Computer Science, Education, and Social Welfare.

Feedback and Lessons Learned

A survey was administered to participants after each of the three days of the Summer Institute in order to assess each session’s usefulness and pace, instructor preparedness, and how well the sessions complemented each other. A comments section also enabled participants to provide any additional feedback they had regarding the sessions and the Institute as a whole. The following is a summary of the key findings from these survey responses.

The Institute included 16 sessions. The average session rating, on a scale from 1-5 (5 being the most useful), was 3.91, a drop from last year’s Institute where the average session rating was a 4.22. However, the average instructor rating was very high at 4.43. The five classes rated highest and found to be the most useful among participants were (1) Data De-Identification, (2) P-Hacking, P-Curve, Specification Curve, (3) Replication, (4) Version Control w/ Cmd Line Git, and (5) Git and Github Part II.

How would you rate the Summer Institute in terms of:	Average Score
--	---------------



Overall curriculum quality?	4.43
Overall faculty quality?	4.52
Improving your overall knowledge on the problems facing social science research transparency and reproducibility?	4.56
Improving your overall knowledge on research transparency and reproducibility best practices?	4.61

The pace of the sessions was recorded at “Just Right” for the majority of the participants in most sessions, with the exception of Lighting Talks: Dynamic Documents, which the majority of participants said was too fast. Two other sessions, Registration and Pre-Analysis Plans: Guidelines and Practice, and Meta-Analysis: Methods and Application were split between being “Just Right” and “Too Fast.”

With the increasing expansion of BITSS’ network, **there is an opportunity to include more female researchers and practitioners involved as faculty.** This year’s faculty included only two women (see Figure 2), and BITSS is exploring ways to improve the diversity of its faculty.

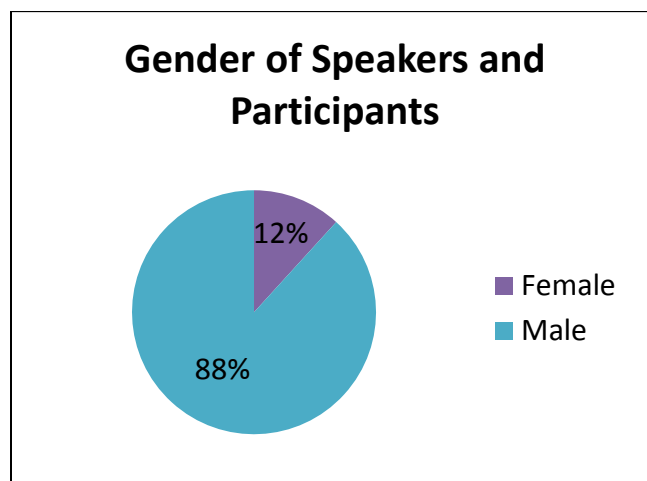


Figure 2: These numbers include individuals that presented multiple times as “1.”

After feedback from participants at last year’s Institute, the 2016 Summer Institute included four breakout sessions, where participants could learn different software, depending on their experience and interest. Two of these sessions, Version Control w/ Cmd Line Git and Git and Github Part II, were amongst the top 5 highest rated sessions, and a few participants listed these hands-on sessions as the highlight of the Institute. Even with these additional breakout sessions, when asked “please describe any ways in which the workshop could be improved?”, **the majority**



of participants requested additional hands-on workshops and practical tools that they could apply and integrate into their workflows.

In addition, to address the varying levels of experience, a number of participants suggested the **Institute offer participants the option to use their own research instead of the provided case studies for some of the sessions.** While not all participants felt comfortable enough with the material to use their own research, some of those who did would have liked an option to complete a pre-analysis plan or other “real-life” deliverable.

Participants also expressed some confusion about the purpose of specific sessions of the workshop. It was suggested both in the open feedback section of the survey and in the final general comments at the Institute that **BITSS include an initial roadmap in the Ted Miguel’s Introduction so participants have a sense of the Institutes’ structure and logical flow of the sessions.** This could include both a description and goals of the sessions, which would foster a better understanding of how all the individual sessions fit together into a cohesive workshop.

