Jurnal Elemen, 11(3), 720-740, July 2025
https:/doi.org/10.29408/jel.v11i3.30220

e- ISSN: 2442 - 4226

Check for updates

How math happens at home: A Rasch-based validation of the home mathematics environment scale in early childhood

Adi Cahya Nugraha *, Ira Rachmadani, Mardiana, Nur Alifa Deviar Refiyanti, Joko Soebagyo

Mathematics Education Department, Graduate School, Universitas Muhammadiyah Prof. DR. HAMKA, Jakarta, Indonesia

* Correspondence: dnugraha37@gmail.com

Abstract

Preschool education is important for developing children's social and cognitive skills, with early mathematical ability being a strong predictor of future success. This study evaluated the psychometric properties of the Preschool Home Mathematics Questionnaire (PHMQ) using Rasch Model Analysis. A total of 95 parents and guardians of preschool-aged children participated in the study. A total of 37 Likert-scale items were analysed. The results showed excellent person reliability (0.93) and item reliability (0.96), indicating strong consistency and well-targeted item difficulty. Of the 37 items, 32 were valid according to the Rasch fit criteria, while five items (Q13, Q15, Q22, Q24, and Q26) showed misfit based on the Outfit mean square (MNSQ), Z-standard (ZSTD), and PTMEA-CORR. The Likert scale functioned well with ordered thresholds and smooth transitions between categories. The unidimensionality test confirmed that the PHMQ measures a single construct, with 44.1% of the variance explained and only 12.8% unexplained in the first contrast. While the PHMQ shows good validity and reliability, this study focused only on basic Rasch validation. Future research should explore differential item functioning (DIF), item bias, and long-term stability to further improve this instrument and support its use in different settings.

Keywords: home mathematics environment; preschool home mathematics questionnaire; Rasch model analysis

How to cite: Nugraha, A. C., Rachmadani, I., Mardiana, Refiyanti, N. A. D., Soebagyo, J. (2025). How math happens at home: A Rasch-based validation of the home mathematics environment scale in early childhood. *Jurnal Elemen, 11*(3), 720-740. https://doi.org/10.29408/jel.v11i3.30220

Received: 8 May 2025 | Revised: 12 June 2025
Accepted: 27 June 2025 | Published: 31 July 2025

Introduction

Preschool education is a critical stage that benefits children’s socio-emotional and intellectual growth (Hahn & Barnett, 2023; Phillips, 2017; Prusinski et al., 2023). Children who attend preschool will reduce their behavioural problems (Yoshikawa et al., 2016) and produce substantial benefits for cognition, including executive function and academic achievement in mathematics (Camilli et al., 2010; Noa & Pennucci, 2014). Additionally, the quality of preschool education matters. Melhuish (2011) found that children who attended high-quality preschools significantly outperformed those who had not attended preschool on literacy and numeracy tests at age 11. Thus, this educational stage is crucial for developing children’s mathematical skills.

Children learn mathematics long before they start school (Reid, 2016), and early mathematics is set as the foundational concept they should obtain during the developmental process. Early mathematics is a key predictor of children's future achievement (Pan et al., 2023; Seitz & Weinert, 2022; Starkey et al., 2004; Zhang & Konstantopoulos, 2025) and plays a central role in daily life, such as shopping, cooking, and adjusting time (Butterworth, 2005; Van Rooijen et al., 2011). Indeed, children are exposed to mathematics in daily activities, even though they are unaware of it (Kvesic et al., 2020; Papic & Papic, 2025). Regarding cognitive skills, early mathematics affects children's written computation, sequences, math language, and comparisons in mathematics (Gashaj et al., 2023; Toll & Van Luit, 2014). In addition, early mathematics skills can be used to identify and monitor the development of preschool children with learning difficulties (Evans et al., 2015; Lopez-Pedersen et al., 2023). Children with this ability are more likely to succeed than those without it (Duncan et al., 2007; Watts et al., 2014). Therefore, this skill is one of the important abilities that children must acquire during their developmental milestones.

Children’s mathematical ability depends on their surroundings (Vygotsky, 1978), and their initial exposure to mathematics is in their home environments. The home mathematics environment has become influential in fostering children's mathematics skills at home (Bonifacci et al., 2021; Lefevre et al., 2009), and parents play a crucial role in enriching this environment (DeFlorio & Beliakoff, 2015; Purnomo et al., 2022; Wolf & McCoy, 2019). In addition, parents can use multiple approaches to develop their children's mathematics skills at home, such as reading storybooks (Petronzi et al., 2023; Saracho & Spodek, 2010; Vandermaas-Peeler et al., 2009), playing board games (Vandermaas-Peeler et al., 2012), puzzles (Kurniati et al., 2022), and block games (Gilligan-Lee et al., 2023; Nadlifah & Latif, 2024). Other factors that can influence how parents provide home numeracy include parental beliefs and expectations (Pesu et al., 2016), socio-economic status (SES) (Girard et al., 2021; Lu et al., 2025), and the use of technology at home (Mantilla & Edwards, 2019; Yuniria et al., 2025). However, there is limited evidence of how research and development enhance early numeracy skills (Nugraha & Muntazhimah, 2024; Seitz & Weinert, 2022). Furthermore, the quantity and quality of the home mathematics environment (HME) should be appropriately measured using significant and relevant measurement tools.

Multiple tools have been used in past research to measure home mathematics environments, such as phone interviews (Blevins-Knabe & Musun-Miller, 1996), parent questionnaires and child tests (Lefevre et al., 2009), and longitudinal studies to examine children's abilities in numeracy (Niklas & Schneider, 2014). Furthermore, Cahoon et al. (2021) developed a rigorous questionnaire by considering child characteristics, psychometric properties, and characteristics of content and activities. However, this and many similar studies have not captured middle-and low SES sufficiently, and a specific context may be considered in future research, since most of the HME questionnaires have been developed and used in developed countries such as Canada, America, and the UK (Cahoon et al., 2021). Therefore, examining the PHMQ questionnaire in a developing country such as Indonesia will enrich the findings, as the country has a variety of SES models and a multicultural community.

