A facet designed test for scientific literacy

Carstensen, Claus H.
IPN Leibniz Institute for Science Education,
Olshausenstr. 62 24098 kiel, Germany,
E-Mail: carstensen@ipn.uni-kiel.de

Keywords: Test development, Item Response Theory; Multidimensional Latent Variables, Large Scale Assessment

Abstract

Germany is participating in the OECD Programme for International Student Assessment PISA. In addition to the instruments used in PISA, national test instruments being closer related to the german curricula were developed and administered. The german PISA science test is designed multidimensionally with each task corresponding to two dimensions at the same time: A topic (from biology, chemistry and physics) and a cognitive component (from retrieving information from graphs, drawing conclusions, applying a mental model, verbalizing a situation, calculating, creative thinking and to form an opinion about something). Each task is assigned to exactly two dimensions, each of which is assigned either to the topic facet or the cognitive component facet. The test may not be split up into one-dimensional subtests and thus a multidimensional model is needed to estimate subject abilities.

An appropriate multidimensional IRT model will presented. It is a generalized Rasch Model defining the probability of a subject *v* with abilities θ_{vk} and θ_{vl} giving response *x* at item *i* by

$$p(X = x|\boldsymbol{\theta}, \boldsymbol{\sigma}) = \exp\left(x\theta_{vk} + x\theta_{vl} + \sigma_{ix}\right)\Big/ d_{vix}$$

with d_{vix} being a norming constant. The abilities θ_{vk} and θ_{vl} correspond to the topics (θ_{vk}) respectively the cognitive components (θ_{vl}). The Model is a submodel of MULTIRA (Rost and Carstensen, Carstensen and Rost), the test model of ConQuest (Wu, Adams and Wilson) or more general item component models.

As the ten dimensions of the test are not defined independently, the ability parameters can not be estimated for all of them independently. But abilities estimates for every dimension are requested to report results of the study: A constraint to obtain estimates for all ten dimensions will be discussed.

Another issue in Large scale Assessment is multi matrix sampling consequence of which there is data missing by design. The data of the german science test from the PISA Field Trial in 2002 was analysed using software that handles missing data (MULTIRA and ConQuest). To conclude the presentation, results on the dimensional

structure of the tests will be reported and the reliability of the test's dimensions will be addressed.

References

Carstensen, C. H., Rost, J. (2002). MULTIRA - A program system for multidimensional Rasch models, Computerprogram (Version 1.64). Kiel: IPN - Institute for Science Education.

Rost, J., & Carstensen, C. H. (2002). Multidimensional Rasch Measurement via Item Component Models and Faceted Designs. *Applied Psychological Measurement, 26*(1), 42 - 56.

Wu, M. L., Adams, R. J., & Wilson, M. (1997). *ConQuest - Generalised Item Response Modelling Software, Draft Release 2*. Camberwell: Australian Council for Educational Research