

Abstract

13

14 This manuscript demonstrates how to use R Markdown and papaja to create an APA
15 conform manuscript. papaja builds on R Markdown, which uses pandoc to turn Markdown
16 into PDF or Word documents. The conversion to Word documents currently supports only a
17 limited set of features.

18

Keywords: APA style, knitr, R, R markdown, papaja

19

Word count: 1,753

20 How to use `papaja`: An Example Manuscript Including Basic Instructions

21 What is `papaja`?

22 Reproducible data analysis is an easy to implement and important aspect of the strive
23 towards reproducibility in science. For *R* users, R Markdown has been suggested as one
24 possible framework for reproducible analyses. `papaja` is a R-package in the making including
25 a [R Markdown](#) template that can be used with (or without) [RStudio](#) to produce documents,
26 which conform to the American Psychological Association (APA) manuscript guidelines (6th
27 Edition). The package uses the \LaTeX document class [apa6](#) and a .docx-reference file, so you
28 can create PDF documents, or Word documents if you have to. Moreover, `papaja` supplies
29 R-functions that facilitate reporting results of your analyses in accordance with APA
30 guidelines.

31 Markdown is a simple formatting syntax that can be used to author HTML, PDF, and
32 MS Word documents (among others). In the following I will assume you know how to use R
33 Markdown to conduct and comment your analyses. If this is not the case, I recommend you
34 familiarize yourself with [R Markdown](#) first. I use [RStudio](#) to create my documents, but the
35 general process works with any text editor.

36 How to use `papaja`

37 Once you have installed `papaja` and all other [required software](#), you can select the
38 APA template when creating a new R Markdown file through the RStudio menus, see
39 [Figure 1](#). When you click RStudio's *Knit* button (see [Figure 2](#)), `papaja`, `bookdown`,
40 `rmarkdown`, and `knitr` work together to create an APA conform manuscript that includes
41 both your text and the output of any embedded R code chunks within the manuscript.

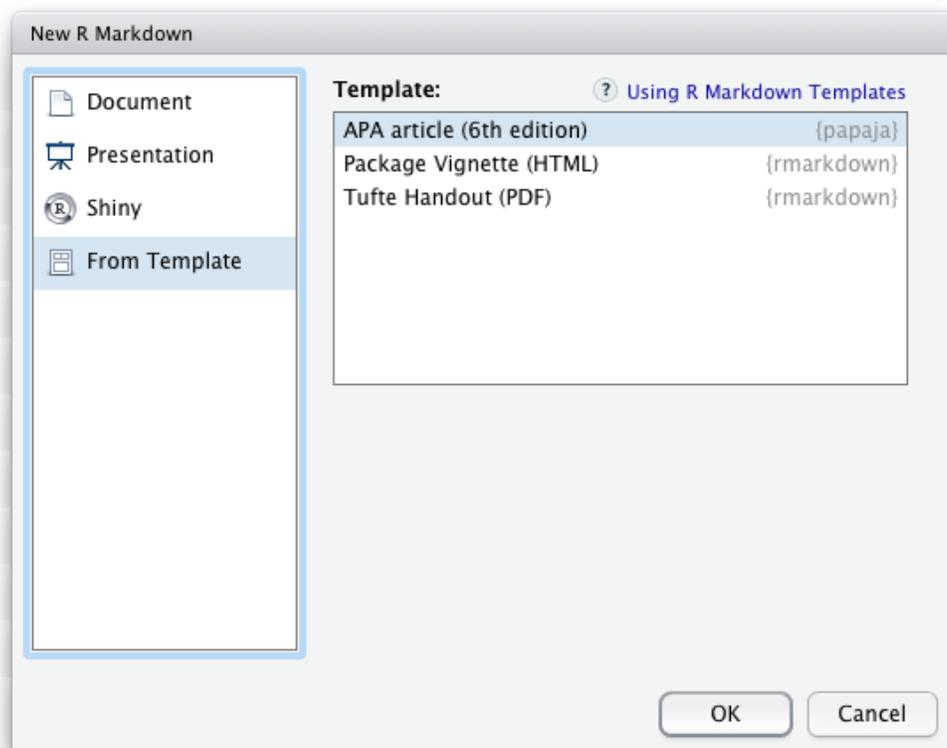


Figure 1. papaja's APA6 template is available through the RStudio menus.



Figure 2. The Knit button in the RStudio.

