

Reproducible Manuscripts in R Markdown

Juli Tkotz

27-05-2020

ZI Mannheim, RG Psychology and Neurobiology of Sleep and Memory

Nice to meet(up) you!



- Juli Tkotz
- PhD student ZI Mannheim
- RG Psychology and Neurobiology of Sleep and Memory

What is a reproducible manuscript?

- A manuscript that directly embeds your **research data** and **analysis code**.
- Any person with the raw data can run the code and **reproduce** your manuscript.
- Interactive stand-alone versions are possible.

Why do we need it?

Analysis of Open Data and Computational Reproducibility in Registered Reports in Psychology

Pepijn Obels¹, Daniel Lakens¹, Nicholas A. Coles², Jaroslav Gottfried³, & Seth Ariel Green⁴

¹ Eindhoven University of Technology, The Netherlands

² University of Tennessee, Knoxville, USA

³ Masaryk University, Brno, Czech Republic

⁴ Code Ocean, New York, USA

Ongoing technological developments have made it easier than ever before for scientists to share their data, materials, and analysis code. Sharing data and analysis code makes it easier for other researchers to re-use or check published research. These benefits will only emerge if researchers can reproduce the analysis reported in published articles, and if data is annotated well enough so that it is clear what all variables mean. Because most researchers have not been trained in computational reproducibility, it is important to evaluate current practices to identify practices that can be improved. We examined data and code sharing, as well as computational reproducibility of the main results, without contacting the original authors, for Registered Reports published in the psychological literature between 2014 and 2018. Of the 62 articles that met our inclusion criteria, data was available for 40 articles, and analysis scripts for 37 articles. For the 35 articles that shared both data and code and performed analyses in SPSS, R, Python, MATLAB, or JASP, we could run the scripts for 31 articles, and reproduce the main results for 20 articles. Although the articles that shared both data and code (35 out of 62, or 56%) and articles that could be computationally reproduced (20 out of 35, or 57%) was relatively high compared to other studies, there is clear room for improvement. We provide practical recommendations based on our observations, and link to examples of good research practices in the papers we reproduced.

263 3.1 Epistemic Trustworthiness

264 Participants placed more epistemic trust in the debaters when reading a neutral debate: Student
265 teachers in the neutral condition ($M = 5.06, SD = 1.00$) perceived the debaters to have more expertise
266 than those in the uncivil condition ($M = 5.06, SD = 1.00$), $t(218.49) = 1.99, p = .047, d = 0.27$.
267 Furthermore, participants reading a neutral debate ($M = 4.76, SD = 1.02$) reported higher ratings of
268 debaters' integrity than those reading an uncivil debate ($M = 4.05, SD = 1.15$), $t(219.41) = 4.87, p <$
269 $.001, d = 0.65$. Additionally, ratings of benevolence were higher in the neutral condition ($M = 4.77,$
270 $SD = 0.98$) than in the uncivil condition ($M = 4.05, SD = 0.89$), $t(214.11) = 5.67, p < .001, d = 0.76$
271 (see Figure 2).

272 We further explored the correlation between the conflict explanation items and the METI subscales,
273 that is, if the perception of various aspects of a conflict was associated with different degrees of
274 epistemic trust. Those who agreed that the debaters in the scenario referred to different research
275 results also thought them to have more expertise, $r(220) = .14, p = .039$. There was no relation with
276 integrity, $r(220) = .07, p = .321$, or benevolence, $r(220) = .03, p = .679$. Assuming personal reasons
277 for the conflict had the strongest relationship with epistemic trust. The more participants perceived
278 the conflict to be personal, the less expertise they assigned to the debaters $r(220) = -.25, p < .001$. In
279 a similar manner, perception of a personal conflict lead to decreased ratings of integrity, $r(220) =$
280 $-.36, p < .001$, and benevolence, $r(220) = -.41, p < .001$. How much participants agreed that the
281 debaters referred to different goals of PAVLOV did not correlate with any of the METI subscales,
282 neither with expertise, $r(220) = .10, p = .122$, nor with integrity, $r(220) = -.00, p = .946$, nor with
283 benevolence $r(220) = -.00, p = .994$. Embracement of the statement that debaters referred to different
284 effects of PAVLOV was not associated with epistemic trust either, neither with expertise, $r(220) =$
285 $.01, p = .863$, nor with integrity, $r(220) = -.06, p = .348$, nor with benevolence $r(220) = -.05, p =$
286 $.475$. Internal consistency of the METI subscales was somewhat lower than initially found by
287 Hendriks et al. (2015), with a Cronbach's α of .87 for expertise, .83 for integrity and .76 for
288 benevolence.

