

# 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

# Welcome!



A few notes before we begin:

- Materials are at [z.umn.edu/latis-summerR](https://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	✓ <b>Reproducible</b>	Replicable
	Different	Robust	Generalizable

# Definitions

Original Article

## Does the Length of a Questionnaire Matter?

Expected and Unexpected Answers From Generalizability Theory

Matthias Ziegler,<sup>1</sup> Arthur Poropat,<sup>2</sup> and Julija Meli<sup>3</sup>

<sup>1</sup>Humboldt Universität zu Berlin, Germany, <sup>2</sup>Griffith University, Australia, <sup>3</sup>Erasmus University Rotterdam, Netherlands

**Abstract:** Short personality questionnaires are increasingly used in research and practice, with some studies including as few as two to five items per personality dimension. Despite the frequency of their use, these short scales are often criticized on the basis of their reduced internal consistency and the potential fallacies across the breadth of latent constructs, such as the Big 5 factors of personality. One reason for this might be the failure of prior studies to use Classical Test Theory (CTT) or, more recently, Generalizability Theory (G) to compare psychometric properties of different scales based on the 1950–2010 and 2010–2015 validity-based personality questionnaire factor lists. Applying both CTT and G to the 1950–2010 and 2010–2015 validity-based personality questionnaire factor lists, we examined the relative validity of short scales and full-length scales. Results showed that shorter scales generally had lower reliability, which in turn affected the validity of the constructs and the internal consistency. Additionally, we found that different validity measures affect the results differently. These findings suggest that the use of short scales might be problematic, and only CTT allowed clear descriptions of these internal consequences, allowing more detailed and detailed reports of the results. Most importantly, the findings highlight the potential for problems in the use of validity and scale length in test construction. Practical as well as theoretical consequences are discussed.

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

### Introduction

Much research exists demonstrating criterion-related evidence for personality questionnaire test scores with regard to academic success and workplace performance (Kerns, Omas, & Sackett, 2010; Poropat, 2009). Consequently, personality questionnaires, especially those assessing the Big 5 (Goldberg, 1990), have been increasingly included in large surveys (e.g., Healey, 2006). Adding personality scales to research questionnaires involves balancing questionnaire length, completion time, and the number of valid assessment 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., Gendling, Kershner, & Swenson, 2003). Construction of these scales has primarily used approaches rooted in Classical Test Theory (CTT; Hershberg & Bollen, 2006). For example, both Gendling et al. (2003), and Kershner 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  $\alpha$ ), test-retest reliability, and intra-rater reliability. For practical purposes, reliability is but only one important aspect. Other aspects of psychometric property are themselves 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

Thalheimer, Senneker, and Elgenbach (2011) also compared scale of different lengths with respect to internal consistency, but used correlations with independently observed criterion variables to assess criterion-related validity and evidence. This approach is consistent with the argument that a personality judgment that consistently predicts independently-observed behavior is more likely to accurately measure an underlying construct than one that is consistent with CTT assumptions. On this basis, the authors of each of these studies concluded that shorter personality

*Journal of Individual Differences* 2016, Vol. 36(1), 15–26  
DOI: 10.1027/1614-0001.a000167

