

Bayesian modeling of comparative judgment data with Stan in R: A tutorial for speech quality researchers

Jose Manuel Rivera Espejo^{a,*}, Tine Daal^a, Sven Maeyer^a, Steven Gillis^b

^a*University of Antwerp, Training and education sciences,*

^b*University of Antwerp, Linguistics,*

Abstract

Comparative judgment (CJ) data is often analyzed using the Bradley-Terry-Luce (BTL) model, which offers a straightforward method for measuring traits and conducting statistical inference. However, despite its usefulness, research has indicated that several core assumptions of the BTL model are rarely satisfied in modern CJ applications. As a result, the model may struggle to capture the complexity of some traits or stimuli, compromising the reliability and accuracy of trait estimates. Additionally, its requirement to separate trait measurement from hypothesis testing can further undermine the accuracy of statistical inferences drawn from such data.

To address these limitations, [Rivera et al.](#) proposed an approach that extends Thurstone's general form using causal and Bayesian inference methods. This approach allows for the construction of a model specifically tailored to the assumptions of the CJ data under study. Furthermore, it integrates measurement and hypothesis testing within a single analytical framework, facilitating precise and accurate inferences from the data.

This tutorial illustrates the application of the proposed approach to a simulated dataset on speech quality. It provides detailed guidance on data simulation, model specification, estimation, and interpretation using the software R and Stan. The tutorial assumes that researchers are familiar with causal and Bayesian inference methods, as well as latent variable models, but may not have prior experience with CJ data or the software. By following the procedures outlined here, researchers can reproduce the analysis and adapt the approach to other CJ studies.

Keywords: causal inference, directed acyclic graphs, structural causal models, bayesian statistical methods, thurstonian model, comparative judgement, probability, statistical modeling

*Corresponding author

Email addresses: JoseManuel.RiveraEspejo@uantwerpen.be (Jose Manuel Rivera Espejo), tine.vandaal@uantwerpen.be (Tine Daal), sven.demaeyer@uantwerpen.be (Sven Maeyer), steven.gillis@uantwerpen.be (Steven Gillis)

1. Introduction

2. Data & Methods

3. Discussion

3.1. Future research directions

3.2. Study limitations

4. Conclusion

Declarations

Funding: The Research Fund (BOF) of the University of Antwerp funded this project.

Financial interests: The authors declare no relevant financial interests.

Non-financial interests: The authors declare no relevant non-financial interests.

Ethics approval: The University of Antwerp Research Ethics Committee confirmed that this study does not require ethical approval.

Consent to participate: Not applicable

Consent for publication: All authors have read and approved the final version of the manuscript for publication.

Data availability: This study did not use any data.

Materials and code availability: A previous version of this manuscript, along with the associated materials and code (see the section titled **CODE LINK**), has been made publicly available at: https://jriverspejo.github.io/paper3_manuscript/.

AI-assisted technologies in the writing process: The authors used various AI-based language tools to refine phrasing, optimize wording, and enhance clarity and coherence throughout the manuscript. They take full responsibility for the final content of the publication.

CRedit authorship contribution statement: *Conceptualization:* J.M.R.E, T.vD., S.DM., and S.G.; *Methodology:* J.M.R.E, T.vD., and S.DM.; *Software:* J.M.R.E.; *Validation:* J.M.R.E.; *Formal Analysis:* J.M.R.E.; *Investigation:* J.M.R.E; *Resources:* T.vD. and S.DM.; *Data curation:* J.M.R.E.; *Writing - original draft:* J.M.R.E.; *Writing - review and editing:* J.M.R.E, T.vD., S.DM., and S.G.; *Visualization:* J.M.R.E.; *Supervision:* S.G. and S.DM.; *Project administration:* S.G. and S.DM.; *Funding acquisition:* S.G. and S.DM.

References

Rivera, J., van Daal, T., De Maeyer, S., Gillis, S., . Let's talk about thurstone & Co.: an information-theoretical model for comparative judgments, and its statistical translation. URL: https://jriverspejo.github.io/paper2_manuscript/.