Journal of Interdisciplinary Perspectives
ISSN Print: 2984-8288, ISSN Online: 2984-8385
Vol. 2, No. 6, pp. 6-20, June 2024

Validation of a Mathematics Teaching Assessment Scale for K-12 Basic Education: A Confirmatory Factor Analysis Approach

Keith A. Madrilejos
Jose V. Palma Senior High School, Quezon City, Philippines

Author email: keith.madrilejos001@deped.gov.ph

Date received: April 8, 2024
Date revised: April 12, 2024
Date accepted: April 16, 2024

Originality: 90%
Grammarly Score: 99%
Similarity: 10%

Recommended citation:
Madrilejos, K. (2024). Validation of a Mathematics teaching assessment scale for K-12 basic education: a confirmatory factor analysis approach. *Journal of Interdisciplinary Perspectives*, 2(6), 6-20.
https://doi.org/10.69569/jip.2024.0081

Abstract. Mathematics teaching in the K-12 Basic Education Curriculum encourages educators to adopt student-centered approaches and emphasize the development of 21st-century skills. Assessing the pedagogical practices of mathematics teachers using a specific tool validated through quantitative approaches is essential to ensure the effectiveness and quality of mathematics instruction. While previous research has focused on instrument validation in various areas, including self-efficacy of math teachers, mathematics teachers' anxiety, and online education, existing classroom observation tools have limitations, particularly in assessing public school teachers' performance specifically intended for mathematics teachers. This study aimed to validate a mathematics teaching assessment scale. The scale, constructed based on the Department of Education's classroom observation tool, initially comprised 22 items across four constructs: Mastery of the K-12 Mathematics, Teaching Methodologies, ICT Integration, and Assessment of Learning. The scale was validated through content validation, reliability testing, inter-item correlation, confirmatory factor analysis (CFA), convergent validity, and discriminant validity. A total of 687 mathematics teachers from four public schools in each of the five municipalities of the National Capital Region participated. Content validity was established by eight experts using the Scale Content Validity Index and Item-Content Validity Index, with indices exceeding 0.833, indicating validity. Each construct demonstrated desirable reliability indices based on Cronbach's alpha and composite reliability coefficients. Confirmatory Factor Analysis removed six indicators, resulting in a final scale of 16 items across four constructs. The validated instrument can effectively evaluate mathematics teachers' classroom teaching, focusing on 21st-century skills and student-centered approaches.

Keywords: Mathematics teaching assessment; 21st-century skills; Content validity; Reliability test; Construct validity.

1.0 Introduction

Student-centered teaching in mathematics shifts the focus from teacher-led instructions to practices that prioritize students' active participation, interests, and learning needs. This approach is grounded in the belief that learning is more meaningful and effective when students actively engage directly with the material, collaborate with peers, and reflect on their understanding (Keiler, 2018). Constructivism, a theory in the psychology of learning advocated by Jean Piaget and other educational psychologists, provides a foundational theoretical basis for student-centered teaching (WGU, 2020). It posits that learners construct their understanding and knowledge of the world, through experiencing things and reflecting on those experiences, which aligns with the emphasis on active engagement and discovery in student-centered mathematics education. Specifically, these student-centered approaches include collaborative activities, project-based exercises, and social constructivist learning (National Council of Teachers of Mathematics, 2020). Moreover, the Mathematics Teachers Association of the Philippines (n.d.) strongly advocates that teachers should motivate students to learn mathematics better through student-centered approaches of collaborative group learning. It is a well-established fact that teaching strategies significantly impact the academic success of learners. Joaquin et al. (2023) of the University of the Philippines NISMED affirmed that teaching styles had a significant relationship with the student's achievement based on the highest mean performance score.

Aside from learner-centered teaching practices, career and life skills are important key points that mathematics teachers should incorporate into his or her teaching. Buckle (2024) of Panorama Education pointed out that the increasing changing demands of the economy compelled many school districts to include 21st-century skills in their strategic plans to prepare students for college, career, and life. These life skills, so-called 21st-century skills, encompass problem-solving skills, information literacy, and critical thinking skills are also stressed in the twin goals of mathematics education in the K-12 curriculum (Department of Education, 2016). In the MATATAG Agenda, one of the supporting components of the framework to facilitate the development of mathematically proficient and critical problem solvers is the pedagogy which necessitates math teachers to become effective by employing relevant teaching strategies and assessment considering the learning phases of the students (Department of Education, 2023). Mathematics is an integral subject in the Philippines Department of Education K-12 curriculum that necessitates critical thinking abilities because it is crucial to daily life along with the development of other sciences (Artuz & Roble, 2021). By implementing engaging activities and real-world examples, teachers are encouraged to demonstrate the practical applications of mathematics, helping students deeply realize its relevance and importance.

Classroom assessment in mathematics is crucial as it provides teachers with valuable insights into students' understanding and helps tailor instruction to meet their needs. Additionally, assessing teaching methodologies ensures that educators are using effective strategies to promote learning and critical thinking in mathematics. Birgin and Yilmaz (2023) discussed in their paper that assessment of learning is essential to measure student achievement but it should include classroom assessment to uphold teaching and learning. Several studies stressed the effectiveness of teachers is an important ingredient in achieving quality learning (Lazarides et al., 2020; Yin, 2022; Nieto et al., 2023). This makes sense as Calaguas (2012) stated in his paper that it is important to consistently motivate teachers to do their best and be effective at all times. Kumar et al. (2024) emphasized that to measure the effectiveness of math teachers as this is crucial to student achievement, a tool must be developed. To determine the efficacy and effectiveness of teachers in their contributions to the educational process, there must be a validated instrument to address this concern.

One key measure of teaching and learning quality in basic education is the National Achievement Test (NAT) results for Grades 6 and Grade 12. However, NAT results from SY 2017-2018 up to SY 2022-2023 have indicated that learners in many schools had declining achievement and were still in the process of developing 21st-century skills such as problem-solving, information literacy, and critical thinking (Branzuela et al., 2023; Casildo et al., 2023; Philippine Star, 2024). This is a significant concern, especially considering the Philippines' 81st ranking in the 2022 PISA results, which highlights the challenges in students' critical thinking abilities. To address these issues, educators must adopt innovative teaching strategies that focus on developing these essential skills. These results suggest the need for assessing mathematics teachers on their content knowledge and teaching practices, particularly in fostering 21st-century skills among students. In this study, the researcher considered the entire National Capital Region (NCR) as the locale of the study with a sufficient number of participants.

In recent years up to now, the Department of Education Philippines followed the system of the Philippine Professional Standard for Teachers (PPST) involving the use of classroom observation tools to assess public school teachers' performance, determine their needs, and support them in professional enhancement. In general, this consists of domains such as content knowledge and pedagogy, learning environment, diversity of learners, curriculum and planning, and assessment and reporting (Department of Education, 2019). These were structured in a broad sense to assess and evaluate public school teachers in particular but not specifically for mathematics teachers. This classroom observation tool was only developed and validated for pre-service teachers using qualitative approaches (Lucero et al., 2024). In the field of mathematics education, measuring the teachers' self-efficacy was one of the areas in development and validation studies (Aldhafri et al., 2017). Another validated instrument to assess mathematics teachers' frequency of using classroom assessment tools was conducted by Birgin and Yilmaz (2023). The professional competence of mathematics teachers was investigated in Russia on a large scale; however, this study did not aim to create and validate a scale for evaluating professional competence

(Podkhodova et al., 2020). Some studies conducted before involved developing and validating a teaching effectiveness scale for college professors (Shahzad & Mehmood, 2019; Calaguas, 2012; Jislan et al., 2018). Hence, there is a need for an instrument to assess the teaching practices of the mathematics teachers in the basic education level. Due to the challenges and problems in the education sector as well as the lack of scale to assess mathematics teachers' pedagogical practices, the researcher intends to contextualize and validate an instrument intended for evaluating mathematics teachers in basic education in terms of teaching strategies in the 21st century.