Additional suggestions for improvement or use in next year’s Summer Institute include the following:

	Suggestion	Number of Mentions
Logistics	-Send out the session materials beforehand, including instructions on how to install any necessary software	6
	-Send additional trainings and materials after the conclusion of the Institute so participants can learn more about concepts presented.	
	-Have an initial group activity so participants can get to know each other	4
Diversity	-Foster more collaboration between participants, including printed participant list with bios.	5
	-Outline the various transparency tools available and their pros and cons. Some participants felt overwhelmed by the number of tools and software available to them: a one-page fact sheet would be helpful.	
	-Increase the number of female speakers to foster an inclusive environment	1
	-Include panelists and speakers from a variety of disciplines, including medicine and public health.	3

Long Term Impact and Broader Implications



Participants commented on the overall value of the workshop and how it informed their future work. The responses revealed 1) the great opportunity and potential for future collaboration and partnership development between participants, 2) a high demand for more BITSS' trainings, and 3) the desire for BITSS to outline clear, actionable steps for participants to take to spread awareness of research transparency at their home institutions.

Participants indicated they would like to further the transparency movement in their home institutions or engage in future collaborations with BITSS. Six attendees would like BITSS to keep in touch with them via email, listserv, newsletter, or other forum. Furthermore, participants appreciated the open feedback discussion at the end of the Institute, and would like BITSS to further outline actionable steps they can take in their home institution to increase research transparency (3 participants).

Another interesting suggestion by one participant who elaborated on the need for BITSS to follow up with participants, was to survey attendees a few months following the conclusion of the Institute to see if they have implemented the learned methods into their own workflow, and if not, why not. This could pose very useful for BITSS to continually improve the curriculum to address barriers researchers face in practically implementing learned transparent methods.

Lastly, participants would like to keep in touch not only with BITSS, but with other participants (via online forum or other platform). Several attendees remarked on the Institute providing a great networking opportunity to meet others of similar interests in their disciplines during coffee and lunch. Participants would like to continue to develop these relationships on a communal platform. This was a suggestion that was first raised last year, when participants suggested **BITSS develop an online forum that would provide researchers with a venue to ask questions about transparency.**

We are excited to see another group of researchers and academics that are passionate about using the knowledge and skills that they developed at the Institute to increase research transparency. Overall, participants were incredibly grateful for the work that BITSS does, and the attention to detail and commitment of the BITSS team. 10 participants expressed their thanks for the workshop, and one added that it was an opportunity to learn something that he or she otherwise would never have been able to.



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES



THANK YOU!



Appendix

JUNE 8

CRISIS OF REPRODUCIBILITY + EMERGING METHODS

Brower Center – 2150 Allston Way

09:00 am	Introduction <u>Edward Miguel</u> (UC Berkeley)
09:20	Crisis in Social Sciences: Scientific Misconduct <u>Daniele Fanelli</u> (Stanford University)
10:30	Coffee Break
11:00	P-Hacking, P-Curve, Specification Curve <u>Leif Nelson</u> (UC Berkeley)
12:30 pm	Lunch
01:30	Registration and Pre-Analysis Plans: Guidelines and Practice <u>Maya Petersen</u> and <u>Fiona Burlig</u> (UC Berkeley)
03:00	Coffee Break
03:30	Transparent Reporting and Disclosure: Guidelines and Practices <u>Sean Grant</u> (RAND) and <u>Arnaud Vaganay</u> (LSE)
05:00 – 06:00	Pre-Analysis Plan Extensions: Hands-On, Observational <u>Garret Christensen</u> (BITSS/BIDS); <u>Sean Tanner</u> and <u>Fiona Burlig</u> (UC Berkeley)

JUNE 9

EMERGING METHODS: PART II

Brower Center – 2150 Allston Way

09:00 am	Meta-Analysis: Methods and Application <u>Tom Stanley</u> (Hendrix College)
10:30	Coffee Break
11:00	Data De-Identification <u>Simson Garfinkel</u> (NIST) and <u>Eric Vance</u> (Virginia Tech)
12:30 pm	Lunch (<i>Working Lunch for Faculty</i>)