R Markdown to the rescue

```
## R Markdown to the rescue
```

```
```{r intext_stats, echo = TRUE}  
nerd <- read.csv("./data/nerd.csv", sep = "\t")
```
```

```
```{r copy_paste_hell}  
include_graphics("./pics/slide_inception.png")
```
```

This example dataset consists of $N = \text{r nrow(nerd)}$ participants with an age range between $\text{r min(nerd[["age"]])}$ and $\text{r max(nerd[["age"]])}$ years. Overall, $\text{r sum(nerd$age > 100)}$ participants reported to be older than 100, so we probably can't trust this data set a lot.

This example dataset consists of $N = 14955$ participants with an age range between 13 and 38822 years. Overall, 8 participants reported to be older than 100, so we probably can't trust this data set a lot.

Data retrieved from <https://openpsychometrics.org/>

R Markdown to the rescue

```
## R Markdown to the rescue
```

```
```{r intext_stats, echo = TRUE}  
nerd <- read.csv("../data/nerd.csv", sep = "\t")
```
```

```
```{r copy_paste_hell}  
include_graphics("../pics/slide_inception.png")
```
```

This example dataset consists of $N =$ ``r nrow(nerd)`` participants with an age range between ``r min(nerd[["age"]])`` and ``r max(nerd[["age"]])`` years. Overall, ``r sum(nerd$age > 100)`` participants reported to be older than 100, so we probably can't trust this data set a lot.

This example dataset consists of $N =$ **14955** participants with an age range between **13** and **38822** years. Overall, **8** participants reported to be older than 100, so we probably can't trust this data set a lot.

Detecting errors

263 3.1 Epistemic Trustworthiness

264 Participants placed more epistemic trust in the debaters when reading a neutral debate: Student
265 teachers in the neutral condition ($M = 5.06, SD = 1.00$) perceived the debaters to have more expertise
266 than those in the uncivil condition ($M = 5.06, SD = 1.00$), $t(218.49) = 1.99, p = .047, d = 0.27$.
267 Furthermore, participants reading a neutral debate ($M = 4.76, SD = 1.02$) reported higher ratings of
268 debaters' integrity than those reading an uncivil debate ($M = 4.05, SD = 1.15$), $t(219.41) = 4.87, p <$
269 $.001, d = 0.65$. Additionally, ratings of benevolence were higher in the neutral condition ($M = 4.77,$
270 $SD = 0.98$) than in the uncivil condition ($M = 4.05, SD = 0.89$), $t(214.11) = 5.67, p < .001, d = 0.76$
271 (see Figure 2).

272 We further explored the correlation between the conflict explanation items and the METI subscales,
273 that is, if the perception of various aspects of a conflict was associated with different degrees of
274 epistemic trust. Those who agreed that the debaters in the scenario referred to different research
275 results also thought them to have more expertise, $r(220) = .14, p = .039$. There was no relation with
276 integrity, $r(220) = .07, p = .321$, or benevolence, $r(220) = .03, p = .679$. Assuming personal reasons
277 for the conflict had the strongest relationship with epistemic trust. The more participants perceived
278 the conflict to be personal, the less expertise they assigned to the debaters $r(220) = -.25, p < .001$. In
279 a similar manner, perception of a personal conflict lead to decreased ratings of integrity, $r(220) =$
280 $-.36, p < .001$, and benevolence, $r(220) = -.41, p < .001$. How much participants agreed that the
281 debaters referred to different goals of PAVLOV did not correlate with any of the METI subscales,
282 neither with expertise, $r(220) = .10, p = .122$, nor with integrity, $r(220) = -.00, p = .946$, nor with
283 benevolence $r(220) = -.00, p = .994$. Embracement of the statement that debaters referred to different
284 effects of PAVLOV was not associated with epistemic trust either, neither with expertise, $r(220) =$
285 $.01, p = .863$, nor with integrity, $r(220) = -.06, p = .348$, nor with benevolence $r(220) = -.05, p =$
286 $.475$. Internal consistency of the METI subscales was somewhat lower than initially found by
287 Hendriks et al. (2015), with a Cronbach's α of .87 for expertise, .83 for integrity and .76 for
288 benevolence.