Rasch model analysis is necessary to examine the PHMQ in different contexts. These properties are realised when the fundamental assumption of one-dimensionality is satisfied, that is, when the data align with the model's requirements (Tabatabaee-Yazdi et al., 2018). The Rasch model is also useful for estimating individuals' abilities based on their responses to test items (McCamey, 2014) and enhances the precision and quality of tests and surveys while also enabling the development of multiple versions of measurement instruments (Al Ali & Shehab, 2020). This model is a robust method for assessing the validity and reliability of an instrument's constructs (Mofreh et al., 2014). However, no PHMQ analysis using the Rasch model has yet been conducted. Therefore, this study aimed to apply the Rasch Model to analyse the reliability and validity of the PHMQ. By doing so, we sought to provide a deeper understanding of how well the questionnaire captures the multidimensional nature of the home mathematics environment and to identify areas for improvement with the following questions:

1. How reliable are the PHMQ items and respondents in measuring the home mathematics environment using the Rasch model?
2. How do the items fit the Rasch model in terms of reliability and difficulty levels?
3. How well do the Likert scale categories function in distinguishing the levels of home mathematics engagement among respondents?
4. To what extent does the PHMQ demonstrate unidimensionality in measuring the HME construct?

Methods

This quantitative study used a psychometric approach to analyse the PHMQ administered to the parents and guardians of preschool children. The method employed is the Rasch Model, supported by the Winstep 5.10.0 version software, a measurement model within the Item Response Theory (IRT). This method provides an in-depth understanding of item characteristics, including validity, reliability, and potential item bias (Mallinckrodt et al., 2016). Additionally, this model aids in analysing item difficulty, person ability, item and person reliability, fit statistics, and unidimensionality of the scale. This approach was chosen because the main objective of the study was to measure the instrument's consistency and accuracy in evaluating the home learning environment for mathematics.

Participants

This study involved participants consisting of 95 parents or guardians of preschool-aged children, specifically those aged 3–5 years (see Table 1). Parents were considered respondents capable of providing information about their children's mathematics activities at home. Participants were selected using a voluntary sampling technique to choose parents or guardians willing to participate in the research process and who were actively involved in their children's home-based mathematics learning activities.

Table 1. Participant in the study

Respondent		Frequency	Percent (%)
Relationship	Mother	73	76.8
	Father	19	20
	Grandparents	1	1.1
	Aunt	1	1.1
	Relatives	1	1.1
Educational Qualification	Junior High School	1	1.1
	Senior High School	14	14.7
	Diploma	2	2.2
	Bachelor	65	68.4
	Master	11	11.6
	Doctor	2	2.1
Child's Gender	Male	44	46.3
	Female	51	53.7

The participants in this study were 95 parents or guardians of preschool children. Demographic data of the participants were collected through an additional questionnaire that included information on marital status, ethnic background, educational qualifications, employment status of the parents or guardians, and the child's gender. This demographic diversity was considered necessary because previous research has shown that factors such as the educational level of parents or caregivers can influence how they teach mathematical concepts to children (Missall et al., 2015; Purpura & Lonigan, 2015), and such an intervention could be measured by Rasch model analysis (Ling et al., 2018). Therefore, by including participants from diverse demographic backgrounds, this study aims to gain a broader and more holistic understanding of home-based mathematics activities.

Instrument

The instrument used in this research was the Preschool Home Mathematics Questionnaire developed by Cahoon et al. (2021). From all the open and closed statements in the questionnaire, only the items that utilised a Likert scale were selected, resulting in 37 items. Criterion validity was evaluated using contrasting cases identified by calculating the total scores across the five subscales with all positive statements for each participant. These scores, ranging from 1 to 5, were based on a 5-point Likert scale. Participants with a score of 1 indicated that the activity did not occur, two reflected that the activity occurred a few times a month, three indicated about once a week, four meant a few times a week, and five represented that the activity occurred almost daily.

Data analysis

This study used the Rasch Model approach to evaluate the validity and reliability of the Pre-School Home Mathematics Questionnaire (PHMQ). The Rasch Model provides information about the validity of both items and persons. The Rasch Model was used for evaluation and can also be utilised to determine the validity and reliability of a mathematics questionnaire using the Winsteps 5.10.0 version software. Validity and reliability were determined based on the following indicators:

Table 2. Reliability in Rasch analysis (Sumintono & Widhiarso, 2014)

Statistics	Fit Index	Description
Cronbach's Alpha (α) (KR-20)	< 0.5	Poor
	0.5 – 0.6	Weak
	0.6 – 0.7	Fair
	0.7 – 0.8	Good
	> 0.8	Excellent
Item and Person Reliability	< 0.67	Poor
	0.67 – 0.80	Weak
	0.81 – 0.90	Fair
	0.91 – 0.94	Good
	> 0.94	Excellent

In several references, the criteria used for item fit analysis state that an item is deemed fit if it meets the specified criteria. Boone et al. (2014) suggest the following criteria for examining item fit or misfit:

Table 3. Index for item and person fits

Statistics	Fit Index
Outfit Mean Square Values (MNSQ)	$0.5 < x < 1.5$
Outfit Z – Standardized Value (ZSTD)	$-2.0 < x < +2.0$
Point Measure Correlation (PTMEA –CORR)	$0.4 < x < 0.85$

These indicators offer complementary insights: Outfit MNSQ identifies unusual response patterns, ZSTD provides a standardised measure of fit, and PTMEA CORR reflects how well an item aligns with a respondent's overall ability. When an item or person failed two or more of these criteria, it was considered a misfit and examined for possible revision or exclusion. This strategy supports a more balanced evaluation, avoiding overly rigid judgments based on single statistics.

Results

Person and ıtem reliability

Based on the results of Rasch analysis using Winsteps 5.10.0 version, person reliability was found to be 0.93, and item reliability was 0.96 (see Table 4). According to the instrument quality guidelines (Fisher, 2007), a person's reliability of 0.93 (Very Good: 91 – 94) indicates excellent consistency in measuring respondents' abilities. Meanwhile, an item reliability of 0.96 (Excellent: > 0.96) suggests that the item difficulty level is very well-constructed to cover the range of respondents' abilities.

Table 4. Person and item reliability

Categories	N	Mean	SD	Reliability	Decision
Person	95	115.6	27.1	0.93	Very Good
Item	37	296.8	52.7	0.96	Excellent

Item analysis

Table 5 shows the distribution of test items considered misfit or fit in the Rasch model. As seen in Table 5, for the first criterion, based on the Outfit Mean Square (MNSQ) index, all items fell within the acceptable range of 0.5 to 1.5, indicating an overall good fit in terms of internal consistency among the items. However, when evaluated against the second criterion, which refers to the Outfit Z – Standardised Value (ZSTD) index, several items exceeded the threshold of -2.0 to 2.0. Specifically, items Q26 (2.82), Q22 (2.01), Q24 (-2.09), Q15 (-2.12), and Q13 (-2.96). This indicates that these items may have attracted unexpected or inconsistent responses, particularly from respondents at the extreme ends of the ability continuum. Such high outfit values suggest that the items might contain ambiguous wording, lack relevance, or fail to engage the respondents effectively.

