"Has the Credibility of the Social Sciences Been Credibly Destroyed?"

17 October 2024

ReproducibiliTea journal club

Newcastle University

Dr Chris Moreh

Lecturer (Assistant Professor) in Sociology

Newcastle University (UK)

@ chris.moreh@ncl.ac.uk

chrismoreh.com

/CGMoreh

@CGMoreh

Chris Moreh

Newcastle University

School of Geography, Politics and Sociology

1. My journey to open research, reproducible workflows and replication
2. Open research and reproducibility in quantitative social science pedagogy
3. Reproducibility in “sociology”: practices, challenges and limits

Auspurg, K, and Josef B. (2021) “Has the Credibility of the Social Sciences Been Credibly Destroyed? Reanalyzing the ‘Many Analysts, One Data Set’ Project.” *Socius* 7:23780231211024421. doi: 10.1177/23780231211024421

Journey into open research and reproducible workflows

- 2018 – 2021: teaching intro quantitative methods to sociology students at York St John University
- How to bridge the gap between “classroom statistics” and applied, “published research”?

- Analysis workflows in sociology are not transparent
- Lack of transparency limits reproducibility
- Analysis workflows and reproducibility are not taught
- Lack of pedagogical focus reinforces the cycle

BJS THE BRITISH JOURNAL OF SOCIOLOGY LSE

Original Article | Full Access

An investigation of social class inequalities in general cognitive ability in two British birth cohorts†

Roxanne Connelly ✉ Vernon Gayle

First published: 19 December 2017 | https://doi.org/10.1111/1468-4446.12343 | Citations: 15

Supporting Information

Filename	Description
bjos12343-sup-0001-suppinfo1.ipynb application/x-ipynb+json, 821.7 KB	**Appendix S1:** Jupyter Notebook file
bjos12343-sup-0002-suppinfo2.pdf PDF document, 9.3 MB	**Appendix S2:** Jupyter Notebook file in PDF
bjos12343-sup-0003-suppinfo3.docx Word document, 38.7 KB	**Online Supplement:** Additional Analyses

Please note: The publisher is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing content) should be directed to the corresponding author for the article.

Reproducible workflows and pedagogy

Teaching quantitative methods reproducibly

Chris Moreh

Lecturer in Sociology, School of Geography, Politics and Sociology.

Introduction

Sociology is lagging behind other social science disciplines – such as political science or economics – when it comes to taking replicability standards seriously (1,2). Among the more benign reasons behind this is that quantitative sociologists rarely use experimental methods (3) and mostly rely on large-scale survey data to target ill-defined estimands that generally result in low p-values (and low effect sizes). As many have argued, that's not ideal for the purposes of scientific advancement (4–6). One reason for this, which applies to other (social) sciences as well, is the lack of methodological socialisation in replicability practices. Potential solutions have long been identified and implemented in some trailblazing centres of excellence – like Harvard's Institute for Quantitative Social Science (7) – but have only recently started gaining wider adoption through teaching reproducibility practices as part of applied research methods training at both undergraduate (8,9) and postgraduate levels (10). In this case study I present one of my own contributions to placing replicability practices at the heart of quantitative methodology teaching.

The Context

When I joined Newcastle University in November 2021, I had recently developed an MA-level *Quantitative Research Methods* course at my previous institution but could only teach it once. The idea behind it was simple, albeit unorthodox: take students on a journey across one popular research theme – why not "social trust"? – deconstructing the data and methods underpinning selected articles and incrementally reproducing (parts of) the original analyses.

At Newcastle, I applied this idea in my undergraduate module on Researching Social Life, and later in my Faculty-level PGR course on Quantitative Analysis. I will focus on the former, which, being an introductory module, presented more elementary challenges.

https://www.ncl.ac.uk/mediav8/library/file-downloads/research-services/open-research/moreh.pdf

Reproducible workflows and pedagogy

RESEARCHING SOCIAL LIFE 1
SOC2069

SOC2069 Datasets

- World Values Survey, Wave 7 (WVS7)
  - WVS7 variable search
  - WVS7 codebook
- European Values Study, 2017 (EVS2017)
- European Social Survey, Round 10, 2020 (ESS10)
- International Social Survey Programme, 2020 (ISSP2020)
- British Social Attitudes Survey, 2020 (BSA2020)

World Values Survey, Wave 7 (WVS7)