© 2016 by Hogrefe Publishing

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



DATA

```
16 orientation testfor 10.88576 37.0461483 51.9813647 AND
17 horizontal post 16.31738 31.5769972 32.30928 AND
18 horizontal post 16.34518 31.5960015 48.873151 AND
19 horizontal pre 16.37131 30.7258892 36.90288 AND
20 horizontal post 7.30466 109.788206 95.705462 AND
21 horizontal pre 5.30006 109.494824 37.0461483 AND
22 horizontal post 7.30466 109.788206 95.705462 AND
23 horizontal pre 5.30006 109.494824 37.0461483 AND
24 vertical post 17.47938 35.35103 25.734848 AND
25 vertical post 17.47938 35.35103 25.734848 AND
26 vertical pre 16.85886 26.545797 23.433023 AND
27 horizontal 26.545797 23.433023 16.85886 AND
28 horizontal 26.545797 23.433023 16.85886 AND
29 horizontal 26.545797 23.433023 16.85886 AND
30 horizontal 26.545797 23.433023 16.85886 AND
31 horizontal 26.545797 23.433023 16.85886 AND
32 horizontal 26.545797 23.433023 16.85886 AND
33 horizontal 26.545797 23.433023 16.85886 AND
34 horizontal 26.545797 23.433023 16.85886 AND
35 horizontal 26.545797 23.433023 16.85886 AND
36 horizontal 26.545797 23.433023 16.85886 AND
37 horizontal 26.545797 23.433023 16.85886 AND
38 horizontal 26.545797 23.433023 16.85886 AND
39 horizontal 26.545797 23.433023 16.85886 AND
40 horizontal 26.545797 23.433023 16.85886 AND
41 horizontal 26.545797 23.433023 16.85886 AND
42 horizontal 26.545797 23.433023 16.85886 AND
43 horizontal 26.545797 23.433023 16.85886 AND
44 horizontal 26.545797 23.433023 16.85886 AND
45 horizontal 26.545797 23.433023 16.85886 AND
46 horizontal 26.545797 23.433023 16.85886 AND
47 horizontal 26.545797 23.433023 16.85886 AND
48 horizontal 26.545797 23.433023 16.85886 AND
49 horizontal 26.545797 23.433023 16.85886 AND
50 horizontal 26.545797 23.433023 16.85886 AND
51 horizontal 26.545797 23.433023 16.85886 AND
52 horizontal 26.545797 23.433023 16.85886 AND
53 horizontal 26.545797 23.433023 16.85886 AND
54 horizontal 26.545797 23.433023 16.85886 AND
55 horizontal 26.545797 23.433023 16.85886 AND
56 horizontal 26.545797 23.433023 16.85886 AND
57 horizontal 26.545797 23.433023 16.85886 AND
58 horizontal 26.545797 23.433023 16.85886 AND
59 horizontal 26.545797 23.433023 16.85886 AND
60 horizontal 26.545797 23.433023 16.85886 AND
61 horizontal 26.545797 23.433023 16.85886 AND
62 horizontal 26.545797 23.433023 16.85886 AND
63 horizontal 26.545797 23.433023 16.85886 AND
64 horizontal 26.545797 23.433023 16.85886 AND
65 horizontal 26.545797 23.433023 16.85886 AND
66 horizontal 26.545797 23.433023 16.85886 AND
67 horizontal 26.545797 23.433023 16.85886 AND
68 horizontal 26.545797 23.433023 16.85886 AND
69 horizontal 26.545797 23.433023 16.85886 AND
70 horizontal 26.545797 23.433023 16.85886 AND
71 horizontal 26.545797 23.433023 16.85886 AND
72 horizontal 26.545797 23.433023 16.85886 AND
73 horizontal 26.545797 23.433023 16.85886 AND
74 horizontal 26.545797 23.433023 16.85886 AND
75 horizontal 26.545797 23.433023 16.85886 AND
76 horizontal 26.545797 23.433023 16.85886 AND
77 horizontal 26.545797 23.433023 16.85886 AND
78 horizontal 26.545797 23.433023 16.85886 AND
79 horizontal 26.545797 23.433023 16.85886 AND
80 horizontal 26.545797 23.433023 16.85886 AND
81 horizontal 26.545797 23.433023 16.85886 AND
82 horizontal 26.545797 23.433023 16.85886 AND
83 horizontal 26.545797 23.433023 16.85886 AND
84 horizontal 26.545797 23.433023 16.85886 AND
85 horizontal 26.545797 23.433023 16.85886 AND
86 horizontal 26.545797 23.433023 16.85886 AND
87 horizontal 26.545797 23.433023 16.85886 AND
88 horizontal 26.545797 23.433023 16.85886 AND
89 horizontal 26.545797 23.433023 16.85886 AND
90 horizontal 26.545797 23.433023 16.85886 AND
91 horizontal 26.545797 23.433023 16.85886 AND
92 horizontal 26.545797 23.433023 16.85886 AND
93 horizontal 26.545797 23.433023 16.85886 AND
94 horizontal 26.545797 23.433023 16.85886 AND
95 horizontal 26.545797 23.433023 16.85886 AND
96 horizontal 26.545797 23.433023 16.85886 AND
97 horizontal 26.545797 23.433023 16.85886 AND
98 horizontal 26.545797 23.433023 16.85886 AND
99 horizontal 26.545797 23.433023 16.85886 AND
100 horizontal 26.545797 23.433023 16.85886 AND
```

SCRIPT

```
## Load the data
data <- read.csv("data.csv")

## Compute the reliability
reliability <- alpha(data[,1:100])

## Compute the validity
validity <- validity(data[,1:100])

## Print the results
print(reliability)
print(validity)
```

