Transparent and Reproducible Research with R

Daniel Anderson[1] & Joshua Rosenberg[2]

[1] University of Oregon
[2] University of Tennessee-Knoxville

Reproducibility of research findings is critical to the validity of inferences from studies. If an independent evaluator with access to the study data is unable to reproduce the published findings *exactly*, the trustworthiness of the findings are called into question, as highlighted by several prominent examples(e.g., the Duke crisis; see Peng, 2015). Particularly in the age of "big data", the majority of analyses include a multitude of decisions and, despite the best efforts of the individual researchers, all decisions are often not communicated through the published journal article. In this training, we provide an overview of reproducibility (and ideas related to open science) and introduce participants to tools that increase the likelihood of a reproducible and transparent analysis workflows. We emphasize tools from the R software environment to weave text with analysis code (e.g., R Markdown), version control with *git*/*GitHub* to documents the entire history of a project, and platforms for sharing analysis workflows publicly (e.g., basic static websites).

In the first hour of this four-hour training, we introduce participants (some familiarity with R assumed) to the idea and importance of open and reproducible research to further educational research. In the second and third hours, we discuss R Markdown and the various formats to which documents can be rendered. Finally, in the fourth hour, we provide a primer on version control using *GitHub* to document the history of a project and collaborate with others. Our target audience is early-career scholars, including doctoral students, or any researcher looking for tools to help increase the likelihood of their work being reproducible. The format will include part lecture and part hands-on applied work. We will assume all participants have a laptop.

Correspondence concerning this article should be addressed to Daniel Anderson, 5262 University of Oregon. E-mail: daniela@uoregon.edu

References

Peng, R. (2015). The reproducibility crisis in science: A statistical counterattack. *Significance*, *12*(3), 30–32.