Wait a minute ...

Detecting errors

Epistemic Trustworthiness

Participants placed more epistemic trust in the debaters when reading a neutral debate: Student teachers in the neutral condition `r print_mean_sd(meti\$expertise_sum[meti\$condition == "neutral"])` perceived the debaters to have more expertise than those in the uncivil condition `r print_mean_sd(meti\$expertise_sum[meti\$condition == "neutral"])`, `r print_ttest(t_meti_exp, d_meti_exp)`. Furthermore, participants reading a neutral debate `r print_mean_sd(meti\$integrity_sum[meti\$condition == "neutral"])` reported higher ratings of debaters' integrity than those reading an uncivil debate `r

There it is!

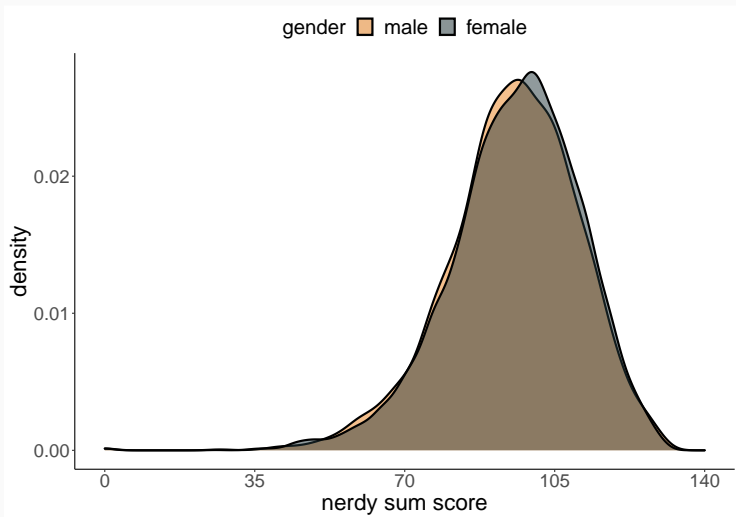
How about some stats?

```
nerd_ttest <- t.test(sum_score ~ gender, data = nerd)
nerd_effsize <- cohen.d(sum_score ~ gender, data = nerd)
```

```
In this dataset, men `r print_mean_sd(nerd[["sum_score"]][nerd[["gender"]] == "male")` have a
significantly lower nerd score than women `r print_mean_sd(nerd[["sum_score"]][nerd[["gender"]]
== "female")`, `r print_ttest(nerd_ttest, nerd_effsize)`.
```

In this dataset, men ($M = 95.18$, $SD = 15.27$) have a significantly lower nerd score than women ($M = 95.82$, $SD = 15.16$), $t(9800.06) = -2.39$, $p = .017$, $d = -0.04$.

Yeah, plots!



Yeah, references!

```
## Yeah, references!
```

```
``{r citation}  
include_graphics("./pics/citation.png")  
``
```

```
If I want to cite a paper, I can do this [san_martin_1968].  
This also works if I cite @san_martin_1968 as an in-text citation.
```

If I want to cite a paper, I can do this (San-Martin et al. 1968).
This also works if I cite San-Martin et al. (1968) as an in-text
citation.

How to get bibtex references

The screenshot shows a web page for a research article titled "ASPECTS OF REPRODUCTION IN THE ALPACA". The page has a green header with the word "Reproduction" in a large font. Below the title, it lists authors: M. SAN-MARTIN, M. COPAIRA, J. ZUNIGA, R. RODRIGUEZ. The article type is "Research Article" and it is available for "Free access". The DOI is <https://doi.org/10.1530/jrf.0.0160395>. The page number is 395-399, and it is from Volume 16, Issue 3. The article is published in the *J. Reprod. Fert.* (1968) 16, 395-399. A modal window titled "Preview Citation" is open, showing the citation in APA format and a "Download" button. Below the citation, there is an "Export Citation" section with three buttons: ".ris", ".bib", and ".enw". A red arrow points to the ".bib" button, which is labeled "BibTeX" and "Zotero".