This dataset comes from the World Values Survey Wave 7 (2017-2022). The raw data (version 5.0, in SPSS format) and documentation was downloaded to a local folder, `Data/raw/wvs7/`:

path	size
../Data/raw/wvs7/F00010733-WVS_Cross-National_Wave_7_spss_v5_0	0
../Data/raw/wvs7/F00010733-WVS_Cross-National_Wave_7_spss_v5_0.zip	24.42M
../Data/raw/wvs7/F00010738-WVS-7_Master_Questionnaire_2017-2020_English.pdf	864.96K
../Data/raw/wvs7/F00011055-WVS_7_Codebook_Variables_report.pdf	4.93M

The SOC2069 version of the dataset can be reproduced from the raw data file with the `R` code shown in the expandable field below.

▶ Show the code

If you use the WVS7 data in your work, cite it as follows:

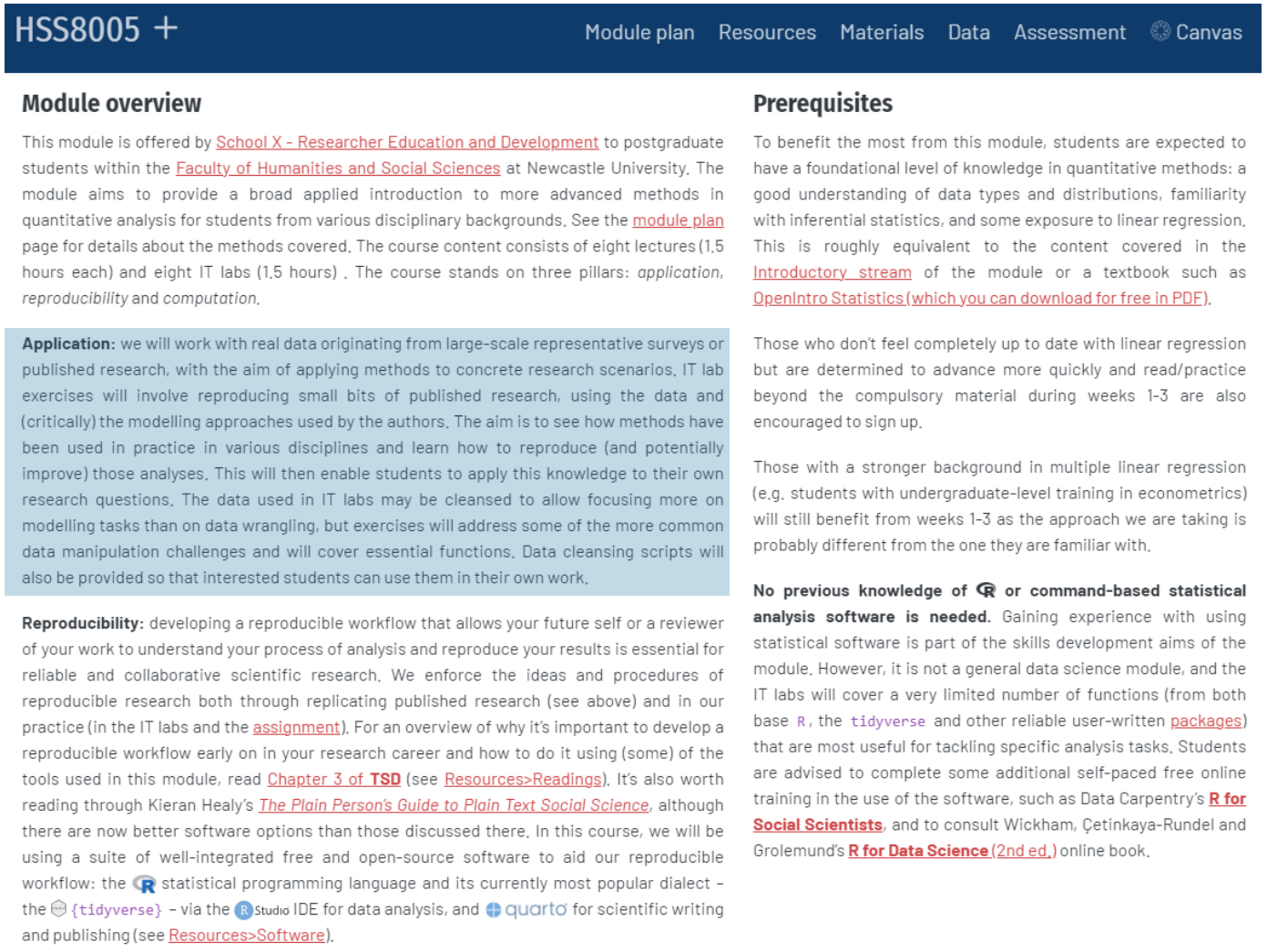
> Haerpfer, C., Inglehart, R., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano J., M. Lagos, P. Norris, E. Ponarin & B. Puranen (eds.). 2022. World Values Survey: Round Seven - Country-Pooled Datafile Version 5.0. Madrid, Spain & Vienna, Austria: JD Systems Institute & WVSA Secretariat. doi:10.14281/18241.20

