JURNAL PENDIDIKAN MATEMATIKA
http://e-journal.uingusdur.ac.id/index.php/circle

Psychometric Evaluation of an Extended Basic Mathematics Skills Measure

Harris Shah Abd Hamid[1], Shariah Bagam Abdul Rahman[2], Siti Masitah Ibrahim[3], Shafinah Abu Hasan[4]
[1,2,3,4]Faculty of Management, Education & Humanities, University College of MAIWP International (UCMI), Malaysia
[1]Faculty of Medicine, Universiti Kuala Lumpur, Royal College of Medicine, Malaysia
[1]Fakultas Psikologi, Universitas Negeri Malang, Indonesia
Corresponding Author: shafinah@ucmi.edu.my[4]

Article history
Received: *October 7, 2024*
Revised: *March 27, 2025*
Accepted: *March 29, 2025*

Keywords:
Basic mathematics skills
Mathematics self-concepts
Psychometric evaluation
Simple rasch model
Test adaptation

Abstract
Basic mathematics skills (BMS) and mathematics self-concepts are related to achievement in mathematics courses. However, the validity evidence of a BMS measure is limited. Of particular concern are its unidimensionality, the difficulty of the items, and the functioning of the distractors. Therefore, this study attempts to validate a bilingual and extended version of a BMS measure and examine the scores' correlation with past mathematics achievement, expected grades in a current maths-related course, and a Mathematics Self-Concept (MSC). The BMS was extended by adding five new items and replacing three items. The BME-Ext and MSC questionnaires were given to students at a college taking a quantitative methods course using Google Forms. Data from 123 students were analysed using Winstep and Jamovi. Evidence for unidimensionality is inconclusive while the person's reliability and separation are not sufficient. Four items are too easy and may need to be replaced with more difficult questions. In six items, there is one distractor that was not chosen at all by the students. The BMS scores have a significant positive correlation, in descending order, with past mathematics achievement, expected grades for the current course, and MSC. Improvements to the BMS and further testing are suggested for measurement and practical purposes.

Kata Kunci:
Kemampuan Matematika Dasar
Konsep Diri Matematika
Evaluasi Psychometric
Model rasch sederhana
Adaptasi tes

Abstrak
Terdapat keterkaitan keterampilan matematika dasar (BMS) dan konsep diri matematika (MSC) dengan prestasi dalam mata kuliah matematika. Akan tetapi, bukti validitas pengukuran BMS terbatas. Yang menjadi perhatian khusus adalah unidimensionalitasnya, tingkat kesulitan butir soal, dan fungsi pengalih perhatian. Oleh

p-ISSN 2776-6268
e-ISSN 2777-1008

karena itu, penelitian ini berupaya memvalidasi versi dwibahasa dan diperluas dari pengukuran BMS dan memeriksa korelasi skor dengan prestasi matematika sebelumnya, nilai yang diharapkan dalam mata kuliah terkait matematika saat ini, dan MSC. BMS diperluas dengan menambahkan lima butir soal baru dan mengganti tiga butir soal. Kuesioner BME-Ext dan MSC diberikan kepada mahasiswa di sebuah perguruan tinggi yang mengambil mata kuliah metode kuantitatif menggunakan Google Forms. Data dari 123 mahasiswa dianalisis menggunakan Winstep dan Jamovi. Bukti unidimensionalitas tidak meyakinkan sementara reliabilitas dan pemisahan orang tersebut tidak memadai. Empat butir soal terlalu mudah dan mungkin perlu diganti dengan soal yang lebih sulit. Dalam enam butir soal, ada satu pengalih perhatian yang sama sekali tidak dipilih oleh mahasiswa. Nilai BMS memiliki korelasi positif yang signifikan, dalam urutan menurun, dengan pencapaian matematika sebelumnya, nilai yang diharapkan untuk mata kuliah saat ini, dan MSC. Perbaikan pada BMS dan pengujian lebih lanjut disarankan untuk tujuan pengukuran dan praktis.

INRODUCTION

Concerns about students' performance and mastery of mathematics are not just an academic curiosity: they have real consequences in terms of students' achievement at university and, subsequently, employment prospects. In Malaysia, while students normally take mathematics for their Malaysian Certificate of Education (SPM) examination, a pass is not a compulsory requirement for some degree programmes. Nevertheless, those programmes may have courses requiring mathematics skills. Assessing university students' mathematics skills beyond SPM grades is indeed relevant for teaching and learning.

