

PSY 503: Foundations of Statistical Methods in Psychological Science

Reproducibility, Papaja

















Suyog Chandramouli

311 PSH (Princeton University)















5th November, 2025


Reproducibility vs Replicability


Computational reproducibility

	Computational reproducibility	Direct replicability	
Population			
Study design			
Experimenter			
Data	01100 10110 11110	01100 10110 11110	 Same  Different
Analysis plan			
Analyst			
Code			
Estimate			

















Computational reproducibility

	Computational reproducibility	Direct replicability	
Population			
Study design			
Experimenter			
Data	01100 10110 11110	01100 10110 11110	
Analysis plan			
Analyst			
Code			
Estimate			


Same


Different

Computational reproducibility

	Computational reproducibility	Direct replicability	
Population			
Study design			
Experimenter			 Same
Data	01100 10110 11110	01100 10110 11110	 Different
Analysis plan			
Analyst			
Code			
Estimate			

Computational reproducibility

- NSF subcommittee on replicability in science:
 - “[Computational] Reproducibility is a minimum necessary condition for a finding to be believable and informative.”
(p. 4, Cacioppo, Kaplan, Krosnick, Olds, & Dean, 2015; also see Peng, 2011)

[Computational] Reproducibility Crisis

- Full reproduction attempts of published analyses had high failure rate.

Journal	Source	Failure rate
<i>Journal of Cognition</i>	Hardwicke et al. (2018)	37%
<i>Quarterly Journal of Political Science</i>	Eubank (2016)	58%
<i>Strategic Management Journal</i>	Bergh et al. (2017)	30%
<i>Science</i>	Stodden, Seiler, & Ma (2018)	41%

[Computational] Reproducibility Crisis

- Full reproduction attempts of published analyses had high failure rate.

Field	Source	Failure rate
Psychology	Artner et al. (2020)	30%
Psychology (RR)	Obels et al. (2020)	42%
Economics	Vilhuber (2020)	39-51%
Organismal biology	Andrew et al. (2015)	35%
Genetics	Gilbert et al. (2012)	30%
Geosciences	Konkol, Kray & Pfeiffer (2019)	56%
RCT (primary outcome)	Naudet et al. (2018)	12%

Research Transparency & Open Data

- A minimum requirement for computational reproducibility
- Bridges the gap between reproducibility and replicability
- Creates a verifiable research chain from data to conclusions
- Enables scientific self-correction through community validation

Research Transparency & Open Data

- Common Artifacts / implementation methods
 - Basic
 - “e-mail authors for code/data”
 - Better practices
 - Public Repositories (e.g. OSF, GitHub, etc.)
 - Standardized data formats
 - Detailed documentation
 - Version-controlled scripts
 - Raw as well as processed data

Computational irreproducibility despite transparency

- Eubank (2016)
 - Only 16.7% (4 out of 24) of studies could be reproduced without any modifications to the original code
 - This low success rate was despite data and code sharing requirements
- Konkol, Kray & Pfeiffer (2019)
 - Seed values in code need to be explicitly set to ensure consistency across runs
 - Complex submissions may require multiple seed values for different packages
- Artner et al. (2020)
 - 11 out of 35 articles reproducible
 - 11 more needed interaction with authors to avail more code, datasets, and details
 - 13 had errors, and couldn't be reproduced despite authors' assistance

Computational irreproducibility

- Common Causes

- Typos
- Copy-paste errors
- Incorrect rounding
- File paths
- Incomplete data
- Incomplete scripts
- “Outdated” results
- Outdated libraries

- Implications

- is wide-spread
- undermines credibility of the literature
- wastes resources
- is unethical
- is not caused by simple sloppiness

- Solutions

- Automation
- Code review
- Archiving

Dynamic Documents

Dynamic Documents

oh my god, now the journal no longer accepts .PDF files,
so I have to find a way to blow this out to a .docx.