Reproduction

Home Browse

ASPECTS OF REPRODUCTION IN THE ALPACA

Free access

in Reproduction

Authors: M. SAN-MARTIN, M. COPAIRA, J. ZUNIGA, R. RODRIGUEZ

DOI: <https://doi.org/10.1530/jrf.0.0160395> Article Type: Research Article

Page(s): 395-399 Online Publication Date: 2020

Volume/Issue: Volume 16, Issue 3

Abstract/Excerpt PDF

J. Reprod. Fert. (1968) 16, 395-399

ASPECTS OF REPR

Preview Citation

Format:

APA Download

SAN-MARTIN, M., COPAIRA, M., ZUNIGA, J., RODRIGUEZ, R., BUSTINZA, G., & ACOSTA, L. (1968). ASPECTS OF REPRODUCTION IN THE ALPACA, *Reproduction*, 16(3), 395-399. Retrieved May 20, 2020, from [https://rep.bioscientifica.com/view/journals/rep/16\(3\)/jrf_16_3_009.xml](https://rep.bioscientifica.com/view/journals/rep/16(3)/jrf_16_3_009.xml)

Export Citation

.ris .bib .enw

ProCite RefWorks Reference Manager BibTeX Zotero EndNote

Fully formatted articles

The R-package `papaja` offers you documents that are formatted according to APA (6) style.

<https://github.com/crsh/papaja>

1How to use papaja: An Example Manuscript Including Basic Instructions

2Frederik Aust¹

3¹ University of Cologne

HOW TO USE PAPAJA

6

Table 1

Descriptive statistics of correct recall by dosage.

| | Mean | Median | SD | Min | Max |
|--|------|--------|------|-------|-------|
| | 4.19 | 14.00 | 4.45 | 5.00 | 25.00 |
| | 3.50 | 14.00 | 5.15 | 4.00 | 22.00 |
| | 9.19 | 19.00 | 3.52 | 13.00 | 25.00 |

table was created with `apa_table()`

References

145

146Allaire, J., Cheng, J., Xie, Y., McPherson, J., Chang, W., Allen, J., ... Hyndman, R. (2016).

147*Rmarkdown: Dynamic documents for r*. Retrieved from

148<https://CRAN.R-project.org/package=rmarkdown>

149

150Aust, F., & Barth, M. (2015). *Papaja: Create apa manuscripts with rmarkdown*.

151

152Bates, D., & Maechler, M. (2016). *Matrix: Sparse and dense matrix classes and methods*.

153Retrieved from <https://CRAN.R-project.org/package=Matrix>

Or a whole book?

R for Data Science

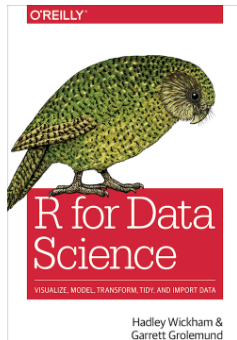
build passing

This repository contains the source of [R for Data Science](#) book. The book is built using [bookdown](#).

The R packages used in this book can be installed via

```
devtools::install_github("hadley/r4ds")
```

This is the website for “**R for Data Science**”. This book will teach you how to do data science with R: You’ll learn how to get your data into R, get it into the most useful structure, transform it, visualise it and model it. In this book, you will find a practicum of skills for data science. Just as a chemist learns how to clean test tubes and stock a lab, you’ll learn how to clean data and draw plots—and many other things besides. These are the skills that allow data science to happen, and here you will find the best practices for doing each of these things with R. You’ll learn how to use the grammar of graphics, literate programming, and reproducible research to save time. You’ll also learn how to manage cognitive resources to facilitate discoveries when wrangling, visualising, and exploring data.



