Analyzing Mathematics 10 First Quarter Examination and Students' Test Taking Skill Through the Rasch Model: A Comprehensive Study

Jay A. Balasabas[1], Alvic A. Arnado[2]

[1]Bayugan National Comprehensive High School, Philippines

[2] Caraga State University, Philippines

jaybalasabas95@gmail.com

Abstract: The study aimed to facilitate a comprehensive understanding of how well the Grade 10 First Quarter Mathematics Examination aligns with students' abilities, allowing insights into item difficulty in the subject matter. This test was administered to 51 Grade 10 students from a regular class, 23 males and 28 females, at Bayugan National Comprehensive High School. The results show that two items are considered very easy for the students, while the other two are considered very difficult for their abilities. It also indicates that seventeen students were misfits (outliers), which could be interpreted as just guessing in answering the test items. At the same time, some of them could have misunderstood the items. Lastly, some of these students may have test anxiety or stress. The researcher listed some recommendations that the teacher may promote a positive attitude and belief to overcome test anxiety and build self-efficacy. The teacher could also equip students with test-taking strategies like eliminating incorrect answers and estimating answers. The teacher may review the individual response patterns of the students and look for inconsistencies or unexpected responses. Individual interviews with students are also encouraged to understand possible causes of underperformance.

Keywords: item characteristics, mathematics, person characteristics, Rasch model, test assessment

INTRODUCTION

Mathematics is connected with everyday problems and using imagination, intuition, and reasoning to find new ideas and solve puzzling problems (Khan, 2015). In the study conducted by Al Ali and Rehab (2020), Al Ali stated that, according to the National Center for Assessment in Higher Education, there is a significant weakness in the basic mathematical principles of most students. Many students fear and worry about mathematics and consider it complex, abstract, and uncreative

(Mutodi & Ngirande, 2014). According to Ngunjiri (2022), students need to know how they do as they learn. This is because knowledge of current understanding gives students a source of awareness of their achievement, which may motivate them to learn more. Thus, mathematics teachers must assess students' learning and give them immediate feedback (Anderson, 1993). Furthermore, assessments should be given to students as frequently as possible using written and oral tests, observation, and projects where possible with feedback given accordingly (Gore, 2000).

According to Isnaini et al. (2019), student learning assessment activities are among the essential tasks teachers must perform. In education, an assessment of student learning achievements is conducted to determine students' progress in the curriculum that has been taught. One effort to assess students is to give examinations. However, sometimes, giving questions that are too difficult or too easy makes it easier for lecturers to distinguish students' abilities. Therefore, an analysis of exam questions is needed in the hope that the exam results present students' abilities.

Assessment in education is a fundamental process of educational activities. With that, it is easier to know whether the progress of learning has been achieved. According to Sumintono (2017), instrument testing and determining students' abilities in educational assessments are fundamental. An analysis that can result in more precise measurements (produce an equal-interval scale) will determine the quality of the investigation results and the improvement of the educational process to help students learn. The Rasch model can help teachers assess and improve the quality of the analysis performed because it applies the appropriate basic principles of data processing. This is because the Rasch model addresses five objective measurement requirements.

Rasch model, according to Isaini et al. (2019), is a dichotomous scoring model that merely has two categories, namely the correct answer with a score of 1 and the incorrect answer with a score of 0. According to Al Ali and Shehab (2020), the Rasch model is a probabilistic unidimensional model that confirms that the more manageable the question, the more likely the student will respond correctly to it, and also, the more able the student, the more likely he will pass the question. In constructing tests using this model, one frequently discards those items that do not fit and that do not meet the assumptions of the Rasch model (Wright & Stone, 1979). Rasch analysis gives the reliability coefficient for each person and item (El-Korashy, 1995). A study by Kurniawan and Mardapi (2015) showed that the Rasch model provides complete information about test items, including their difficulty level.

A study by Ramadhani, Saragih, and Napitupulu (2022) demonstrated the practical use of the Rasch model in evaluating students' statistical reasoning abilities within real classroom contexts. Their research highlights how Rasch analysis provides detailed diagnostics on item difficulty and student performance, supporting its value in developing meaningful and adaptive instruction, a goal shared by the current study.

In the Philippines, teachers and lawmakers are making ways to improve the educational system through continued assessments. Magno (2010) highlights the diverse applications of educational assessment, measurement, and evaluation, leading to a field shaped by numerous influences. The passage of the Philippine Psychology Act has further clarified the roles of psychologists and teachers in school-based assessment. Current research suggests that the field is poised for continued growth and development in the future. Filipino researchers utilized a modern test theory approach with various model parameters, while other social science studies employed a simpler one-parameter Rasch model. Within the Philippines' educational assessment landscape, professionals are becoming increasingly vital due to the growing emphasis on quality assurance in schools, especially concerning instruction and program effectiveness. This necessitates stronger partnerships between teachers and psychometricians to create and interpret assessments that effectively measure student learning.

In the context of Grade 10 mathematics examinations, the Rasch model helps to assess how well the test matches the students' levels of ability and how suitable each test item is. Although the Rasch model is widely used in educational research internationally, there are few studies that apply it in secondary schools in the Philippines, especially in regular classes where student performance varies significantly. This limits the understanding of how Rasch-based analysis can support teaching and assessment in typical classrooms. Therefore, this study aims to provide teachers with useful information on the level of difficulty of each test item and the performance of students, which can help improve instructional planning, support student needs, and guide future assessment development.

Research Questions

This study explored how the Rasch model can assess the extent to which a standardized mathematics test reflects student ability and how effectively each test item functions. Specifically, the study addressed the following research questions:

1. How well does the Grade 10 Mathematics First Quarter Examination align with students' abilities based on Rasch model analysis?
2. What item characteristics—such as difficulty level and fit statistics—emerge from the Rasch model, and how do they inform the quality of the test?
3. How do students' response patterns (e.g., misfitting or guessing behavior) reflect their test-taking skills and affect the interpretation of performance data?
4. What instructional strategies can be developed to address issues such as test anxiety, motivation, and response inconsistency identified through the Rasch analysis?
5. How can the Rasch model be used to improve assessment design and teaching practices in regular Grade 10 mathematics classrooms?

