Reproducible Manuscripts in R Markdown

Juli Tkotz

27-05-2020

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.

Beyond reproducible analyses

3.1 Epistemic Trustworthiness

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

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

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.

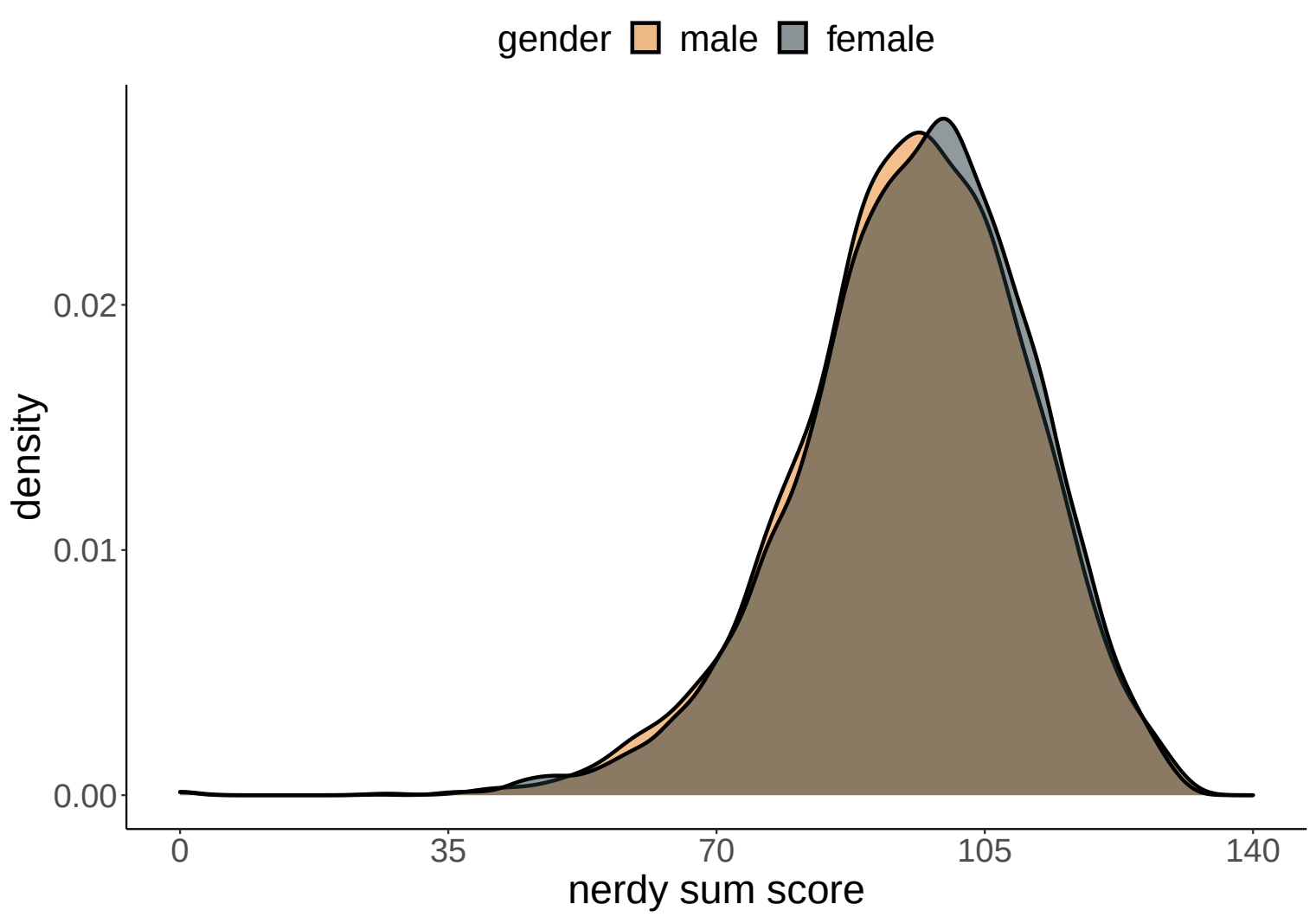
How about some stats?

```
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!


gender
male
female
density
0.02
0.01
0.00
0
35
70
105
140
nerdy sum score

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.
```

