Reviewer’s report

Title: Mokken Scale Analysis of Health and Well-being Questionnaire item responses: a Non-Parametric IRT Method in Empirical Research for Applied Health Researchers

Version: 2 Date: 20 December 2011

Reviewer: Per Bech

Reviewer’s report:

My issue was to evaluate how the authors were guiding the readers of the journal on the use of the Mokken test in a psychometric analysis. My main concern on this background was the problem with example 1, i.e. the GHQ-12. From a clinical point of view (clinical validity), the GHQ-12 contains two subscales, namely the GHQ6 distress subscale with negatively worded items (symptoms of anxiety or depression) and the GHQ6 well-being scale with the positively worded items (positive well-being). In their response to this situation the authors now say that they are not aware of a single study in this respect. However, the authors should take the study in JAD 2009;119:43-51 into consideration. In this paper the Mokken analysis has been carried out on the two GHQ subscales and the consequence of using both scales in a placebo-controlled study with clomipramine in patients with chronic idiopathic pain disorder is illustrated.

The other part of the title about the Mokken analysis ,”empirical research with health and well-being questionnaires” was also a platform for my concern about using the GHQ-12 as an example for the readers of the journal.

The total score of the GHQ-12 is not a sufficient statistic. Factor analysis is not a method for identifying dimensions (total score being a sufficient statistic), but a method for identifying structures.

The authors certainly need to revise their manuscript taking these issues into account.

I would especially emphasize the need to correct the fatal error that the Mokken analysis can be used as a kind of factor analysis. Time has come to be much more concerned about the nature of item response theory models when compared to factor analysis!