Reproducible Research in R

Welcome! We will get started at 10:15am.

What should I be doing now?

- Review the **Reproducible Research** section on Canvas: https://z.umn.edu/latis-summerR
- Download the materials for today's workshop
- Make sure you have the **pacman** and **rmarkdown** packages installed
- Ask a question by unmuting or typing in the chat

COLLEGE OF LIBERAL ARTS | LATIS

Welcome!

A few notes before we begin:

- Materials are at z.umn.edu/latis-summerR
  - Workshops Navigation → Reproducible Research in R
  - Scroll down to "Live Workshop Materials"
- Main session will be recorded

Agenda

This will be a slightly different format than previous workshops:

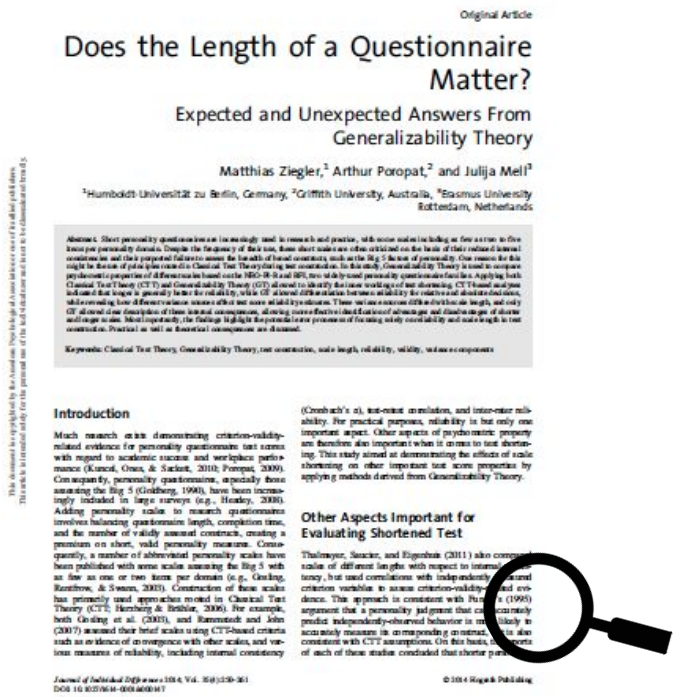
1. Overview of reproducibility
2. Running someone else's script
    - Activity with breakout groups
3. Tips and tools for reproducible workflows in R

What are we talking about?

		Data	
		Same	Different
Analysis	Same	✓ **Reproducible**	Replicable
	Different	Robust	Generalizable

The Turing Way Community, Becky Arnold, Louise Bowler, Sarah Gibson, Patricia Herterich, Rosie Higman, … Kirstie Whitaker. (2019, March 25). The Turing Way: A Handbook for Reproducible Data Science (Version v0.0.4). Zenodo. http://doi.org/10.5281/zenodo.3233986

Definitions



Original Article

Does the Length of a Questionnaire Matter?

Expected and Unexpected Answers From Generalizability Theory

Matthias Ziegler,[1] Arthur Poropat,[2] and Julija Mell[3]

[1]Humboldt-Universität zu Berlin, Germany, [2]Griffith University, Australia, [3]Erasmus University Rotterdam, Netherlands

Abstract. Short personality questionnaires are increasingly used in research and practice, with scale lengths as few as one to five items per personality domain. Despite the frequency of their use, these short scales are often criticized on the basis of their reduced internal consistencies and the purported failure to assess the breadth of broad constructs, such as the Big 5 factors of personality. One reason for this might be the use of principles rooted in Classical Test Theory during test construction. In this study, Generalizability Theory is used to compare psychometric properties of different scales based on the NEO-FF-R and BFI, two widely used personality questionnaires found in...

Keywords: Classical Test Theory, Generalizability Theory, test construction, scale length, reliability, validity, variance components