METHODS

This study was descriptive and diagnostic in nature. No instructional interventions were administered during the research period. The Rasch model was employed to evaluate the psychometric properties of an existing classroom assessment. The results are intended to inform future revisions to test design and instructional planning.

Participants

The participants of the study are 51 Grade 10 students, selected through purposive sampling, from a regular class with 23 males (45.10%) and 28 females (54.90%) at Bayugan National Comprehensive High School. The students had varied learning styles and levels of mastery of the subject matter.

Instrument

The first quarter mathematics examination for grade 10 students at Bayugan National Comprehensive High School was developed by the researcher and aligned to a Table of Specifications that an expert and highly competent professor validated. The examination consisted of 50 multiple-choice items and was composed of 60% low-level thinking skill questions (remembering and understanding), 30% moderate-level thinking skill questions (applying and analyzing), and 10% high-level thinking skill questions (evaluating and creating). These levels were based on the Revised Bloom's Taxonomy of Cognitive Objectives by Anderson and Krathwohl (2001).

Procedure

The researcher developed and administered a first-quarter mathematics examination to collect the data for this study. The researcher acquired permission from the school principal, students, and parents to ensure the willingness and participation of the participants.

Analysis

The data collected were tabulated and treated according to the problems presented in this study. Data was cleaned through descriptive statistics by examining the valid N and missing values. Responses were coded as 1 if the student answered the item correctly and 0 if incorrect. The data were analyzed on Jamovi version 2.3.28 using its free module on snowIRT on the Dichotomous Rasch Model.

Rasch analysis was employed to assess the test's structural validity, item fit, person fit, and difficulty distribution. The primary statistics examined included:

- Infit and outfit mean square values are used to evaluate how well individual items and students align with model expectations.
- Person reliability index, which reflects the consistency of student responses across varying item difficulties.
- Item difficulty measures, expressed in logits, indicate the cognitive demand required for each item.
- MADaQ3 statistic, used to assess the overall model fit through residual correlations.

A Wright Map was generated to visually compare student ability against item difficulty. Additionally, response patterns were reviewed to identify students exhibiting high outfit values (potential guessing behavior) or low values (possible anxiety or disengagement). Misfitting items and students were flagged for further instructional analysis.

RESULTS

Model Fit

Model Fit is a measure of how well the model fits the data. A good model fit is crucial because it means the model accurately represents the relationships between the variables. Person Reliability measures how reliably the model can estimate the person's ability. A higher person reliability indicates that the model is more reliable. MADaQ3 is a measure of model fit, with lower values indicating a better fit. The range of values a MADaQ3 can take on can be between 0 and infinity. The MADaQ3 is calculated by taking the mean of the absolute values of the centered Q_3 statistic. The Q_3 statistic measures how well the person's responses fit the model. P-value is the probability of obtaining a MADaQ3 as extreme or more extreme than the observed MADaQ3, assuming that the model is correct. A p-value less than 0.05 is generally considered statistically significant, meaning that the model fit is unlikely to have occurred by chance.

	Person Reliability	**MADaQ3**	**p**
scale	0.552	**0.136**	**0.001**

Note. MADaQ3= Mean of absolute values of centered Q_3 statistic with p-value obtained by Holm adjustment; H_0= the data fit the Rasch model.

Table 1. Model fit

Table 1 shows the results of a model fit for a person reliability model. The model fit is good, with a person reliability of 0.552, demonstrating that the students have consistently correctly answered questions and a MADaQ3 of 0.136. The MADaQ3 measures model fit, with lower values indicating a better fit. The p-value for the MADaQ3 is 0.001, which indicates that the model fit is statistically significant.

Item Characteristics

Item characteristics are a property of an item in a test or assessment that describes how well the item can distinguish between individuals with different ability levels or traits being measured. A complete summary of item characteristics is presented in Appendix A. The proportion column shows the proportion of people who answered each item correctly. The measured column shows the estimated measure for each item, which measures the item's difficulty. The S.E. Measure column shows the standard error of the measure. The infit and outfit columns show the infit and outfit mean square errors, which measure how well each item fits the model. Sumintono and Widhiarso (2015) guide assessing these items into four categories, namely (1) Measure value < -1 = very easy item, (2) Measure value -1 to 0 = easy item, (3) Measure value 0 to 1 = difficult item, and (4) Measure value > 1 = very difficult item.

The result reflects the characteristics of each item in the first quarter examination in Mathematics 10. Twelve items are considered very difficult. Subsequently, it shows in Appendix A that items 39, 28, and 11 are some of the items that are considered very difficult, with 2%, 4%, and 10%, respectively, of the number of students who got the answer correctly. Seven items are considered difficult, with 31% (0.31) to 43% (0.43) of the students getting the answer correctly. Eleven out of 50 items are considered easy questions with measure coefficients ranging from -0.04 to -0.91 and a proportion value of 0.51 to 0.71 (or 51% to 71% of the students got the answer correctly, respectively. Lastly, there are 20 items interpreted as very easy for the student's level of subject knowledge.

Boone et al. (2014) suggest the following criterion to check whether an item is fit (item fit) or not fit (outlier or misfit). This criterion is the OUTFIT value. If the OUTFIT value is greater than 0.5 and less than 1.5, then these items are considered misfits since they undermine the Rasch model's assumptions of unidimensionality, local independence, and monotonicity. Hence, if the outfit value is greater than 1.5 and less than 0.5, the item is considerably unfit (outlier). The result shows that the items' outfit coefficient lies between the ranges 0.00 to 1.38, with only items 1 and 3 being considerably more outliers than any other items because of its coefficient of 0.00, which is out of the range of a fit item. With this, items 1 and 3 must be improved or changed. Items 1 and 3 on the test were too easy, as every student answered them correctly. These items covered common knowledge that was directly taught from the textbook and lecture notes, suggesting the teacher may have focused heavily on this material. Additionally, the questions may have lacked effective distractors, making the correct answers obvious. The teacher should carefully review the test questions, particularly items 1 and 3, to determine if they were too simple and direct, given that all students answered them correctly, to increase the difficulty.