For the third criterion, based on the Point Measure Correlation (PT Measure CORR) index, which assesses the degree to which each item correlates with the underlying construct being measured, only item Q26 fails to meet the standard, indicating that it does not discriminate well between higher- and lower-ability respondents and may not be aligned with the intended construct.

Table 5. Item fit statistics

Item	Outfit MNSQ	Outfit ZSTD	PT-Measure Corr	Decision
Q1	1.22	1.09	0.43	Fit
Q2	0.78	-1.05	0.50	Fit
Q3	1.29	1.66	0.47	Fit
Q4	0.99	-0.01	0.54	Fit
Q5	1.19	1.04	0.51	Fit
Q6	1.08	0.58	0.53	Fit
Q7	0.91	-0.62	0.53	Fit
Q8	1.15	1.02	0.50	Fit
Q9	0.83	-1.26	0.62	Fit
Q10	0.88	-0.86	0.63	Fit
Q11	0.74	-1.99	0.62	Fit
Q12	0.99	-0.03	0.55	Fit

Item	Outfit MNSQ	Outfit ZSTD	PT-Measure Corr	Decision
Q13	0.63	-2.96	0.65	Misfit
Q14	0.93	-0.46	0.64	Fit
Q15	0.70	-2.12	0.65	Misfit
Q16	0.92	-0.52	0.58	Fit
Q17	0.91	-0.58	0.61	Fit
Q18	0.97	-0.19	0.63	Fit
Q19	0.90	-0.71	0.58	Fit
Q20	0.87	-0.82	0.62	Fit
Q21	1.08	0.55	0.51	Fit
Q22	1.31	2.01	0.48	Misfit
Q23	1.29	1.43	0.54	Fit
Q24	0.73	-2.09	0.67	Misfit
Q25	1.29	1.34	0.41	Fit
Q26	1.48	2.82	0.36	Misfit
Q27	0.87	-0.81	0.68	Fit
Q28	1.13	0.91	0.59	Fit
Q29	0.96	-0.24	0.60	Fit
Q30	1.14	0.99	0.51	Fit
Q31	0.99	0.02	0.59	Fit
Q32	0.78	-1.50	0.62	Fit
Q33	0.91	-0.63	0.62	Fit
Q34	1.33	1.80	0.43	Fit
Q35	0.80	-1.38	0.64	Fit
Q36	1.12	0.83	0.50	Fit
Q37	1.08	0.51	0.54	Fit

Given these issues, several steps are recommended to improve the overall quality of this instrument. First, the misfit items should be reviewed and revised to improve clarity and ensure alignment with the home numeracy construct. This can include rewording items to eliminate ambiguity and ensure age-appropriateness for the target respondents (Suryani, 2018). Expert validation is also advisable to check the content and relevance of each item. Following revision, these items should undergo pilot testing with a smaller group to re-evaluate their fit statistics (Piussi et al., 2025). If the items continue to exhibit poor fit or low correlation with the construct, it may be necessary to remove them from the final questionnaire (Fischer et al., 2021; Johansson, 2025) to preserve the validity and reliability of the instrument.

To visualise how the item fits the expected model, the solid red "model" line is generated by the relevant Rasch model. The empirical blue line is formed by connecting the average ratings within each segment of the variable, represented by the "x" markers. These "x" points indicate the mean values of both the item measures (on the x-axis) and the observed ratings (on the y-axis) for all data within each interval, while the data outside the range represent data that are beyond the scope or do not conform to the Rasch Model. The figure illustrates that, in general, most of the data follow the Rasch Model, while others fall outside the range or do not conform to the model. This indicates that the item fits the model.

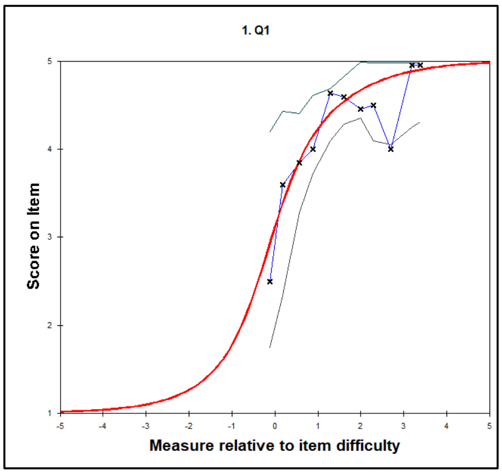

Figure 1. Expected score ICC graph

The Wright Map illustrates the distribution of respondents' ability levels to the item difficulty levels. In the analyzed data, 95 respondents and 37 items were measured on the Rasch scale. The map shows that Q25 ("Maths related websites (e.g. coolmaths.com)") and Q23 ("Asking shape related questions (e.g. "how many sides does a circle have?") have the highest difficulty levels (above +3 logits), while Q1 ("In the past month, how often did you and your child engage in reading?") and Q2 ("Counting") have the lowest difficulty levels (below -2 logits). This variation in difficulty levels indicates a broad measurement range, which is essential for the instrument's validity.

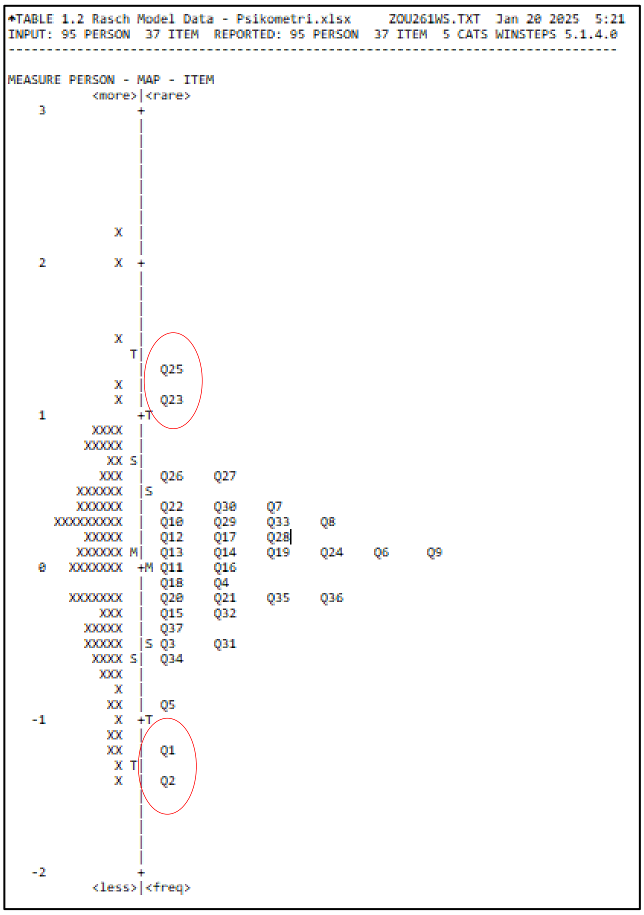

Figure 2. Wright Map output

Item function of Likert scale

The Likert Scale items function shows how the scale is used by respondents, for example, the probability of choosing a particular category (Nemoto & Beglar, 2014; Yamashita, 2022). To determine whether the functions of Likert scale options are effective or not, one can refer to Table 6.