Introduction

Much research exists demonstrating criterion-validity-related evidence for personality questionnaire test scores with regard to academic success and workplace performance (Kuncel, Ones, & Sackett, 2010; Poropat, 2009). Consequently, personality questionnaires, especially those measuring the Big 5 (Goldberg, 1990), have been increasingly included in large surveys (e.g., Headey, 2008). Adding personality scales to research questionnaires involves balancing questionnaire length, completion time, and the number of validly measured constructs, creating a premium on short, valid personality measures. Consequently, a number of abbreviated personality scales have been published with some scales assessing the Big 5 with as few as one or two items per domain (e.g., Gosling, Rentfrow, & Swann, 2003). Construction of these scales has primarily used approaches rooted in Classical Test Theory (CTT; Hartwig & Bröhler, 2006). For example, both Gosling et al. (2003), and Rammstedt and John (2007) assessed their brief scales using CTT-based criteria such as evidence of convergence with other scales, and various measures of reliability, including internal consistency (Cronbach's α), test-retest correlation, and inter-rater reliability. For practical purposes, reliability is but only one important aspect. Other aspects of psychometric property are therefore also important when it comes to test shortening. This study aimed at demonstrating the effects of scale shortening on other important test score properties by applying methods derived from Generalizability Theory.

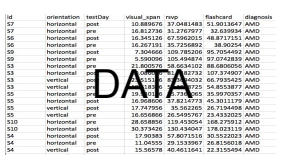
Other Aspects Important for Evaluating Shortened Test

Thalmayer, Saucier, and Eigenhuis (2011) also compared scales of different lengths with respect to internal consistency, but used correlations with independently measured criterion variables to assess criterion-validity- and evidence. This approach is consistent with Pu[illegible] (1995) argument that a personality judgment that can accurately predict independently-observed behavior is [illegible] likely to accurately measure its corresponding construct. It is also consistent with CTT assumptions. On this basis, [illegible] reports of each of these studies concluded that shorter per[illegible]

Journal of Individual Differences 2014; Vol. 35(1):230–261
DOI: 10.1027/1614-0001/a000147

Contribution to knowledge

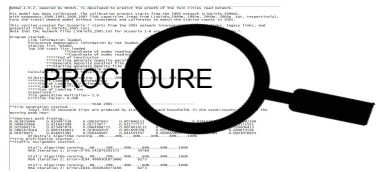
Computational Reproducibility: Given the author's data & statistical code, can someone produce the same results?



DATA

SCRIPT

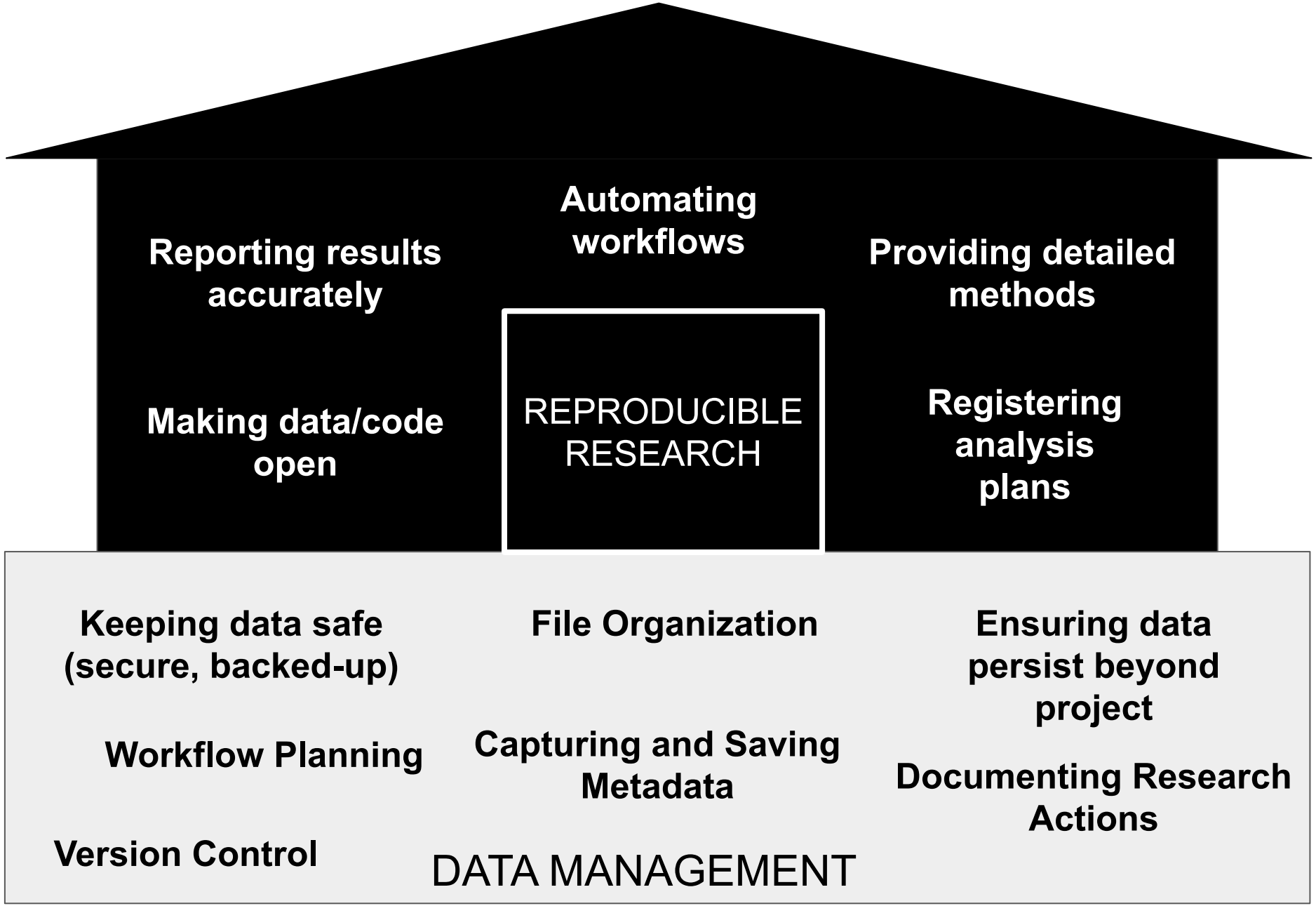
Empirical Reproducibility: Is there enough information for someone else to do the study exactly the same way?