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

01:30	Replication <u>Jesse Rothstein</u> (UC Berkeley), Justin McCrary, (UC Berkeley), Joshua Kalla (UC Berkeley)
03:00	Coffee Break
03:30	Interactive Session: Reproducible Workflow Case Studies <u>Garret Christensen</u> (BITSS/BIDS), <u>Cyrus Dioun</u> (UC Berkeley)
05:00 – 06:00pm	Reception

JUNE 10 SOFTWARE FOR REPRODUCIBLE WORKFLOW

	Brower Center – 2150 Allston Way	BIDS – 190 Doe Library
09:00 am	Version Control w/ Github App Dillon Niederhut (D-Lab)	Version Control w/ Cmd Line Git <u>Justin Kitzes</u> (BIDS)
10:30	Coffee Break	Coffee Break
11:00	OSF Andre Wang (UC Davis)	Git and GitHub Part II <u>Justin Kitzes</u> (BIDS)
12:30 pm	Lunch	Lunch

Brower Center – 2150 Allston Way

01:30 pm	Lighting Talks: Dynamic Documents (R Markdown/KnitR in R Studio, Markdoc in Stata), Jupyter Notebook, Docker for Cross-Platform Reproducibility) <u>Garret Christensen</u> (BITSS/BIDS), Nick Adams (BIDS) <u>Cyrus Dioun</u> (BIDS),
03:00	Coffee Break
03:15	New Initiatives in Transparency: TOP, Peer Reviewer Openness, Results-Blind Review, SSMART, and Leamer-Rosenthal Prizes <u>Garret Christensen</u> (BITSS/BIDS)
03:45	Wrap-Up <u>Edward Miguel</u> (UC Berkeley)



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Summer Institute 2016 Participants			
Name	Institution	Position	Discipline
Aisling Scott	UC Berkeley	PhD Student	Public Policy
Alasdair Cohen	UC Berkeley	PhD Student	Public Health
Amanda Charbonneau	UC Berkeley	PhD Student	Public Policy
Christian Kamala KAGHOMA	Université Catholique de Bukavu	Head of the Department of Economics/Lecturer	Economics
Deepak Singhania	UC Riverside	PhD Student	Economics
Eric Vance	Virginia Tech, Department of Statistics	Associate Research Professor/ Directs Laboratory for Interdisciplinary Statistical Analysis	Statistics
Felix Hartman	University of Gothenburg	PhD Student	Political Science
Ifeoluwapo Amao	National Horticultural Research Institute	Principal Research Officer	Economics
Jaclyn Leaver	UC Berkeley	Program Director	Economics
Jason Wong	Columbia University	PhD Student	International Development
Joan Mutyoba	Makerere University	Lecturer	Public Health
Justin Abraham	Princeton University/ Busara Center for Behavioral Economics	Research Assistant	Economics
Kathryn Asher	University of New Brunswick	PhD Student/Part-time Research Director	Sociology
Keesler Welch	J-PAL	Trial Registry Administrator	International Development
Kristin Porter	MDRC	MDRC	Statistics
Maura Liévano	UC Berkeley	PhD Student	Public Policy
Maxwell Hong	UC San Francisco	Staff Researcher	Psychology
Molly King	Stanford University	PhD Student	Sociology
Morufat Olakojo	University of Ibadan, Nigeria	PhD Student/Assistant Lecturer+C33	Education
Nochi Faha Dief Reagen	University of Yaounde II	PhD Student	Economics



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Summer Institute 2016 Participants (Continued)			
Name	Institution	Position	Discipline
Robert Wright	UC Riverside	PhD Student	Psychology
Rodrigo Lins	Federal University of Pernambuco, Recife	PhD Student	Political Science
Sofia Herrera	Fuller Graduate School of Psychology	Courtesy Faculty	Psychology
Tahora H. Nazer	Arizona State University	PhD Student	Computer Science
Timothy Foreman	Columbia University	PhD Student	International Development
Woojin Jung	UC Berkeley	PhD Student	Social Welfare
Xi Pan	Clemson University	PostDoc	Public Health