Table 6. Category function

Scale	Observe Count (in percent)	Measure
1	18	-2.06
2	21	-0.77
3	11	-0.07
4	31	0.70
5	19	2.31

Respondents selected scale categories 1, 2, 3, 4, and 5, with at least 11% of respondents choosing scale 3 (about once a week). Furthermore, the size of the categories measure consistently increases from the lowest to the highest. Thus, it can be concluded that the functions of the Likert scale items work effectively.

Figure 3 illustrates the Andrich Rating Scale Model, which is used to examine whether each response option in a scale functions effectively in measuring the concept or trait intended to be assessed (Van Zile-Tamsen, 2017). This model provides information on whether each response option contributes meaningfully to the measurement. In other words, it helps ensure that the scale used is truly valid and reliable.

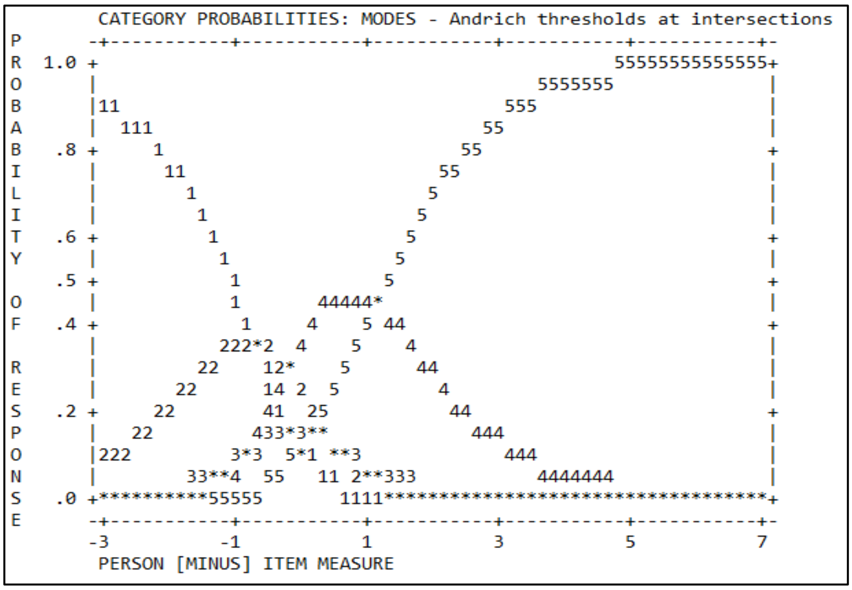

Figure 3. Scale category visualization

The visualization output also represents the probability of responses for each Likert scale category based on the difference between respondent ability (Person Measure) and item difficulty (Item Measure) (Massof, 2004). The curve shows the probability of an individual choosing a specific category at various levels of ability. The highest probability for each category is indicated by the curve at the top within a specific area along the X-axis.

The output shows five categories (1, 2, 3, 4, and 5), each represented by a probability curve. This indicates that respondents use each category according to their level of ability. No category is overlooked or unused. The category thresholds (points where two curves intersect) appear in sequence. The thresholds are in a logical order (ordered thresholds). This shows that the Likert scale functions well in distinguishing between categories. No disordered thresholds exist, meaning respondents can consistently understand and differentiate between the categories.

The range of respondent abilities covers the entire X-axis with smooth transitions between categories. All Likert categories function within a specific range of abilities without gaps. This Likert scale reflects appropriate levels of ability. There is some overlap between category curves, but the overlap is reasonable. Another entirely dominates no single category curve. These characteristics are consistent with Colledani et al. (2025) and (Tennant & Küçükdeveci, 2023) who emphasize that effective rating scales should show distinct ordered thresholds and reasonable overlap between adjacent categories and underscores the importance of examining monotonicity and avoiding disordered thresholds in polytomous items.

Dimensionality test

The dimensionality of the measurement tool was tested using Principal Component Analysis of Residuals (PCAR) in Winsteps. Based on the analysis results (see Figure 7), the total raw variance explained by the model was 44.1%, with raw unexplained variance in the first contrast at 12.8%. This indicates that the Likert scale used is still not entirely satisfactory and is influenced by other factors.

According to Tennant and Pallant (2006), one-dimensionality is considered adequate if the raw variance explained by the model exceeds 40% and the unexplained variance in the first contrast is less than 15%. Based on these results, the measurement tool can be considered unidimensional and measure only one main construct.

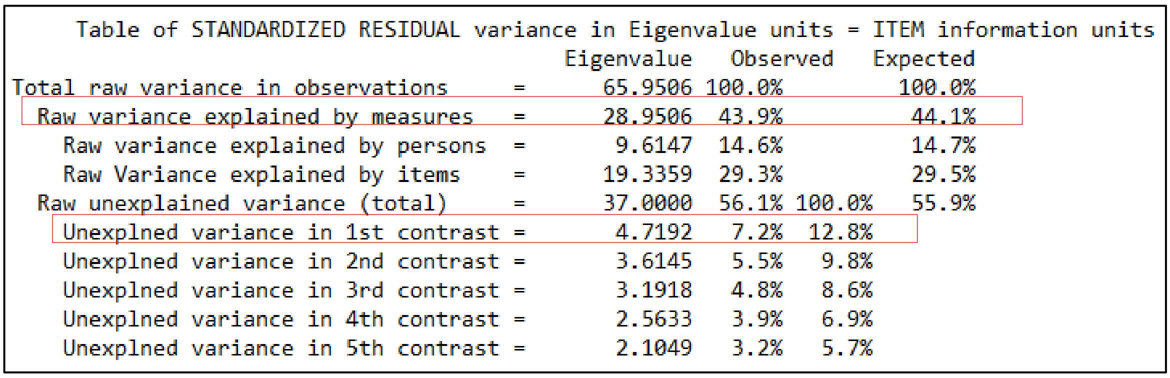

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = ITEM information units

		Eigenvalue	Observed		Expected
Total raw variance in observations	=	65.9506	100.0%		100.0%
Raw variance explained by measures	=	28.9506	43.9%		44.1%
Raw variance explained by persons	=	9.6147	14.6%		14.7%
Raw Variance explained by items	=	19.3359	29.3%		29.5%
Raw unexplained variance (total)	=	37.0000	56.1%	100.0%	55.9%
Unexplned variance in 1st contrast	=	4.7192	7.2%	12.8%	
Unexplned variance in 2nd contrast	=	3.6145	5.5%	9.8%	
Unexplned variance in 3rd contrast	=	3.1918	4.8%	8.6%	
Unexplned variance in 4th contrast	=	2.5633	3.9%	6.9%	
Unexplned variance in 5th contrast	=	2.1049	3.2%	5.7%	

Figure 4. Dimensionality map

Furthermore, the residual analysis showed no significant patterns indicating additional dimensions. The eigenvalues of unexplained variance in the second to fifth contrasts were also below the critical threshold, further supporting the assumption of one-dimensionality in the measurement tool.

