

Multilevel multidimensional item response modeling for complex survey data with application to large-scale educational assessments

Li Cai, Departments of Education and Psychology, CSE/CRESST, UCLA

The importance of large-scale educational surveys continues to rise in educational research. Surveys such as the National Assessment of Educational Progress (NAEP) or Program for International Student Assessment (PISA) typically employ complex survey sampling (of individual students and schools) in conjunction with a randomized incomplete blocks item sampling design to maximize efficiency of data collection. The survey instruments also tend to exhibit multidimensionality, requiring the use of multidimensional item response theory modeling. The combination of multidimensionality with multilevel data arising from complex survey design creates substantial challenges for the underlying item response theory models used to scale the outcome variables of interest. This research focuses on a particular topic, differential item functioning (DIF) testing under multidimensionality with complex survey data. Extending a Wald-test based DIF procedure proposed by Cai, Yang, & Hansen (2011) for simple random samples, a complex sample adjusted Wald DIF test is proposed based on pseudo maximum likelihood estimation with covariance matrix of item parameters produced by Taylor series linearization. The proposed procedure is illustrated with analysis of data from PISA.

Cai, L., Yang, J. S. & Hansen, M. (2011). Generalized full-information item bifactor analysis. *Psychological Methods*, 16, 221–248.