Contribution ID: **155** Type: **Oral Presentation**

Empirical Evaluation of Psychometric Properties in Forced-Choice Tests: Comparing Binary and Graded Preference Formats Across Test Designs

Wednesday 23 July 2025 10:00 (15 minutes)

Oral presentation

Empirical Evaluation of Psychometric Properties in Forced-Choice Tests: Comparing Binary and Graded Preference Formats Across Test Designs

Author

Diego F. Graña

Affiliation

Universidad Autónoma de Madrid

Abstract

Graded preference (polytomous forced-choice) tests have been proposed as a solution to the lower reliability observed in binary forced-choice tests compared to traditional Likert-format assessments. This format retains key advantages, such as larger robustness to social desirability, while increasing the available information by expanding the number of options. However, despite favorable evidence from simulation studies, empirical validation with real data remains scarce. This is particularly relevant because certain response patterns (e.g., if respondents tend to make binary decisions despite the graded forced-choice format), may undermine the expected benefits of graded preferences. Moreover, it remains unclear whether the format change affects test blocks differently depending on whether they are homopolar (equally-keyed) or heteropolar (unequally-keyed). Accordingly, the present study compares graded and binary forced-choice formats against the Likert format under two forced-choice test designs: one using only homopolar blocks and another incorporating 30% heteropolar blocks. Empirical data are analyzed using a Thurstonian Item Response Theory model. Forced-choice tests were developed from a Likert-based personality item bank rated for social desirability, with optimal pairing to maximize expected reliability. We present findings on reliability (marginal and conditional), structural validity (factor loading patterns and model fit), convergent validity, discriminant validity, and criterion-related validity. Results indicate that the graded preference format outperforms the binary format in reliability, particularly in tests composed solely of homopolar blocks. In terms of validity, higher correlations with established Big Five inventories and criterion variables emerged for the graded preference format compared to binary forced choice. In conclusion, the graded preference format enhances the reliability and validity of forced-choice questionnaires, especially when using only homopolar blocks. Given that homopolar blocks naturally facilitate better matching for social desirability, adopting the graded preference format becomes particularly advantageous. Finally, we provide general recommendations for constructing forced choice and graded preference tests.

Keywords

Graded-Preference, Forced-Choice, Item-Response-Theory, Reliability, Ipsativity

Primary authors: Mr GRAÑA, Diego F. (Universidad Autónoma de Madrid); Dr ABAD, Francisco J. (Universidad Autónoma de Madrid); Dr SORREL, Miguel A. (Universidad Autónoma de Madrid); Dr KREITCHMANN, Rodrigo S. (Universidad Nacional de Educación a Distancia)

Presenters: Mr GRAÑA, Diego F. (Universidad Autónoma de Madrid); Dr ABAD, Francisco J. (Universidad Autónoma de Madrid); Dr SORREL, Miguel A. (Universidad Autónoma de Madrid); Dr KREITCHMANN, Rodrigo S. (Universidad Nacional de Educación a Distancia)

Session Classification: Session 8 : "Psychometric evaluation in forced-choice tests"

Track Classification: Measurement: Measurement