Discussion

The application of the Rasch model in this study provided robust psychometric evaluation of the Preschool Home Mathematics Questionnaire (PHMQ). The results revealed a high level of internal consistency, as reflected in the person reliability of 0.93 and item reliability of 0.96. These values exceed the benchmark of 0.90, which is considered indicative of excellent reliability in educational assessments (Boone et al., 2014; Cook & Wind, 2024; Fisher, 2007). High person and item reliability indicate that the instrument effectively discriminates among different respondent abilities and that items are well-targeted along the latent trait continuum (Bond, 2015; Othman & Zaini, 2019), in this case, the construct of the home mathematics environment (HME).

The item fit analysis confirmed that most items performed well within the expected Rasch model parameters, supporting the unidimensionality and construct validity of the instrument. However, a few items (Q13, Q15, Q22, Q24, and Q26) were flagged as misfits based on outfit ZSTD and PTMEA-CORR statistics. Misfitting items can arise due to multidimensionality, ambiguous wording, or inconsistencies in respondent interpretation (Bond, 2015; Boone et al., 2014). As Bond (2015)explain, items with significant misfit may compromise measurement precision and should be reviewed for content alignment and clarity. For instance, item Q26 (“Maths-related websites”) may be less relevant in low-income settings with limited digital access, suggesting a need for contextual adaptation. To address these issues, it is recommended that misfitting items be carefully reviewed and revised for clarity, cultural relevance, and construct alignment (Suryani, 2018). This includes conducting expert validation and pilot testing of revised items (Piussi et al., 2025). If, after these steps, the items continue to show misfit, they should be considered for removal (Fischer et al., 2021; Johansson, 2025) to maintain the scale's psychometric integrity.

The functioning of the Likert scale categories also showed positive results. All five categories were used appropriately, exhibited increasing measures, and demonstrated ordered thresholds. This aligns with Andrich (1988) Rating Scale Model, which emphasizes that properly functioning categories should reflect monotonic increases along the latent continuum (Colledani et al., 2025). The smooth transitions between category curves and absence of disordered thresholds suggest that the scale effectively differentiates respondent engagement levels. These findings are supported by more recent studies (Colledani et al., 2025; Van Zile-Tamsen, 2017), which emphasize the importance of scale optimization in instrument development.

The Wright map further corroborated the comprehensive item targeting, illustrating a broad distribution of item difficulties matched with respondent abilities. This distribution is essential for ensuring measurement precision across the full ability spectrum (Boone et al., 2014; Linacre, 2002). Items Q1 and Q2, which appeared at the easier end of the continuum, assess more frequent and general activities, while Q23 and Q25, located at the more difficult end, capture specific or less common activities. This range is beneficial for identifying both low and high levels of home mathematics engagement, particularly in heterogeneous populations like those found in developing countries.

Regarding dimensionality, the Principal Component Analysis of Residuals (PCAR) showed that 44.1% of the total variance was explained by the Rasch dimension, and the first contrast accounted for only 12.8% of unexplained variance. These values meet the thresholds suggested by Tennant & Pallant (2006), confirming that the PHMQ measures a single dominant construct. Bond and Fox (2015) also highlight that the strength of the Rasch model lies in its capacity to isolate and quantify latent variables, which is clearly reflected in this finding.

Finally, this study underscores the strength of Rasch analysis not only for item refinement but also for evaluating response patterns and scale function. As Boone et al. (2014)emphasized, Rasch modelling provides diagnostic information that is often absent in classical test theory, especially in evaluating the appropriateness of rating scale usage and detecting subtle item misfits. The methodological rigor afforded by Rasch analysis enhances the interpretability and generalizability of findings, especially when instruments are adapted for culturally diverse and socioeconomically varied settings like Indonesia.

Conclusion

This study applied the Rasch model to evaluate the psychometric properties of the Preschool Home Mathematics Questionnaire (PHMQ). The findings demonstrated high person and item reliability, effective functioning of the Likert scale, and support for unidimensionality. These results indicate that the PHMQ is a valid and reliable tool for assessing the home mathematics environment among preschool-aged children in a developing country context.

However, this study has several limitations. The analysis was limited to basic Rasch procedures, primarily focusing on item and person reliability, item fit, Likert scale functionality, and unidimensionality. No further analysis such as Differential Item Functioning (DIF), multidimensional Rasch modeling, or longitudinal Rasch calibration was conducted. Additionally, the study was cross-sectional and did not incorporate qualitative insights that might help explain misfitting items or contextual challenges in implementation.

Given the simplicity of this validation-focused research, future studies should explore the PHMQ more extensively using Rasch analysis. This includes conducting DIF analysis to assess fairness across demographic groups, examining item bias, refining or reconstructing misfitting items, and investigating the instrument's performance over time or across different cultural settings. Such efforts would contribute to a more comprehensive understanding of the home mathematics environment and enhance the utility of the PHMQ in both research and practical applications.

Acknowledgment

The authors would like to express sincere gratitude to all the parents of preschool children who participated as respondents in this study. Their time, openness, and valuable insights greatly contributed to the success of this research.

Conflicts of Interest

We declare no conflict of interest regarding the publication of this manuscript. In addition, we have completed the ethical issues, including plagiarism, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies.

Funding Statement

This work received no specific grant from any public, commercial, or not-for-profit funding agency.

Author Contributions

Adi Cahya Nugraha: Conceptualization, writing - original draft, editing, and visualization; **Ira Rachmadani:** Writing - review & editing, formal analysis, and methodology; **Mardiana:** Writing - review & editing, formal analysis, and methodology; **Nur Alifa Deviar Refiyanti:** Writing - review & editing, formal analysis, and methodology; **Joko Soebagyo:** Validation, formal analysis, and methodology.