2.0 Methodology

2.1 Research Design

This study used a quantitative-correlational research design because it utilized confirmatory factor analysis as a primary statistical technique which involves measuring the inter-correlations of constructs of mathematics teaching assessment and regression analyses between the constructs and their respective indicators. Studies that use factor analysis involve complex relationships among a large number of conceptually similar variables that apply to correlational design (Price et. al., 2015). Furthermore, this is supported by Tavakol and Wetzel (2020), who comprehensively discussed factor analysis, explaining the relationships between multiple variables/items of a scale.

2.2 Research Participants

The researcher, being a mathematics teacher from Quezon City, considered selecting the participants of this validation study in the National Capital Region. This research includes 687 math teacher-respondents selected from the following cities, each with its corresponding number: Quezon City (147), Makati City (133), Valenzuela City (140), Pasig City (137), and Marikina City (130), during the school year 2022-2023. A multi-stage random sampling approach was employed for respondent selection. Clusters were formed from municipalities within the National Capital Region (NCR), with five cities chosen randomly. Within each city, four public high schools were selected at random. The study included all mathematics teachers from these selected schools as respondents. Random sampling was conducted because it is an acceptable type of sampling procedure for quantitative studies. Mananghaya (2020) of the University of the Philippines Los Banos pointed out that random samples are only valid for statistical inferences. The sample size was determined based on the recommendation of White (2021) of 620 respondents for quantitative validation studies and Myers et al. (2011) emphasizing at least 300 participants after they conducted a Monte Carlo method for validation studies.

2.3 Research Instrument

Based on the indicators of the classroom observation tool developed by the Department of Education following the Philippine Professional Standards (PPST), the researcher contextualized the items of the scale. This classroom observation tool is used to assess the teaching performance in classroom-based settings of public school teachers. This developed tool was the primary basis for crafting the items and grouping these items into domains or constructs. The classroom observation tool originally consisted of domains such as content knowledge and pedagogy, learning environment, diversity of learners, curriculum and planning, and assessment and reporting (Lucero et al., 2024). In this study, the researcher simplified and contextualized these constructs into four such as mastery of the k-12 mathematics, teaching methodologies, ICT integration, and assessment of learning. Furthermore, the researcher consulted eight (8) experts (master teachers, head teachers, college professors, and high school principals) from the field of mathematics education in crafting the items and ensuring that these are aligned with the domains of 21st-century skills.

Mastery of the Content was considered a construct of mathematics teaching assessment based on the studies of George (2021) and Baumert and Kunter (2013). Teaching Methodology, being taken as one of the constructs, was included with specific indicators highlighting the relevance of developing critical thinking, higher-order thinking skills, differentiated instruction, developing career life skills, cooperative learning, and diversified teaching. Items under the teaching methodology domain were crafted from the studies of Sachdeva and Eggen (2021), Asmar et al. (2020), Bal (2023), Patnode and Lee (2021), Motseki and Kakoma (2023), Kaya and Kesan (2023), and the Mathematics Teachers Association of the Philippines (n.d.).

Moreover, the integration of ICT in teaching mathematics was considered a construct based on the existing literature (Arrieta, 2020; Turel et al, 2017; Ghavifekr et al., 2015; Gebremedhin et al., 2015). The researcher formulated the indicators of the Assessment of Learning based on DepEd Order No. 8 s. (2015), Chigonga (2019), and Teaching at UNSW (2023).

Unlike the template of the classroom observation tool in IPCRF (Individual Performance Commitment and Review Form), this instrument is structured on a Likert Scale form with five options. This may be used to specifically evaluate the classroom teaching performance of mathematics teachers under Basic Education emphasizing 21st-century skills.

Initial Development of the Scale

Based on the classroom observation tool for public school teachers in the IPCRF, the scale was primarily structured with four domains with the corresponding number of items such as Mastery of the K-12 Mathematics (5), Teaching Methodologies (6), ICT Integration (5), and Assessment of Learning (6). After conducting intensive discussions with the subject matter experts and a thorough examination of the related studies, the following indicators with respective codes were primarily constructed before the data analysis procedures of validation and reliability testing.

I. Mastery of the K-12 Mathematics (MC)

MC1. Demonstrated mastery of the topics in the K-12 mathematics curriculum.
MC2. Utilized knowledge of the mathematical content across various curriculum teaching areas.
MC3. Offered classroom-based examples during the lesson execution within the class's time frame.
MC4. Incorporated real-world examples to illustrate mathematical concepts.
MC5. Ability to explain complex mathematical concepts clearly and effectively.

II. Teaching Methodologies (T)

T1. Employed various teaching strategies that improve learner achievement in literacy and numeracy skills.
T2. Applied multiple teaching strategies to enhance student understanding of mathematical concepts.
T3. Utilized a variety of teaching approaches to cultivate critical and creative thinking, alongside other advanced cognitive abilities.
T4. Promoted cooperative and collaborative learning among students to make learning mathematics enjoyable.
T5. Implemented diverse teaching in mathematics lessons to accommodate different learning styles and individual variances.
T6. Integrated learning activities into mathematics instruction placing a focus on developing career and life skills.

III. ICT Integration (ICT)

ICT1. Used technology to enhance mathematical concepts and learning experiences.
ICT2. Demonstrated proficiency in using ICT for mathematical problem-solving and exploration.
ICT3. Incorporated technology-based assessments to evaluate student understanding and progress in mathematics.
ICT4. Facilitated online collaborative activities and discussions related to mathematical concepts.
ICT 5. Integrated interactive multimedia presentations to engage students in learning.

IV. Assessment of Learning (A)

A1. Designed, selected, organized, and applied diagnostic, formative, and summative assessment strategies aligned with curriculum standards.
A2. Provided timely and constructive feedback to students on their performance in mathematics.
A3. Employed diversified assessment methods to evaluate student understanding of mathematical concepts.
A4. Implemented strategies to address student misconceptions and promote a deeper understanding of mathematical concepts.
A5. Alignment of assessments with learning objectives and standards in mathematics.
A6. Incorporated technology-based assessments to evaluate student understanding and progress in mathematics.

2.4 Data Gathering Procedure

To accommodate time constraints, the survey was structured using Google Forms and distributed to selected schools' respondents via QR codes. To ensure understanding and compliance, virtual orientations and face-to-face meetings were held with school leaders and mathematics teachers. Clear instructions and the survey's purpose were explained. Before administering the survey, all necessary protocols were observed, including obtaining approval from Schools Division Offices in Quezon City, Makati City, Pasig City, Marikina City, and Valenzuela City. After a month, formal letters were sent to school principals for survey administration. Responses in Google Forms were monitored for completion, and follow-ups were conducted through various communication channels such as phone calls, emails, chat messages, and text messages. Data inspection and cleaning were performed to avoid missing entries before tabulation. The finalized data was then organized and summarized using Microsoft Excel and analyzed using SPSS (Statistical Package for Social Sciences) Version 26, with additional scale validation and analysis conducted using SPSS AMOS (Analysis of Moment Structures).

2.5 Ethical Considerations

Data gathering and storage of data were under strict conformance with ethical standards such as letter of informed consent, privacy, and confidentiality of the respondents. The proponent of this study attended a research ethics webinar and training before the data gathering started. The researcher only communicated with the respondents using their DepEd email accounts only. The researcher respected the respondents' privacy and ensured they were not pressured to participate. Upon submission of the accomplished informed consent, participants were designated with a unique respondent number. Names or any personal information was not collected from the respondents. Correspondences directly sent to the respondents who chose not to participate were deleted from the personal files of the researcher. Digital copies of the questionnaire, consent letter, demographic profile, and the respondents' answers were protected in a password-templated file saved in the researcher's designated folder.

