AN INVESTIGATION OF ASPECTS OF A MATHEMATICS LITERACY PRELIMINARY PAPER THROUGH AN APPLICATION OF RASCH MEASUREMENT THEORY

[1]Caroline Long, [2]Sarah Bansilal and [3]Rajan Debba

[1]University of Pretoria, [2,3]University of KwaZulu-Natal

Extended abstract

Mathematical Literacy is a relatively new school subject that is examined in the final grade in South African schools. An investigation of the 2009 preliminary examination written by matric pupils in KwaZulu-Natal (KZN) was conducted by Debba (2012). In this subsequent study we supplement the qualitative investigation with an application of Rasch measurement theory (RMT) in order to investigate the validity and reliability of current mathematical literacy testing practices as exemplified in this test. The test data from the subgroup of 73 learners from one school were analysed. This very small sample is used to illustrate the principles of measurement.

An assumption underlying RMT is that underpinning any assessment procedure, and any coherent testing region, there exists a somewhat unidimensional construct (in this case competence in mathematical literacy) that may be gauged along a continuum indicating more or less proficiency in the construct (in the case of learners) and more or less difficulty in the construct (in the case of items). The items are understood here to be operationalisations of the construct. In the ideal test each of the items contributes some unique component to the measurement-like properties of the test, collectively establishing a plausible continuum of proficiency avoidance of undue repetition of one or other aspect of the construct.

In an application of the Rasch model, checks are made on the test as a whole, the items, and the learner responses, to ensure fairness of the instrument within the particular reference group, in this case a group of mathematical literacy students in a KZN high school.

In this paper we consider the implications for improving the construction of rubrics so that they work coherently in the interests of attaining valid measurement-like properties and consequently provide reliable information for teachers and a fair assessment for learners. In so doing we hope to add some methods for providing insight into the domain of ML, and some observations about the assessment of ML.

We focus on the scoring of polytomous items, which may be allocated scores of more than 1, for example 0, 1, 2 … m. For some items it was found that the scoring allocation unduly advantaged some learners and disadvantaged others, for example a score of 4 may be allocated where the empirical responses are either correct or

Long, C., Bansilal, S. & Debba, R. (2013). An investigation of aspects of a Mathematics Literacy preliminary paper through an application of Rasch measurement theory. (Extended abstract) In Z. Davis & S. Jaffer (Eds.), *Proceedings of the 19th Annual Congress of the Association for Mathematics Education of South Africa*, Vol. 1. (pp. 100 – 101). Cape Town: AMESA.

incorrect. This allocation means that a learner is substantially disadvantaged by "losing" 4 marks, as opposed to the usual loss of 1 mark, for responding to the item (a single component) incorrectly. Through the qualitative investigation of the item and the associated scoring logic, and the results of the Rasch analysis, rescoring was investigated. We report here on the overall result prior to rescoring, the analysis and rescoring of the items, and the post rescore analysis. The purpose of the paper is firstly to illustrate the importance of appropriate scoring in the interests of validity (fitness for purpose) and fairness (relative performance extremes), and secondly to show how an application of the Rasch model enables one to check validity of assessments and to verify adjustments where required.

In essence we answer the question, "How may detailed attention to the scoring of the items in a mathematics literacy test through theoretical investigation and the application of the Rasch model contribute to a more informative and fair outcome?"

The process described in this paper, involved identifying items which did not fit the model, and therefore do not exhibit measurement-like principles. We describe the items and the original scoring rubric. The Rasch output is provided and the anomalies and inconsistencies, indicative of inadequate data, that emerged in the analysis are discussed.

Before initiating any changes in the scoring we looked for theoretical verification that the rescoring made sense. The process that was conducted involved systematically rescoring each of the assessment items. This process required input from the mathematics literacy specialist, in this case the mathematically proficient and experienced teacher. Thus the Rasch analysis and rescoring processes were guided at all times by the qualitative analysis that was performed by the ML specialist teacher.

The process has revealed the important role of scoring rubrics when designing valid assessments. Examiners need to ensure that each mark that is allocated can be justified educationally as an indication of some future ability. Marks should not be allocated on the basis of time that is required to be spent during the examination. This post-hoc analysis of the questions and the scoring rubrics highlights the importance of ensuring that the items are carefully constructed and verified by specialist teachers with due attention to the appropriate allocation of scores.

Without the assurance of measurement-like properties, then the total score is not statistically robust. If the total score is to be indicative of a position on a unidimensional scale, then the differences between the score must reflect differences between proficiency levels of learners, and the total score must be a valid representation of relative performance.

REFERENCE

Debba, R. (2012). An exploration of the strategies used by grade 12 mathematical literacy learners when answering mathematical literacy examination questions based on a variety of real-life contexts. Unpublished M.Ed. dissertation, University of KwaZulu-Natal.