Advanced modeling of comparative judgment data: Applications to speech quality

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 provides a straightforward method for measuring traits and conducting statistical inference. Despite its usefulness, research has shown that the BTL model can struggle to capture the complexity of some traits or stimuli, compromising the reliability and accuracy of trait estimates. Moreover, its requirement to separate trait measurement from hypothesis testing can further weaken the accuracy of statistical inferences. These limitations are especially evident in modern CJ applications, where several core assumptions of the model are rarely satisfied and robust statistical inference is often needed.

To address these shortcomings, Rivera et al. proposed an approach that extends Thurstone's general form using causal and Bayesian inference methods. For measurement, the approach combines Thurstone's core theoretical principles with key CJ assessment design features to build a model consistent with assumptions about the data-generating process. For inference, it integrates measurement and hypothesis-testing within a single analytical framework, allowing for precise and accurate inferences from CJ data.

This tutorial illustrates the application of the proposed approach to a simulated dataset on speech quality. It offers detailed guidance on model specification, estimation, and interpretation using the software `R` and `Stan`. The tutorial assumes familiarity with causal and Bayesian inference methods, as well as latent variable models. But does not require prior experience with CJ data or the software. By following the procedures here described, 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)

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: The `CODE LINK` section at the top of the digital document located at: https://jriveraespejo.github.io/paper3_manuscript/ provides access to all materials and code.

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., S.G.; *Writing - original draft:* J.M.R.E.; *Writing - review and editing:* 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://jriveraespejo.github.io/paper2_manuscript/.