Multiscale Measurement of Extreme Response Style

Bolt D. M. and Newton J. R.

This study extended a model-based approach for the study and control of Extreme Response Style (ERS) in rating data. The extension mainly consider the multiple scale structure of data where the correlations between the latent traits exist. The results show that the simultaneous analysis of item response across the scales could improve the accuracy of estimation of ERS and thus improve the estimation of latent traits.

Questions:

1. In simulation, by comparing the bias from M1 (one substantive trait and one ERS trait), M2(two latent traits, no ERS trait) versus M3 (true model; two substantive traits, one ERS trait), it was found that the bias is larger when M2 is used than that when M1 is used. It means that ignoring ERS trait (M2) has more serious effect than ignoring the added scale (M1). Thus, it was concluded that the improvement of estimation of latent traits depend more on the ERS than the added scale.

I think, the conclusion is too arbitrary. The bias is larger when M2 is used than that when M1 is used because the true model is M3 in which ERS is included. Suppose the true model is M2. We use M3 and M1a (one latent trait, no ERS trait) to analysis the simulated data. If M3 will have smaller bias in the estimation of 1 compared to the estimation of M1a (it is very possible). In this case, shall we conclude that the improvement due to the added scale is bigger than the ERS?

1. Whether ERS in the proposed model represent extreme response style exactly or, actually, a propensity that measures all kind of unusual response style (such as extreme responding, midpoint responding, etc)? If ERS truly only measures extreme response style, how to model other response style? Is it possible to add another latent trait such as MID in the model? But I have no idea about the difference between the extreme responding and midpoint responding in the psychological mean.
2. How to discriminate the extreme response because of the extreme high or low capability and because of a propensity?
3. The percentage of extreme response will affect the estimation of ERS. It is not surprising that the recovery of ERS is poor.
4. For the results in Table 1, when the empirical data which came from two subscales, it was found that three-dimensional NRM is superior according to the different information criteria. Since the model is the most complicate one among the three models, the information criteria is certainly lower than other two models. However, it was not necessary to indicate that the model is the best model. Therefore, if a forth model in which four dimensions (three latent traits, one ERS) is introduced and compared with the three-dimensional NRM and found that the former one is not superior than the latter one, then the conclusion that the three-dimensional NRM is the best one could be drawn.