Talking about reproducibility ...

```
> anticlusters <- anticlust::anticlustering(  
+   iris[, -5],  
+   K = 3,  
+   objective = "variance",  
+   method = "exchange"  
+ )  
Error in loadNamespace(name) : there is no package called 'anticlust'
```


Code capsules

The screenshot displays the CodeOcean web interface for a project named "MOSM_talk". The interface is divided into several sections:

- Top Bar:** Includes the CodeOcean logo, a "Private" status indicator, and navigation links for "Capsule", "File", "Edit", "View", "Tabs", "Settings", and "Help". On the right, there are buttons for "Launch Cloud Workstation" and "Collaborate", along with a user profile icon.
- Left Panel:** Contains a "Files" sidebar with a tree view of the project structure. The "Core Files" section lists:
 - metadata (66 B)
 - environment (1.24 KB)
 - code (4.77 KB)
 - 00reproducible_manuscripts... (4.64 KB)
 - run.sh (127 B)
 - data Manage Datasets (5.46 MB)
 - pics (1.15 MB)
 - codebook.txt (4.64 KB)
 - nerd.csv (4.31 MB)
 - rep_manuscripts.bib (757 B)
 - gignore (40 B)The "Results" section shows a link to "View latest results". The "Other Files" section lists a "files" folder (1.64 KB).
- Main View:** Displays the rendered R Markdown document titled "Reproducible Manuscripts in R Markdown". The content includes the author's name "Juli Tkotz", affiliation "Z1 Mannheim, RG Psychology and Neurobiology of Sleep and Memory", and the date "27-05-2020".
- Right Panel:** Features a "Reproducible Run" section with options to "or launch a cloud workstation" using various providers (lab, status, docker, etc.). Below this is a "Timeline" showing the run history:
 - Run 406321: Includes files like 00reproducible_man... (1.04 MB), buildLog (176.1 KB), and output (7.3 KB).
 - Commit: "Juli Tkotz committed May 25, 2020" with a note "reduced to beamer".
 - Run: "Juli Tkotz ran May 25, 2020" with a duration of 00:08:50.

CodeOcean

Code capsules

The screenshot displays the RStudio IDE with a project named "reproducible_manuscripts.Rmd". The editor shows the following R Markdown code:

```
1: ---
2: title: "Reproducible Manuscripts in R Markdown"
3: author: "Julia Kott"
4: date: 27-05-2020
5: institute: "Zi Mannheim, RG Psychology and Neurobiology of Sleep and Memory"
6: bibliography: ["rep_manuscripts.bib"]
7: output: bibtex:metropolis
8: ---
9:
10: ```{r setup, include = FALSE}
11: library(knitr)
12: library(purrr)
13: library(efsize)
14: library(tidyverse)
15:
16: julis_theme <- theme(legend.position = "top",
17:   plot.title = element_text(size = 22, hjust = 0.5),
18:   plot.subtitle = element_text(hjust = 0.5),
19:   axis.title = element_text(size = 22),
20:   axis.text = element_text(size = 18),
21:   axis.line = element_line(colour = "black"),
22:   legend.text = element_text(size = 20),
23:   legend.title = element_text(size = 20),
24:   panel.background = element_blank(),
25:   panel.grid = element_blank(),
26:   strip.background = element_rect(fill = NA, size = 1),
27:   strip.text = element_text(size = 20))
28: ```
```

The console output shows the R version (3.6.3) and the R Foundation's copyright notice. The Files pane on the right lists the project files:

| Name | Size | Modified |
|--------------------------------|--------|-----------------------|
| gsgispro | 40 B | May 26, 2020, 9:12 AM |
| 00reproducible_manuscripts.pdf | 1.9 MB | May 26, 2020, 9:12 AM |
| 00reproducible_manuscripts.Rmd | 5.4 KB | May 26, 2020, 9:12 AM |
| data | | |
| Dockerfile | 482 B | May 26, 2020, 9:12 AM |
| install.R | 207 B | May 26, 2020, 9:12 AM |
| litematic | | |
| MOSM_talk.Rproj | 205 B | May 26, 2020, 9:12 AM |
| pics | | |
| README.md | 1.8 KB | May 26, 2020, 9:12 AM |
| rep_manuscripts.bib | 757 B | May 26, 2020, 9:12 AM |

Binder

There will be pain

```
Error: LaTeX failed to compile /results/00reproducible_manuscripts.tex. See
https://yihui.org/tinytex/r/#debugging for debugging tips. See
00reproducible_manuscripts.log for more info.
@@reproducible Warning message:
In system2(..., stdout = if (use,
error in running command
Execution halted
```



MiKTeX Console

Something went wrong

Another MiKTeX program is running for updates.
Remedy: Close running MiKTeX programs and try again.