To enhance data protection in the Google Form Survey, the researcher mandated that only participants with DepEd Gmail Accounts could complete the electronic survey form. After collecting all the necessary information, the researcher closed the Google Form for responses by adjusting the privacy settings. The original copy of all responses was retained solely in the Excel File of the organized dataset; following data analysis, results, and findings production, the researcher deleted both the spreadsheets and the survey to prevent unauthorized access. The tabulated data of the participant's responses were deleted after one year. None of the respondents' information or actual survey responses were stored in any public data storage. The findings from this study will be disseminated in scientific and international research journals, utilizing anonymous data.

2.6 Data Analysis

The content validity through the scale content validity (S-CVI) formula and item content validity (I-CVI) formula were first utilized in the study. The main reason for using validation statistics like the content validity index formula and content validity ratio is to assess the extent to which the items in a scale represent the content domain of the construct being measured. The researcher adhered to the procedures and acceptable value of the convent validity index suggested by Yusoff (2019). Since there were eight (8) experts being consulted, the required range of content validity index must be at least 0.833 (Lynn, 1986; Yusoff, 2019). Each item was evaluated by the experts using the rating scale from 1 to 4: 1 – The item is not relevant to the measured domain, 2 – The item is somewhat relevant to the measured domain, 3 – The item is relevant to the measured domain, and 4 – The item is highly relevant to the measured domain.

The internal consistency of the items per construct was determined using Cronbach's Alpha Coefficient. Based on the framework of Adamson and Prion (2013), and Taber (2018), the reliability index of a scale must be at least 0.70. As suggested by Magno (2021), the Mathematics Teaching Assessment, a multidimensional variable, requires the calculation of a reliability coefficient for each construct.

Inter-item correlations were computed to determine the degree of association or relationship between different items per construct in the Mathematics Teaching Assessment, aiding in understanding the internal consistency and reliability of the instrument. This was also determined to inspect if there are highly correlated items because it could indicate redundancy in the assessment tool, suggesting that some items may be measuring the same underlying construct and could potentially be simplified or combined. If inter-item correlations are 0.70 or higher, it suggests that the items are measuring nearly the same thing, and it may be advisable to remove one of them (Röschel and Wagner, 2021). Correlations between 0.30 and 0.70 are considered desirable, as lower values may suggest a lack of consistency among items, while higher values could indicate that some items are redundant (Hawker, 2008; Streiner & Norman, 1991).

Confirmatory Factor Analysis (CFA) was used to examine the model fit of the instrument, composite reliability, convergent validity, and discriminant validity. The scale model fit must have a Chi-Square p-value of greater than 0.05 and $2 < \chi^2/df \leq 3$ (Muller et al., 2003). However, if the Chi-Square result is not significant and the results of other fit measures are desirable, then the model attains an acceptable fit (PARSSU, 2022). Comparative Fit Index (CFI) must be 0.90 or higher (Capinding, 2023). The Tucker-Lewis Index must have values greater than 0.95 (Magno, 2021). On the other hand, Finster and Feldman (2019) indicated that TLI must have values of at least 0.90. The Normed Fit Index (NFI) must be at least 0.90 or between 0.90 and 0.95 inclusive (Muller et al., 2003; Byrne, 1994). The Root Mean Square Error of Approximation (RMSEA) value must be 0.05 and 0.08 (Muller et al., 2003; Fabrigar et al, 1999; Magno, 2021). A model is considered an acceptable fit if its Relative Fit Index (RFI) is closer to 1, its Incremental Fit Index (IFI) must be greater than 0.90, and its Parsimony-Adjusted Measures Index (PNFI) should be higher than 0.50 (Mattan et al., n.d.).

3.0 Results and Discussion

3.1 Content Validity Results

Table 1. Content Validity of the Mastery of the K-12 Mathematics Construct

Item Code	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Rater 6	Rater 7	Rater 8	Number of Agreement	I-CVI
MC1	3	3	3	4	4	4	3	2	7	0.875
MC2	4	4	4	3	4	4	4	3	8	1
MC3	4	4	4	3	4	4	3	4	8	1
MC4	4	4	4	4	4	4	4	4	8	1
MC5	2	2	2	3	3	1	4	2	3	0.375
									S-CVI	0.85
									Total Agreement	4

Table 1 presents the content validation of the Mastery of the K-12 Mathematics Construct. As shown in the table, the overall scale-content validity index of 0.85 indicates the scale's validity, as it falls within the recommended range of at least 0.833, as discussed by Yusoff (2019). However, it is suggested that MC5 (Ability to explain complex mathematical concepts clearly and effectively) be revised or deleted from the scale, as only three raters agreed that it is relevant to mastery of the content. Conversely, MC1, MC2, MC3, and MC4 had higher item-content validity indices of 0.875 to 1.00, indicating their validity. This evaluation highlights the importance of thorough item analysis in scale development, ensuring that each item contributes meaningfully to the overall construct.

Table 2. Content Validity of the Teaching Methodologies Construct

Item Code	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Rater 6	Rater 7	Rater 8	Number of Agreement	I-CVI
T1	2	2	1	2	2	3	4	4	3	0.375
T2	3	3	3	4	3	4	3	4	8	1
T3	3	4	4	4	4	4	3	4	8	1
T4	4	4	4	4	4	4	4	4	8	1
T5	4	4	4	3	4	4	4	3	8	1
T6	4	4	4	3	3	4	4	3	8	1
									S-CVI	0.896
									Total Agreement	5

Shown in Table 2 are the results of the content validation of the items under Teaching Methodologies. The overall scale-content validity index of 0.896 indicates strong validity, meeting the required threshold based on Lynn (1986) and Yusoff (2019). However, T1 (Employed various teaching strategies that improve learner achievement in literacy and numeracy skills) stands out as needing improvement or deletion due to its limited perceived relevance by only three raters. Conversely, items T2 to T6 demonstrated high item-content validity indices ranging from 0.896 to 1.00, suggesting their strong alignment with the construct of Teaching Methodologies. This content analysis underscores the importance of reviewing and refining individual items to ensure the overall quality and validity of the scale.

Table 3. Content Validity of the ICT Integration Construct

Item Code	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Rater 6	Rater 7	Rater 8	Number of Agreement	I-CVI
ICT 1	4	3	3	3	4	4	3	3	8	1
ICT 2	4	4	4	3	3	4	3	4	7	0.875
ICT 3	4	3	4	3	4	4	3	4	8	1
ICT 4	4	4	3	4	4	3	3	4	8	1
ICT 5	2	2	2	3	2	2	4	4	3	0.375
									S-CVI	0.85
									Total Agreement	4

Table 3 shows the results of the content validation of the scale under the ICT Integration construct. The overall scale-content validity index of 0.85 ensures the validity of most items (Yusoff, 2019), yet ICT 5 (Integrated interactive multimedia presentations to engage students in learning) was deemed relevant by only three raters, suggesting a need for its revision or removal. Conversely, ICT 1, ICT 2, ICT 3, and ICT 4 were rated as valid, with item-content validity indices ranging from 0.875 to 1.00 based on expert ratings.

Table 4. Content Validity of the Assessment of Learning Construct

Item Code	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Rater 6	Rater 7	Rater 8	Number of Agreement	I-CVI
A1	4	4	3	4	4	3	4	4	8	1
A2	4	3	4	4	3	4	3	4	8	1
A3	4	4	3	4	3	4	3	4	8	1
A4	2	1	3	3	3	2	2	1	3	0.375
A5	2	3	3	2	3	2	2	2	3	0.375
A6	1	2	2	3	3	2	2	3	3	0.375
									S-CVI	0.688
									Total Agreement	3

Presented in Table 4 are the content validation results of the Assessment of Learning Construct taken from the ratings of the eight experts. The overall scale-content validity index of 0.688 falls short of the standard required by Yusoff (2019) and Lynn (1986). Three items, namely A4, A5, and A6, exhibited a low item content validity index of 0.375, prompting the researcher to revisit and delete these items. On the other hand, items A1, A2, and A3 were retained.