PROCEDURE

Definitions: Center for Open Science

Good data management is the foundation for reproducibility



Components of Reproducibility in R

1. **Directory & Files**: Is it clear what script reads in and where it is?
2. **Script Organization**: Packages at top, comments?
3. **Script Descriptions**: Is it clear what the script does?
4. **Running the script**: Does it run?
5. **Software & Package Documentation**: What versions were used?
6. **Version Control**: Is this the actual final version?

Let's try it

Pretend you are helping a colleague check over a script they created to share with their data for a paper. The script does some basic data cleaning, creates a figure, and creates a table for the paper.

Using the **ReproducibilityFramework_R Workshop.pdf** document as a guide, go through the script (Pattern_script.R) and the data file (data_final_2.csv)

- Does it meet the suggestions in the framework?
- What improvements would you suggest?

We will put you into breakout rooms, and come back in ~15 minutes

The curated final versions

LIBRARIES digital conservancy

Search Browse Help Sign in

Data Repository for U of M

CORE TRUST SEAL

University Digital Conservancy Home / University of Minnesota / Data Repository for U of M (DRUM) / View Item

Risk prioritization of pork supply movements during an FMD outbreak in the US - Data and Materials

Patterson, Gilbert R; Hofelich Mohr, Alicia; Snider, Tim; Lindsay, Thomas A; Davies, Peter; G
Tim; Sampedro, Fernando (2016)

Persistent link to this item
http://doi.org/10.13020/D6NC7R
http://hdl.handle.net/11299/181833

Published Date
2016-08-22

Author Contact
Patterson, Gilbert R (patte606@umn.edu)

Type
Dataset
Survey Data-Quantitative

Abstract
In the event of a Foot and Mouth Disease (FMD) ou
state, and federal authorities will implement a foreig
emergency response plan restricting the pork supply
likely disrupting the continuity of the swine industry
disruptions of the food supply while providing an eff

Show full item record

View/Download file

File View/Open	Description	Size	Format
Pattersonetal2016_FMD_Survey.pdf	Survey Instrument	151.4Kb	PDF
Pattersonetal2016_FMD_Data.csv	Data File	31.72Kb	CSV file
Pattersonetal2016_FMD_Analysis_Code.R	R script for analysis	7.913Kb	Text file
Pattersonetal2016_FMD_DataDictionary.txt	Data Dictionary	17.04Kb	Text file
Pattersonetal2016_FMD_Movements.html	Interactive graph of main findings	1.593Mb	HTML

http://doi.org/10.13020/D6NC7R

Tools in R Studio

- **rm(list=ls())** - clear environment at start of script
- **sessionInfo()** - reports current environment and package versions
- **pacman** - package loading/installing
- **R Projects** - contained wording directory & file history
- **git/github** - version control
- **R Markdown** - reporting + results woven together

Thank you

Questions? Email or ask in the Reproducible Research page on Canvas!

hofelich@umn.edu

hammell@umn.edu

latisresearch@umn.edu