References

Al Ali, R. M. A., & Shehab, R. T. (2020). Psychometric properties of social perception of mathematics: Rasch model analysis. *International Education Studies*, *13*(12), 102–110. https://doi.org/10.5539/ies.v13n12p102

Andrich, D. (1988). *Rasch Models for Measuremen*. SAGE Publications.

Blevins-Knabe, B., & Musun-Miller, L. (1996). Number use at home by children and their parents and its relationship to early mathematical performance. *Early Development and Parenting*, *5*(1), 35–45. https://doi.org/10.1002/(SICI)1099-0917(199603)5:1<35::AID-EDP113>3.0.CO;2-0

Bond, T. (2015). *Applying the Rasch model*. Routledge. https://doi.org/10.4324/9781315814698

Bonifacci, P., Compiani, D., Affranti, A., & Peri, B. (2021). Home literacy and numeracy interact and mediate the relationship between socio-economic status and early linguistic and numeracy skills in preschoolers. *Frontiers in Psychology*, *12*(September), 1–16. https://doi.org/10.3389/fpsyg.2021.662265

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences*. Springer Netherlands. https://doi.org/10.1007/978-94-007-6857-4

Butterworth, B. (2005). The development of arithmetical abilities. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *46*(1), 3–18. https://doi.org/10.1111/j.1469-7610.2004.00374.x

Cahoon, A., Cassidy, T., Purpura, D. J., & Simms, V. (2021). Developing a rigorous measure of the pre-school home mathematics environment. *Journal of Numerical Cognition*, *7*(2), 172–194. https://doi.org/10.5964/jnc.6373

Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *Teachers College Record: The Voice of Scholarship in Education*, *112*(3), 579–620. https://doi.org/10.1177/016146811011200303

Colledani, D., González Pizzio, A. P., Devita, M., & Anselmi, P. (2025). Investigating the Functioning of Rating Scales With Rasch Models. *Assessment*, *32*(3), 434–446. https://doi.org/10.1177/10731911241245792

Cook, R. M., & Wind, S. A. (2024). Item response theory: A modern measurement approach to reliability and precision for counseling researchers. *Measurement and Evaluation in*

Counseling and Development, *57*(2), 116–135. https://doi.org/10.1080/07481756.2023.2301284
DeFlorio, L., & Beliakoff, A. (2015). Socioeconomic status and preschoolers' mathematical knowledge: The contribution of home activities and parent beliefs. *Early Education and Development*, *26*(3), 319–341. https://doi.org/10.1080/10409289.2015.968239
Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428
Evans, T. M., Kochalka, J., Ngoon, T. J., Wu, S. S., Qin, S., Battista, C., & Menon, V. (2015). Brain structural integrity and intrinsic functional connectivity forecast 6 year longitudinal growth in children's numerical abilities. *Journal of Neuroscience*, *35*(33), 11743–11750. https://doi.org/10.1523/JNEUROSCI.0216-15.2015
Fischer, L., Rohm, T., Carstensen, C. H., & Gnambs, T. (2021). Linking of Rasch-scaled tests: Consequences of limited item pools and model misfit. *Frontiers in Psychology*, *12*, 1–10. https://doi.org/10.3389/fpsyg.2021.633896
Fisher, W. P. (2007). Rating scale instrument quality criteria. *Rasch Measurement Transactions*, *21*, 1095.
Gashaj, V., Thaqi, Q., Mast, F. W., & Roebers, C. M. (2023). Foundations for future math achievement: Early numeracy, home learning environment, and the absence of math anxiety. *Trends in Neuroscience and Education*, *33*, 1–9. https://doi.org/10.1016/j.tine.2023.100217
Gilligan-Lee, K. A., Fink, E., Jerrom, L., Davies, M. P., Dempsey, C., Hughes, C., & Farran, E. K. (2023). Building numeracy skills: Associations between DUPLO® block construction and numeracy in early childhood. *Journal of Intelligence*, *11*(8), 1–18. https://doi.org/10.3390/jintelligence11080161
Girard, C., Bastelica, T., Léone, J., Epinat-Duclos, J., Longo, L., & Prado, J. (2021). The relation between home numeracy practices and a variety of math skills in elementary school children. *PLoS ONE*, *16*(9 September), 1–25. https://doi.org/10.1371/journal.pone.0255400
Hahn, R. A., & Barnett, W. S. (2023). Early childhood education: Health, equity, and economics. *Annual Review of Public Health*, *44*, 75–92. https://doi.org/10.1146/annurev-publhealth-071321-032337
Johansson, M. (2025). Detecting item misfit in Rasch models. *Educational Methods and Psychometrics*, *3*(2025), 1–58. https://doi.org/10.61186/emp.2025.5
Kurniati, A., Yuniati, S., & Rahmi, D. (2022). Media puzzle angka: Pengenalan angka pada anak tahap praoperasional (Teori Piaget) [Number puzzle media: Introduction to numbers for children in the preoperational stage (Piaget's theory)]. *Jurnal Cendekia: Jurnal Pendidikan Matematika,* *6*(3), 2846–2856.
Kvesic, L., Brkic, S., & Imre, A. (2020). Mathematical abilities of preschool children. *World Journal of Educational Research*, *7*(1), 167–186. https://doi.org/10.22158/wjer.v7n1p167
Lefevre, J. A., Kwarchuk, S. L., Smith-Chant, B. L., Fast, L., Kamawar, D., & Bisanz, J. (2009). Home numeracy experiences and children's math performance in the early school years. *Canadian Journal of Behavioural Science*, *41*(2), 55–66. https://doi.org/10.1037/a0014532
Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, *3*(1), 85–106.
Ling, M.-T., Pang, V., & Ompok, C. C. (2018). Measuring change in early mathematics ability of children who learn using games: Stacked analysis in Rasch measurement. In *Pacific Rim Objective Measurement Symposium (PROMS) 2016 Conference Proceedings* (pp. 215–226). Springer Singapore. https://doi.org/10.1007/978-981-10-8138-5_17
Lopez-Pedersen, A., Mononen, R., Aunio, P., Scherer, R., & Melby-Lervåg, M. (2023). Improving numeracy skills in first graders with low performance in early numeracy: A randomized controlled trial. *Remedial and Special Education*, *44*(2), 126–136. https://doi.org/10.1177/07419325221102537

Lu, L., Vasilyeva, M., & Laski, E. V. (2025). Home math environment as a mediator of socioeconomic differences in early math skills: A study of Chinese families from disparate backgrounds. *Developmental Psychology*, *61*(3), 417–431. https://doi.org/10.1037/dev0001918