Research has shown that basic mathematics skills (BMS) and mathematics self-concept (MSC) have a strong correlation with performance in mathematics courses. BMS incorporates fundamental mathematical abilities obtained during primary education, which are necessary for future success in mathematics. In the context of the United States, BMS refers to skills "that the majority of high school graduates would be able to perform successfully after exposure to the typical mathematics curriculum in the educational system" (Trombley & Weiss, 1993, p23). These foundational skills allow students to solve basic problems and understand core concepts (Sa'di et al., 2023). BMS had been operationally defined in terms of tests with various mathematics topics. In the context of the present study, BMS is not to be equated with similar constructs like numeracy, mathematics literacy, and quantitative reasoning. As reviewed by Kaarali (2016), these three terms have common content, but BMS explicitly refers to the skills gained in primary education which are expected to be applied in the secondary education.

While BMS refers to skills, MSC refers to cognitive and emotional experiences. Self concept is an important construct in the field of pscyhology. Generally, it refers to "the set of feelings that the subject has about him/herself" (Palenzuela-Luis et al, 2022, p2). More narrowly, academic self concept refers individual's perceptions of his or her academic ability and this can be made more specific by focusing on domains like mathematics (Mejía-Rodríguez et al., 2021). MSC has significant roles in education; it has been shown to positively influence mathematical performance, with higher self-concept correlating with better achievement (Passiatore et al., 2023; Kania & Juandi, 2023). This positive correlation indicates that adopting a strong MSC can improve students' engagement and persistence in mathematics.

Apart from course-related math summative assessments, researchers had used quizzes or tests to measure basic math skills. For example, a 10-item MCQ math quiz (Ballard & Johnson, 2004) was later expanded into 15-item Basic Math Quiz (Johnson & Kuennen, 2006). Another example is a 7-item open ended quiz covering "operations of arithmetic as well as the hierarchy of math rule, decimal and percent conversions, algebra, and numerical substitution into and evaluation of formulas" (p 381) and targeted 10th grade math competency (Jones et al, 2011). Basic Math Skills Test (BMST) is a longer open-ended test (50 questions) design with grade 8 math topics (Latif, 2002). The validity of current BMS measures remains an important scientific endeavour not just for research but also for practical purposes. In the publications cited earlier, scant psychometric properties of the quizzes of rests were reported. As noted by Alcívar-Castro et al. (2023), ensuring the accuracy of assessments is necessary for supporting students' mathematical understanding. Therefore, there is a need for validated measurement tools that can accurately assess BMS in a bilingual context, accommodating the diverse linguistic backgrounds of Malaysian students.

This study aims to address this need by validating a bilingual and extended version of the BMS measure. The study involves the addition of five new items and the replacement of three existing items, with the BMQ-Ext and MSC questionnaires administered to students enrolled in a quantitative methods course via Google Forms. Data from respondents were analyzed using Winstep and Jamovi to explore the correlation between BMS scores, past mathematics achievement, expected grades in current mathematics-related courses, and MSC. The findings will contribute to the development of a more effective assessment tool that not only indicates

students' mathematical proficiencies but also supports their academic and future employment in a competitive environment.

Problem Statement

There is limited validity evidence for current BMS measures despite of the established relationship between basic mathematics skills (BMS) and mathematics self-concept (MSC) with academic achievement in mathematics courses. In the article describing the first use of BMQ (Johnson & Kuennen, 2006), no psychometric properties were reported. According to Passiatore et al. (2023) and Núñez-Peña et al. (2024), concerns regarding the unidimensionality of mathematics ability measures, the appropriateness of item difficulty, and the effectiveness of distractors in multiple-choice questions limit their applicability in educational settings. These limitations can be overcome using Simple Rasch Model which is the analytical tool used in the present study.

In Malaysia, studies examining university students' mathematics performance tend to use course's test scores and final grade (e.g. Ismail et al., 2012). One study used Basic Mathematics Quiz but did not report its psychometric properties (Abd Hamid & Sulaiman, 2014). Thus, this study aims to address these gaps by validating a bilingual and extended version of a BMS measure. The research will examine the correlation of BMS scores with previous mathematics achievement, expected grades in current mathematics-related courses and MSC. This would contribute to a clearer understanding of how these constructs influence on the student achievement in mathematics education.