3.2 Reliability Test Results

Table 5. Reliability Indices of the Scale

Constructs	Number of Items	Cronbach's Alpha	Interpretation
Mastery of the K-12 Mathematics	4	0.850	Good
Teaching Methodologies	5	0.834	Good
ICT Integration	4	0.871	Good
Assessment of Learning	3	0.811	Good

Table 5 reveals the reliability coefficient of each construct. The computation of the reliability index per construct was done after deleting the non-significant items from the scale. The matrix above illustrates the computed Cronbach alpha coefficients for each construct, with values exceeding 0.80 indicating strong internal consistency among items. Specifically, coefficients of 0.850, 0.834, 0.871, and 0.811 fall within the 0.80 to 0.89 range, categorizing them as "good" items according to Arof et al. (2018). These coefficients signify the reliability of the constructs, suggesting that the items consistently measure the intended concepts within each construct of the Mathematics Teaching Assessment Scale.

3.3 Inter-Item Correlations

Shown in Table 6 are the inter-item correlations of the pairs of items under Mastery of the K-12 Mathematics. This inter-item correlation was computed after performing content validation and confirmatory factor analysis. Notably, no correlation coefficients exceeding 0.70 were observed, indicating that none of the items are overly redundant or measuring the same underlying construct to a high degree. This aligns with the recommendation by Röschel and Wagner (2021) to retain these items in the scale, as they each contribute distinct information.

Table 6. Inter-Item Correlation Matrix for Mastery of the K-12 Mathematics

Variable	1	2	3	4
1. MC1	1.00	-	-	-
2. MC2	0.598	1.00	-	-
3. MC3	0.486	0.578	1.00	-
4. MC4	0.550	0.668	0.641	1.00

Furthermore, the computed correlation coefficients among the items ranged from 0.498 to 0.668, suggesting moderate to strong relationships between some items. These findings underscore the scale's coherence and the complementary nature of its items in measuring the intended construct, further supporting its validity and utility in assessing mastery of K-12 mathematics.

Table 7. Inter-Item Correlation Matrix for Teaching Methodologies

Variable	1	2	3	4	5
1. T2	1.00	-	-	-	-
2. T3	0.579	1.00	-	-	-
3. T4	0.530	0.566	1.00	-	-
4. T5	0.482	0.562	0.666	1.00	-
5. T6	0.360	0.395	0.408	0.463	1.00

Revealed in Table 7 the inter-item correlations between the pairs of items of the Teaching Methodologies Construct. The calculation of this inter-item correlation followed content validation and confirmatory factor analysis. Results in this matrix demonstrate that akin to the findings in Table 6, no highly correlated items were observed, as none of the correlation coefficients exceeded 0.70. This aligns with the recommendations of Hawker (2008), Streiner and Norman (1991), and Röschel and Wagner (2021), supporting the retention of these indicators in the scale due to their unique contributions. The computed inter-item correlations, ranging from 0.360 to 0.666, indicate moderate to strong relationships between some items. These results highlight the scale's robustness and the complementary nature of its items, affirming its effectiveness in measuring the Teaching Methodologies Construct.

Table 8. Inter-Item Correlation Matrix for ICT Integration

Variable	1	2	3	4
1. ICT1	1.00	-	-	-
2. ICT2	0.620	1.00	-	-
3. ICT3	0.601	0.627	1.00	-
4. ICT4	0.626	0.637	0.657	1.00

Table 8 shows the inter-item correlation results among the pairs of items in the ICT Integration Construct. After completing content validation and confirmatory factor analysis, this inter-item correlation was computed. The correlation coefficients for each item, ranging from 0.601 to 0.675, indicate moderate to strong relationships between these items. This suggests that the items within the ICT Integration Construct are related but not redundant, highlighting their collective ability to capture different facets of ICT integration in education. These findings support the construct's internal consistency and coherence, reinforcing its validity in assessing the integration of information and communication technology in educational settings. The results affirmed the recommendations made by Röschel and Wagner (2021).

Table 9. Inter-Item Correlation Matrix for Assessment of Learning

Variable	1	2	3
1. A1	1.00	-	-
2. A2	0.525	1.00	-
3. A3	0.628	0.620	1.00

Table 9 presents the inter-item correlation between the pairs of items under Assessment of Learning Construct. This inter-item correlation was derived following the completion of content validation and confirmatory factor analysis procedures. The correlation coefficients, ranging from 0.525 to 0.628, indicate moderate to strong relationships between these items. These findings suggest that while the items are related, they each contribute unique information to the Assessment of Learning Construct. This supports the construct's validity and indicates that the items collectively assess various aspects of learning assessment effectively. As suggested by Hawker (2008), Streiner and Norman (1991), and Röschel and Wagner (2021), inter-item correlations below 0.70 indicate that items are not redundant and can therefore be retained in the scale.

3.4 Removal of the Non-Significant Items from the Scale

Table 10. Deleted Items from the Scale

Item Code	Indicator	Factor Loadings	p-value
MC5	Ability to explain complex mathematical concepts clearly and effectively	-0.049	0.882
T1	Employed various teaching strategies that improve learner achievement in literacy and numeracy skills	-0.034	0.328
ICT5	Integrated interactive multimedia presentations to engage students in learning	0.115	0.732
A4	Implemented strategies to address student misconceptions and promote a deeper understanding of mathematical concepts	0.232	0.992
A5	Alignment of assessments with learning objectives and standards in mathematics	0.225	0.323
A6	Incorporated technology-based assessments to evaluate student understanding and progress in mathematics	0.189	0.644

Table 10 reveals the deleted items from the scale. After conducting the content validity, reliability test, and confirmatory factor analysis, it was found that the above items were not significant and did not contribute to their respective constructs based on factor loadings meaningfully.

In terms of scale development, this means that the above items failed to meet the threshold for significant factor loadings and p-values do not sufficiently align with the underlying construct they are intended to measure. These items were dropped from the scale as a result of low factor loadings ($r < 0.50$) to their corresponding latent variables. Hair et al. (2009), Ximenez (2015), and Oswald (2019) stressed that the p-values associated with the indicator loadings must be less than or equal to 0.05 and the factor loadings must be higher than or equal to 0.50. The results obtained above decreased the total number of items on the scale from 22 to 16. Six non-significant indicators were removed from the instrument.

Table 11. Factor Loadings in CFA

Item Code	Indicator	Factor Loading
MC1	Demonstrated mastery of the topics in the K-12 mathematics curriculum.	0.632
MC2	Utilized knowledge of the mathematical content across various curriculum teaching areas.	0.760
MC3	Offered classroom-based examples during the lesson execution within the class's time frame.	0.770
MC4	Incorporated real-world examples to illustrate mathematical concepts.	0.854
T2	Applied multiple teaching strategies to enhance student understanding of mathematical concepts.	0.636
T3	Utilized a variety of teaching approaches to cultivate critical and creative thinking, alongside other advanced cognitive abilities.	0.725
T4	Promoted cooperative and collaborative learning among students to make learning mathematics enjoyable.	0.777
T5	Implemented diverse teaching in mathematics lessons to accommodate different learning styles and individual variances.	0.798
T6	Integrated learning activities into mathematics instruction placing a focus on developing career and life skills.	0.592
ICT1	Used technology to enhance mathematical concepts and learning experiences.	0.773
ICT2	Demonstrated proficiency in using ICT for mathematical problem-solving and exploration.	0.795
ICT3	Incorporated technology-based assessments to evaluate student understanding and progress in mathematics.	0.788
ICT4	Facilitated online collaborative activities and discussions related to mathematical concepts.	0.815
A1	Designed, selected, organized, and applied diagnostic, formative, and summative assessment strategies aligned with curriculum standards.	0.743
A2	Provided timely and constructive feedback to students on their performance in mathematics.	0.740
A3	Employed diversified assessment methods to evaluate student understanding of mathematical concepts.	0.831

3.5 Confirmatory Factor Analysis (CFA)

Table 11 reveals the results of conducting the Confirmatory Factor Analysis in terms of the factor loadings. Factor loadings in Confirmatory Factor Analysis (CFA) are similar to regression coefficients (Cambridge University Press, 2021). In this study, these represent the relationships between observed variables (items of the Mathematics Teaching Assessment scale) and latent variables (factors or constructs) in the model. These loadings indicate the strength and direction of the linear relationship between the observed variables and the underlying latent factors such as Mastery of the K-12 Mathematics, Teaching Methodologies, ICT Integration, and Assessment of Learning. As observed in the table, each factor loading exceeds 0.50, with computed p-values less than 0.05, indicating the retention of items in the scale. Hair et al. (2009), as discussed by PARSSU (2022), emphasized that factor loadings must surpass the threshold of 0.50 and be statistically significant.