Mallinckrodt, B., Miles, J. R., & Recabarren, D. A. (2016). Using focus groups and Rasch item response theory to improve instrument development. *The Counseling Psychologist*, *44*(2), 146–194. https://doi.org/10.1177/0011000015596437

Mantilla, A., & Edwards, S. (2019). Digital technology use by and with young children: A systematic review for the Statement on Young Children and Digital Technologies. *Australasian Journal of Early Childhood*, *44*(2), 182–195. https://doi.org/10.1177/1836939119832744

Massof, R. (2004). Likert and Guttman scaling of visual function rating scale questionnaires. *Ophthalmic Epidemiology*, *11*(5), 381–399. https://doi.org/10.1080/09286580490888771

McCamey, R. (2014). A primer on the one-parameter Rasch model. *American Journal of Economics and Business Administration*, *6*(4), 159–163. https://doi.org/10.3844/ajebasp.2014.159.163

Melhuish, E. C. (2011). Preschool matters. *Science*, *333*(6040), 299–300. https://doi.org/10.1126/science.1209459

Missall, K., Hojnoski, R. L., Caskie, G. I. L., & Repasky, P. (2015). Home numeracy environments of preschoolers: Examining relations among mathematical activities, parent mathematical beliefs, and early mathematical skills. *Early Education and Development*, *26*(3), 356–376. https://doi.org/10.1080/10409289.2015.968243

Mofreh, S. A. M., Ghafar, M. N. A., Omar, A. H. H., Mosaku, M., & Ma'ruf, A. (2014). Psychometric properties on lecturers' beliefs on teaching function: Rasch model analysis. *International Education Studies*, *7*(11), 47–55. https://doi.org/10.5539/ies.v7n11p47

Nadlifah, & Latif, M. A. (2024). Creative freedom and holistic growth: Implementation of the STEAM approach in Islamic early childhood education. *Golden Age: Jurnal Ilmiah Tumbuh Kembang Anak Usia Dini*, *9*(4), 689–702. https://doi.org/10.14421/jga.2024.94-09

Nemoto, T., & Beglar, D. (2014). Developing Likert-scale questionnaires campus reference data. *JALT2013 Conference Proceedings*.

Niklas, F., & Schneider, W. (2014). Casting the die before the die is cast: the importance of the home numeracy environment for preschool children. *European Journal of Psychology of Education*, *29*(3), 327–345. https://doi.org/10.1007/s10212-013-0201-6

Noa, K., & Pennucci, A. (2014). *Early childhood education for low-income students: A review of the evidence and benefit-cost analysis*. Washington State Institute for Public Policy.

Nugraha, A. C., & Muntazhimah, M. (2024). Analisis konten penelitian kemampuan numerasi awal pada anak usia dini: Tinjauan metodologi penelitian [Content analysis of research on early numeracy skills in young children: A review of research methodology]. *Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini*, *8*(5), 1174–1184. https://doi.org/10.31004/obsesi.v8i5.6027

Othman, N., & Zaini, M. A. Y. M. (2019). Reliability testing on instrument development measuring the attitude and perception towards academic and personality: Application of the Rasch model. *International Journal of Academic Research in Business and Social Sciences*, *9*(1), 191–209. https://doi.org/10.6007/IJARBSS/v9-i1/5388

Pan, Y., Hu, B. Y., Hunt, J., Wu, Z., Chen, Y., & He, M. (2023). Chinese preschool children's home numeracy experiences and their mathematical abilities. *Journal of Early Childhood Research*, *21*(1), 31–45. https://doi.org/10.1177/1476718X221125583

Papic, M. M., & Papic, C. (2025). Is confidence in mathematics pedagogy enough? Exploring early childhood teachers' mathematics beliefs and confidence. *Early Childhood Education Journal*. https://doi.org/10.1007/s10643-024-01840-4

Pesu, L., Viljaranta, J., & Aunola, K. (2016). The role of parents' and teachers' beliefs in children's self-concept development. *Journal of Applied Developmental Psychology*, *44*, 63–71. https://doi.org/10.1016/j.appdev.2016.03.001

Petronzi, D., Schalkwyk, G., & Petronzi, R. (2023). A pilot math anxiety storybook approach to normalize math talk in children and to support emotion regulation. *Journal of Research* *in Childhood Education*, *00*(00), 1–19. https://doi.org/10.1080/02568543.2023.2214591

Phillips, D. A. (2017). Commentary: Beyond more of the same – sustaining the benefits of preschool education – reflections on Bierman et al. (2016). *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *58*(2), 138–139. https://doi.org/10.1111/jcpp.12674

Piussi, R., Thomeé, R., Samuelsson, K., & Hamrin Senorski, E. (2025). Measurement properties of the Swedish version of the Knee self-efficacy scale: A Rasch analysis. *Journal of Experimental Orthopaedics*, *12*(2). https://doi.org/10.1002/jeo2.70306

Prusinski, E., Mahler, P. P., Collins, M., & Couch, H. (2023). Strengthening early childhood education and care in a "Childcare Desert." *Early Childhood Education Journal*, *51*(7), 1317–1333. https://doi.org/10.1007/s10643-022-01375-6

Purnomo, Y. W., Ainun, T. R., Nina, P. La, Utami, K., Wijayanti, R., & Ismail, S. N. (2022). Mother as a teacher at home: Challenges and opportunities for parental involvement in online mathematics learning for elementary school students. *New Educational Review*, *69*, 130–140. https://doi.org/10.15804/tner.2022.69.3.10

Purpura, D. J., & Lonigan, C. J. (2015). Early numeracy assessment: The development of the preschool early numeracy scales. *Early Education and Development*, *26*(2), 286–313. https://doi.org/10.1080/10409289.2015.991084

Reid, K. (2016). Counting on it: Early numeracy development and the preschool child. *Australian Council for Educational Research*, *2*, 1–16. https://research.acer.edu.au/learning_processes/19%0Awww.acer.edu.au

Saracho, O. N., & Spodek, B. (2010). Parents and children engaging in storybook reading. *Early Child Development and Care*, *180*(10), 1379–1389. https://doi.org/10.1080/03004430903135605

Seitz, M., & Weinert, S. (2022). Numeracy skills in young children as predictors of mathematical competence. *British Journal of Developmental Psychology*, *40*(2), 224–241. https://doi.org/10.1111/bjdp.12408

Starkey, P., Klein, A., & Wakeley, A. (2004). Enhancing young children's mathematical knowledge through a pre-kindergarten mathematics intervention. *Early Childhood Research Quarterly*, *19*(1), 99–120. https://doi.org/10.1016/j.ecresq.2004.01.002

Sumintono, B., & Widhiarso, W. (2014). *Aplikasi model Rasch untuk penelitian ilmu-ilmu sosial (edisi revisi) [Application of the Rasch model to social science research (revised edition)]*. Trim Komunikata Publishing House.

