Technical Details of Mokken Scale Analysis (MSA)

Dimensionality Assessment

In MSA, scalability coefficients (Loevinger's *H*) are used to form unidimensional scales and describe the strength of the scales (Loevinger 1947). There are three types of scalability coefficients: H_{ij}, which is the normed covariance between items *i* and *j* and is expressed as the ratio of the observed covariance and maximum possible covariance; H_i, which represents how well item *i* separates subjects relative to other items and the distribution of the latent ability; and *H*, which represents how well the scale can order subjects with respect to the latent ability. In a monotone homogeneity model (MHM), all of the H_{ij}, H_i, and *H* coefficients must range from 0 to 1 (Hemker et al. 1995). Higher values of *H* mean that ordering of subjects on the scale by total score is a more accurate reflection of the ordering on the latent variable.

The automated item selection procedure (AISP) is an algorithm available in MSA that begins by selecting the two items that have the largest value of H_{ij} that is significantly different from 0 (Mokken 1971). Then, items that have positive covariance with the items that have already been selected (also with H_i that is significantly different from 0) are added one by one until no items remain that have H_i larger than a prespecified cutoff value. Next, the algorithm will try to form a second scale from any remaining items, and so on. The algorithm stops when no more items are scalable, potentially leaving some items unselected. The genetic algorithm (GA) improves on the AISP by considering all possible groupings of subsets of the items, or partitions (Straat et al. 2013). In the AISP, an item could initially be selected that no longer meets the scaling criteria after more items are selected for inclusion on the scale. By considering all possible partitions, the GA confirms that the items selected satisfy the scaling conditions even after including additional items. The result of the GA, and the AISP, is that unidimensional groups of items are formed.

Invariant Item Ordering (IIO)

MSA can provide evidence to support another useful result, which is related to the second model originally defined by Mokken (1971), the double monotonicity model (DMM). The DMM shares the three assumptions of the MHM and also requires a fourth assumption of non-intersection of the item step response functions (ISRFs; Sijtsma et al. 2011). Let X_i be the score on item i with values $x_i = 0, ..., m$, where $m + 1$ is the number of response categories. The ISRF gives the probability of obtaining an item score of at least x_i for a person with a given level of the latent variable. For polytomous items, the item response function (IRF) is the expected value of X_i as a weighted sum of the ISRFs. When the items are dichotomous, the ISRF is equivalent to the IRF.

For dichotomous items, the DMM establishes invariant item ordering (IIO), which means that the order of items by popularity is the same across different levels of the latent ability, or that the items can be ordered by their mean score (Sijtsma et al. 2011). For polytomous items, the DMM itself does not establish IIO and the manifest IIO (MIIO) method should be used to assess whether the items can be ordered (Ligtvoet et al. 2010).

REFERENCES

Hemker, B. T., Sijtsma, K., Molenaar, I. W. (1995). Selection of unidimensional scales from a multidimensional item bank in the polytomous Mokken IRT model. *Appl Psychol Meas, 19*, 337-352.

Ligtvoet, R., van der Ark, L. A., te Marvelde, J. M., et al. (2010). Investigating an invariant item ordering for polytomously scored items. *Educ Psychol Meas, 70*, 578-595.

Loevinger, J. (1947). A systematic approach to the construction and evaluation of tests of ability. *Psychol Monogr, 61*.

Mokken, R. J. (1971). *A theory and procedure of scale analysis: With applications in political research*. Berlin, Germany: De Gruyter Mouton.

Sijtsma, K., Meijer, R. R., van der Ark, L. A. (2011). Mokken scale analysis as time goes by: An update for scaling practitioners. *Pers Individ Dif, 50*, 31-37.

Straat, J. H., van der Ark, L. A., Sijtsma, K. (2013). Comparing optimization algorithms for item selection in Mokken scale analysis. *J Classif, 30*, 75-99.