Table 12. Model Fit Statistics

Fit Measure	Value	Threshold for Acceptable Fit	Verbal Interpretation
CMIN	314.863, $p < 0.001$	$2\text{df} < \chi^2 \leq 3\text{df}$	Significant
Degrees of Freedom (df)	96	-	-
CMIN/df	3.280	$2 < \chi^2/\text{df} \leq 3$	Not A Good Fit
NFI	0.943	$0.90 \leq \text{NFI} < 0.95$	Acceptable
RFI	0.928	Closer to 1.00	Acceptable
IFI	0.960	$\text{IFI} \geq 0.90$	Acceptable
TLI	0.949	$\text{TLI} \geq 0.90$	Acceptable
CFI	0.959	$0.95 \leq \text{CFI} < 0.97$	Acceptable
PNFI	0.754	$\text{PNFI} > 0.50$	Acceptable
RMSEA	0.058	$0.05 < \text{RMSEA} \leq 0.08$	Acceptable

Table 12 shows the model fit indices. The obtained values revealed that the majority of the required fit measures were satisfied. However, it was noticed that the CMIN p-value is less than 0.05 and CMIN/df reached 3.280 which does not fall between the range of $\mathbf{2 < \chi^2/df \leq 3}$. This suggests that the model does not fit the data based on the Chi-Square test. This discrepancy could be attributed to the Chi-Square's sensitivity to large sample sizes.

Nonetheless, other fit statistics were utilized to evaluate the data. The normed Fit Index (NFI) of 0.943 is higher than 0.90, the Relative Fit Index (RFI) is closer to 1.00, the Incremental Fit Index (IFI) of 0.960 is above the threshold of 0.90, the Tucker-Lewis Index (TLI) of 0.949 is greater than 0.90, Comparative Fit Index (CFI) of 0.959 falls within the range of [0.95,0.97), Parsimony-Adjusted Measures Index (PNFI) of 0.754 is greater than 0.50, and the Root Mean Square Error of Approximation (RMSEA) of 0.058 falls within the ranges of values between 0.05 and 0.08. These findings collectively suggest that while the Chi-Square test indicates poor fit, the model fits reasonably well based on the other fit indices.

Figure 1 illustrates the Four-Factor Model Diagram of the Mathematics Teaching Assessment Scale for K-12 Basic Education, displaying standardized estimates. The inter-factor correlations observed were as follows: Mastery of the K-12 Mathematics and Teaching Methodologies (r = 0.482), Mastery of the K-12 Mathematics and ICT Integration (r = 0.517), Mastery of the K-12 Mathematics and Assessment of Learning (r = 0.566), Teaching Methodologies and ICT Integration (r = 0.682), Teaching Methodologies and Assessment of Learning (r = 0.584), and Assessment of Learning and ICT Integration (r = 0.468).

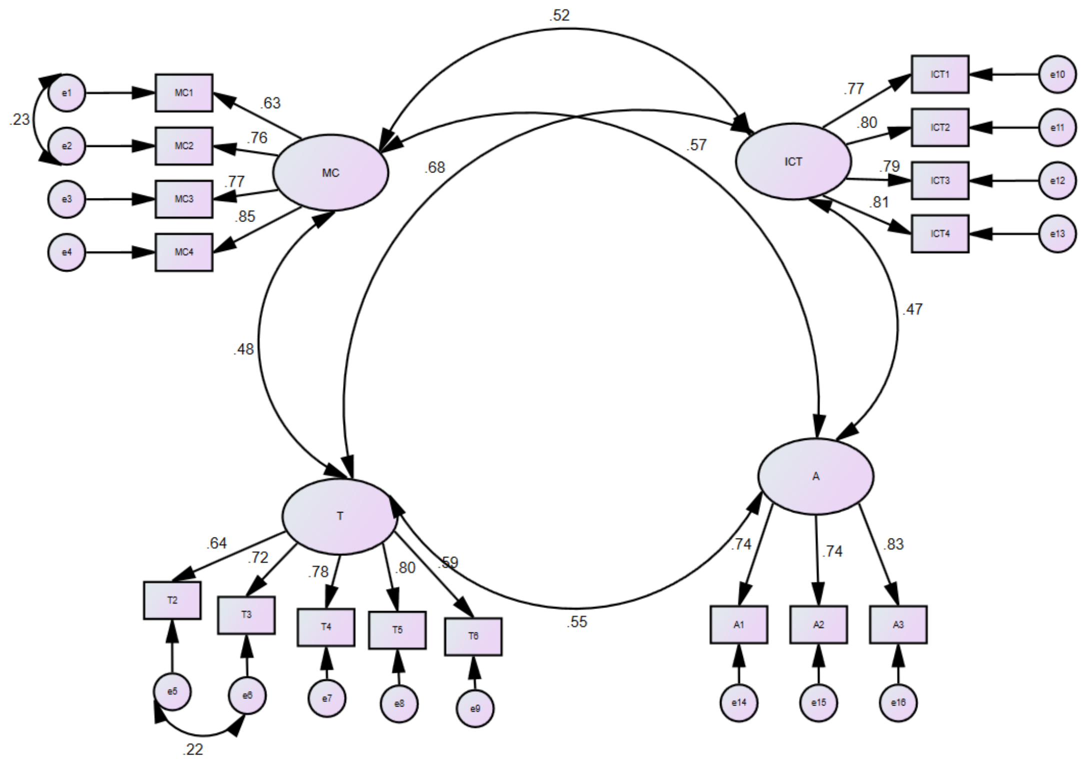

Figure 1. Four-Factor Measurement Model of Mathematics Teaching Assessment for K-12 Basic Education

Furthermore, the error terms of items MC1 and MC2, as well as items T2 and T3, were found to be correlated. This correlation can be attributed to shared method variance and the common wording of the statements. The above correlation values indicate that the constructs are not too correlated with each other. The moderate correlations (between 0.30 and 0.70) between related constructs are common, as they indicate some degree of overlap or shared variance without indicating redundancy (Berry et al., 2006). This range suggests that the constructs are related but distinct, providing evidence of convergent validity without raising concerns about discriminant validity—that is, the constructs are not so highly correlated that they appear to measure the same underlying dimension (Lesia et al., 2024).

Table 13. Composite Reliability

Latent Variable	Number of Items	Composite Reliability
Mastery of the K-12 Mathematics	4	0.834
Teaching Methodologies	5	0.842
ICT Integration	4	0.871
Assessment of Learning	3	0.816

Shown in Table 13 are the coefficients of the composite reliability. This is another measure of reliability to consider when conducting Confirmatory Factor Analysis. Based on the computed values, each construct in the scale demonstrates desirable internal consistency, as indicated by coefficients exceeding 0.70. This aligns with the conservative criterion set by Shamim (2022), Nunally and Berstein (1994), as discussed by PARSSU (2022). Specifically, Mastery of the K-12 Mathematics has a composite reliability index of 0.834, Teaching Methodologies has 0.842, ICT Integration has 0.871, and Assessment of Learning has 0.816.