Person Characteristics

According to Sumintono (2017), the Rasch model can measure an individual's ability more precisely. Based on the accuracy of the response given, the pattern will show individual tendency regarding how one performs solving test items.

The total score is the sum of the scores on all the items. It is a measure of the student's overall ability on the test. The measure is the student's ability estimate on the Rasch scale. It is a more precise measure of the student's ability than the total score because it considers the items' difficulty. The S.E. is the standard error of the measure. It is a measure of the uncertainty in the measure. A lower S.E. indicates a more precise measure. The infit measures how well the student's responses fit the Rasch model. A value close to 1 indicates that the student's responses fit the model well. A value above 1 indicates that the student's responses are more variable than expected, and a value below 1 indicates that the student's responses are less variable than expected. The misfit measures how well the student's responses fit the Rasch model when the student is removed from the analysis. In general, both infit and outfit values between 0.7 and 1.3 are acceptable. Values outside this range may indicate that the student is a misfit or an outlier.

The student participants in the results represent a wide range of abilities. The students with the lowest total scores are likely to struggle with the material, while those with the highest scores are likely to excel. The students with negative measurement values are likely to underperform on the assessment, while those with positive ones will likely overperform. In the result in Appendix B, it can be seen that student 28 has the lowest total score of 21, and students 8 and 13 have the highest total score of 37. Note that students 10, 7, and 30 have an outfit value greater than 1.3, with 1.47, 1.58, and 1.53, respectively, which could mean that these students may be just guessing when answering the test items. In addition, fourteen (14) students have an outfit value lower than 0.7, which could mean that students misunderstood the items and were not motivated to take the examination. These students may have been careless or speeding in answering the examination. They may not pay close attention to the test items, which leads to random or guessing. Apart from this, these misfit students may have experienced test anxiety or stress as students might second-guess themselves, which leads to guessing due to stress. Hence, the teacher may review the individual response patterns of the misfit students and look for inconsistency or unexpected responses. Interviewing the students individually may help the teacher understand why their students might have struggled with the exam.

Wright Map

The Wright map shows the difficulty of the items in a test, as well as the ability of the test-takers. The items are arranged in order of difficulty, from the most difficult at the top to the easiest at the bottom. The test-takers are arranged in order of ability, from the highest-ability test-takers at the top to the lowest at the bottom.

To find out the distribution of items and the ability of the student to respond to items in general, the person-item map can be seen in Figure 1. The left side is the distribution of the students' abilities, while the right side is the distribution of items. From the map, the easiest items are items 1 and 3, which are in the lowest position, while the most difficult item is item 39. The map allows us to see which students are likely to have successfully answered which items. Generally, the test questions are lower (easier) than the student's ability. In other words, the questions could be more challenging for test takers who generally have high abilities.

