

Detecting careless responding in ipsative data

Friday 25 July 2025 09:15 (15 minutes)

Background. To prevent response styles associated with the use of rating scales, test items may be presented in so-called ipsative (or relative to self) formats including popular ‘forced choice’, and also ‘graded preferences’ or ‘proportions-of-total’. Like any other questionnaires, ipsative questionnaires can be subject to careless responding when respondents are not sufficiently motivated to give their full attention to the questions. However, detecting such responding can be more challenging than when using Likert scales because ipsative response formats usually involve comparisons between items measuring different traits and their modelling is inherently multidimensional. Moreover, the comparative nature of ipsative responses makes challenging the use of a method factor (latent variable) to control careless responding.

Objectives. This presentation will describe and evaluate two alternative strategies for dealing with careless responses in ipsative data: (1) identifying (and ultimately removing from the sample) careless responders using “person fit” indices designed for ipsative formats; and (2) controlling for careless responding using method factors specifically designed for Thurstonian IRT and factor models (Brown & Maydeu-Olivares, 2012).

Methods. The two approaches are illustrated on a sample of $N=1,338$ volunteers who participated in a trial of an assessment measuring 24 non-cognitive skills with 276 multidimensional graded response pairs. Under Approach 1, two “person fit” indices were computed for each respondent. The first index summarized the discrepancies between a person’s observed responses and responses expected under the fitted Thurstonian measurement model, thus resembling the lco index (Ferrando, 2010). The second index summarized the concordance between a person’s observed and expected responses by computing a correlation coefficient between them. Under Approach 2, a random intercept was added to the Thurstonian measurement model to control carelessness expressed as overusing one rating scale category.

Results and Conclusions. The concordance index had a median of 0.532 and a long left tail, identifying at least 10% of aberrant responders. The discrepancy index had a median of 0.770 and a long right tail, again identifying at least 10% of aberrant responders. The Thurstonian model with the random intercept factor fitted better than the baseline model (SRMR were .055 and .068, respectively), and the random intercept explained between 1% and 2% in the variances of observed responses. However, at the individual level the discrepancy, concordance and random intercept agreed only for careful responders. For careless responders, each index provided unique information about the nature of carelessness. We conclude with recommendations for the use of such indices in practice.

Primary author: BROWN, Anna

Presenter: BROWN, Anna

Session Classification: Symposium : ”Understanding, detecting and managing careless responding in survey research.”