WVS7 variable search

Show [ ] entries Search:

	Name	Label	Type	Levels	N (%) missing
1	B_COUNTRY	ISO 3166-1 numeric country code	categorical	64	0 (0.0%)
2	G_TOWNSIZE	Settlement size_8 groups	numeric	8	1279 (1.4%)

Reproducible workflows and pedagogy



HSS8005 +

Module plan Resources Materials Data Assessment Canvas

Module overview

This module is offered by School X - Researcher Education and Development to postgraduate students within the Faculty of Humanities and Social Sciences at Newcastle University. The module aims to provide a broad applied introduction to more advanced methods in quantitative analysis for students from various disciplinary backgrounds. See the module plan page for details about the methods covered. The course content consists of eight lectures (1.5 hours each) and eight IT labs (1.5 hours) . The course stands on three pillars: *application*, *reproducibility* and *computation*.

Application: we will work with real data originating from large-scale representative surveys or published research, with the aim of applying methods to concrete research scenarios. IT lab exercises will involve reproducing small bits of published research, using the data and (critically) the modelling approaches used by the authors. The aim is to see how methods have been used in practice in various disciplines and learn how to reproduce (and potentially improve) those analyses. This will then enable students to apply this knowledge to their own research questions. The data used in IT labs may be cleansed to allow focusing more on modelling tasks than on data wrangling, but exercises will address some of the more common data manipulation challenges and will cover essential functions. Data cleansing scripts will also be provided so that interested students can use them in their own work.

Reproducibility: developing a reproducible workflow that allows your future self or a reviewer of your work to understand your process of analysis and reproduce your results is essential for reliable and collaborative scientific research. We enforce the ideas and procedures of reproducible research both through replicating published research (see above) and in our practice (in the IT labs and the assignment). For an overview of why it's important to develop a reproducible workflow early on in your research career and how to do it using (some) of the tools used in this module, read Chapter 3 of **TSD** (see Resources>Readings). It's also worth reading through Kieran Healy's *The Plain Person's Guide to Plain Text Social Science*, although there are now better software options than those discussed there. In this course, we will be using a suite of well-integrated free and open-source software to aid our reproducible workflow: the R statistical programming language and its currently most popular dialect – the {tidyverse} – via the RStudio IDE for data analysis, and quarto for scientific writing and publishing (see Resources>Software).

Prerequisites

To benefit the most from this module, students are expected to have a foundational level of knowledge in quantitative methods: a good understanding of data types and distributions, familiarity with inferential statistics, and some exposure to linear regression. This is roughly equivalent to the content covered in the Introductory stream of the module or a textbook such as OpenIntro Statistics (which you can download for free in PDF).

Those who don't feel completely up to date with linear regression but are determined to advance more quickly and read/practice beyond the compulsory material during weeks 1-3 are also encouraged to sign up.

Those with a stronger background in multiple linear regression (e.g. students with undergraduate-level training in econometrics) will still benefit from weeks 1-3 as the approach we are taking is probably different from the one they are familiar with.

No previous knowledge of R or command-based statistical analysis software is needed. Gaining experience with using statistical software is part of the skills development aims of the module. However, it is not a general data science module, and the IT labs will cover a very limited number of functions (from both base `R`, the `tidyverse` and other reliable user-written packages) that are most useful for tackling specific analysis tasks. Students are advised to complete some additional self-paced free online training in the use of the software, such as Data Carpentry's **R for Social Scientists**, and to consult Wickham, Çetinkaya-Rundel and Grolemund's **R for Data Science** (2nd ed.) online book.