Literature Review

Recent studies highlight a strong positive correlation among BMS scores, past achievement, predicted grades, and MSC, emphasizing the importance of these constructs in mathematics education. For example, Passiatore et al., (2023) found significant positive correlation between mathematics self concept and two mathematics tasks namely operations and number judgement. Woldemichael et al., (2023) tested a structural model for predicting mathematics achievement and found self-concept, together with self-efficacy and self-determination, is a significant predictor. Other studies highlighted the importance of having strong mathematical abilities for academic performance in college. Mathematics is a necessary subject for STEM education and encourages early engagement and achievement in math in order to improve academic performance in the future (Ben-Jacob, 2019). Moreover, there is a strong correlation between mathematics proficiency and academic success. Through regression analyses,

Delaney and Devereux (2020) found that mathematical skills is a stronger predictor of university achievement compared to verbal skills. This is especially true for degree completion and earning first-class honours, particularly in STEM courses.

Moreover, mathematical skills are not only about numerical and quantitative reasoning; they also enhance broader cognitive capabilities such as logical reasoning, problem-solving, and critical thinking. These cognitive skills are valuable across various disciplines, highlighting the importance of mathematics education even for students outside of STEM fields. Cresswell and Speelman (2020) showed a positive correlation between increased mathematics training and improved performance on logical reasoning tasks like Cognitive Reflection Test and the Wason Selection Task,.

Additionally, mathematical skills are necessary for career readiness, serving as the basis for problem-solving, critical thinking, and analytical abilities that are highly valued in today's job market (Hafizi & Kamarudin, 2020). As the job market evolves, mathematical proficiency becomes gradually more crucial, with 72% of executives classifying critical thinking and problem-solving as top skills (Moore, 2020). The ability to think both logically and creatively is essential for navigating the challenges of a data-driven and automated work environment. Moreover, proficiency in high school mathematics is linked to a greater likelihood of pursuing math-intensive STEM careers, such as engineering and computer science, which often lead to higher-paying jobs (Yea & Kaci, 2023). These results emphasize the significance of developing effective mathematical abilities for future career accomplishment.

Instruments like Basic Math Skills Test (BMST) supports various aspects of education purposes by being a vital tool to diagnose a student's current level of mathematical understanding. According to Rylands and Shearman (2022), the diagnostic tests are used to assess students' mathematical abilities upon university entry which positively impacting student learning. Besides, it also has been implemented in regular assessment which allows educators to monitor student progress and fine-tune teaching methods. These tests can detect students' weaknesses in problem-solving skills that affect academic performance, suggesting the need for improved teaching approaches, including the use of ICTs (Alcívar-Castro et al., 2023). Moreover, the data obtained from basic math skills tests also helps curriculum design and development. A study conducted by Lin and Chen (2023) found that there is a significant improvement in student performance, particularly in foundational math skills when schools

utilized assessment data for curriculum adjustments. These studies collectively highlight the utility of the BMST in evaluating and addressing basic math skills in education.

In addition, PISA's framework reflects real-world scenarios, including complex problem-solving and application-based questions. PISA's validity is sustained through consistent updates and field testing, focusing on ensuring that the test items are culturally neutral and relevant to students' daily lives. The assessment's comprehensive design allows it to measure students' ability to apply mathematical concepts in numerous contexts (OECD, 2019). The study by Chan Choon Tak (2021) found that an instrument measuring self-efficacy and metacognitive awareness in mathematics had a Cronbach Alpha coefficient above 0.9, indicating high reliability. Meanwhile, Mohd Dzin et al. (2023) confirmed the validity of adapted self-efficacy scales through PLS-SEM, achieving satisfactory reliability and convergent validity. However, Ing et al. (2024) highlighted the need for rigorous validity evidence in elementary math education measures, with only one in ten measures complying with this standard. Thus, the study suggests that mathematics education researchers take action to improve the validity of evidence for quantitative measures used in the field.

Research Objectives

The aim of this study is to provide additional validity evidence for an existing measure of basic mathematics skills, namely the BMQ, using Rasch model and convergent validity approach. This aim is relevant to address the research problem stated in the earlier section. To achieve this aim, the following specific objectives were formulated:

1. To adapt and extend the BMQ
2. To evaluate the psychometric properties of the extended BMQ using Simple Rasch Model
3. To examine the correlation of the extended BMQ scores with past mathematics achievement, expected mathematics achievement, and mathematics self-concept.

Research Questions

The research question are as follows:

1. What items needs to be modified from the BMS?
2. How good are the psychometric properties of the extended BMS based on Simple Rasch model?
3. Are there significant positive correlations among extended BMS scores, past mathematics achievement, expected mathematics achievement, and mathematics self-concept?

METHOD

Research Design

A quantitative research design is utilized, applying a cross-sectional survey approach to collect data from students enrolled in a quantitative methods course. This study allows for the measurement of relationships between BMS, MSC, and academic achievement.

