

The Reproducibility Crisis

The "Reproducibility Crisis"

RESEARCH ARTICLE

Estimating the reproducibility of psychological science

Open Science Collaboration^{*,†}

+ See all authors and affiliations

Science 28 Aug 2015:
Vol. 349, Issue 6251, aac4716
DOI: 10.1126/science.aac4716

Note: It's not just Psychology! Neuroscience, Political Science, Public Health, etc... Nearly every scientific discipline has issues!

The "Reproducibility Crisis"

How did the authors actually do their analyses?

- Did they transform their variables? If so, how?
- How did they handle missing data?
- Did they do what they said they were going to do in the Methods section of the paper?

Can we check that they did everything correctly?

- We need to be able to *reproduce* their findings
- In order to reproduce a figure, you need the data and the code!
 - Code needs to be in a format that **other humans** can read!

Make Your Life Prettier

- Format code and text so that it's easier to read
 - Comments in a script are great, but it can get cumbersome
- This is great for homework assignments, showing your research advisor or boss your findings, and publishing code along with your manuscripts!
- Some examples:

Neuroimaging of Individual Differences: A Latent Variable Modeling Perspective

Supplement 1

This supplement provides the basic code for all SEM analyses run in the corresponding manuscript. This is meant to provide a basic skeleton for future studies to build upon. The following assumes that users have access to the Human Connectome Project (HCP) dataset, and have already applied the Gordon et al. (2016) parcellation algorithm. Each row should be a subject, and each column is one of the parcels. Code for generating path diagrams is not provided here, but available upon request.

Please go [here](#) to learn more about how to get access to the HCP. The procedures for applying the parcellation algorithm can be found in the Gordon et al. (2016) manuscript. Exact scripts are available upon request.

Note: results of each model are provided in the manuscript, and are therefore not printed here to save space. To view the actual model output, find the `#summary(fit.model, standardize = TRUE)` line of code in each chunk, and simply remove the `#` to un-comment the line.

Example 1 - SEM vs. Averaging

```
rm(list=setdiff(ls(), c("nbk3")))  
  
model.la <- '  
# measurement model  
FPN =~ 1*L_FrontoParietal_ID108 +  
1*L_FrontoParietal_ID109 +  
1*L_FrontoParietal_ID148 +  
1*L_FrontoParietal_ID149 +  
1*L_FrontoParietal_ID24 +
```

Study 4: R Code & Output

```
library(dplyr)  
library(multicon)  
library(lm.beta)  
library(sjPlot)  
library(apaTables)  
library(ggplot2)  
library(gridExtra)  
  
vars <- read.csv("~/Desktop/study4.csv")
```

1) Creativity: Agreement between self and informant

```
cor.test(vars$creative_i, vars$creative_s)  
  
##  
## Pearson's product-moment correlation  
##  
## data: vars$creative_i and vars$creative_s  
## t = 5.1193, df = 377, p-value = 4.9e-07  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.1582745 0.3467732  
## sample estimates:  
## cor  
## 0.2549444
```

How do we do this?

To make our code pretty and readable, we're going to use something called **Markdown**. It's a way to take plain text and convert it into a pretty, formatted document like HTML, PDF, or even MS Word.

- Specifically, it's a way to write formatted text without writing full-blown HTML.
- It is not specific to **R**! Lots of things use Markdown! For example, the [README files on Github like this one](#) (scroll down to formatted text). Learning this will help even if you decide to leave academia.
- **RMarkdown** is simply Markdown but within **R**. This will, obvs, be our focus.

Why use RMarkdown?

...but my script has lots of comments and it's fine, right?

NO!

- .R scripts are monotonous
- Make a professional document or report
 - Headers
 - Bullet lists
 - **Clearly differentiate between text and code**
 - Link to websites, insert pictures etc.
- MUCH easier to read!
- MUCH easier to reproduce!

Before we begin

You will need the following packages. I'm 95% sure that R automatically installs the first two packages when you install R. The third package you should have already installed as part of one of the practice sets. However, please check your Packages tab now -- if you do not have any of these, please install them!

- `rmarkdown`
- `knitr`
- `kableExtra`

Before we begin

When you go to `File > New File >` you will see options for RMarkdown & RNotebook.

We are going to work with RNotebook.

- Frankly, they are nearly identical, including having the same file extension (`.Rmd`), so it doesn't matter much tbh
- It's slightly easier to preview an RNotebook, so we'll stick to that.
- *I will use RMarkdown and RNotebook interchangeably from here on out, because they are seriously that similar!*
- Those of you coming from Python 🐍, this is the equivalent of a Jupyter notebook

This section

There are 2 sections (in addition to this one) spanning reproducibility with RMarkdown.

1. The first will introduce you to the skeleton of a `.Rmd` document. This is a way to get familiar with the basics.
2. The second will show you all the parameters you can change to make your document even prettier. This is to give you a taste of all of the flexibility available at your fingertips.