For more information, visit <https://miktex.org/kb/via-package-manager>.
Do you want to see the error details?

```
LaTeX Error: File 'beamer.cls' not found.
```

Commits on Jan 22, 2020

fixed stupid mistake

 einGlasRotwein committed 25 seconds ago

calculate d prime

 einGlasRotwein committed 8 minutes ago

fixed stupid naming bug

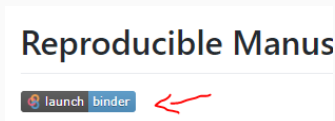
 einGlasRotwein committed 3 hours ago

But it's worth it

- Saves time for others and for future you.
- Mistakes are easier to spot and easier to correct.
- Your data and your manuscript will survive longer.
- Others can learn from your analyses.

Thank you!

Find this presentation on [GitHub](#).



And definitely check the binder button! Thank you, Philipp Zumstein!

Or on [CodeOcean](#).

Or on the [OSF](#).


Useful Ressources

- [Tutorial for a reproducible workflow](#) with R Markdown, Git, Make and Docker by Peikert and Brandmaier (2019)
- [Example capsule](#) for papaja on CodeOcean (also see the [papaja manual](#) and give [Frederik Aust](#) a cookie)
- The computational reproducibility part (week 5) of Daniel Laken's coursera course "[Improving Statistical Questions](#)"
- R Markdown - [The Definitive Guide](#) by Yihui Xie, J. J. Allaire and Garrett Golemund



Questions

- Collaborative writing?
- Journals?
- Why not only code?

Collaborative writing

 einGlasRotwein committed 22 seconds ago 1 parent f7408a3 commit cdd

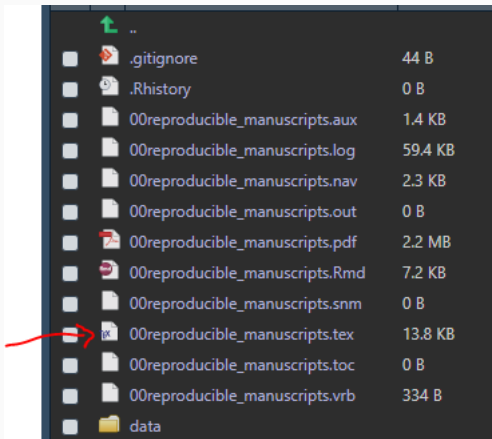
Showing 1 changed file with 1 addition and 1 deletion.

2  00reproducible_manuscripts.Rmd 

| | | |
|-----|---|--|
| ↑ | @@ -254,7 +254,7 @@ Or on the [OSF](https://osf.io/gzf9h/). | |
| 254 | ## Collaborative writing | 254 ## Collaborative writing |
| 255 | | 255 |
| 256 | Here is an example of how I am editing this slide. | 256 Here is an example of how I am editing this slide. |
| 257 | - I made a typo here . | 257 + I made a typo here . |
| 258 | In the commit history, you can see how I fixed it. | 258 In the commit history, you can see how I fixed it. |
| 259 | | 259 |

Journals

papaja e.g. also renders Word files (sometimes with formatting errors). But a lot of journals also accept tex files, which are also generated by R Markdown if you create a PDF.



Why not only code?

- It would still have to be in a capsule.
- A whole manuscript is readable for a broader audience and you can see where things “belong”.
- But yes - the code for the manuscript is probably the most difficult part, so trade-offs have to be made.

References

- Obels, Pepijn, Daniel Lakens, Nicholas A Coles, Jaroslav Gottfried, and Seth A Green. 2019. "Analysis of Open Data and Computational Reproducibility in Registered Reports in Psychology." PsyArXiv. <https://doi.org/10.31234/osf.io/fk8vh>.
- Peikert, Aaron, and Andreas M Brandmaier. 2019. "A Reproducible Data Analysis Workflow with R Markdown, Git, Make, and Docker." PsyArXiv. <https://doi.org/10.31234/osf.io/8xzqy>.
- San-Martin, M., M. Copaira, J. Zuniga, R. Rodreguez, G. Bustinza, and L. Acosta. 1968. "Aspects of Reproduction in the Alpaca." *Reproduction* 16 (3): 395–99. https://rep.bioscientifica.com/view/journals/rep/16/3/jrf_16_3_009.xml.