Participants

The study involved college students enrolled in a quantitative methods course. Convenience sampling using voluntary response technique was used to obtain the study sample. The resulting sample size of 123 allows for stability of estimate at ± ½ logit and 95% confidence (Linacre, 1994). As the Rasch model is sample-independent, the issue of representativeness to the population of students is not an issue. The emphasis of the present study is students' mathematics ability and the obtained sample size is sufficient for this purpose.

Data Collection and Analysis

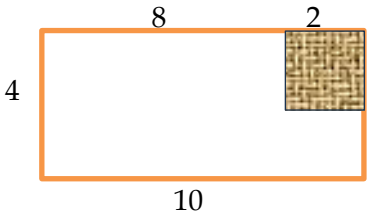
Data are collected using Google Forms, allowing for effective distribution and collection of the questionnaires. Participants are informed regarding the study's purpose, and consent is acquired prior to participation. The questionnaires are administered at the end of the course to ensure students have effectively engaged with the information. The extended basic mathematic skills test and mathematic self-concept questions have been utilized in this study.

Extended Basic Mathematics Skills Test

The BMQ-Ext is based on a Basic Math Quiz (BMQ) used by Johnson and Kuennen (2006). The original BMQ comprises 15 questions with five answer options each and covers basic concepts of arithmetics, algebra, and geometry. Wordings were revised where appropriate (e.g. the term 'yard' was replaced with 'meter'). Questions 12 to 15 from the BMQ were removed due to redundancy of topic (percentage, fraction, and square root). In their place, three questions with similar content were added. In accordance to the guideline of having 20 items for a dichotomous test (Kruyen, 2012), five new items were added. They tap into areas of mathematics that are relevant to college-level statistics and quantitative analysis courses.

12. Which of the following is equal to 16?
    a) $4^2 \times 6^1$ b) $2^2 + 2^2$ c) $5^2 - 2^2$ d) $8^0 + 8^0$ e) $16^1 \times 2^0$

13. A seller reduced the price of durian from RM30 to RM20 per kilogram. By what percent has the price decreased?
    a) 10% b) 15% c) 33% d) 40% e) 60%



14. The shaded region is a square. What is the area of the unshaded region?
    a) 4 b) 36 c) 20 d) 32 e) 40

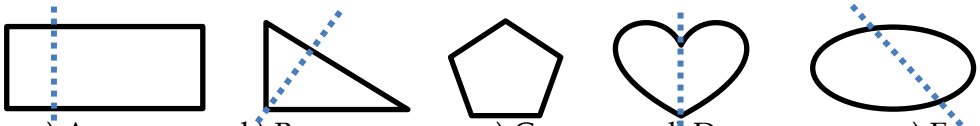
15. What is the percentage value for 0.638?
    a) 638% b) 63.8% c) 6.38% d) .638% e) 0.0638%

16. 7% ____ 0.7
    a) = b) > c) < d) ≥ e) ≤

17. In a bowl, there are 8 red marbles, 3 green marbles, and 9 blue marbles. If one marble is chosen randomly, what is the probability of getting a blue marble?
    a) 90% b) 9/11 c) 9/20 d) 20/9 e) 0.9

18. In a box, there are 5 purple shirts, 12 yellow shirts, and 8 black shirts. If a shirt is chosen randomly, what is the probability of getting a purple OR yellow shirt?
    a) 17% b) 12/17 c) 7/25 d) 17/25 e) 8/25

19. Which diagram is symmetrical along the marked axis?



    a) A b) B c) C d) D e) E

20. What is the nearest value to √(2&2,740) ?
    a) 2,000 b) 1,370 c) 100 d) 50 e) 27

Mathematics Self Concept

The Mathematics Self-Concept (MSC) survey is a vital tool in measuring students' perceptions of their mathematical abilities and confidence, especially in the context of PISA studies. Study indicates that mathematics self-concept significantly influences students' performance and learning behaviours. For instance, a study found that a poorer mathematics self-concept in female students led to more cautious response patterns in examinations, while males

exhibited risk-taking behaviours, highlighting the affective component of self-concept in academic performance (Núñez-Peña et al., 2024). Additionally, another study emphasized the protective role of self-concept in moderating the relationship between cognitive functions and mathematical skills, indicating that a positive self-concept can improve mathematical abilities (Passiatore et al., 2023). Furthermore, a study by Yuliani and Zaenal (2023) mentioned that self-concept is shown to correlate with mathematical connection abilities, where students with positive self-concepts performed better in mathematical tasks These findings highlight the importance of the MSC Questionnaire in understanding and improving students' mathematical self-perceptions and outcomes (Mukhibin et al., 2023).

