



psyphr.read: wrangle and tidy psychophysiological data in R

Kathleen E. Wendt [1], Siqi Zhang, Mallory J. Feldman [2], G. Brooke Anderson [1], & J. Douglas Coatsworth [1]

[1] Colorado State University; [2] University of North Carolina at Chapel Hill



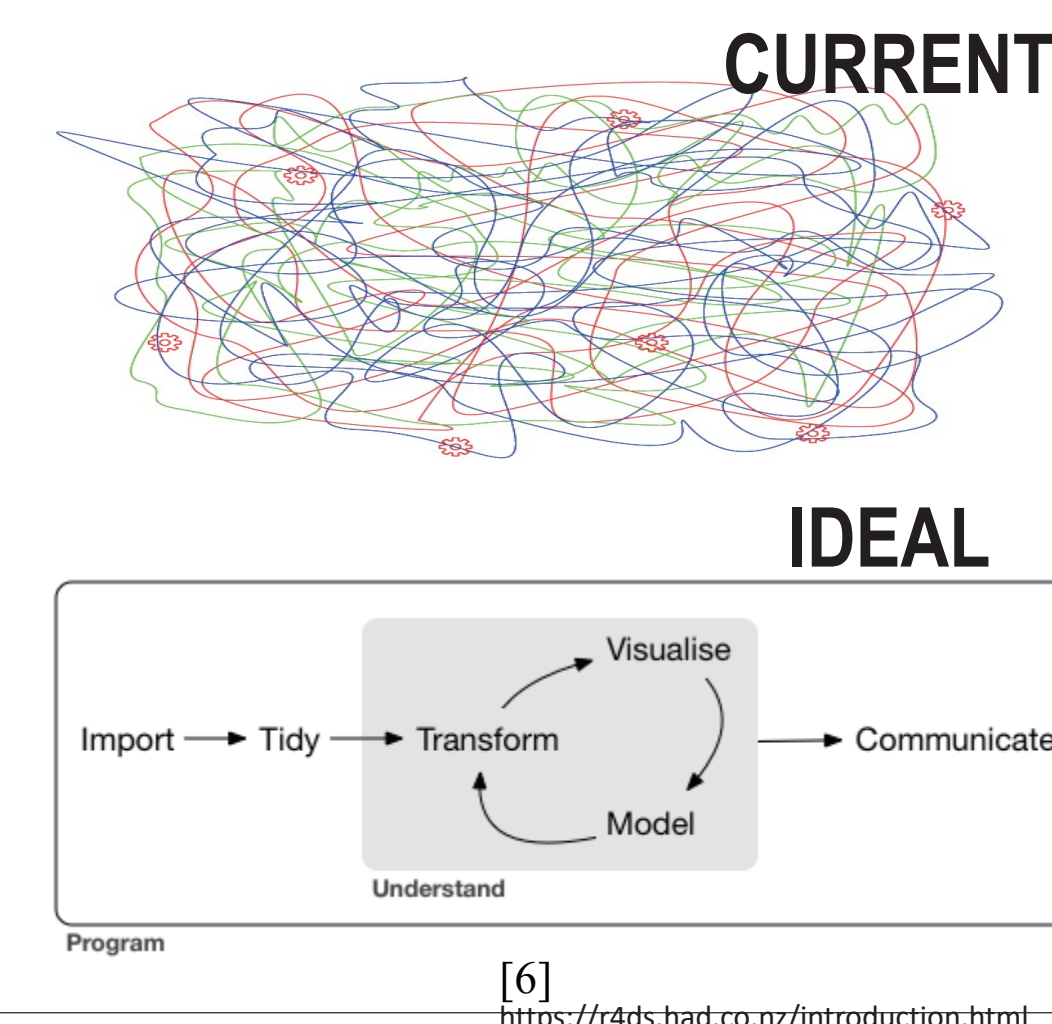
1 a burgeoning field

Physiological measures are increasingly popular in social, behavioral, and psychological research but yield unwieldy amounts of data and bring their own challenges [1].

To date, few tools have been developed [2] to support reproducibility, efficiency, and accuracy [3-5] in psychophysiological data processing and management.

2 workflow concerns

Output files from proprietary software are often compiled or aligned by hand (e.g., copy-paste). Manual processes introduce opportunities for bias and error. Error-ridden data produces error-driven inferences. Lack of standardization obstructs shareability and reproducibility.