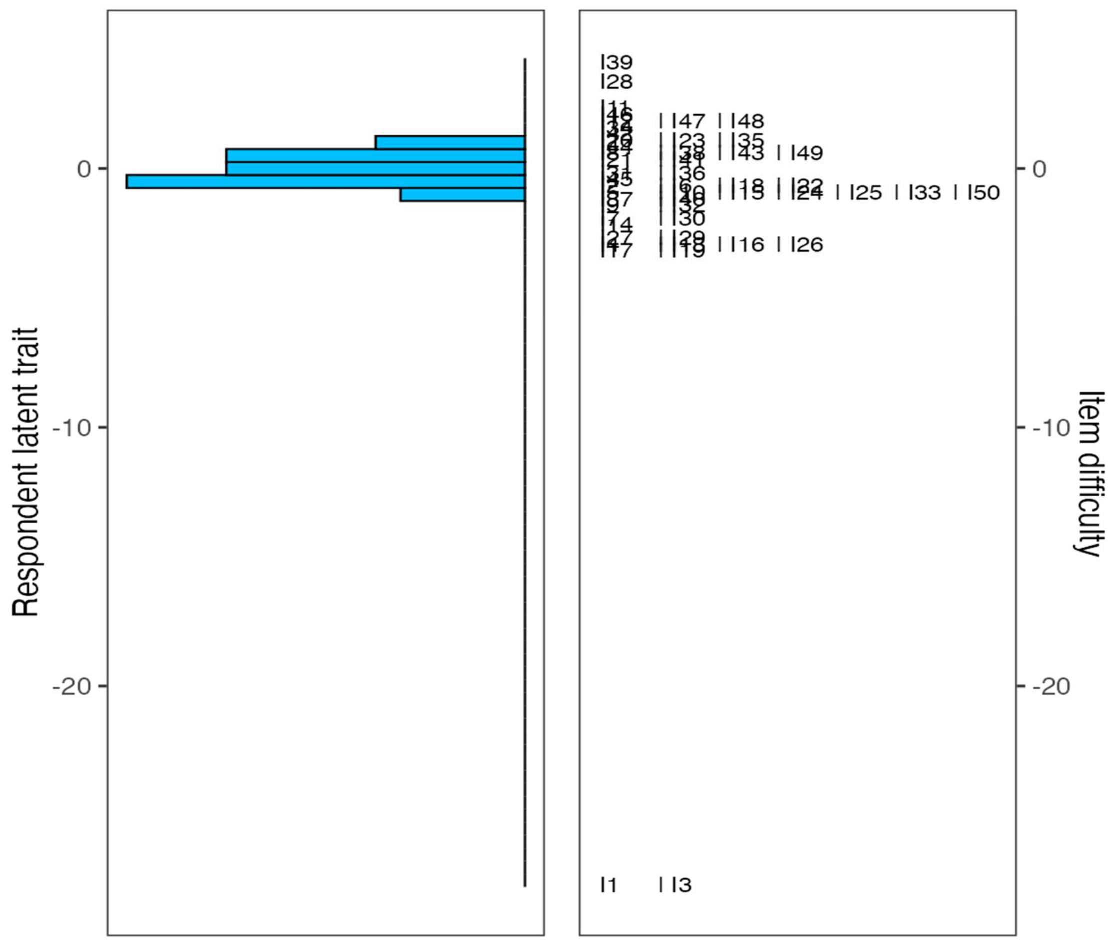

Figure 1. Wright map of students' latent trait and item difficulty

DISCUSSION

Based on the model fit results, the first quarter examination for Grade 10 in Mathematics to assess students' mathematics skills, gauge their understanding, identify areas for improvement, and enhance overall academic achievement represents an instrument that can measure the Grade 10 students' mathematics skills. Additionally, the result accurately represents the relationships between the variables, and the model fit is statistically significant. Based on the item characteristics (Appendix A) and the Wright Map (Figure 1), items 1 and 3 are very easy, and all students got both items correctly. This would mean that all students understand this item or subject matter, which needs to be higher than their abilities. Note that items 1 and 3 were made with low-level thinking skills, such as specifically remembering the facts and definitions of the subject matter. Subsequently, the students understand item 1 in finding the next term given a series of terms under the learners' learning competency, which generates a pattern. Also, the students understand item 3, which is remembering in the Revised Bloom's taxonomy of cognitive level of Anderson and Krathwohl (2001), which asks about the formula for finding the nth term of an arithmetic sequence. Some reasons why all students may have answered items 1 and 3 correctly are if the question was very basic or directly repeated information from the textbook or class notes. This highlights a potential flaw in the question, as it does not effectively differentiate between students who understand the material and those who do not. Alternatively, it could be that if the teacher discussed the specific answer or concept during class or made it very clear what the expected response was, students might all answer correctly for that reason.

A possible reason is that both items are general knowledge, given that remembering level and their easy calculations or straightforward applications. Maybe the construction of the test item has obvious distractors and its alignment with the learning objectives. This could also be that the teacher has built a strong foundation of essential knowledge and skills to ensure students have the necessary tools to tackle such questions. Moreover, items 1 and 3 must be improved or changed because they considerably outfit or misfit the students' abilities and assess their mathematics skills. On the characteristics of the students, it can be seen in Appendix B that three (3) students are merely guessing when answering the examinations, which may be because they need more preparation. Lacking preparation due to insufficient understanding of the concepts of the subject matter or procrastination and lack of time management skills are reasons why these results were seen. However, it could also be attributed to unclear or ambiguous questions, or they feel rushed to complete the exam since it was given quickly. In addition, fourteen (14) students needed to have understood the items and were not motivated to take the examination. This could be because the students have mathematics anxiety. Difficulty in mathematics is related to cognitive components of social perception, which are highly varied among students. Anxiety in mathematics is related to the emotional component of social perception, which varies slightly among students (Al Ali, 2016). It could also be because they need more confidence and believe they are not good at math. Other factors could be the lack of support at home or school, or personal issues like stress, anxiety, and other problems that can also impact their motivation and ability to learn.

Contribution to the Broader Literature

Although Rasch model analysis has been utilized in various international studies, its application in evaluating classroom-level assessments within the Philippine secondary school setting remains underexplored. This study contributes to the wider body of literature by showcasing how item-level diagnostics and student response patterns derived from Rasch analysis can inform both assessment quality and teaching strategies in actual classroom environments. The emphasis on identifying misfit items, understanding student behavior such as guessing and disengagement, and aligning instruction with ability profiles presents a practical framework for educators in similar contexts. This localized approach provides evidence that supports the development of responsive assessment practices not only in mathematics but also in other subjects and grade levels where student diversity and emotional factors affect performance.

LIMITATIONS

This study has several drawbacks. First, the most evident limitation was the tiny data set, which reduced the validity of the inferences derived from the data. The data analysis used in this study was based on a sample of data gathered only in one section of Bayugan National Comprehensive High School with a limited number of respondents (N = 51), which impacted generalization to other groups. Second, the study is limited to the characteristics of the participants. The school is located in the heart of Bayugan City and is one of the biggest schools, hence, the school has more varied students. The study participants were students in a regular class whose level of understanding of the mathematics content varied from one another. They were slower than how science classes learn. It is important to note that some items in higher-level thinking are considered very easy for the students, with 72.55% of them answering it correctly. This could be because the students understand more about this subject matter. Additionally, female participants outnumbered male participants in this research. It is possible to study gender bias with this instrument. The test questionnaire was made for grade 10 students taking Mathematics 10 for the first quarter, hence, it cannot be inferred for other quarters and other groups. Future studies should use bigger random samples from other groups of students and look for evidence of different backgrounds.

CONCLUSION

In conclusion, the model fit is good, as evidenced by a personal reliability of 0.552 and a MADaQ3 value of 0.136. The MADaQ3 serves as an indicator of model fit, with lower values signifying a more favorable fit. The p-value associated with the MADaQ3 is 0.001, indicating statistical significance in the model fit. Analysis of the first quarter Mathematics 10 examination reveals that among the 50 items, twelve are classified as very difficult, seven as difficult, eleven as easy, and thirty as very easy. The Wright map illustrates that items 1 and 3 are the easiest, located at the bottom, while item 39 is the most challenging. Both items 1 and 3 are outliers or misfits and are suggested to be improved or changed. Overall, the test questions are relatively easier than the students' abilities, suggesting that the questions may be too simple for test takers with generally high aptitudes. Thus, the items of the first-quarter mathematics examination for grade 10, developed by the researcher, are effective in assessing students' mathematics skills, especially in regular classes. The results of this research can be directly applied to classroom instruction in several impactful ways. Teachers can use item-level analysis to revise ineffective questions, aligning difficulty levels with students demonstrated proficiencies. The Wright Map provides diagnostic insights that allow teachers to differentiate instruction by grouping students according to ability, tailoring lesson complexity, and designing appropriate scaffolds. Additionally, identifying students with response behavior enables the implementation of targeted support strategies, such as confidence-building exercises, mindfulness practices, and individualized feedback. Since some students are merely guessing in answering the test items and with the given factors, the teacher may promote a positive attitude and belief to overcome test anxiety and build self-efficacy. Also, provide clear and well-designed items like using unambiguous language and applying appropriate time constraints to minimize confusion and reduce the likelihood of students needing to guess. The teacher could also equip students with test-taking strategies like eliminating incorrect answers, estimating answers, and checking their work to enhance their confidence and accuracy. Most importantly, the teacher may talk to students individually to understand and find ways to address test anxiety. The students should acknowledge that test anxiety is normal and can be managed through test-taking strategies. Introduce students to deep breathing exercises and mindfulness practices. The teacher may help students with positive affirmations to replace negative thoughts. During the test, the teacher may create a calm and supportive environment and minimize time pressure by avoiding frequent announcements about time limits, as this can increase anxiety.