Past Mathematics Achievement

Students were asked to state the grade that they get in the Malaysian Education Certificate (Sijil Pelajaran Malaysia – SPM) examination. The letter grade (A, B, C and so on) were converted into numbers (10, 9, 8 and so on) with higher numbers indicating better letter grade.

Future (Expected) Mathematics Achievement

Students were also asked to state the grade that they expect to get in their current math-related course. Similar recoding of the letter grade as mentioned above was applied.

Data Analysis

Data analysis is conducted using Winstep and Jamovi software. Rasch analysis in Winstep is employed to evaluate the unidimensionality of the BMS measure, analyzing item fit statistics and overall model fit. The procedure and cut off scores for the Rasch analysis are based on recommendation in Tennant and Küçükdeveci (2023). Pearson correlation coefficients are calculated to examine the relationships between BMS scores, past mathematics achievement, expected grades, and MSC scores.

RESULTS AND DISCUSSION

Reserach Result

Data from 123 students were received for analysis. The students' age is between 18 and 22 years (M=19.70, SD=.92). There were more female (n=77, 62.6%) than male (n=46, 37.4%) students. Person reliability and separation are poor at .65 and 1.36 respectively. These figures are the same even when 3.3% (4 students) achieved full marks (ceiling effect). Item reliability

and separation are very good at .93 and 3.60 respectively.

The primary dimension has minimal supporting evidence. The variance explained by measures is 25.8% (eigenvalue = 6.97) with 16.4% explained by persons and 9.5% explained by items. Meanwhile, the first contrast explained 8.8% of the variance with eigenvalue = 2.37. The contribution of the secondary dimension is not large enough to be confirmed as a real separate dimension from the primary Rasch dimension. Disattenuated correlation between clusters 1 and 3 is larger than .5, indicating strong enough relationship between those clusters of items. Moreover, only one item (Item12) has a loading to the first cluster exceeding .5. Local dependence was not observed among the items: the largest standardised residual correlation is only .36 (Item12 and Item19). Thus, overall, a secondary dimension is not supported by the evidence; however, the primary dimension is quite weak (not achieving 40% variance explained). Therefore, the unidimensionality of the BMQ-Ext would benefit from further development work.

All except two students have fit statistics larger than 2.0 (mean Infit Mean Square = 1.00, SD =.16, mean Outfit Mean Square =.93, SD=.35). With only two misfitting students, their effects to person measure are negligible. Thus, data from all students were included in the analyses reported in this section. The students have a high level of average ability (M=1.51, SD=1.24, average SE=.67) relative to the items' difficulty. However, the spread of the ability is large, indicating heterogenous mathematics ability among the students in the sample.

Item statistics as presented in Table 1 support the homogeneity requirements of a Rasch model (similar hierarchical ordering for each level of the scores) (Tennant, & Küçükdeveci, 2023). The items have a lower variability than persons (SD= 1.06). The fit statistics are in the acceptable range for all items (Infit Mean Square M=0.99, SD=.16; Outfit Mean Square M=0.92, SD=.34). The items also have better precision than the persons (average SE=0.26). The polarity of all items is in the positive region with Item19 having the smallest value (.19). Order of presentation effect (e.g. earlier items are easier) is not evident; Spearman correlation between the item order and item measure is not significant, r(18)=.334, p=.150.

BMQ-Ext item statistics

Table 1. Homogeneity Requirements of a Rasch Model

Measure	Score	SE	IN.MSQ	IN.Z	OUT.MSQ	OUT.Z	PTMA	Item
-1.77	115.00	0.39	0.89	-0.29	0.50	-0.88	0.37	1
0.22	90.00	0.23	0.97	-0.25	0.85	-0.76	0.47	2
-2.57	119.00	0.53	0.91	-0.07	0.36	-0.78	0.31	3
-0.50	102.00	0.27	0.92	-0.46	0.80	-0.66	0.45	4
0.52	84.00	0.22	1.12	1.20	1.26	1.60	0.35	5
-1.27	111.00	0.33	0.84	-0.65	0.56	-1.04	0.44	6
0.52	84.00	0.22	0.89	-1.10	0.81	-1.21	0.53	7
0.89	76.00	0.21	0.84	-1.88	0.74	-2.14	0.59	8
-0.30	99.00	0.25	1.11	0.77	1.10	0.46	0.34	9
0.66	81.00	0.22	1.07	0.74	1.11	0.75	0.40	10
-0.97	108.00	0.30	0.89	-0.54	0.58	-1.21	0.46	11
0.94	75.00	0.21	1.16	1.80	1.21	1.56	0.35	12
1.11	71.00	0.21	0.92	-0.93	0.86	-1.18	0.53	13
1.41	64.00	0.21	1.03	0.39	0.95	-0.37	0.47	14
0.06	93.00	0.24	1.04	0.37	1.24	1.15	0.39	15
1.49	62.00	0.21	1.20	2.48	1.37	2.81	0.31	16
-0.30	99.00	0.25	0.86	-0.95	0.69	-1.30	0.51	17
-0.50	102.00	0.27	0.90	-0.58	1.15	0.60	0.43	18
0.38	87.00	0.22	1.44	3.63	1.70	3.45	0.12	19
-0.05	95.00	0.24	0.79	-1.66	0.60	-2.06	0.58	20