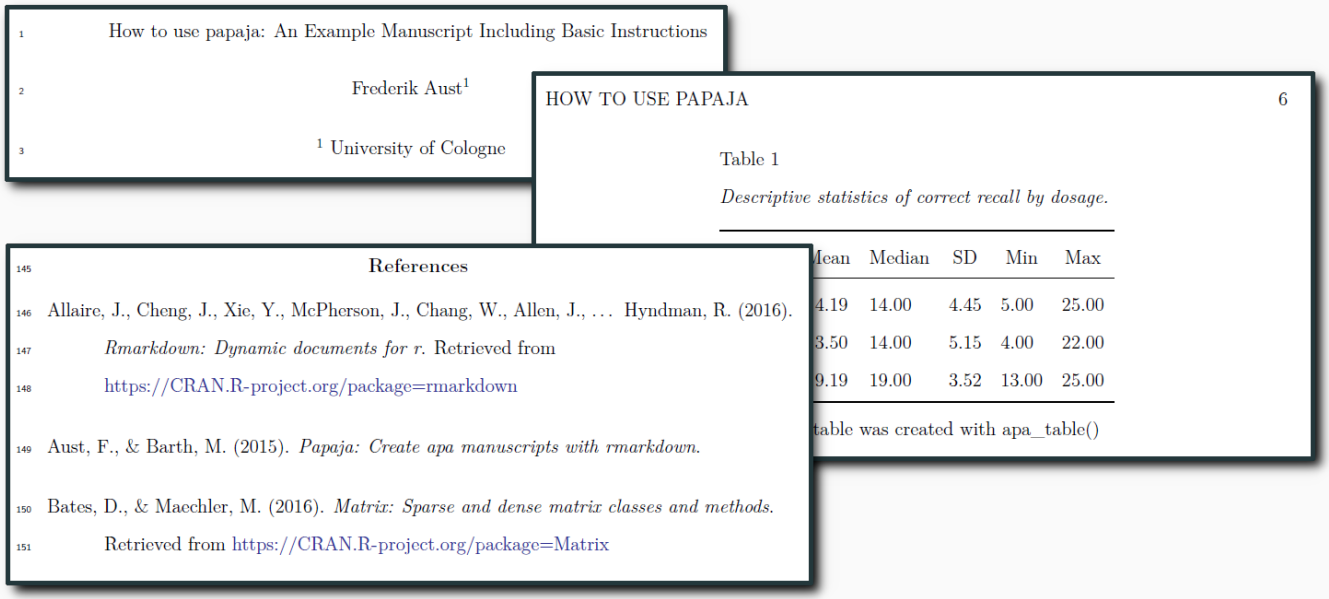
If I want to cite a paper, I can do this (San-Martin et al. 1968).
This also works if I cite San-Martin et al. (1968) as an in-text citation.

How to get bibtex references

Fully formatted articles

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

https://github.com/crsh/papaja



How to use papaja: An Example Manuscript Including Basic Instructions

Frederik Aust[1]

[1] University of Cologne

HOW TO USE PAPAJA 6

Table 1

Descriptive statistics of correct recall by dosage.

	Mean	Median	SD	Min	Max
	14.19	14.00	4.45	5.00	25.00
	13.50	14.00	5.15	4.00	22.00
	19.19	19.00	3.52	13.00	25.00

table was created with apa_table()

References

Allaire, J., Cheng, J., Xie, Y., McPherson, J., Chang, W., Allen, J., ... Hyndman, R. (2016). *Rmarkdown: Dynamic documents for r*. Retrieved from https://CRAN.R-project.org/package=rmarkdown

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

Bates, D., & Maechler, M. (2016). *Matrix: Sparse and dense matrix classes and methods*. Retrieved from https://CRAN.R-project.org/package=Matrix

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

On Code Ocean.

References

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.