— **Joe Hilgard, data guy (@JoeHilgard)** July 1, 2015

Dynamic Documents

Dreaming of a day when resubmitting an article to a different journal doesn't mean hair-pulling reformatting, especially when tables and figures go to the end with separate captions ... #academictimewasting 🤯

— **Alexa Morcom (@alexa_morcom)** August 16, 2019

Dynamic Documents

- Discuss:
 - Are you familiar with any tools for generating dynamic documents?

Dynamic Documents

- Different packages for different languages and outputs
 - rmarkdown / knitr / Quarto
 - StatTag
 - tidystats
 - Jupyter
 - org-mode
 - ...

Dynamic Documents

- Features
 - Combine manuscript and analysis scripts into a single document
 - Automates reporting of results
 - Ensures statistical consistence
 - Documentation is easier
 - Reproduction is easier
 - ..
- Papaja
 - Works with R-markdown

papaja

papaja

- Preparing **APA** Journal **Articles**

1. Designed for APA-style manuscripts

- <https://apastyle.apa.org/6th-edition-resources>
- https://owl.purdue.edu/owl/research_and_citation/apa6_style/apa_formatting_and_style_guide/apa_changes_6th_edition.html

2. Templates for PDF and DOCX documents

3. Functions to report results, e.g.

1. `apa_print()`
2. `apa_table()`
3. `apa_factorial_plot()`, `theme_apa()`

papaja : getting started

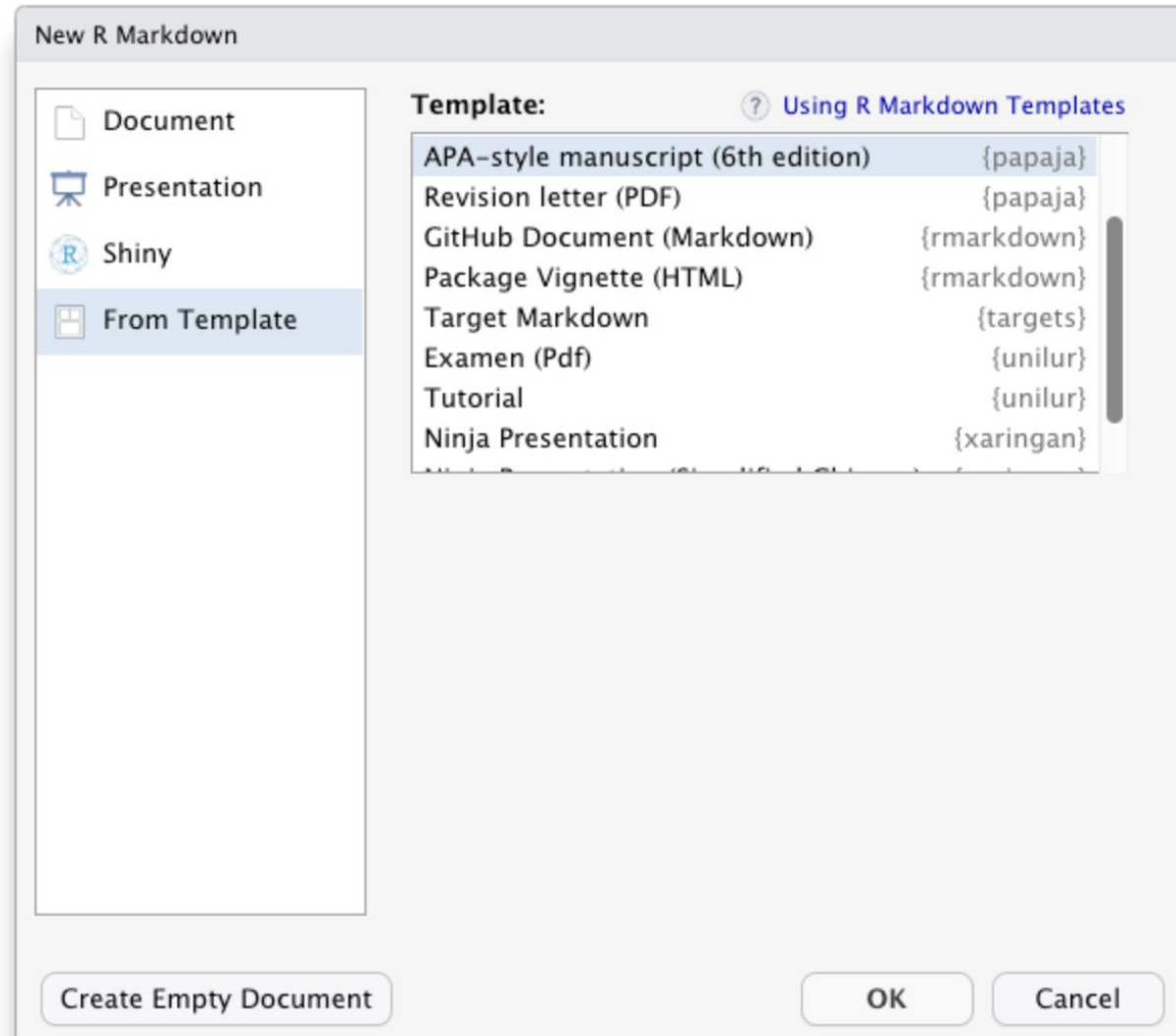
- It's a good practice to visit documentation and see instructions
 - https://frederikaust.com/papaja_man/introduction.html#getting-started
 - <https://github.com/crsh/papaja>

```
# Install package from CRAN
install.packages("papaja")

# If you want a potentially more recent version
remotes::install_github("crsh/papaja")
```

```
# Install TinyTeX if needed
install.packages("tinytex")
tinytex::install_tinytex()
```