Note: SE = Standard error; IN.MNSQ = Infit Mean Square; IN.Z = Infit Z-standard, OUT.Z = Outfit Z-standard, PTMA = Point measure correlation

The correct option in all questions have the highest average measure except for item 19 (about symmetrical shape). Revision to Item 19 may be necessary to ensure students with higher ability are able to choose the correct option. Overall, the distractors are functioning adequately. The limitations of the distractors are observed in six items where only four options were used. In other words, one option in those six items were not selected by any of the students. Additionally, there were very low number of students choosing some distractors. For example, nine distractors were chosen only once. The suitability of these distractors needs to be reevaluated and perhaps replaced with more suitable options. Alternatively, the number of options could be reduced to four from the existing five.

The BMQ-Ext scores obtained from the Rasch analysis were correlated with three variables. All correlations were positive and significant as shown in Table 2. BMQ-Ext is most strongly correlated with past achievement, and followed by future achievement and Math Self-Concept. Math Self-Concept is most strongly correlated with actual past performance (grade obtained in SPM). Overall, the BMQ-Ext and MSC are more strongly correlated with actual (past) performance than with predicted (future) performance.

Table 2. *Correlation Matrix of BMQ-Ext and Selected Variables*

		1		2		3	
1.	BMQ-Ext	—					
2.	Expected Grade	0.342	***	—			
3.	SPM	0.366	***	0.498	***	—	
4.	MSC	0.242	**	0.439	***	0.584	***

Note. * p < .05, ** p < .01, *** p < .001

Discussion

The present study extended the Basic Math Quiz into a 20-item bilingual BMQ-Ext. The findings support estimates of item difficulties which are useful for its further development. In its current form, the BMQ-Ext shows convergent validity in terms of its correlation with variables identified from the literature.

Unidimensionality is an issue and a limitation of the study. The percentage of variance explained by the measure is below 40%. However, this result is not unique to this study. MCQ tests seem to have low variance explained by measure: 27.1% for an Indonesian language test (Azizah et al. 2022), 23.7% for a science test (Mohd Dzin & Lay, 2021), and 23.3% for a mathematics test (Adi et al., 2022). Some studies that used Simple Rasch Model to analyse data from MCQ test did not even report unidimensionality statistics (e.g. Almubarak et al., 2023; Jowinies & Siew, 2024; Winarti & Mubarak, 2019). The variability of the math topics covered in the BMQ-Ext may contribute to the low variance explained value.

Evidence of convergent validity was in line with those reported in the literature. While correlation between BMQ scores and actual achievement in a statistics course was shown by Johnson and Kuennen (2006), this study shows a significant correlation with expected achievement in a quantitative data analysis course. Adlawon et al (2022) found a significant concurrent relationship between basic math skills and math performance. In comparison, the present study found similar relationship with retrospective and anticipative math achievement. Thus, the convergent validity findings from this study adds to the existing body of evidence.

In the present study, test administration was done in class using printed copies of the BMQ-Ext. Adapting the BMQ-Ext for online administration would be desirable in terms of cost, ease of data input, scoring, and providing feedback. However, further research needs to be done to ensure comparability of paper-based and computer-based test. Measurement invariance

between the two modes of administration cannot be taken for granted (Lynch, 2022). For an English achievement test, test takers preferred computer-based test over paper-based test but they actually performed better in the paper-based condition (Hosseini, & Toroujeni, 2017). Moreover, evidence of measurement invariance between online and face-to-face administration was found for non-cognitive scales (Zhang et al., 207). For a cognitive test like BMQ-Ext, empirical data are needed to justify the interpretation of data obtained from online administration.

CLOSING

Conclusion

The present study was successful in constructing a basic mathematics skills test with potentials to be refined further. The psychometric properties of the BMQ-Ext, while less than perfect, show adequate performance in terms of item quality. The BMQ-Ext scores correlate in the expected direction with mathematics self-concept, past math achievement and future (expected) math grades.