Additionally, assessment should not only measure knowledge—it should inform teaching practice. Teachers are encouraged to use this data to create a more inclusive and responsive classroom environment by promoting a positive attitude, building self-efficacy, and teaching effective test-taking strategies like eliminating distractors and estimating answers. Emotional and behavioral factors, such as anxiety, must also be addressed through affirmations, calming techniques, and reducing time-related stress during exams. Post-assessment feedback must emphasize perseverance, effort, and growth mindsets to cultivate resilience and long-term academic success.

Beyond the immediate classroom context, the findings of this study have broader implications for mathematics education and assessment practice. The use of Rasch analysis offers a data-driven method that educators in similar grade levels or regions can adopt to improve test quality, evaluate student performance, and identify instructional gaps. The patterns observed—such as item misfit, varying difficulty levels, and inconsistent student responses—can inform curriculum planners and teachers in designing more balanced assessments that align with diverse learner abilities. Furthermore, addressing non-cognitive factors such as test anxiety and engagement through evidence-based strategies can enhance learning outcomes across various school settings. This study contributes to a growing body of educational research advocating for diagnostic assessment tools that support effective teaching and equitable learning environments in Philippine secondary education and beyond.

REFERENCES

[1] Al Ali, R. M., & Shehab, R. T. (2020). Psychometric properties of social perception of mathematics: Rasch model analysis. *International Education Studies*, *13*(12), 102–111. https://doi.org/10.5539/ies.v13n12p102

[2] Anderson, A. (1993). Assessment: A means to empower children. In N. L. Webb & A. L. Oxford (Eds.), *Assessment in the mathematics classroom*. NCTM.

[3] Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.* Longman.

[4] Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Routledge.

[5] Clements, D. H., Sarama, J. H., & Liu, X. H. (2008). Development of a measure of early mathematics achievement using the Rasch model: The research-based early maths assessment. *Educational Psychology*, *28*(4), 457–482. https://doi.org/10.1080/01443410701777272

[6] El-Korashy, A. (1995). Applying the Rasch model to the selection of items for a mental ability test. *Educational and Psychological Measurement*, *55*(5), 753–763. https://doi.org/10.1177/0013164495055005006

[7] Isnaini, Utami, W. B., Susongko, P., & Lestiani, H. T. (2019). Estimation of college students' ability on real analysis course using Rasch model. *Research and Evaluation in Education*, *5*(2), 95–102. https://doi.org/10.21831/reid.v5i2.20924

[8] Khan, L. A. (2015). What is mathematics – an overview. *International Journal of Mathematics and Computational Science*, *1*(3), 98–101. https://doi.org/10.13140/RG.2.1.3626.8967

[9] Kurniawan, D. D., & Mardapi, D. (2015). Penyetaraan vertikal tes matematika SMP dengan teori respons butir model Rasch [Vertical equating of junior high school mathematics tests using the Rasch model item response theory]. *Jurnal Evaluasi Pendidikan*, *3*(1), 12–25. http://journal.student.uny.ac.id/ojs/index.php/jep/article/view/1221/1093

[10] Magno, C. (2010). A brief history of educational assessment in the Philippines. *Educational Measurement and Evaluation Review*, *1*, 140–149. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1638490

[11] Mutodi, P., & Ngirande, H. (2014). The influence of students' perceptions on mathematics performance: A case of a selected high school in South Africa. *Mediterranean Journal of Social Sciences*, *5*(3), 431–438. http://doi.org/10.5901/mjss.2014.v5n3p431

[12] Ngunjiri, M. (2022). The role of assessment in mathematics classrooms: A review. *International Journal of Advanced Research*, *5*(1), 156–160. https://doi.org/10.37284/ijar.5.1.887

[13] Ramadhani, R., Saragih, S., & Napitupulu, E. (2022). Exploration of students' statistical reasoning ability in the context of ethnomathematics: A study of the Rasch model. *Mathematics Teaching-Research Journal*, *14*(1), 138–168. https://eric.ed.gov/?id=EJ1361651

[14] Sumintono, B. (2017). Rasch model measurements as tools in assessment for learning. *Advances in Social Science, Education and Humanities Research*, *173*, 57–61. https://doi.org/10.2991/icei-17.2018.11

[15] Sumintono, B., & Widhiarso, W. (2015). *Application of Rasch modeling on educational assessment*. Cimahi, Indonesia: Trim Komunikata Publishing House.

[16] Wright, B. D., & Stone, M. H. (1979). *Best test design: Rasch measurement*. Chicago, IL: MESA Press.

