The effect of erroneous R code on student performance

James Bartlett[1]

[1] University of Glasgow, United Kingdom

Author Note

Enter author note here.

The authors made the following contributions. James Bartlett: Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing.

Correspondence concerning this article should be addressed to James Bartlett, 62 Hillhead Street, Glasgow. E-mail: james.bartlett@glasgow.ac.uk

Abstract

Blah blah blah.

Keywords: learning, R, statistical programming, error-full learning

Word count: 1200

The effect of erroneous R code on student performance

Data skills are increasingly recognised as a key component of psychological literacy. To promote reproducible data preparation and analysis workflows, educators have highlighted the role of teaching students how to use statistical programming languages instead of point-and-click software (McAleer et al., 2022). However, programming is rare in UK psychology curricula (TARG Meta-Research Group, 2022) and offers unique challenges such as how to prepare students to debug their code. Debugging code is a separate problem solving skill to learn alongside statistics, so it is important to understand how best to teach students debugging skills.

Hoffman and Elmi (2021) reported a small pilot study using SAS where they compared a traditional error-free course structure to an error-full course focusing on debugging errors alongside key concepts. 80% of students preferred the error-full course but the study only included 18 participants and just 4 students completed assignments following each course, meaning they could not compare performance. Therefore, in our study, we want to apply these methods to the programming language R and recruit a larger sample.

We hypothesise that students who complete the error-full lecture will score higher on a data skills assignment than students who complete the error-free lecture.

Methods

Participants

Before collecting data, we performed an *a priori* power analysis to calculate how many participants we would need. Hoffman and Elmi (2021) did not report any performance data, so we used Bebermeier and Hagemann (2019) to set our smallest effect size of interest. They investigated the effect of creating statistics exercises based on research article. The researchers found students performed better on a class assignment when they completed these exercises than when students did not complete the exercises ($d = 0.55$). The authors

did not comment on the effect size, so we chose a more conservative estimate based on the small telescopes approach (Simonsohn, 2015) for the effect size the original study had 33% power to detect. Using an effect size of d = 0.38, we aimed to recruit 149 participants per group for an independent samples t-test ($\alpha = 0.05$, power = 0.90).

We finished with two groups of 145 and 130 participants (N = 275), slightly fewer than our initial target.

Material

In the error-free lecture, students heard a one hour presentation on data wrangling, showing how to use Tidyverse functions like mutate, filter, and select.

In the error-full lecture, students heard a one hour presentation on data wrangling, the same as in the error-free group. However, in this group we also guided students through an error interpretation session to demonstrate common errors when using these data wrangling functions.

Both groups of students completed the same data skills assignment on data wrangling where students had to write code to solve problems and debug errors. Scores could range from 0-100.

Procedure

We offered participants an additional bonus lecture outside their normal course curriculum. Students could register interest on their course page and provide informed consent. On sign up, students were randomly allocated to attend the error-free lecture or error-full lecture. In the hour immediately following the lecture, students completed the data skills assignment and were debriefed. We provided students who did not receive the error-full lecture a link to the lecture recording to ensure they received the debugging guidance. We demonstrate the procedure as a diagram in Figure 1.