Reproducible workflows and pedagogy

The British Academy

Dr Chris Moreh

SRG24\242020

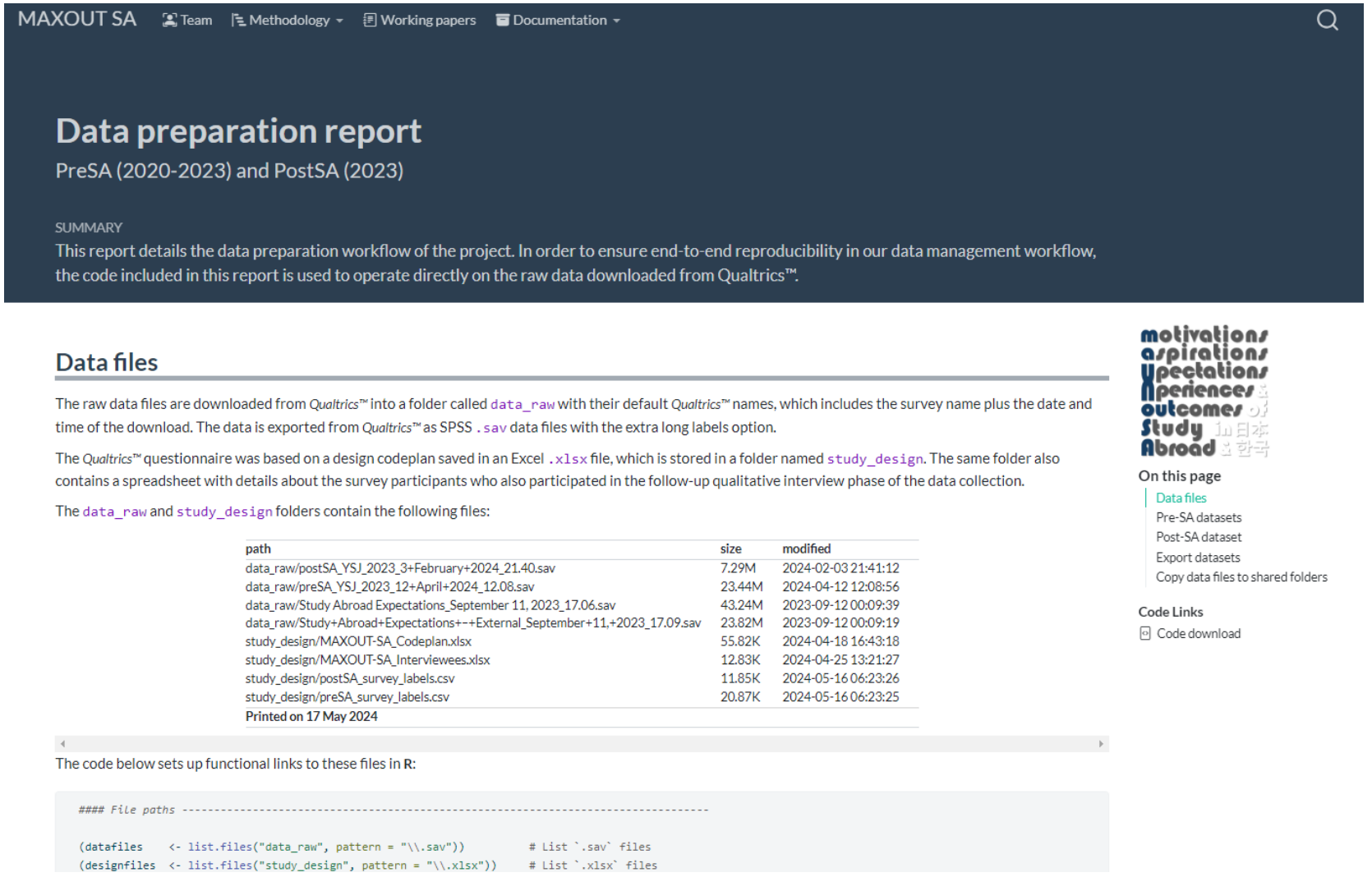
Trust Research Methods Database (TReMeDa)

Newcastle University

Value Awarded: £9,434.00

The proposed project aims to create a curated database of secondary quantitative replication data on the topic of 'social trust'. Its main deliverables will contribute both to improving practice in quantitative research pedagogy in sociology and related social science disciplines, and to advancing the research reproducibility agenda. The main outcome of the project with be the Trust Research Methodology Database (TReMeDa), which will serve as the basis for a planned sociological methodology textbook and various pedagogical resources. The texbook proposal will be submitted at the end of the first year of the project, based on a preliminary first-stage development of the TReMeDa. In the second year, alongside work completing the database and the underlying computer code - all of which is destined for the public domain -, the project also delivers a pedagogical research article documenting the implementation of the dataset in class over two academic years in two modules.