APPENDIX

Appendix A

Item	Proportion	Measure	S.E.Measure	Infit	Outfit	Interpretation
I39	0.02	4.00	1.01	1.01	1.02	very difficult
I28	0.04	3.28	0.72	0.97	0.71	very difficult
I11	0.10	2.29	0.47	1.04	1.17	very difficult
I46	0.12	2.08	0.44	1.05	1.15	very difficult
I47	0.14	1.90	0.41	1.02	1.05	very difficult
I12	0.16	1.74	0.39	1.03	1.07	very difficult
I48	0.16	1.74	0.39	1.12	1.38	very difficult
I34	0.18	1.60	0.37	1.12	1.28	very difficult
I42	0.22	1.34	0.35	1.05	1.10	very difficult
I35	0.24	1.23	0.34	1.05	1.09	very difficult
I20	0.27	1.01	0.32	1.14	1.22	very difficult
I23	0.27	1.01	0.32	1.09	1.14	very difficult
I44	0.31	0.82	0.31	1.00	1.03	difficult
I38	0.35	0.63	0.30	1.09	1.10	difficult
I43	0.35	0.63	0.30	0.88	0.86	difficult
I49	0.35	0.63	0.30	0.98	0.99	difficult
I8	0.37	0.55	0.30	1.00	0.99	difficult
I21	0.43	0.29	0.29	0.99	0.99	difficult
I41	0.43	0.29	0.29	1.00	0.99	difficult
I31	0.51	-0.04	0.29	1.00	1.00	easy
I36	0.51	-0.04	0.29	0.99	0.99	easy
I45	0.57	-0.28	0.29	0.96	0.95	easy
I2	0.67	-0.72	0.30	1.00	0.99	easy
I6	0.67	-0.72	0.30	0.98	0.95	easy
I18	0.67	-0.72	0.30	0.91	0.88	easy
I22	0.67	-0.72	0.30	0.91	0.88	easy
I10	0.69	-0.81	0.31	0.92	0.89	easy
I5	0.71	-0.91	0.31	1.02	1.01	easy
I25	0.71	-0.91	0.31	0.98	0.97	easy
I33	0.71	-0.91	0.31	0.98	0.95	easy
I15	0.73	-1.00	0.32	0.94	0.91	very easy
I24	0.73	-1.00	0.32	0.96	0.92	very easy

I50	0.73	-1.00	0.32	0.99	0.96	very easy
I40	0.75	-1.11	0.33	0.94	0.89	very easy
I37	0.76	-1.22	0.34	0.91	0.84	very easy
I9	0.78	-1.33	0.35	0.96	0.91	very easy
I32	0.78	-1.33	0.35	1.04	1.08	very easy
I7	0.86	-1.89	0.41	1.03	1.06	very easy
I30	0.86	-1.89	0.41	1.00	1.02	very easy
I14	0.88	-2.07	0.44	1.07	1.23	very easy
I27	0.92	-2.53	0.52	0.99	0.89	very easy
I29	0.92	-2.53	0.52	1.02	1.03	very easy
I4	0.94	-2.84	0.60	1.00	0.94	very easy
I13	0.94	-2.84	0.60	1.01	0.96	very easy
I16	0.94	-2.84	0.60	0.98	0.82	very easy
I26	0.94	-2.84	0.60	1.00	0.93	very easy
I17	0.96	-3.27	0.72	1.00	0.91	very easy
I19	0.96	-3.27	0.72	1.01	1.02	very easy
I3	1.00	-27.52	123953.03	NaN	**0.00**	very easy
I1	1.00	-27.52	123953.03	NaN	**0.00**	very easy

Table 2. Item characteristics

Appendix B

Student	Total score	Measure	SE	Infit	Outfit
28	21	-1.03	0.36	0.98	0.84
31	22	-0.90	0.36	1.14	1.30
44	22	-0.90	0.36	0.92	0.95
27	23	-0.78	0.36	1.10	1.24
48	23	-0.78	0.36	1.04	1.12
3	24	-0.65	0.36	1.05	1.16
33	24	-0.65	0.36	1.11	1.23
10	25	-0.52	0.36	1.45	**1.47**
21	25	-0.52	0.36	0.86	0.75
37	25	-0.52	0.36	1.23	1.28
14	26	-0.39	0.36	0.91	0.78
29	26	-0.39	0.36	1.20	1.12
36	26	-0.39	0.36	1.08	1.01
39	26	-0.39	0.36	1.26	1.30
49	26	-0.39	0.36	1.09	1.22
6	27	-0.26	0.36	0.90	1.00
7	27	-0.26	0.36	1.47	**1.58**
15	27	-0.26	0.36	0.76	0.77
19	27	-0.26	0.36	1.58	1.71
26	27	-0.26	0.36	1.26	1.11
50	27	-0.26	0.36	1.05	1.01
20	28	-0.13	0.36	0.86	0.83
34	28	-0.13	0.36	1.04	1.02
51	28	-0.13	0.36	1.18	1.13
2	29	0.00	0.36	0.58	**0.42**
5	29	0.00	0.36	0.94	0.73
23	29	0.00	0.36	0.78	**0.64**
25	29	0.00	0.36	1.12	0.89
30	29	0.00	0.36	1.32	**1.53**
35	29	0.00	0.36	1.18	1.17
1	30	0.13	0.37	0.74	**0.56**
11	30	0.13	0.37	0.69	**0.60**
40	30	0.13	0.37	1.13	0.95

4	31	0.26	0.37	0.91	1.01
16	31	0.26	0.37	1.03	1.05
42	31	0.26	0.37	0.92	0.90
9	32	0.40	0.37	0.54	**0.40**
12	32	0.40	0.37	1.22	1.10
24	32	0.40	0.37	0.87	**0.68**
45	32	0.40	0.37	0.92	**0.68**
46	32	0.40	0.37	0.92	0.83
47	32	0.40	0.37	0.77	**0.57**
18	34	0.68	0.38	0.68	**0.52**
41	34	0.68	0.38	0.49	**0.32**
43	34	0.68	0.38	1.12	0.99
17	35	0.82	0.38	0.66	0.70
22	35	0.82	0.38	0.58	**0.38**
32	36	0.97	0.39	0.85	0.92
38	36	0.97	0.39	0.81	**0.51**
8	37	1.12	0.40	0.58	**0.35**
13	37	1.12	0.40	0.75	**0.54**

Table 3. Person characteristics

Appendix C

First Quarter Examination in Mathematics 10

1. Which pattern follows the sequence: 2, 4, 8, 16, __?
   A. 24 C. 64
   B. 32 D. 12

2. What is an arithmetic sequence?
   A. A sequence where each term is the product of the preceding term and a fixed number.
   B. A sequence where each term is the sum of the preceding term and a fixed number.
   C. A sequence where each term is a random number.
   D. A sequence where each term alternates between odd and even numbers.