Recommendations

Further development of the test should focus on the distractors and would benefit from test administration procedure with better controls. Using a Simple or Dichotomous Rasch Model is highly instructive and informative in developing an accurate and precise instrument to measure students' basic mathematic skills.

REFERENCES

Abd Hamid, H. S., & Sulaiman, M. K. (2014). Statistics anxiety and achievement in a statistics course among psychology students. *International Journal of Behavioral Science, 9*(1), 55-66.

Adi, N. R. M., Amaruddin, H., Adi, H. M. M., & A'yun, I. L. Q. (2022). Validity and reliability analysis using the Rasch model to measure the quality of mathematics test items of vocational high schools. Journal of Research and Educational Research Evaluation, 11(2), 103-113.

Adlawon, C. M. C., Valdevarona, A. M. G, & Miranda, A. T. (2022). The relationship of self-concept and mathematics performance of college students in the new normal. Asian Journal of Education and Social Studies, 39-47. https://doi.org/10.9734/ajess/2022/v37i3805

Alcívar-Castro, E. J., Tamayo-Batista, M., Bravo-Andrade, B. D., Alcívar-Santander, T. L., & Pinargote-Jiménez, J. A. (2023). Development of mathematical skills in students of basic.

International Journal of Social Sciences and Humanities, 7(1), 14–23. https://doi.org/10.53730/ijssh.v7n1.13818

Almubarak, A., Saadi, P., Prayogi, R., & Maldini, P. P. (2023). Assessing students understanding of chemical bonds material by Rasch modeling. Indonesian Journal on Learning and Advanced Education (IJOLAE), 5(3), 217-232.

Azizah, N., Suseno, M., & Hayat, B. (2022). Item analysis of the Rasch model items in the final semester exam Indonesian language lesson. *World*, *12*(1).

Ben-Jacob, M. (2019). The importance of mathematics in STEM education. In EDULEARN19 Proceedings (pp. 2496-2500). https://doi.org/10.21125/edulearn.2019.0683

Chan, C. T., Zulnaidi, H., & Leong, K. E. (2021). Analysis validity and reliability of Self - Efficacy and Metacognitive Awareness instrument toward mathematical reasoning. Türk Bilgisayar Ve Matematik Eğitimi Dergisi, 12(9), 3332–3344. https://doi.org/10.17762/turcomat.v12i9.5739

Cresswell, C., & Speelman, C.P. (2020). Does mathematics training lead to better logical thinking and reasoning? A cross-sectional assessment from students to professors. PLoS ONE, 15.

Delaney, J.M., & Devereux, P.J. (2020). Math matters! the importance of mathematical and verbal skills for degree performance. Economic Letters, 168, 108850. https://doi.org/10.1016/j.econlet.2019.108850

Dzin, N. H. M., & Lay, Y. F. (2021). Validity and reliability of Adapted Self-Efficacy Scales in Malaysian context using PLS-SEM approach. Education Sciences, 11(11), 676. https://doi.org/10.3390/educsci11110676

Hafizi, M.H., & Kamarudin, N. (2020). Creativity in mathematics: Malaysian Perspective. *Universal Journal of Educational Research, 8*, 77-84.

Hosseini, M., & Toroujeni, S. M. H. (2017). Replacing paper-based testing with an alternative for the assessment of Iranian undergraduate students: Administration mode effect on testing performance. *International Journal of Language and Linguistics*, *5*(3), 78-87.

Ismail, N. A., Nopiah, Z. M., Asshaari, I., Othman, H., Tawil, N. M., & Zaharim, A. (2012). Mathematical performance of engineering students in Universiti Kebangsaan Malaysia (UKM). *Procedia-Social and Behavioral Sciences, 60*, 206-211.

Jones, T. W., Price, B. A., & Randall, C. H. (2011). A comparative study of student math skills: perceptions, validation, and recommendations. Decision Sciences Journal of Innovative Education, 9(3), 379–394. https://doi.org/doi:10.1111/j.1540-4609.2011.00314.x

Jowinis, A., & Siew, N. M. (2024). Validity and reliability of multiple-choice biology test: a Rasch analysis perspective. International Journal of Modern Education, 6(21), 131-139. https://doi.org/10.35631/IJMOE.621010

Karaali, G. (2016). What's in a name? A critical review of definitions of quantitative literacy, numeracy, and quantitative reasoning. Numeracy (Journal of the National Numeracy Network), Volume 9, (1), Article 6.

Kania, N., & Juandi, D. (2023). Does self-concept affect mathematics learning achievement? Journal of Education and Learning, 17(3), 455–461. https://doi.org/10.11591/edulearn.v17i3.20554

Kruyen, P. M. (2012). Using short tests and questionnaires for making decisions about individuals: When is short too short. Ridderprint.

