

Psychometric Validation and Mastery Profiling of a Grade 7 Mathematics Achievement Test: Integrating Classical Test Theory and Rasch Analysis Under the MATATAG Curriculum

Student: Von Lorenz A. Chavez

Adviser: Professor Su-Wei Lin

Doctor of Philosophy in Measurement and Statistics,
Department of Education, National University of Tainan

ABSTRACT

This study aimed to develop and validate the MAT-G7, a Grade 7 Mathematics Achievement Test aligned with the Philippine Department of Education's MATATAG Curriculum, employing both Classical Test Theory (CTT) and Item Response Theory (IRT). The research addressed the urgent need for psychometrically sound, curriculum-based assessment tools to support diagnostic, formative, and summative evaluation of learner performance in mathematics within the Philippine basic education system.

Following an instrumentation-based design, the study was implemented in two phases. The first phase involved a pilot tryout with 703 Grade 7 students from four public secondary schools and two private secondary schools using CTT to initially evaluate item characteristics such as difficulty index, discrimination index, and reliability. The second phase comprised the final test administration involving 324 Grade 7 learners from four additional public schools. This phase utilized both CTT and IRT (1PL-Rasch Model) to comprehensively evaluate item and test functioning, establish mastery thresholds, and assess reliability and validity evidence.

The MAT-G7 is comprised of 34 multiple-choice items constructed based on a validated table of specifications grounded on MATATAG content standards. After expert panel validation using the Content Validity Ratio (CVR), iterative item analysis, and psychometric screening, several items of the 34 items were revised and paraphrased to increase their psychometric

properties for the final version of the MAT-G7. CTT analyses confirmed the test's high internal consistency ($KR-20 = 0.861$), acceptable item difficulty range ($P = 0.23$ to 0.79), and strong discrimination indices ($D = 0.34$ to 0.62 and $r_{PB} = 0.25$ to 0.53). The IRT analysis confirmed that all 34 items exhibited acceptable model-data fit, with mean square (MNSQ) values ranging from 0.77 to 1.22 (infit) and 0.89 to 1.16 (outfit) – well within the recommended 0.80 to 1.30 range for Rasch-conforming items. The person and item reliability indices were 0.853 and 0.868 , respectively, demonstrating strong measurement precision across a broad spectrum of learner abilities.

The IRT results also uncovered mastery thresholds based on person ability (θ) values: Foundational (-1.56 to -0.86), Developing/Core (-0.74 to -0.37), Proficient (-0.17 to 0.36), and Advanced (0.64 to 1.44). These thresholds established interpretive benchmarks that can guide instruction and learner support. Moreover, the analysis revealed six distinct content domains within the MAT-G7: Basic Geometry, Angle Concepts, Polygon Concepts, Number Sense, Financial Literacy, and Measurement. These subdomains exhibited unique psychometric profiles and item difficulty distributions. Notably, item difficulties were skewed toward mid-level ability, with fewer items measuring extremely low or high proficiency levels. This finding underscores the need to expand item coverage at both extremes to improve test precision across the full ability spectrum.

The psychometric properties, structure, and empirical rigor of the MAT-G7 support its use as a valid, reliable, and fair assessment instrument aligned with the MATATAG Curriculum. It holds substantial potential as a diagnostic tool to inform instruction, assess learner progress, and support curriculum-based educational decision-making in Philippine schools. The study contributes to the body of literature on localized assessment development, highlighting the

relevance of integrating both CTT and IRT approaches in constructing robust achievement tests grounded in curriculum standards.

Keywords: *MAT-G7, mathematics achievement test, MATATAG Curriculum, Classical Test Theory, Item Response Theory, Rasch Model, psychometric validation, Philippine education*