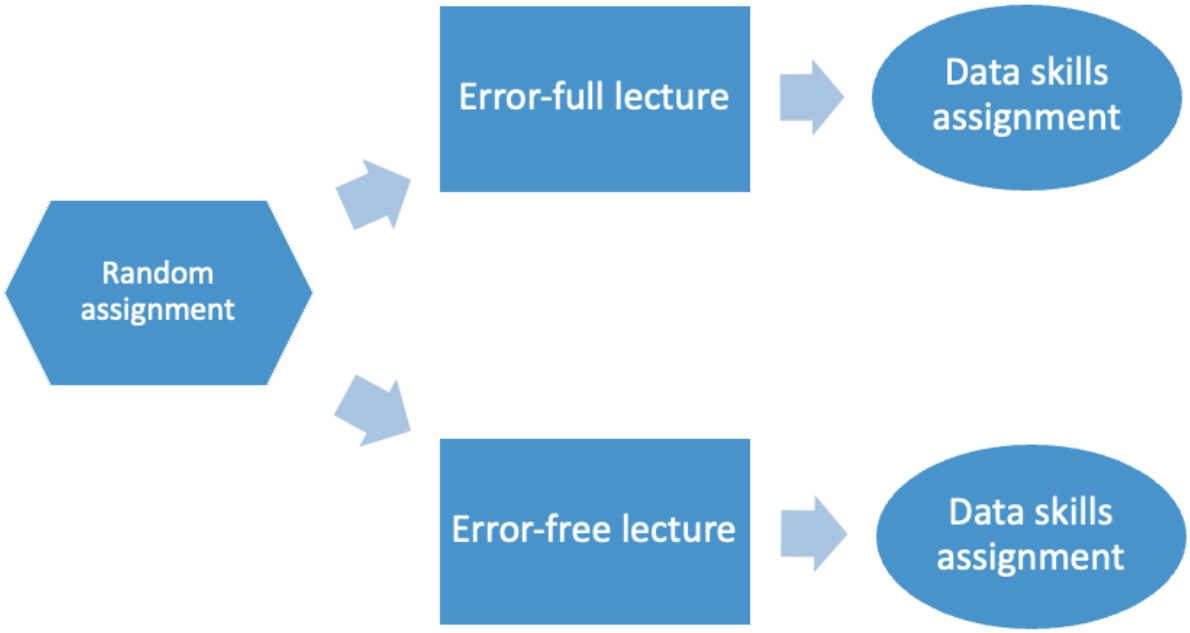

Figure 1. Procedure diagram showing how students were randomly allocated to an error-free or error-full version of a lecture before completing a data skills assignment.

Design and data analysis

We used R (Version 4.1.3; R Core Team, 2022) and the R-packages *dplyr* (Version 1.0.10; Wickham, François, et al., 2022), *forcats* (Version 0.5.1; Wickham, 2021), *ggplot2* (Version 3.3.6; Wickham, 2016), *papaja* (Version 0.1.1; Aust & Barth, 2022), *purrr* (Version 0.3.4; Henry & Wickham, 2020), *pwr* (Version 1.3.0; Champely, 2020), *readr* (Version 2.1.2; Wickham, Hester, et al., 2022), *shiny* (Version 1.7.3; Chang et al., 2022), *stringr* (Version 1.4.1; Wickham, 2022), *tibble* (Version 3.1.8; Müller & Wickham, 2022), *tidyr* (Version 1.2.0; Wickham & Girlich, 2022), *tidyverse* (Version 1.3.2; Wickham et al., 2019), and *tinylabels* (Version 0.2.3; Barth, 2022) for all our analyses.

Results

Discussion

References

Aust, F., & Barth, M. (2022). *papaja: Prepare reproducible APA journal articles with R Markdown*. https://github.com/crsh/papaja

Barth, M. (2022). *tinylabels: Lightweight variable labels*. https://cran.r-project.org/package=tinylabels

Bebermeier, S., & Hagemann, A. (2019). Creating Statistics Exercises on the Basis of Research Articles. *Teaching of Psychology*, *46*(3), 240–245. https://doi.org/10.1177/0098628319853938

Champely, S. (2020). *Pwr: Basic functions for power analysis*. https://CRAN.R-project.org/package=pwr

Chang, W., Cheng, J., Allaire, J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson, J., Dipert, A., & Borges, B. (2022). *Shiny: Web application framework for r*. https://CRAN.R-project.org/package=shiny

Henry, L., & Wickham, H. (2020). *Purrr: Functional programming tools*. https://CRAN.R-project.org/package=purrr

Hoffman, H. J., & Elmi, A. F. (2021). Do Students Learn More from Erroneous Code? Exploring Student Performance and Satisfaction in an Error-Free Versus an Error-full SAS® Programming Environment. *Journal of Statistics and Data Science Education*, *0*(0), 1–13. https://doi.org/10.1080/26939169.2021.1967229

McAleer, P., Stack, N., Woods, H., DeBruine, L., Paterson, H., Nordmann, E., Kuepper-Tetzel, C. E., & Barr, D. J. (2022). *Embedding Data Skills in Research Methods Education: Preparing Students for Reproducible Research*. PsyArXiv. https://doi.org/10.31234/osf.io/hq68s

Müller, K., & Wickham, H. (2022). *Tibble: Simple data frames*. https://CRAN.R-project.org/package=tibble

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Simonsohn, U. (2015). Small Telescopes: Detectability and the Evaluation of Replication Results. *Psychological Science*, *26*(5), 559–569. https://doi.org/10.1177/0956797614567341

TARG Meta-Research Group. (2022). Statistics Education in Undergraduate Psychology: A Survey of UK Curricula. *Collabra: Psychology*, *8*(1), 38037. https://doi.org/10.1525/collabra.38037

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wickham, H. (2021). *Forcats: Tools for working with categorical variables (factors)*. https://CRAN.R-project.org/package=forcats

Wickham, H. (2022). *Stringr: Simple, consistent wrappers for common string operations*. https://CRAN.R-project.org/package=stringr

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., ... Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., François, R., Henry, L., & Müller, K. (2022). *Dplyr: A grammar of data manipulation*. https://CRAN.R-project.org/package=dplyr

Wickham, H., & Girlich, M. (2022). *Tidyr: Tidy messy data*. https://CRAN.R-project.org/package=tidyr

Wickham, H., Hester, J., & Bryan, J. (2022). *Readr: Read rectangular text data*. https://CRAN.R-project.org/package=readr