Suryani, Y. E. (2018). Aplikasi Rasch model dalam mengevaluasi intelligenz structure test (IST). *Psikohumaniora: Jurnal Penelitian Psikologi*, *3*(1), 73. https://doi.org/10.21580/pjpp.v3i1.2052

Tabatabaee-Yazdi, M., Motallebzadeh, K., Ashraf, H., & Baghaei, P. (2018). Development and validation of a teacher success questionnaire using the Rasch model. *International Journal of Instruction*, *11*(2), 129–144. https://doi.org/10.12973/iji.2018.11210a

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*. https://doi.org/10.3389/fresc.2023.1208670

Tennant, A., & Pallant, J. (2006). Unidimensionality Matters. *Rasch Measurement Transactions*, *20*, 1048–1051.

Toll, S. W. M., & Van Luit, J. E. H. (2014). Explaining numeracy development in weak performing kindergartners. *Journal of Experimental Child Psychology*, *124*(1), 97–111. https://doi.org/10.1016/j.jecp.2014.02.001

Van Rooijen, M., Verhoeven, L., & Steenbergen, B. (2011). Early numeracy in cerebral palsy: Review and future research. *Developmental Medicine and Child Neurology*, *53*(3), 202–209. https://doi.org/10.1111/j.1469-8749.2010.03834.x

Van Zile-Tamsen, C. (2017). Using Rasch analysis to inform rating scale development. *Research in Higher Education*, *58*(8), 922–933. https://doi.org/10.1007/s11162-017-9448-0

Vandermaas-Peeler, M., Ferretti, L., & Loving, S. (2012). Playing the ladybug game: Parent guidance of young children's numeracy activities. *Early Child Development and Care*, *182*(10), 1289–1307. https://doi.org/10.1080/03004430.2011.609617

Vandermaas-Peeler, M., Nelson, J., Bumpass, C., & Sassine, B. (2009). Numeracy-related exchanges in joint storybook reading and play. *International Journal of Early Years Education*, *17*(1), 67–84. https://doi.org/10.1080/09669760802699910

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (86th ed.). Harvard University Press.

Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. E. (2014). What's past is prologue: Relations between early mathematics knowledge and high school achievement. *Educational Researcher*, *43*(7), 352–360. https://doi.org/10.3102/0013189X14553660

Wolf, S., & McCoy, D. C. (2019). Household Socioeconomic Status and Parental Investments: Direct and Indirect Relations With School Readiness in Ghana. *Child Development*, *90*(1), 260–278. https://doi.org/10.1111/cdev.12899

Yamashita, T. (2022). Analyzing Likert scale surveys with Rasch models. *Research Methods in Applied Linguistics*, *1*(3), 100022. https://doi.org/10.1016/j.rmal.2022.100022

Yoshikawa, H., Weiland, C., & Brooks-Gunn, J. (2016). When does preschool matter? *Future of Children*, *26*(2), 21–36. https://doi.org/10.1353/foc.2016.0010

Yuniria, A. R., Atikah, C., & Asmawati, L. (2025). Interactive educational games: Boosting numeracy skills in 4-5 year-olds through multimedia. *AL-ISHLAH: Jurnal Pendidikan*, *17*(1). https://doi.org/10.35445/alishlah.v17i1.5503

Zhang, B., & Konstantopoulos, S. (2025). Short- and long-term effects of approaches to learning on reading and mathematics achievement in childhood: a fixed effects approach. *Educational Research and Evaluation*, 1–32. https://doi.org/10.1080/13803611.2024.2441785

Appendix

Questionnaire

No of items	**Statements**				
MATHS LITERACY					
1	In the past month, how often did you and your child engage in reading?				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
FREQUENCY OF HOUSEHOLD ACTIVITIES In the past month, how often did you and your child engage in the following?					
2	Counting				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
3	Write numbers				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
4	Scenarios number games (e.g. "If I have two toy cars and I take one away, how many cars do I have?")				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
5	Counting on fingers/hands				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
6	Watching number related TV shows (e.g. Number Jacks or Numtums)				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
7	Teaching about measurements (e.g. baking, height)				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
8	Sticker books				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
9	Sorting shapes				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
10	Rhyming TV shows involving numbers (e.g. Number Jacks)				

No of items	Statements				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
11	Play with jigsaws				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
12	Watch educational programs (e.g. Dora the Explorer)				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
13	Sorting objects by size				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
14	Comparing sets of objects (e.g. brother has more than mum)				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
15	Pairing/matching games				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
16	Playing with building blocks				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
17	Identifying names of written numbers				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
18	Counting out food, dinner plates, knifes and forks				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
19	Creating patterns with objects (e.g. arranging blocks into shapes)				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
20	Counting objects (e.g. ducks in bath, blocks, new toys, books)				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
21	Teaching about money (e.g. informal – playing shop or formal – buying sweeties)				

No of items	Statements				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
22	Time terminology (e.g. big hand, little hand)				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
23	Asking shape related questions (e.g. "how many sides does a circle have?")				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
TECHNOLOGY In the past month, how often did your child engage in the following?					
24	Maths applications (e.g. Number Jacks)				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
25	Maths related websites (e.g. coolmaths.com)				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
26	Racing games (e.g. the faster they complete sums, the faster the boat moves)				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
27	Size/matching apps (e.g. "put the big skirt on the small girl")				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
28	Add and subtraction games				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
29	Filling in the gap number games (e.g. what is next in the sequence?)				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
30	Maths related YouTube videos (e.g. NumTums)				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
SIBLINGS When your children are doing activities together that involve maths, what types of activities are they most likely to do together? Keeping this in mind, in the past month, how often have you and your child engage in the following?					
31	Counting objects together				

No of items	Statements				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
32	Arranging objects by size, shape or colour				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
33	Watching number related TV shows together (e.g. Number Jacks or Numtums)				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
34	Sing rhyming songs together (e.g. "1, 2, 3, 4, 5 once I caught a fish alive")				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
35	Reading books together that involve numbers (e.g. Hungry Caterpillar)				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
36	Timed games (e.g. hide and seek)				
	activity did not occur	few times a month	about once a week	few times a week	almost daily
37	Everyday activities that involve number (e.g. using money while shopping)				
	activity did not occur	few times a month	about once a week	few times a week	almost daily