3. What is the formula for the nth term of an arithmetic sequence?
   A. $a_n = a_1 + (n-1)d$ C. $a_n = a_1 \times n \times d$
   B. $a_n = a_1 \times d^{n-1}$ D. $a_n = a_1 + n \times d$

4. What defines a geometric sequence?
   A. A sequence where each term is the sum of the preceding term and a fixed number.
   B. A sequence where each term is a random number.
   C. A sequence where each term is the product of the preceding term and a fixed number.
   D. A sequence where each term alternates between odd and even numbers.

5. What distinguishes a geometric sequence from an arithmetic sequence?
   A. In a geometric sequence, each term is found by adding a constant value to the previous term, whereas in an arithmetic sequence, each term is found by multiplying a constant value to the previous term.
   B. In a geometric sequence, each term is found by multiplying a constant value to the previous term, whereas in an arithmetic sequence, each term is found by adding a constant value to the previous term.
   C. Geometric sequences are entirely random in their progression, while arithmetic sequences follow a predictable pattern.
   D. Geometric sequences involve both addition and multiplication operations, while arithmetic sequences only involve addition operations.

6. What is the formula to find the nth term (a_n) of a geometric sequence?
   A. $a_n = a_1 + (n-1)d$ C. $a_n = \frac{a_1+a_n}{2}$
   B. $a_n = a_1 r^{n-1}$ D. $a_n = \frac{n}{2}(a_1 + a_n)$

7. Which of the following sequences is an arithmetic sequence?
   A. 3, 6, 9, 12, 14 C. 1, 3, 9, 27, 81
   B. 2, 4, 8, 16, 32 D. 5, 10, 15, 20, 25

8. Which method among the following is specifically used for dividing polynomials by linear factors or divisors?
   A. Synthetic division C. Horner's method
   B. Long division D. Newton's method

9. Which theorem states that if a polynomial *f*(x) is divided by x+c, the remainder will be *f*(-c)?
   A. Remainder theorem C. Rational root theorem
   B. Factor theorem D. Intermediate value theorem

10. Factor the polynomial $(x^2 - 4)$.
   A. $(x-2)(x-2)$ C. $(x+2)(x+2)$
   B. $(x+2)(x-2)$ D. $(x-4)(x+4)$

11. Identify the type of polynomial equation represented by the expression: $3x^4 - 2x^2 + 5x + 1 = 0$
   A. Quadratic equation C. Quartic equation
   B. Cubic equation D. Quintic equation

12. What is the degree of the polynomial equation represented by $4x^3 - 2x^2 + 5x - 1 = 0$
   A. Degree 1 C. Degree 3
   B. Degree 2 D. Degree 4

13. In the sequence 3, 6, 12, 24, __, which rule or pattern is being followed?
   A. Addition by 12 C. Subtraction by 2
   B. Multiplication by 2 D. Division by 3

14. If the first term in an arithmetic sequence is 7 and the common difference is 4, what is the fifth term in the sequence?
   A. 19 C. 27
   B. 23 D. 31

15. What is the formula to find the sum of the first *n* terms of an arithmetic sequence when the nth term in unknown?
   A. $S_n = \frac{n}{2}(a_1 + a_n)$ C. $S_n = \frac{n}{2}(2a_1 + (n-1)d)$
   B. $S_n = n\, x\, a_1\, x\, d$ D. $S_n = a_1 + (n-1)d$

16. In a geometric sequence where the first term (a_1) is 5 and the common ratio (r) is 3, what are the first four terms of the sequence?
   A. 5, 8, 13, 18 C. 5, 15, 30, 60
   B. 5, 10, 15, 20 D. 5, 15, 45, 135

17. Differentiate between an arithmetic sequence and a geometric sequence based on their defining characteristics.
   A) Arithmetic sequences have a common ratio, while geometric sequences have a common difference.

B) Arithmetic sequences have a common difference, while geometric sequences have a common ratio.
C) Arithmetic sequences have an exponential growth, while geometric sequences have a linear growth.
D) Arithmetic sequences have a common ratio, while geometric sequences have an exponential growth.

18. What is the formula for the sum of the first n terms (S_n) of a finite geometric sequence when a_n is known?
   A. $S_n = \frac{n}{2}(a_1 + a_n)$ C. $S_n = \frac{a_1(1-r^n)}{1-r}$
   B. $S_n = \frac{n}{2}(2a_1 + (n-1)d)$ D. $S_n = a_1 r^n$

19. If the 7th term in an arithmetic sequence is 31 and the common difference is 5, what is the 15th term of this sequence?
   A. 51 C. 61
   B. 65 D. 71

20. When should synthetic division be used instead of long division for polynomial division?
   A. When the dividend is a higher-degree polynomial than the divisor.
   B. When the divisor is a higher-degree polynomial than the dividend.
   C. When the divisor is a linear polynomial.
   D. When the divisor is a quadratic polynomial.

21. Which theorem helps in determining the possible roots of a polynomial equation and the relationship between the roots and the factors of a polynomial?
   A. Remainder theorem C. Rational root theorem
   B. Factor theorem D. Intermediate value theorem

22. Factor the polynomial $x^2 + 5x + 6$.
   A. $(x+3)(x+2)$ C. $(x-3)(x-2)$
   B. $(x+1)(x+6)$ D. $(x-1)(x-6)$

23. Identify the degree of the polynomial equation: $6x^3 - 4x^2 + 2x - 1 = 0$
   A. Degree 1 C. Degree 3
   B. Degree 2 Degree 4

24. Solve the equation $2x^2 + 5x - 3 = 0$ using factorization.
   A. $x = 1/2, x = -3$ C. $x = -1/2, x = -3$
   B. $x = 1/2, x = 3$ D. $x = -1/2, x = 3$

25. Using a pattern or rule, extend the following sequence: 5, 10, 20, __.
   A. 25 C. 40
   B. 35 D. 50

26. In an arithmetic sequence, if the first term a_1 is 3 and the common difference d is 7, what would be the tenth term of this sequence?
   A. 66 C. 83
   B. 76 D. 97