Reproducible workflows and research



MAXOUT SA Team Methodology Working papers Documentation

Data preparation report

PreSA (2020-2023) and PostSA (2023)

SUMMARY

This report details the data preparation workflow of the project. In order to ensure end-to-end reproducibility in our data management workflow, the code included in this report is used to operate directly on the raw data downloaded from Qualtrics™.

Data files

The raw data files are downloaded from *Qualtrics™* into a folder called `data_raw` with their default *Qualtrics™* names, which includes the survey name plus the date and time of the download. The data is exported from *Qualtrics™* as SPSS `.sav` data files with the extra long labels option.

The *Qualtrics™* questionnaire was based on a design codeplan saved in an Excel `.xlsx` file, which is stored in a folder named `study_design`. The same folder also contains a spreadsheet with details about the survey participants who also participated in the follow-up qualitative interview phase of the data collection.

The `data_raw` and `study_design` folders contain the following files:

path	size	modified
data_raw/postSA_YSJ_2023_3+February+2024_21.40.sav	7.29M	2024-02-03 21:41:12
data_raw/preSA_YSJ_2023_12+April+2024_12.08.sav	23.44M	2024-04-12 12:08:56
data_raw/Study Abroad Expectations_September 11, 2023_17.06.sav	43.24M	2023-09-12 00:09:39
data_raw/Study+Abroad+Expectations+-+External_September+11,+2023_17.09.sav	23.82M	2023-09-12 00:09:19
study_design/MAXOUT-SA_Codeplan.xlsx	55.82K	2024-04-18 16:43:18
study_design/MAXOUT-SA_Interviewees.xlsx	12.83K	2024-04-25 13:21:27
study_design/postSA_survey_labels.csv	11.85K	2024-05-16 06:23:26
study_design/preSA_survey_labels.csv	20.87K	2024-05-16 06:23:25
Printed on 17 May 2024		

The code below sets up functional links to these files in **R**:

```
#### File paths ----------------------------------------------------------------------------

(datafiles    <- list.files("data_raw", pattern = "\\.sav"))          # List `.sav` files
(designfiles  <- list.files("study_design", pattern = "\\.xlsx"))     # List `.xlsx` files
```