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



PROCEDURE

```
## Load the data
data <- read.csv("data.csv")

## Compute the reliability
reliability <- alpha(data[,1:100])

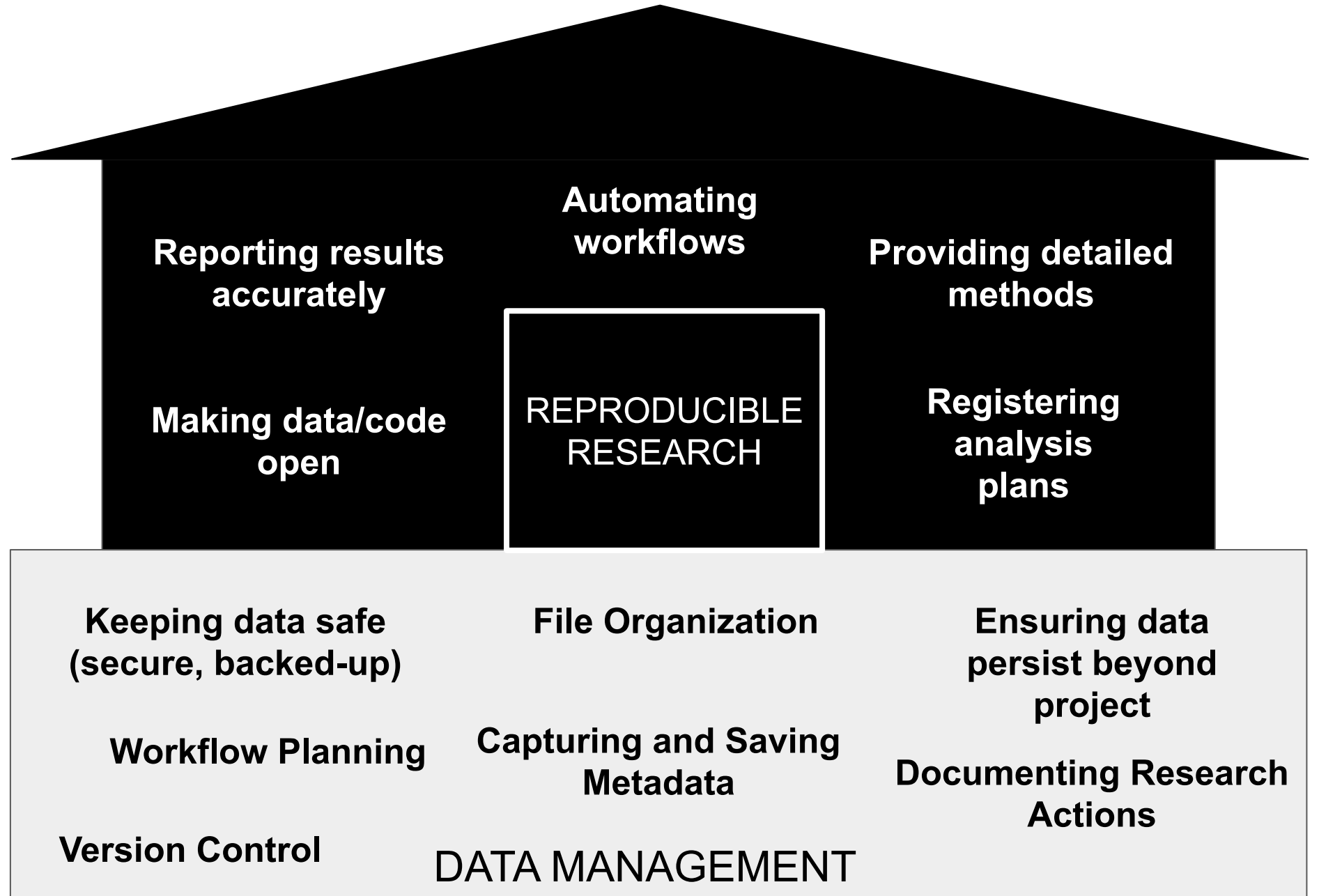
## Compute the validity
validity <- validity(data[,1:100])

## Print the results
print(reliability)
print(validity)
```



Contribution to knowledge

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



<http://doi.org/10.13020/D6NC7R>

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) outbreak, state, and federal authorities will implement a foreign emergency response plan restricting the pork supply, likely disrupting the continuity of the swine industry and disruptions of the food supply while providing an effective

Show full item record

### View/Download file

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

# 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 working 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](mailto:hofelich@umn.edu)

[hammell@umn.edu](mailto:hammell@umn.edu)

[latisresearch@umn.edu](mailto:latisresearch@umn.edu)