27. In an arithmetic sequence, if the first term (a_1) is 7 and the common difference (d) is 4, what is the 15th term (a_{15}) and the sum of the first 15 terms in this sequence?
   A. $a_{15} = 63,\ S_{15} = 435$ C. $a_{15} = 63,\ S_{15} = 525$
   B. $a_{15} = 63,\ S_{15} = 465$ D. $a_{15} = 63,\ S_{15} = 635$

28. In a geometric sequence, if the 5th term is 81 and the 3rd term is 9, what is the common ratio of this sequence?
   A. 3 C. 27
   B. 9 D. 36

29. Identify the following sequence as either arithmetic or geometric: 2, 4, 8, 16, 32, ...
   A. Arithmetic Sequence C. Neither
   B. Geometric Sequence D. Both

30. In a geometric sequence where the first term (a_1) is 3 and the common ratio (r) is 2, what is the third term (a_3) of the sequence?
   A. 6 C. 12
   B. 9 D. 15

31. In a geometric sequence, if the first term (a_1) is 8 and the common ratio (r) is 3, what is the 6th term (a_6) and the sum of the first 5 terms (S_5)?
   A. a_6=1944, S_5=968 C. a_6=729, S_5=3280
   B. a_6=729, S_5=1092 D. a_6=1944, S_5=19683

32. Divide $3x^3 - 7x^2 + 5x - 9$ by $x - 2$ using the most suitable method. Which method should be employed for this division?
   A. Long division C. Horner's method
   B. Synthetic division D. Newton's method

33. Given a polynomial $f(x)=2x^3-5x^2+7x+3$. When divided by $(x-2)$, the remainder is 0. What can be concluded using this information?
   A. $f(2)=0$ C. $f(2)=3$
   B. $f(2)=7$ D. $f(2)=2$

34. Factor the following polynomial completely: $x^3 - x^2 - 6x$
   A. $x(x-3)(x+2)$ C. $x(x+3)(x+2)$
   B. $x(x+3)(x-2)$ D. $x(x-3)(x-2)$

35. Determine the values of k that satisfies the equation $2k^2 - 5k + 2 = 0$.
    A. $k = 2, k = \frac{1}{2}$ C. $k = 2, k = 1$
    B. $k = 1, k = 3$ D. $k = 3, k = 2$

36. Solve the equation $x^3 - 5x^2 + 8x - 4 = 0$ given that one root is $x = 1$. What is the quadratic equation formed by the remaining roots?
    A. $x^2 - 4x + 4 = 0$ C. $x^2 - 4x + 6 = 0$
    B. $x^2 - 6x + 8 = 0$ D. $x^2 - 6x + 10 = 0$

37. In an arithmetic sequence, if the 5th term is 23 and the 10th term is 43, what is the common difference of the sequence?
    A. 3 C. 5
    B. 4 D. 6

38. In an arithmetic sequence, if the 7th term is 32 and the 15th term is 72, what is the sum of the first 15 terms in this sequence?
    A. 445 C. 635
    B. 555 D. 685

39. If in a geometric sequence the first term (a_1) is 5 and the common ratio (r) is 1/3, what is the sum of the first 5 terms of this sequence?
    A. 325/81 C. 545/81
    B. 495/81 D. 605/81

40. In a sequence, the first term (a_1) is 10, and the second term (a_2) is 16. If this sequence is an arithmetic sequence, what is the 6th term (a_6)?
    A. 28 C. 36
    B. 32 D. 40

41. Factor the polynomial $x^2 + 7x + 12$.
    A. $(x + 3)(x + 4)$ C. $(x - 3)(x - 4)$
    B. $(x + 2)(x + 5)$ D. $(x + 6)(x + 2)$

42. Given the equation $x^3 - 4x^2 + 5x - 2 = 0$, what is the sum of the roots?
    A. 2 C. 4
    B. 3 D. 5

43. Consider the equation $3x^3 - 11x^2 + 8x = 0$. What conclusion can be drawn about the roots based on the coefficients?
    A. The equation has one real and two complex roots.
    B. The equation has real roots.
    C. the equation has all complex roots.
    D. The equation has one real and one complex conjugate pair of roots.

44. In an geometric sequence, if the first term (a_1) is 3 and the common ratio (r) is 5, what is the 4th term (a_4) and the sum of the first 4 terms?
    A. $a_{15} = 375,\ S_{15} = 438$
    B. $a_{15} = 375,\ S_{15} = 468$
    C. $a_{15} = 375,\ S_{15} = 528$
    D. $a_{15} = 375,\ S_{15} = 638$

45. In an arithmetic sequence, the 10th term is 72, and the 15th term is 102. What is the 20th term of this sequence?
    A. 132
    B. 112
    C. 142
    D. 152

46. Factor the polynomial $2x^2 + 5x - 3$.
    A. $(2x - 1)(x + 3)$
    B. $(2x + 3)(x - 1)$
    C. $(2x - 3)(x + 1)$
    D. $(2x + 1)(x + 3)$

47. Solve the equation $2x^2 - 7x + 3 = 0$. Identify the nature of the roots.
    A. $x = \frac{1}{2}, x = 3$, Real and unequal
    B. $x = 1, x = 3$, Real and unequal
    C. $x = 3, x = -\frac{1}{2}$, Real and unequal
    D. $x = 3, x = 3$, Real and equal

48. Solve the equation $x^3 - 6x^2 + 11x - 6 = 0$. Determine the nature of the roots.
    A. Real and unequal
    B. Real and equal
    C. Two real and one complex
    D. One real and two complex

49. Create a geometric sequence where the first term (a_1) is 4 and the common ratio (r) is 3. Then, determine the 5th term (a_5) and the sum of the first 4 terms (S_4).
    A. a_5=324, S_4=160
    B. a_5=243, S_4=260
    C. a_5=243, S_4=160
    D. a_5=324, S_4=260

50. Create a polynomial expression that can be factored by difference of squares.
    A. $x^2 - 4$
    B. $x^2 - 5$
    C. $x^2 - 6$
    D. $x^2 - 7$