42 Printing R output

43 Any output from R is included as you usually would using R Markdown. By default
 44 the R code will not be displayed in the final documents. If you wish to show off your code
 45 you need to set `echo = TRUE` in the chunk options. For example, to include summary
 46 statistics of your data you could use the following code:

```
summary(mixed_data[, -1])
```

```
47 ##      Subject  Gender Dosage Task  Valence      Recall
48 ##  A      : 6    F:54   A:36   C:54   Neg:36   Min.    : 4.00
49 ##  B      : 6    M:54   B:36   F:54   Neu:36   1st Qu.:13.00
50 ##  C      : 6                C:36                Pos:36   Median  :15.00
51 ##  D      : 6                                Mean    :15.63
52 ##  E      : 6                                3rd Qu.:19.00
53 ##  F      : 6                                Max.    :25.00
54 ## (Other):72
```

55 But, surely, this is not what you want your submission to look like.

56 **Print tables.** For prettier tables, I suggest you try `apa_table()`, which builds on
 57 `knitr`'s `kable()`, and `apa_num()`, which can be used to properly round and report numbers.

```
descriptives <- mixed_data %>%
  group_by(Dosage) %>%
  summarize(
    Mean = mean(Recall)
    , Median = median(Recall)
    , SD = sd(Recall)
    , Min = min(Recall)
    , Max = max(Recall)
```

Table 1

Descriptive statistics of correct recall by dosage.

Dosage	Mean	Median	SD	Min	Max
A	14.19	14.00	4.45	5	25
B	13.50	14.00	5.15	4	22
C	19.19	19.00	3.52	13	25

Note. This table was created with`apa_table()`.

```

)
descriptives[, -1] <- apa_num(descriptives[, -1])

apa_table(
  descriptives
  , caption = "(ref:descriptives-caption)"
  , note = "(ref:descriptives-note)"
)

```

58 Of course popular packages like `xtable`¹ or `tables` can also be used to create tables
59 when knitting PDF documents. These packages, however, cannot be used when you want to
60 create Microsoft Word documents because they rely on L^AT_EX for typesetting. `apa_table()`
61 creates tables that conform to APA guidelines and are correctly rendered in PDF and Word
62 documents. But don't get too excited; table formatting is somewhat limited for Word
63 documents due to missing functionality in pandoc (e.g., it is not possible to have cells or
64 headers span across multiple columns).

¹ When you use `xtable()`, table captions are [set to the left page margin](#).

65 As required by the APA guidelines, tables are deferred to the final pages of the
66 manuscript when creating a PDF. Again, this is not the case in Word documents due to
67 limited pandoc functionality. To place tables and figures in your text instead, set the
68 `figsintext` parameter in the YAML header to `yes` or `true`, as I have done in this document.

69 The bottom line is, Word documents will be less polished than PDF. The resulting
70 documents should suffice to enable collaboration with Wordy colleagues and prepare a
71 journal submission with limited manual labor.

72 **Embed plots.** As usual in R Markdown, you can embed R-generated plots into your
73 document, see Figure 3.

```
apa_beeplot(  
  mixed_data  
  , id = "Subject"  
  , dv = "Recall"  
  , factors = c("Task", "Valence", "Dosage")  
  , dispersion = conf_int  
  , ylim = c(0, 30)  
  , las = 1  
  , args_points = list(cex = 1.5)  
  , args_arrows = list(length = 0.025)  
)
```

74 Again, as required by the APA guidelines, figures are deferred to the final pages of the
75 document unless you set `figsintext` to `yes`.

76 **Referencing figures and tables.** `papaja` builds on the `bookdown` package, which
77 provides limited cross-referencing capabilities within documents. By default you can insert
78 figure and table numbers into the text using `\@ref(fig:chunk-name)` for figures or
79 `\@ref(tab:chunk-name)` for tables. Note that for this syntax to work chunk names cannot