motivations aspirations expectations experiences outcomes Study Abroad

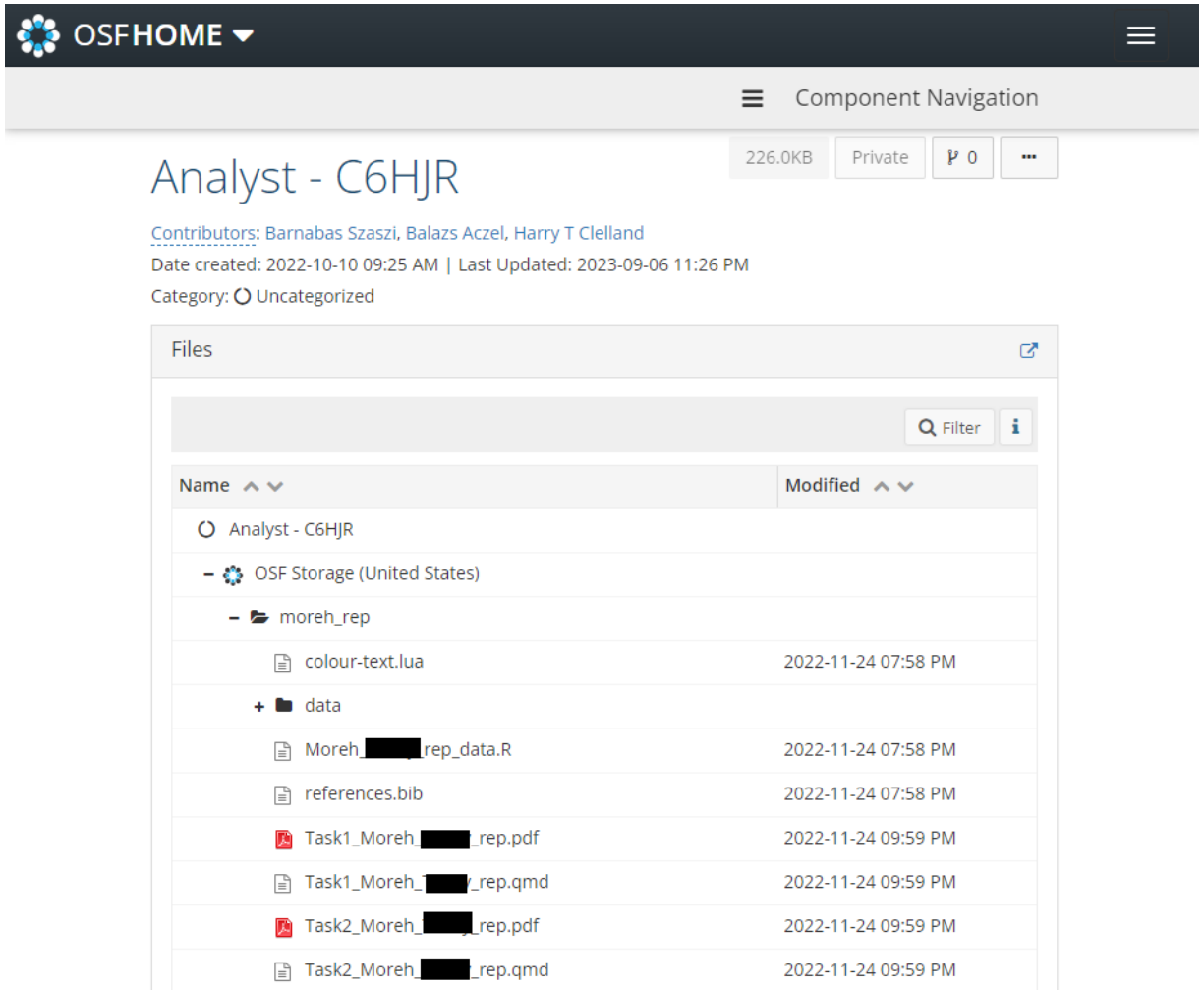
On this page

- Data files
- Pre-SA datasets
- Post-SA dataset
- Export datasets
- Copy data files to shared folders

Code Links

- Code download

Reproducible workflows and replication exercises



Reproducibility in sociology:

practices, challenges and limits

- Reproducibility exercises themselves can attract criticism and fail to replicate.
- The criticism of OSC (2015) by Gilber et al. (2016) and the ensuing discussion among a wider group of researchers with strong track records in promoting open science and championing reproducibility practices has demonstrated this in the context of psychological research and experimental data.
- Disciplines where the most commonly employed quantitative data come from large-scale observational studies – like sociology – have largely stayed on the margins of the reproducibility agenda.
- Silberzahn et al. (2018) was an important contribution in this respect, not only in terms of the approach and size of the conducted replication exercise, but also in its ability to cast doubts on the credibility of social research with observational data.
- Auspurg and Brüderl (2021), however, have challenged these findings, putting forward that at fault was the unclear research question given to the replicators. They argue that crowdsourcing (and then simply averaging) many studies of lower quality is not a useful approach to reproducibility exercises in the social sciences, and that a better strategy is to increase the quality of individual studies by calling for more precise definitions of the theoretical (research question) and empirical (parameter of interest) estimands driving the analysis, a plea made forcefully by Lundberg, Johnson, and Stewart (2021).

“Has the Credibility of the Social Sciences Been Credibly Destroyed?”

Original Article

Has the Credibility of the Social Sciences Been Credibly Destroyed? Reanalyzing the “Many Analysts, One Data Set” Project

Katrin Auspurg[1] and Josef Brüderl[1]

