A quantitative enquiry into the fairness and equity in Norwegian secondary school assessment practices: Project proposal

Tony C. A. Tan

Centre for Educational Measurement

University of Oslo

Attachment 3 to UV PhD application form

Prof Rolf V. Olsen & Dr Astrid M. J. Sandsør

27 August 2021

Introduction

Purpose Statement

Research Questions

Methods

Sample

For this study, students' GPA records will be captured from the Norwegian registry covering the decade between 2010 and 2020. GDPR registration is lodged through the NSD Portal and the UiO ethics approval is also obtained. All data import, storage, and analyses are to be conducted within the secured infrastructure TSD provided by the UiO Central IT Division. TSD logs all activities and no data or results can be copied out of the restricted system without prior approval from project leaders.

Under the advisory of He and Stockford (2015), subjects with fewer than 1,000 candidates and students taking fewer than two GPA subjects will be excluded from subsequent analyses. Each year's record (score matrix) will contain N rows representing the number of valid candidates and L columns reflecting the usable number of GPA subjects in that year. Since no student took all the GPA subjects, a large proportion of the score matrices will remain missing by design. The existence of missing data does not pose any problems for using the Rasch model as the model functions at the individual item and items as long as there is sufficient overlap across subjects in the score matrix. The ability to deal with incomplete data is one major advantage of using the Rasch model for studying inter-subject comparability.

Rasch Model

The Rasch model was developed in the 1960s for establishing measurement scales and for improving test development (Rasch, 1980). In a simple Rasch model, the underlying ability or latent trait of the person (θ) and the item characteristics (δ_j) are specified; a logistic function (Λ) is then used to describe the probability that the person will successfully pass a subject ($x_j = 1$) given their ability θ and the item characteristics δ_j (de Ayala, 2009):

$$\mathbb{P}(x_j = 1|\theta, \delta_j) = \Lambda(\theta - \delta_j) = \frac{1}{1 + e^{-(\theta - \delta_j)}}. \tag{1}$$

Equation (1) is suitable for modelling subjects with dichotomous (pass/fail) outcomes. GPAs on the other hand are polytomous in nature (§3-5, *Forskrift til opplæringslova*), therefore requires models capable of accommodating more than two achievement outcomes. Multiple extensions have been put forward over the decades such as rating scale models (Rasch, 1980), partial credit models (Masters, 1982), and generalised partial credit models (Muraki, 1992). Master's (1982) partial credit models (PCM) are particularly parsimonious and flexible for studying the structure of GPA data. A PCM states that, for a subject with $m + 1$ available grades ($m = 5$ for Norway's GPA system), the probability of a candidate with ability θ receiving grade x can be expressed as:

$$\mathbb{P}(x|\theta) = \begin{cases} \dfrac{1}{1 + \sum_{l=1}^{m} \exp\left\{\sum_{k=1}^{l}(\theta - \delta_k)\right\}} & \text{for } x_j = 0 \\ \dfrac{\exp\left\{\sum_{k=1}^{x}(\theta - \delta_k)\right\}}{1 + \sum_{l=1}^{m} \exp\left\{\sum_{k=1}^{l}(\theta - \delta_k)\right\}} & \text{for } x_j = 1, 2, \ldots, m \end{cases} \tag{2}$$

where δ_k is the location of the k-th step on the latent trait continuum and is referred to as the item step parameter associated with a grade category (also referred to as step difficulty or threshold). Model parameters in Equation (2) can be solved using joint maximum likelihood estimation or the conditional maximum likelihood estimation.

Data Analysis

Hypothesised Results

References

de Ayala, R. J. (2009). *The theory and practice of item response theory*. Guilford.

He, Q., & Stockford, I. (2015). *Inter-subject comparability of exam standards in GCSE and A Level* (ISC Working Paper 3). Office of Qualifications and Examinations Regulation.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, *47*(2), 149–174. https://doi.org/10.1007/BF02296272

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *ETS Research Report Series*, *1992*(1), 1–30. https://doi.org/10.1002/j.2333-8504.1992.tb01436.x

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. University of Chicago Press.