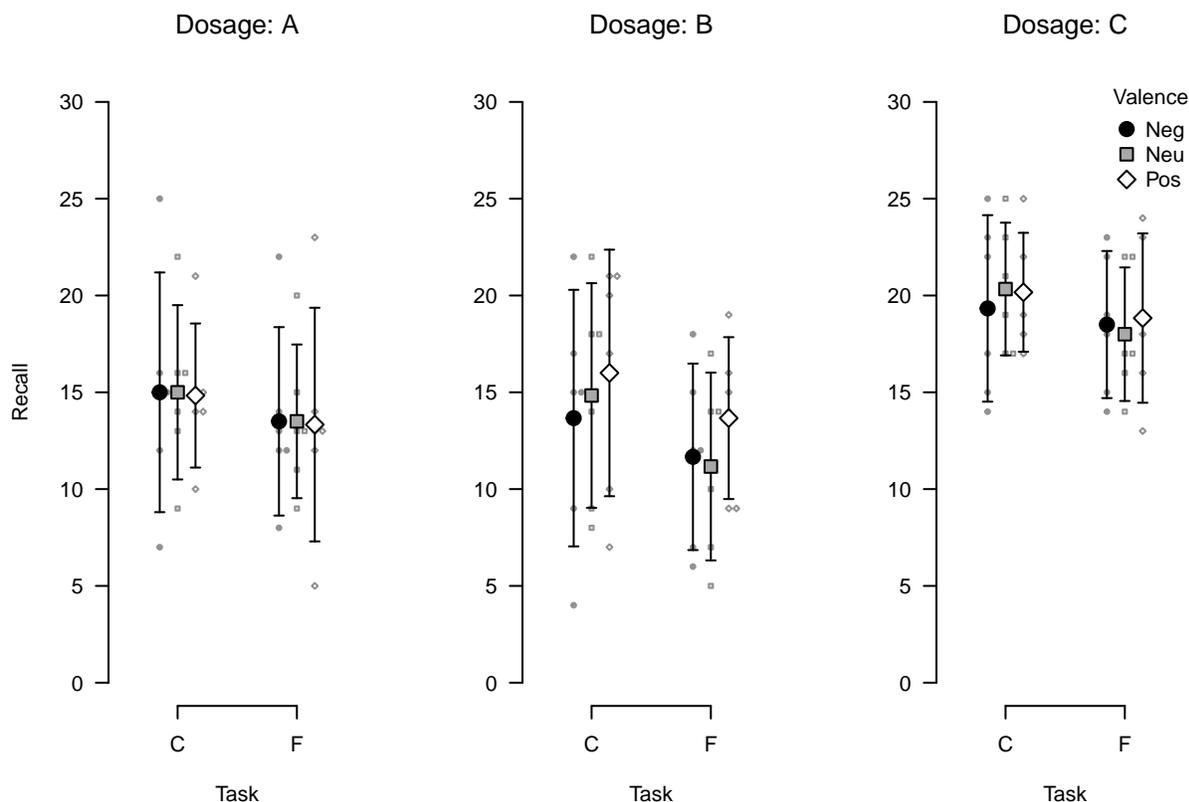


Figure 3. Bee plot of the example data set. Small points represent individual observations, large points represent means, and error bars represent 95% confidence intervals.

80 include `_`. If you need to embed an external image that is not generated by R use the
 81 `knitr::include_graphics()` function. See the [great book on bookdown](#) for details.
 82 Cross-referencing is currently not available for equations in `bookdown`. However, as anywhere
 83 in R Markdown documents you can use \LaTeX commands if the functionality is not provided
 84 by `rmarkdown/bookdown` and you don't need to create Word documents.

85 **Report statistical analyses.** `apa_print()` will help you report the results of your
 86 statistical analyses. The function will format the contents of R objects and produce readily
 87 reportable text.

```
recall_anova <- afex::aov_car(
  Recall ~ (Task * Valence * Dosage) + Error(Subject/(Task * Valence)) + Dosage
```

```

, data = mixed_data
, type = 3
)
recall_anova_results <- apa_print(recall_anova, es = "pes")
recall_anova_results_p <- apa_print(recall_anova, es = "pes", in_paren = TRUE)

```

88 Now, you can report the results of your analyses like so:

```

Item valence (`r recall_anova_results_p$full$Valence`) and the task
affected recall performance, `r recall_anova_results$full$Task`; the dosage,
however, had no effect on recall, `r recall_anova_results$full$Dosage`.
There was no significant interaction.

```

89 Item valence ($F[1.62, 24.36] = 3.46$, $MSE = 2.62$, $p = .056$, $\hat{\eta}_p^2 = .187$) and the
90 task affected recall performance, $F(1, 15) = 43.13$, $MSE = 2.23$, $p < .001$,
91 $\hat{\eta}_p^2 = .742$; the dosage, however, had no effect on recall, $F(2, 15) = 2.97$,
92 $MSE = 117.17$, $p = .082$, $\hat{\eta}_p^2 = .283$. There was no significant interaction.