Latif, D. A. (2002). The relationship among pharmacy students' basic math scores, traditional preadmission indicators, and performance in a pharmaceutical calculations course. *Journal of Pharmacy Teaching*, *10*(1), 17-29.

Lin, C. H., & Chen, C. H. (2023). Development, testing, and application of a mathematics learning scale of self-direction. *Frontiers in Psychology*, *14*, 1145442. https://doi.org/10.3389/fpsyg.2023.1145442

Linacre, J. M. (1994). Sample size and item calibration stability. Rasch Measurement Transactions, 7(4), 328.

Lynch, S. (2022). Adapting paper-based tests for computer administration: lessons learned from 30 years of mode effects studies in education. Practical Assessment, Research & Evaluation, 27, 22.

Mejía-Rodríguez, A. M., Luyten, H., & Meelissen, M. R. (2021). Gender differences in mathematics self-concept across the world: An exploration of student and parent data of TIMSS 2015. International Journal of Science and Mathematics Education, 19(6), 1229-1250.

Mohd Dzin, N. H., & Lay, Y. F. (2021). Assessing the validity and reliability of science multiple choice test using Rasch Dichotomous Measurement Model. Journal of Baltic Science Education, 20(6), 927-941.

Mukhibin, A., Rusyid, H. K., Lutfi, A., Herman, T., & Suryo Utomo, D. A. (2023). Analysis of Students' Mathematical Self-Efficacy Instruments using Rasch model. *Indonesian Journal of Mathematics Education, 6(2)*, 72–80. https://doi.org/10.31002/ijome.v6i2.994

Núñez-Peña, M. I., Ramon-Casas, M., Cunillera, T., & Campos-Rodríguez, C. (2024). Mathematics self-concept and response pattern in higher education examinations: differences between genders. *Assessment & Evaluation in Higher Education,* 1–14. https://doi.org/10.1080/02602938.2024.2364024

OECD. (2019). *PISA 2018 assessment and analytical framework: Reading, mathematics, and science*. OECD Publishing. https://doi.org/10.1787/b25efab8-en

Palenzuela-Luis, N., Duarte-Clíments, G., Gómez-Salgado, J., Rodríguez-Gómez, J. Á., & Sánchez-Gómez, M. B. (2022). International comparison of self-concept, self-perception and lifestyle in adolescents: a systematic review. *International Journal of Public Health*, *67*, 1604954.

Passiatore, Y., Costa, S., Grossi, G., Carrus, G., & Pirchio, S. (2023). Mathematics self-concept moderates the relation between cognitive functions and mathematical skills in primary school children. *Social Psychology of Education*. https://doi.org/10.1007/s11218-023-09854-3

Rylands, L.J., & Shearman, D. (2022). Diagnostic tests: Purposes and two case studies. *MSOR Connections, 20*(3). https://doi.org/10.21100/msor.v20i3.1281

Sa'di, D. R., Firdaus, N. P. N., Sinaga, R. D. H., & Yonvitra, N. H. (2023). Kemampuan siswa SMP dalam menyelesaikan persoalan matematika dasar. *Jurnal Pendidikan Matematika*, 1(2), 10.

Tennant, A., & Küçükdeveci, A. A. (2023). Application of the Rasch measurement model in rehabilitation research and practice: early developments, current practice, and future challenges. *Frontiers in Rehabilitation Sciences, 4*, 1208670.

Trombley, R. J., & Weiss, D. J. (1993). *Measurement of basic skills in mathematics.* National Center for Research in Vocational Education

Winarti, A., & Mubarak, A. (2019). Rasch modeling: A multiple choice chemistry test. *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, *2*(1), 1-9.

Woldemichael, B., Semela, T., & Tulu, A. (2023). The effect of mathematics self-beliefs on mathematics achievement: the case of undergraduate university students in Ethiopia. *Routledge Open Research, 2*, 41. https://doi.org/10.12688/routledgeopenres.17995.1

Yuliani, N. A., & Zaenal, R. M. (2023). Analisis kemampuan koneksi matematis ditinjau dari self-concept siswa. *Indo-MathEdu Intellectuals Journal, 4(2)*, 243–254. https://doi.org/10.54373/imeij.v4i2.163

Zhang, X., Kuchinke, L., Woud, M. L., Velten, J., & Margraf, J. (2017). Survey method matters: Online/offline questionnaires and face-to-face or telephone interviews differ. *Computers in Human Behavior*, *71*, 172-180.