These results indicate that the items within each construct are highly reliable and consistently measure the underlying construct. This strengthens the overall validity and reliability of the scale for assessing mathematics teaching in K-12 education. These results were also aligned with the obtained Cronbach's Alpha in Table 5.

Table 14. Convergent and Discriminant Validity

	AVE	MSV	MaxR(H)	Teaching Methodologies	Mastery of the K-12 Mathematics	ICT Integration	Assessment of Learning
Teaching Methodologies	0.504	0.465	0.849	**0.710**			
Mastery of the K-12 Mathematics	0.575	0.320	0.861	0.482	**0.758**		
ICT Integration	0.629	0.465	0.872	0.682	0.517	**0.793**	
Assessment of Learning	0.597	0.320	0.824	0.548	0.566	0.468	**0.772**

Table 14 presents the results of performing convergent validity and discriminant validity as the last phase of conducting Confirmatory Factor Analysis (CFA). According to Kock (2020), a research instrument demonstrates good convergent validity when respondents interpret the item statements associated with each latent construct in the same way as intended by the designers. The Average Variance Extracted (AVE) of each constructed was calculated to determine the proportion of variance in the items explained by the latent variable (PARSSU, 2022). As observed from the matrix, each average variance extracted exceeds the required value of greater than 0.50 as indicated by Shamim (2022), and Fornell and Larker (1981). Hence, the model has achieved the desired convergent validity. These results were connected with the reported data in Table 11 which detailed the significant factor loadings higher than 0.70.

In terms of discriminant validity, the square root of the average variance extracted was calculated in each latent variable. To attain discriminant validity, the square root of the average variance extracted should be higher than the correlations with the constructs (Fornell & Larker, 1981). For teaching methodologies, the square root of 0.504 is 0.710 which is higher than the inter-construct correlations such as 0.482, 0.682, and 0.548. For the Mastery of K-12 Mathematics, the square root of the average variance extracted of 0.758 exceeds the inter-construct correlations such as 0.517 and 0.566. For ICT Integration, the square root of the average variance extracted of 0.793 exceeds the inter-construct correlation of 0.468. Lastly, the square root of the average variance extracted of 0.772 for Assessment of Learning is higher than all inter-construct correlations. These results indicate that the model exhibited good discriminant validity, suggesting that respondents were able to differentiate between the indicators of each construct and did not confuse them with the item statements associated with other constructs (Kock, 2020).

4.0 Conclusion

This study offers a validated and reliable instrument for assessing mathematics teachers' pedagogical practices, aligning with the objectives of K-12 education with 21st-century skills. The research utilized the scale-content validity index and item-content validity index to determine whether the research instrument effectively measures the intended constructs based on the evaluation made by the eight experts on the subject matter. This resulted in content validity indices of greater than 0.833 in the three constructs. However, the Assessment of Learning Construct had a scale content validity index below 0.833, prompting the removal of three items to enhance the scale's validity. This instrument is considered reliable based on the computed Cronbach's alpha coefficient of greater than 0.80 in each construct. It makes the items internally consistent as proven by the empirical evidence. The inter-item correlations showed that there were no highly correlated items implying that the items in the scale are not redundant and measure distinct aspects of the construct. The confirmatory factor analysis (CFA) validated the four constructs of the mathematics teaching assessment scale. These four constructs are as follows: Mastery of the K-12 Mathematics, Teaching Methodologies, ICT Integration, and Assessment of Learning. The teaching performance of mathematics teachers at the basic education level can be measured using the mastery of the K-12 mathematics scale and the teaching methodologies scale. The teaching Methodologies scale is composed of items aligned to multiple or diverse teaching strategies, enhancement of creative and critical thinking skills, promotion of collaborative and cooperative learning, differentiated instruction, and emphasis on developing life and career skills. The ICT Integration scale can assess the mathematics teachers' skill to incorporate and utilize technology in their teaching and assessment. The Assessment of Learning Scale can measure mathematics teachers' skills in providing relevant and aligned assessments to standards and competencies, utilizing diverse ways of assessing learners and providing constructive feedback to students.

Public school teachers, specifically the Master Teachers and Head Teachers, can use the validated instrument to determine the level of classroom teaching performance of mathematics teachers and use it as a benchmark to craft training on pedagogical styles and various interventions. This is to make the mathematics teachers well-equipped with the necessary teaching skills for 21st-century learners under the K-12 curriculum. Researchers in the field can utilize this instrument for extended projects. Future researchers can re-assess and evaluate the scale or contextualize it for further investigations.

5.0 Contributions of Authors

This is single-author research.

6.0 Funding

This work received no specific grant from any funding agency.

7.0 Conflict of Interests

The author declares no conflicts of interest about the publication of this paper.

8.0 Acknowledgment

The researcher extends heartfelt gratitude to everyone who supported this endeavour.

9.0 References

15-Day Capacity Development Program on Data Analysis (2021). Sponsored by STAR Training Center. Title of the Topic: Factor Analysis. Speaker: Dr. Carlo P. Magno.

Adamson, K., & Prion, S. (2013). Reliability: Measuring Internal Consistency Using Cronbach's α. Clinical Simulation in Nursing. 9, 179-180, https://doi.org/10.1016/j.ecns.2012.12.001

Ali, S., & Maat, S. M. (2019). Exploring the 21st Century Teaching and Learning Practice among Mathematics Secondary School Teachers. International Journal of Academic Research in Progressive Education and Development, 8(2), 361–378. http://dx.doi.org/10.6007/IJARPED/v8-i2/5990

Alkharusi, B., Aldhafri, S., Al-Hosni, K., Al-Busaidi, S., Ambusaidi, A. & Alrajhi, M. (2017). Development and Validation of a Scale for Measuring Mathematics Teaching Self Efficacy for Teachers in the Sultanate of Oman. International Journal of Instruction, 10(3), 143-158. https://doi.org/10.12973/iji.2017.10310a

Arrieta, G.S. (2020). Assessment of the ICT Integration in Teaching Math and Science in High School: Basis for an ICT Integration Program. Jurnal Pendidikan MIPA. 21(1), 95-108. http://dx.doi.org/10.23960/jpmipa/v21i1.pp95-108

Artuz, J. & Roble, D. (2021). Developing Students' Critical Thinking Skills in Mathematics Using Online-Process Oriented Guided Inquiry Learning (O-POGIL). American Journal of Educational Research. https://www.sciepub.com/EDUCATION/abstract/13518

Asmar, A., Arnellis, A., & Sriningsih, R. (2020). Implementation of the 21st Century Skills to High Order Thinking Students' Mathematical in Senior High School Padang. Journal of Physics: Conference Series. International Conference on Mathematics and Mathematics Education. Volume 1554, https://iopscience.iop.org/article/10.1088/1742-6596/1554/1/012080

Bal, A.P. (2023). Assessing the impact of differentiated instruction on mathematics achievement and attitudes of secondary school learners. South African Journal of Education. 43(1), 1-10. https://orcid.org/0000-0003-1695-9876

Baumert, J., & Kunter, M. (2013). The Effect of Content Knowledge and Pedagogical Content Knowledge on Instructional Quality and Student Achievement. https://link.springer.com/chapter/10.1007/978-1-4614-5149-5_9

Birgin, O. & Yılmaz, M. (2023). Development and Validation of a Scale to Assess the Frequency of Using Classroom Assessment Tools and Methods for Mathematics Teachers. International Journal of Educational Research and Innovation (IJERI). 20, 1-20. https://doi.org/10.46661/ijeri.8075

Branzuela, N., Namoco, S., Duero, J., & Walag, A.P. (2023). Descriptive Analysis of the National Achievement Test of Primary and Secondary Students in Misamis Oriental Philippines. Sci.Int.(Lahore). 35(6), 809-813. https://www.researchgate.net/publication/376613773_Descriptive_Analysis_of_the_NAtional_Achievement_Test_Performanc_e of_Primary_and_Secondary_Schools_Students_in_Misamis_Oriental_Philippines