93 What's even more fun, you can easily create a complete ANOVA table using by passing
94 `recall_anova_results$table` to `apa_table()`, see [Table 2](#).

```

apa_table(
  recall_anova_results$table
, align = c("l", "r", "c", "r", "r", "r")
, caption = "(ref:anova-caption)"
, note = "(ref:anova-note)"
)

```

Table 2

ANOVA table for the analysis of the example data set.

Effect	$\hat{\eta}_p^2$	F	df^{GG}	df_{res}^{GG}	MSE	p
Dosage	.283	2.97	2	15	117.17	.082
Task	.742	43.13	1	15	2.23	< .001
Valence	.187	3.46	1.62	24.36	2.62	.056
Dosage \times Task	.196	1.83	2	15	2.23	.195
Dosage \times Valence	.241	2.38	3.25	24.36	2.62	.090
Task \times Valence	.091	1.50	1.35	20.20	2.67	.242
Dosage \times Task \times Valence	.049	0.39	2.69	20.20	2.67	.743

Note. This is a table created using `apa_print()` and `apa_table()`.

95 Citations

96 No manuscript is complete without citation. In order for citations to work, you need to
 97 supply a .bib-file to the `bibliography` parameter in the YAML front matter. Once this is
 98 done, [e.g., `@james_1890; @bem_2011`] produces a regular citation within parentheses
 99 (e.g., Bem, 2011; James, 1890). To cite a source in text simply omit the brackets; for
 100 example, write `@james_1890` to cite James (1890). For other options see the [overview of the](#)
 101 [R Markdown citation syntax](#).

102 The citation style is automatically set to APA style. If you need to use a different
 103 citation style, you can set in the YAML front matter by providing the `csl` parameter. See
 104 the [R Markdown documentation](#) and [Citation Style Language](#) for further details.

105 If you use RStudio, I have created an [easy-to-use add-in](#) that facilitates inserting
 106 citations into a document. The relevant references will, of course, be added to the documents
 107 reference section automatically. Moreover, the addin can directly access you Zotero database.

108 I think it is important to credit the software we use. A lot of R packages are developed
 109 by academics free of charge. As citations are the currency of science, it's easy to compensate
 110 volunteers for their work by citing the R packages we use. I suspect that, among other
 111 things, this is rarely done because it is tedious work. That's why papaja makes citing R and
 112 its packages easy:

```
r_refs(file = "r-references.bib")
my_citation <- cite_r(file = "r-references.bib")
```

113 `r_refs()` creates a BibTeX file containing citations for R and all currently loaded
 114 packages. `cite_r()` takes these citations and turns them into readily reportable text.
 115 `my_citation` now contains the following text that you can use in your document: R (Version
 116 4.1.2; R Core Team, 2015) and the R-packages *afex* (Version 1.0.1; Singmann, Bolker,
 117 Westfall, & Aust, 2016), *bindrcpp* (Müller, 2017), *boot* (Version 1.3.28; Davison & Hinkley,
 118 1997), *broom* (Version 0.7.12; Robinson, 2016), *dplyr* (Version 1.0.8; Wickham & Francois,
 119 2016), *emmeans* (Version 1.7.2; R. Lenth, 2018), *estimability* (Version 1.3; R. V. Lenth,
 120 2015), *knitr* (Version 1.37; Xie, 2015), *lme4* (Version 1.1.28; Bates, Mächler, Bolker, &
 121 Walker, 2015), *lsmeans* (Version 2.30.0; R. V. Lenth, 2016), *Matrix* (Version 1.3.4; Bates &
 122 Maechler, 2016), *MBESS* (Version 4.9.0; Kelley, 2016), *papaja* (Version 0.2.0; Aust & Barth,
 123 2015), *reshape2* (Version 1.4.4; Wickham, 2007), *rmarkdown* (Version 2.13; Allaire et al.,
 124 2016), *testthat* (Version 3.1.2; Wickham, 2011), and *tinylabels* (Version 0.2.3; Barth, 2022)

125 Math

126 If you need to report formulas, you can use the flexible \LaTeX syntax (it will work in
 127 Word documents, too). Inline math must be enclosed in $\$$ or $\backslash($ and $\backslash)$ and the result will
 128 look like this: $d' = z(H) - z(FA)$. For larger formulas displayed equations are more
 129 appropriate; they are enclosed in $\$\$$ or $\backslash[$ and $\backslash]$,

$$d' = \frac{\mu_{old} - \mu_{new}}{\sqrt{0.5(\sigma_{old}^2 + \sigma_{new}^2)}}.$$

130 Document options

131 This text is set as manuscript. If you want a thesis-like document you can change the
132 `class` in the YAML front matter from `man` to `doc`. You can also preview a polished journal
133 typesetting by changing the `class` to `jou`. Refer to the `apa6` document class [documentation](#)
134 for further `class` options, such as paper size or draft watermarks.