papaja : creating a document



papaja : APA manuscript template

```
---
title      : "The title"
shorttitle : "Title"

author:
- name      : "First Author"
  affiliation : "1"
  corresponding : yes    # Define only one corresponding author
  address     : "Postal address"
  email       : "my@email.com"
  role: # Contributorship roles (e.g., CRediT, https://credit.niso.org/)
    - "Conceptualization"
    - "Writing - Original Draft Preparation"
    - "Writing - Review & Editing"
- name      : "Ernst-August Doelle"
  affiliation : "1,2"
  role:
    - "Writing - Review & Editing"
    - "Supervision"

affiliation:
- id      : "1"
  institution : "Wilhelm-Wundt-University"
- id      : "2"
  institution : "Konstanz Business School"

authornote: |
  Add complete departmental affiliations for each author here. Each new line herein must be indented, like this line.

  Enter author note here.

abstract: |
  One or two sentences providing a basic introduction to the field, comprehensible to a scientist in any discipline.
  Two to three sentences of more detailed background, comprehensible to scientists in related disciplines.
  One sentence clearly stating the general problem being addressed by this particular study.
```

Alternatives

- apathe : APA thesis with R Markdown
 - <https://github.com/crsh/apathe>
 - Built on and abstracted over papaja, primarily for thesis documents
 - simpler, but a very new project with little documentation
- APAQUARTO
 - Quarto version of PAPAJA
 - Still in active development
 - <https://github.com/wjschne/apaquarto>

Live Demo

Lab tasks

1) Create a new papaja document and populate the YAML front matter of with the manuscript metadata.

2) Make your manuscript “preprint-ready” (unlike the default):

- Hide line numbers,
- use the single-column single-spaced document style (i.e. use the "doc" class option), and
- add a “DRAFT” watermark on every page.

Lab tasks

3) Now let's report some results.

- Copy the contents of one of your R Markdown document to your papaja document.
- Report the numerical results of one or more analysis using in-line code chunks. If possible, use `apa_print()`.

4) Add a figure and/or table to your results section and cross-reference it in the text. Use text-references to add a caption.