Socius: Sociological Research for a Dynamic World
Volume 7: 1–14

DOI: 10.1177/23780231211024421
srd.sagepub.com
SAGE

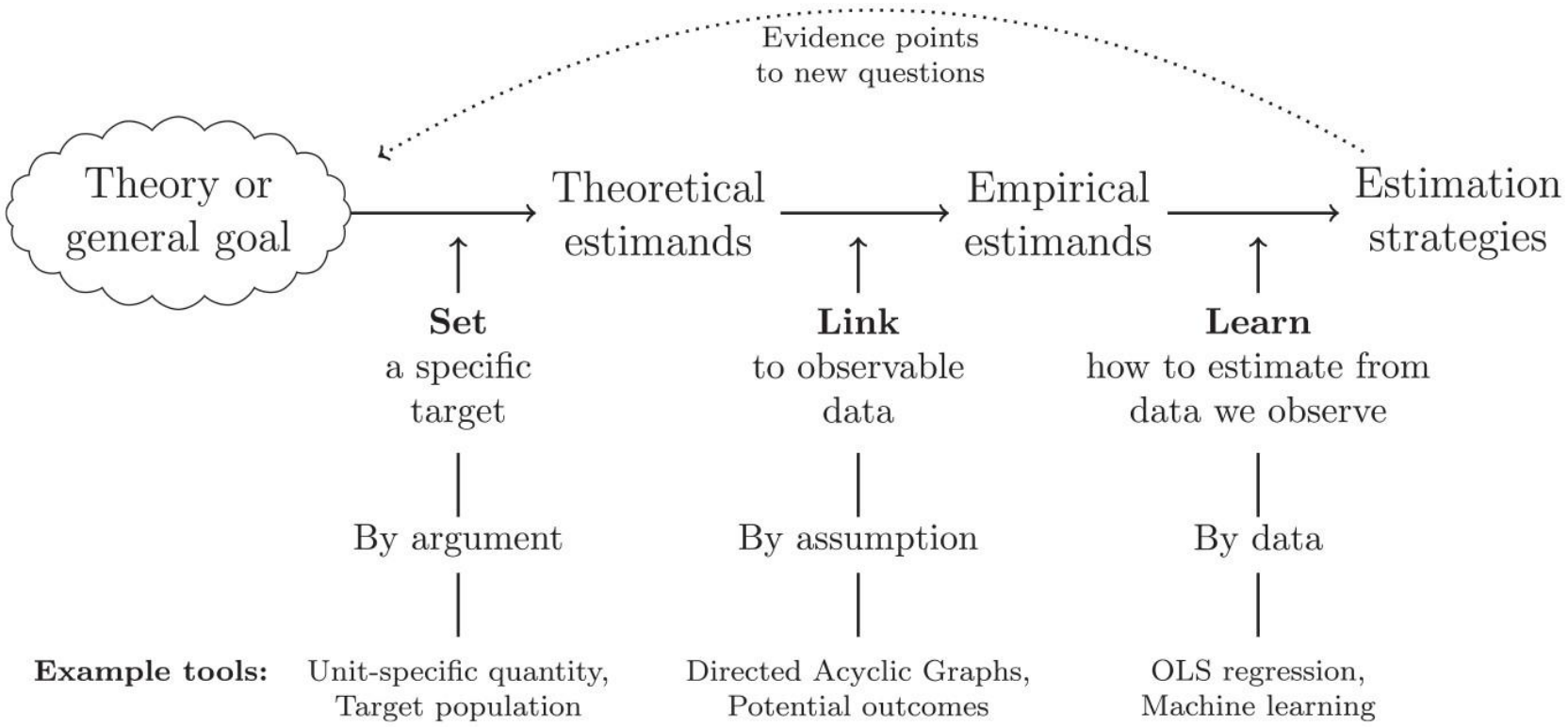
Abstract

In 2018, Silberzahn, Uhlmann, Nosek, and colleagues published an article in which 29 teams analyzed the same research question with the same data: Are soccer referees more likely to give red cards to players with dark skin tone than light skin tone? The results obtained by the teams differed extensively. Many concluded from this widely noted exercise that the social sciences are not rigorous enough to provide definitive answers. In this article, we investigate why results diverged so much. We argue that the main reason was an unclear research question: Teams differed in their interpretation of the research question and therefore used diverse research designs and model specifications. We show by reanalyzing the data that with a clear research question, a precise definition of the parameter of interest, and theory-guided causal reasoning, results vary only within a narrow range. The broad conclusion of our reanalysis is that social science research needs to be more precise in its “estimands” to become credible.

Keywords

credibility crisis, crowdsourcing, reproduction, replication, causal reasoning, robustness analysis, multiverse analysis, sensitivity analysis, estimands

“Has the Credibility of the Social Sciences Been Credibly Destroyed?”



Lundberg Ian, Johnson Rebecca, Stewart Brandon M. 2021“What Is Your Estimand? Defining the Target Quantity Connects Statistical Evidence to Theory.” *American Sociological Review*. OnlineFirst. https://doi.org/10.1177/00031224211004187.

@ chris.moreh@ncl.ac.uk
chrismoreh.com
/CGMoreh
@CGMoreh
Chris Moreh

Newcastle University
School of Geography, Politics and Sociology