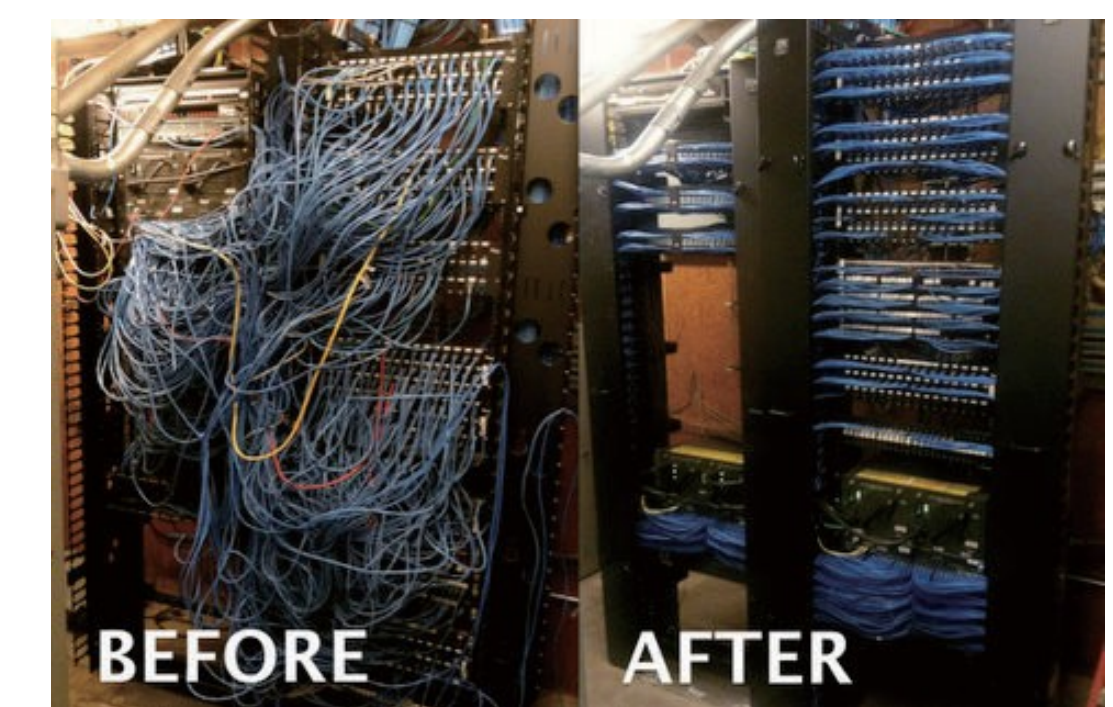


3 tidy data

A main principle of computational reproducibility is tidy data [7].

- Each column as a variable
- Each row as an observation
- Each table as a measure

Our work extends the tidy principles to physiological data.



<https://advwireonline.com/cable-management.html>

01_01.xlsx
01_02.xlsx
01_03.xlsx
02_01.xlsx
02_02.xlsx
02_03.xlsx

=>

my study

subject 1
activity 1

+

subject 1
activity 2

+

subject 1
activity 3

+

subject n
activity m

+ ...

=>

measure A

measure B

summary

, ...

```
my_study <- psyphr.read::MW_study("D:/my_study")
```

4 psyphr.read package

psyphr.read is the first in our suite of free, open source R packages. With one line of code, the user combines and organizes the entire study (see above illustration). In this tidy form, the user can quickly and easily generate summary statistics and exploratory plots. psyphr.read includes simple data validation techniques to catch errors or problems early.

5 psyphr suite; future directions

psyphr.read

+

psyphr.check

+

psyphr.plot

+

psyphr.model

+

psyphr.bids

wrangle and tidy
physiological data

sophisticated data
validation schemes

publication-grade
plots and graphics

common statistical
models in area

compatibility with Brain
Imaging Data Structure
(BIDS) [8]

The full psyphr suite will help researchers spend less time on data management and more time “doing science.” This work will provide the foundation of open science practices for psychophysiological research and promote computational reproducibility in fields with increasingly large and complex datasets. Watch our progress and contribute on github.com/psyphr-dev.



COLORADO STATE UNIVERSITY

references &
poster copy



KEW supported by
National Science Foundation
Graduate Research Fellowship Program
Grant No. 006784-00002