Buckle, J. (2024). A Comprehensive Guide to 21st Century Skills. Panorama Education. https://www.panoramaed.com/blog/comprehensive-guide-21st-century-skills

Byrne, B. M. (1994). Structural equation modeling with EQS and EQS/Windows. Thousand Oaks, CA: Sage Publications. https://www.scirp.org/reference/referencespapers?referenceid=2264937

Calaguas, G. (2012). Teacher effectiveness scale in higher education: Development and psychometric properties. International Journal of Research Studies in Education, 2(2), 3-20. https://www.researchgate.net/publication/272696664_Teacher_effectiveness_scale_in_higher_education_Development_and_psychometric_properties

Cambridge University Press (2021). Chapter 5 - Confirmatory Factor Analysis, retrieved from: https://www.cambridge.org/core/books/abs/assessing-measurement-invariance-for-applied-research/confirmatory-factor-analysis/6244C89BBAC08AC1B33338076F15BEED

Capinding, A.T. (2023). Construction and Validation of Mathematical Satisfaction Questionnaire: An Exploratory and Confirmatory Factory Analysis. Problems of Education in the 21st Century. 81(3), 327-339. http://dx.doi.org/10.2139/ssrn.4507618

Casildo, N., Aguirre, M., Bonifacio, KL., Ayunar, G., Escarlos, G., & Wenceslao, R. (2023). A Decision Support System for Predicting Students' Performance in the National Achievement Test (NAT) of Senior High School Students. ASSEHR 745, 178–190. https://www.atlantis-press.com/proceedings/icetech-22/125989114

Chigonga, B. (2020). Formative Assessment in Mathematics Education in the Twenty-First Century. Theorizing STEM Education in the 21st Century. https://www.intechopen.com/chapters/69139

Demosthenous E, Christou C, Pitta-Pantazi D. (2021) Mathematics Classroom Assessment: A Framework for Designing Assessment Tasks and Interpreting Students' Responses. European Journal of Investigation in Health, Psychology and Education. 11(3), 1088-1106. https://doi.org/10.3390/ejihpe11030081

Department of Education (2019). Classroom Observation Tool, Philippine Professional Standards for Teachers

Department of Education. (2016). K-12 curriculum guide in Mathematics.

Department of Education. (2023). MATATAG curriculum for Mathematics.

DepEd order no. 8, s.2015. Policy Guidelines on Classroom Assessment for the K to 12 Basic Education Program

Engel, K., Moosbrugger, H., and Muller H. (2003). Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness Fit Measures. MRP-Online. 8, 23-74. https://www.researchgate.net/publication/251060246_Evaluating_the_Fit_of_Structural_Equation_Models_Tests_of_Significance_and_Descriptive_Goodness-of-Fit_Measures

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. Psychological Methods, 4(3), 272-299. https://doi.org/10.1037/1082-989X.4.3.272

Finster, M. & Feldman, J. (2019). Examining Construct Validity of a Child Health Interview Instrument. Westat, An Employee-Owned Research Corporation, 1600 Research Boulevard Rockville, Maryland 20850-3129 (301) pages 251-1500, https://files.eric.ed.gov/fulltext/ED599089.pdf

Fornell, C., & Larcker, D.F. (1981). Evaluating Structural Equation Models with Unobserved variables and Measurement error. JMR. Journal of Marketing Research, 18(1), 39. https://www.scirp.org/reference/ReferencesPapers?ReferenceID=1927929

Gebremedhin, M.A., & Fenta, A.A. (2015). Assessing Teachers' Perception on Integrating ICT in Teaching Learning Process: The Case of Adwa College. Journal of Education and Practice. 6(4), 114-124. https://eric.ed.gov/?id=EJ1083759

George, M.P. (2021). Mathematics Teachers' Pedagogical Content Knowledge and its Relation to Student Achievement. Walden University, College of Education. https://scholarworks.waldenu.edu/cgi/viewcontent.cgi?article=12664&context=dissertations

Ghavifekr, S. & Rosdy, W.A.W. (2015). Teaching and learning with technology: Effectiveness of ICT integration in schools. International Journal of Research in Education and Science (IJRES), 1(2), 175-191. https://files.eric.ed.gov/fulltext/EJ1105224.pdf

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). Multivariate data analysis (7th ed.). Prentice-Hall. https://www.drnishikantjha.com/papersCollection/Multivariate%20Data%20Analysis.pdf

Hawker, G.A., Davis, A.M., French, M.R., Cibere, J., Jordan, M.D., March, M.B., Suarez-Almazor, M., Katz, J.M., & Dieppe, P. (2008). Development and Preliminary Psychometric Testing of a New OA Pain Measure – an OARSI/OMERACT Initiative. Osteoarthritis Research Society International, 16, 409-414. https://pubmed.ncbi.nlm.nih.gov/18381179/

Ismail, S. N., Muhammad, S., Omar, M. N., Shanmugam, S. K. S., & Rajoo, M. (2022). The practice of critical thinking skills in teaching mathematics: Teachers' perception and readiness. Malaysian Journal of Learning and Instruction, 19(1), 1-30. https://doi.org/10.32890/mjli2022.19.1

Jislan, F., Jani, S. H. M., Shahid, S. A. M., Thomas, M., Francis, P., & Francis, P. (2018). Using Teaching Effectiveness Scale as Measurement for Quality Teaching. International Journal of Academic Research in Business and Social Sciences, 8(9), 1394–1404. http://dx.doi.org/10.6007/IJARBSS/v8-i9/4704

Joaquin, N.B., Lumbre, A.P., & Monterola, S.C. (2023). Relationship between Mathematics Teachers' Teaching Styles and Students' Achievement in Mathematics. Athens Journal of Sciences, 10(1), 9-30. https://doi.org/10.30958/ajs.10-1-1

Kaya, D., & Keşan, C. (2023). The Connection of Mathematics with Real-Life Situations: Preservice Elementary Mathematics Teachers' Perceptions of Creating and Evaluating Story Problems. 12. 118-135. https://doi.org/10.55020/iojpe.1135191

Keiler, L.S. (2018). Teachers' roles and identities in student-centered classrooms. IJ STEM Ed 5, 34. https://doi.org/10.1186/s40594-018-0131-6

Kock, N. (2020). Using indicator correlation fit indices in PLS-SEM: Selecting the algorithm with the best fit. Data Analysis Perspectives Journal, 1(4), 1-4. https://scriptwarp.com/dapj/2020_DAPJ_1_4/Kock_2020_DAPJ_1_4_XsCorrMatrixIndices.pdf

Lal Kumar, A.C., Dinesh Kumar, A., & Vasuki, M. (2024). A Study on Teaching Effectiveness of Mathematics Teachers, International Journal of Scientific Research and Modern Education, International Peer Reviewed - Refereed Research Journal, 9(1), 33- 37, https://www.researchgate.net/publication/378595477_A_STUDY_ON_TEACHING_EFFECTIVENESS_OF_MATHEMATICS_TEACHERS

Lazarides, R., Watt, H., & Richardson, P. (2020). Teachers' classroom management self-efficacy, perceived classroom management, and teaching contexts from beginning until mid-career. Learning and Instruction. Volume 69. https://doi.org/10.1016/j.learninstruc.2020.101346

Lesia, M.P., Aigbavboa, C.O. & Thwala, W.D. (2024). Factors influencing residential location choice in South Africa: exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). J Hous and the Built Environ, 39, 133–160. https://doi.org/10.1007/s10901-023-10070-w

Lynn, MR. (1986). Determination and quantification of content validity. Nursing Research. 35(6), 381–385, https://www.scirp.org/reference/referencespapers?referenceid=2044975

Lucero, L., Hernandez, N., Grageda, C., & Semilla, J. (2024). Development and Validation of Classroom Observation Tool for Pre-Service Teachers based on Career Stage 1 of the Professional Standard for Teachers. 12(1), 194-203. https://www.russianlawjournal.org/index.php/journal/article/view/3642/2328

Mananghaya, A. K. (2020). It's Not Complicated: Analysis of Data from A Single, Two, or More Populations. *University of the Philippines Los Baños, Institute of Statistics*. Webinar Title: One-STAT at a time Session 1. YouTube.