135 When creating PDF documents, line numbering can be activated by setting the
136 `linenumbers` argument in the YAML front matter to `yes`. Moreover, you can create lists of
137 figure or table captions at the end of the document by setting `figurelist` or `tablelist` to
138 `yes`, respectively. These option have no effect on Word documents.

139 Last words

140 That's all I have; enjoy writing your manuscript. If you have any trouble or ideas for
141 improvements, open an [issue](#) on GitHub or open a pull request. If you want to contribute,
142 take a look at the [open issues](#) if you need inspiration. Other than that, there are many
143 output objects from analysis methods that we would like `apa_print()` to support. Any new
144 S3/S4-method for this function are always appreciated (e.g., `factanal`, `fa`, `lavaan`, `lmer`, or
145 `glmer`).

146 References

- 147 Allaire, J., Cheng, J., Xie, Y., McPherson, J., Chang, W., Allen, J., ... Hyndman, R.
148 (2016). *Rmarkdown: Dynamic documents for r*. Retrieved from
149 <https://CRAN.R-project.org/package=rmarkdown>
150 Aust, F., & Barth, M. (2015). *Papaja: Create APA manuscripts with RMarkdown*.

- 151 Barth, M. (2022). *tinylabls: Lightweight variable labels*. Retrieved from
152 <https://cran.r-project.org/package=tinylabls>
- 153 Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects
154 models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.
155 <https://doi.org/10.18637/jss.v067.i01>
- 156 Bates, D., & Maechler, M. (2016). *Matrix: Sparse and dense matrix classes and*
157 *methods*. Retrieved from <https://CRAN.R-project.org/package=Matrix>
- 158 Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous
159 retroactive influences on cognition and affect. *Journal of Personality and Social*
160 *Psychology*, *100*(3), 407–425. <https://doi.org/10.1037/a0021524>
- 161 Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their applications*.
162 Cambridge: Cambridge University Press. Retrieved from
163 <http://statwww.epfl.ch/davison/BMA/>
- 164 James, W. (1890). *The principles of psychology*. Holt: New York.
- 165 Kelley, K. (2016). *MBESS: The MBESS r package*. Retrieved from
166 <https://CRAN.R-project.org/package=MBESS>
- 167 Lenth, R. (2018). *Emmeans: Estimated marginal means, aka least-squares means*.
168 Retrieved from <https://CRAN.R-project.org/package=emmeans>
- 169 Lenth, R. V. (2015). *Estimability: Tools for assessing estimability of linear*
170 *predictions*. Retrieved from <https://CRAN.R-project.org/package=estimability>
- 171 Lenth, R. V. (2016). Least-squares means: The R package lsmeans. *Journal of*
172 *Statistical Software*, *69*(1), 1–33. <https://doi.org/10.18637/jss.v069.i01>
- 173 Müller, K. (2017). *Bindrcpp: An 'rcpp' interface to active bindings*. Retrieved from
174 <https://CRAN.R-project.org/package=bindrcpp>
- 175 R Core Team. (2015). *R: A language and environment for statistical computing*.
176 Vienna, Austria: R Foundation for Statistical Computing. Retrieved from
177 <http://www.R-project.org/>

- 178 Robinson, D. (2016). *Broom: Convert statistical analysis objects into tidy data*
179 *frames*. Retrieved from <https://CRAN.R-project.org/package=broom>
- 180 Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2016). *Afex: Analysis of factorial*
181 *experiments*. Retrieved from <https://CRAN.R-project.org/package=afex>
- 182 Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical*
183 *Software*, 21(12), 1–20. Retrieved from <http://www.jstatsoft.org/v21/i12/>
- 184 Wickham, H. (2011). Testthat: Get started with testing. *The R Journal*, 3, 5–10.
185 Retrieved from
186 http://journal.r-project.org/archive/2011-1/RJournal_2011-1_Wickham.pdf
- 187 Wickham, H., & Francois, R. (2016). *Dplyr: A grammar of data manipulation*.
188 Retrieved from <https://CRAN.R-project.org/package=dplyr>
- 189 Xie, Y. (2015). *Dynamic documents with R and knitr* (2nd ed.). Boca Raton, Florida:
190 Chapman; Hall/CRC. Retrieved from <http://yihui.org/knitr/>