Mathematics Teachers Association of the Philippines (n.d.). Cooperative Learning in the Teaching of Mathematics, prepared by Sr. Iluminada C. Coronel

Mattan S., Ben-Shachar, Dominique, M., Daniel, L., Indrajeet, P., Brenton, W., Remi, T., & Wagonner, P. (n.d.). Interpret of CFA / SEM Indices of Goodness of Fit. effectsize. Retrieved from: https://easystats.github.io/effectsize/reference/interpret_gfi.html

Mohd Arof, K., Ismail, S., & Saleh, A.L. (2018). Contractor's Performance Appraisal System in the Malaysian Construction Industry: Current Practice, Perception and Understanding. International Journal of Engineering & Technology. https://www.researchgate.net/publication/326698967_Contractor's_Performance_Appraisal_System_in_the_Malaysian_Construction_Industry_Current_Practice_Perception_and_Understanding

Motseki, P. & Kakoma, L. (2023). The influence of different teaching and learning strategies in mathematics - a case study. The Independent Journal of Teaching and Learning. 18(1), 98-117. https://doi.org/10.17159/ijtl.v18i1.17174

Myers, N., Ahn, S., & Jin, Y. (2011). Sample Size and Power Estimates for a Confirmatory Factor Analytic Model in Exercise and Sport: A Monte Carlo Approach. Research Quarterly for Exercise and Sport, American Alliance for Health, Physical Education, Recreation and Dance, 82(3), 412–423, https://pubmed.ncbi.nlm.nih.gov/21957699/

National Council of Teachers of Mathematics (2020). NCTM 2020 Standards for Mathematics Teacher Preparation. https://www.nctm.org/Standards-and-Positions/CAEP-Standards/

Nieto, J.L., Garrido, S., Dominguez-Lara, S., Levican, K., & Chacaltana, M. (2023). Psychometric properties of the Teachers' Sense of Efficacy Scale in a sample of Chilean public-school teachers. Frontiers in Psychology, Volume 14, https://doi.org/10.3389/fpsyg.2023.1272548

Nunnally, J.C., & Bernstein, I.H. (1994). Psychometric Theory. New York. NY: McGraw-Hill

Oswald, F., Donnellan, B., Ozkok, O., Zyphur, M., Barsky A., & Theilacker, M. (2019). Modeling Measurement as a Sequential Process: Autoregressive Confirmatory Factor Analysis (AR-CFA). Frontiers of Psychology. 10| https://doi.org/10.3389/fpsyg.2019.02108

Patnode, L.H., & Lee, H. (2021). Differentiated Instruction to Teach Mathematics: Through the Lens of Responsive Teaching. Mathematics Teacher Education and Development. 23(3), 6-25, https://eric.ed.gov/?id=EJ1320720

Philippine Association of Researchers and Statistical Software Users (PARSSU). (2022). Confirmatory Factor Analysis and Structural Equation Modeling using SPSS AMOS: Statistical Tools for Theory-Testing Research, Topic 3. Presented by Prof. Johnny T. Amora.

Philippine Star (2024, January 23). Delays in testing leave DepEd in the dark about student progress. Authored by Cristina Chi. https://www.philstar.com/headlines/2024/01/23/2327932/delays-testing-leave-deped-dark-about-student-progress

Podkhodova, N., Snegurova, V., Stefanova, N., Triapitsyna, A., & Pisareva, S. (2020). Assessment of Mathematics Teachers' Professional Competence. Journal on Mathematics Education, 11(3), 477-500. http://doi.org/10.22342/jme.11.3.11848.477-500

Price, P.C., Jhangiani, R., & Chiang, I.C. (2015). Complex Correlational Designs. Research Methods in Psychology – 2nd Edition. https://opentextbc.ca/researchmethods/chapter/complex-correlational-designs/

Rakes, C.R., Stites, M.L., Ronau, R.N., Bush, S.B., Fisher, M.H., Safi, F., Desai, S., Schmidt, A., Andreasen, J.B., & Saderholm, J. (2022). Teaching Mathematics with Technology: TPACK and Effective Teaching Practices. Educ. Sci. 12, 133. https://doi.org/10.3390/educsci12020133

Röschel A, & Wagner C, D, M. (2021). Examination of validity, reliability, and interpretability of a self-reported questionnaire on Occupational Balance in Informal Caregivers (OBI-Care) - A Rasch analysis. PLoS One. 16(12), 261-315. https://pubmed.ncbi.nlm.nih.gov/34941966/

Sachdeva, S., & Eggen, P.O. (2021). Learners' Critical Thinking About Learning Mathematics. International Electronic Journal of Mathematics Education. 16(3), 6-44. https://doi.org/10.29333/iejme/11003

Shahzad, S., & Mehmood, N. (2019). Development of Teaching Effectiveness Scale for University Teachers. Journal of Research in Social Sciences, 7(2), 2305-6533. https://www.researchgate.net/publication/336145557_Development_of_Teaching_Effectiveness_Scale_for_University_Teachers

Shun, S., Beck, S., Pett, M., & Berry, P. (2006). Psychometric Testing of Three Chinese Fatigue Instruments in Taiwan. Journal of Pain and Symptom Management. 32, 155-167. https://pubmed.ncbi.nlm.nih.gov/16877183/

Streiner, D., & Norman, G. (1991). Health Measurement Scales: A Practical Guide to Their Development and Use. Oxford University Press.

Taber, K.S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. Research in Science Education, 48, 1273-1296. https://doi.org/10.1007/s11165-016-9602-2

Tavakol M, & Wetzel A. (2020). Factor Analysis: a means for theory and instrument development in support of construct validity. Int J Med Educ. 11, 245-247. https://pubmed.ncbi.nlm.nih.gov/33170146/

Teaching at UNSW, Sydney (2023). Giving Assessment Feedback. https://www.teaching.unsw.edu.au/assessment-feedback

Türel, Y., Özdemir, T., & Elmali, F. (2017). Teachers' ICT Skills Scale (TICTS): Reliability and Validity. Cukurova University Faculty of Education Journal. 46, 503-516. https://dergipark.org.tr/tr/download/article-file/358165

WGU. (2020). What is Constructivism? https://www.wgu.edu/blog/what-constructivism2005.html#:~:text=Constructivism%20is%20based%20on%20the,new%20things%20that%20they%20learn

White M. (2021). Sample size in quantitative instrument validation studies: A systematic review of articles published in Scopus, Heliyon. 8(12), 122-143. https://pubmed.ncbi.nlm.nih.gov/36568672/

Ximénez C. (2016). Recovery of Weak Factor Loadings When Adding the Mean Structure in Confirmatory Factor Analysis: A Simulation Study. Frontiers of Psychology. 6, 1943. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2015.01943/full

Yanuari, N. & Turmudi, T. (2023). Critical thinking in mathematics education: A bibliometric Analysis. International Journal of Trends in Mathematics Education Research. 6(2), pages 191-197. https://doi.org/10.33122/ijtmer.v6i2.241

Yin, M., Cronen, S., Condelli, L., & Ogut, B. (2022). Teacher Effectiveness in Adult Education: The Relationship Between Teacher Characteristics and Student Test Gains and Transitions into Postsecondary Education. Adult Education Quarterly. 72(3), 262–283, https://journals.sagepub.com/doi/abs/10.1177/07417136211044515

Yusoff, M.S. (2019). ABC of Content Validation and Content Validity Index Calculation. Education in Medicine Journal. 11. 49-54. https://doi.org/10.21315/eimj2019.11.2.6