ISSN: 2714-2965 (Print)
ISSN: 2714-3449 (Online)
University of the Gambia supported edition
The African Journal of Behavioural and Scale Development Research
AJB-SDR Vol. 7, No 1, 2025

Establishing the Repeatability and Reliability Integrity of the Mathematics Value Scale (MaVscale)

Flourish Isaac-Oloniyo[1], Joshua Oluwatoyin Adeleke[2].
[2]Institute of Education, University of Ibadan.

Abstract

This study examines the repeatability and reliability of the Mathematics Value Scale (MaVscale), a psychometrically grounded instrument assessing students' perceptions of mathematics through value in mathematics and value for mathematics. MaVscale demonstrated strong measurement stability, with low variation coefficients (0.00 - 0.03) and high reliability indices (0.91 - 0.96), indicating consistent performance across repeated administrations. A comparative analysis with existing instruments, including the Mathematics Motivation Questionnaire, Attitude Toward Mathematics Inventory, Simple Mathematics Motivation Scale, Elementary Mathematics Motivation Inventory, and Visualized Scale of Attitude Toward Mathematics, revealed the MaVscale's theoretical cohesion, conciseness, and contextual flexibility. Unlike prior tools that are lengthy, narrowly validated, or focused solely on attitudes, MaVscale integrates intrinsic and instrumental dimensions of mathematical value, providing a robust and generalisable framework for research and practice. Implications include enhanced diagnostic utility, guidance for targeted instructional strategies, and opportunities for advancing studies on motivation and achievement. Recommendations support its adoption in educational practice, teacher training, policy formulation, and cross-disciplinary research.

Keywords: Mathematics Value Scale (MaVscale), Mathematical motivation, Psychometric reliability, Educational assessment, Value in and for mathematics

Introduction

Assessing the need for the Mathematics Value Scale (MaVscale), about six existing mathematics-related instruments, their purposes, and their limitations were examined. The Mathematics Motivation Questionnaire (MMQ) was designed to measure secondary students' motivational constructs, including intrinsic, utility, and attainment value, drawing on expectancy-value theory. Many existing instruments like MMQ are lengthy and not specifically tailored to mathematics learning contexts; furthermore, they may lack a cohesive theoretical framework (Fiorella et al., 2021). The Simple Mathematics Motivation Scale (Mexico) is a concise (14-item) motivation scale for rural Mexican adolescents, capturing task value, interest, importance, and utility. While practical and reliable (α. ≈ .87), its context-specific validation in rural areas reduces generalisability across broader or diverse populations (Arellano-Garcia et al., 2022). But, MaVscale validation was established in all the regions. Zakariya & Massimiliano (2021), on the Development of Mathematics Motivation Scale (24 items), carried out an exploratory instrument developed for secondary school students, offering five motivational subscales with strong psychometric foundations. Validation was exploratory and needs replication across independent samples for confirming factor structure and applicability, as it was done for MaVscale.

Elementary Mathematics Motivation Inventory (EMMI), through Smart & Linder (2018), is a 17-item tool for elementary-level students assessing constructs like self-efficacy, motivation, and expectancy × value. Reliability was high (α = .91). Designed for younger learners, its constructs may not align with older students' value orientation components (skills versus beliefs) like MaVscale. Visualised Scale of Attitude Towards Mathematics is a visually engaging 14 item attitude scale for measuring enjoyment, self-efficacy, and liking; exhibits strong construct validity (CFA fit indices) and internal consistency (α = .94). Focuses on attitude, not the dual constructs of “value in” versus “value for” mathematics that MaVscale addresses (Kılıç & Bölükbaş (2025).

According to Yáñez-Marquina & Villardón (2017), Attitude and Motivation Scales (ATMI) measures various attitude components, such as confidence, usefulness, and success orientation.

ISSN: 2714-2965 (Print)
ISSN: 2714-3449 (Online)
University of the Gambia supported edition
The African Journal of Behavioural and Scale Development Research
AJB-SDR Vol. 7, No 1, 2025

ATMI covers self-confidence, value, enjoyment, and motivation. Often lengthy and lacking theoretical clarity. Subscales are sometimes selectively used, limiting consistency and comparability. From all these mathematical scales that have been in existence, MaVscale is a unique construct, which targets the dual facets of mathematics value: value in mathematics (mathematics skills) and value for mathematics (belief or orientation). These are underexplored in existing instruments. MaVscale is concise yet comprehensive. Unlike lengthy instruments like Fennema-Sherman and MSLQ, MaVscale balances thoroughness with practicality, suitable for diverse populations. MaVscale is of contextual flexibility, which was designed for generalisability across subpopulations like different regions and school types. This is addressing limitations seen in context-specific scales.

MaVscale has strong psychometric foundations. It incorporates principles of repeatability, reproducibility, reliability, and validity to ensure precision and usability across settings, something many prior scales lack. Moreover, by reviewing existing scales, it was shown clearly that, while many assess aspects of motivation, attitude, or anxiety, none comprehensively measure the dual values that shape students' orientation toward mathematics. MaVscale fills this crucial gap, offering a reliable, valid, and versatile diagnostic instrument for educational research and practice.

Repeatability practices were introduced by scientists Bland and Altman (1986). For repeatability to be established, some conditions must be satisfied. The conditions are that the study must be repeated at the same location of the initial study; the same measurement procedure; the same observer; the same measuring instrument, used under the same conditions; and repetition over a short period of time, while nothing is likely to change significantly (Labmate, 2023; Statbooks, 2023). Repeatability, as defined by Statistics.com (2025), is the variance of results from any process of observation or measurement performed in the same laboratory under the same conditions and by the same operator. The variance of weight values or the closeness of measured values between repeated measurements of the same weight, conducted under the same conditions, such as at the same location, by the same person, by the same method, on the same equipment, and over a brief period, is known as repeatability According to Conjointly (2023), Repeatability in the sense that, the developed scale must consistently establish the purpose for being developed. A scale that does not have repeatability produce scattered results. Reliability is maintaining the level of quality over time. That is, the scale must consistently measure what it should measure. Consistency is making sure every system is as consistent as possible. It can be concluded that the term repeatability means reliability or consistency. So, to measure the repeatability of a scale is the same as measuring the reliability of such scale. Petersen (2025), defines reliability as the proportion of observed-score variance that is attributable to true-score variance, precisely matching your statement: “It is the value of a reliability estimate that gives the proportion of variability in the measure attributable to the true score.”

Repeatability is the same as test-retest reliability (Elsevier, 2023). There are many means of establishing the reliability of a scale. Psychologists consider three types of consistency (reliability) as test-retest reliability (over time), internal consistency (across items) and inter-rater reliability (across different researchers). Test-retest reliability (methods are used to assess the consistency of a measure from one period to another). One approach in test-retest is to look at a split-half correlation. This is splitting the items into two sets, such as first and second halves of the items or the even and odd numbered items. Another one is parallel-forms reliability which is used to assess the consistency of the results of two scales constructed in the same way from the same content domain, while Parallel Alpha and Split half reliability are for measuring the internal consistency of such instrument (Robertson, 2017).

When Reliability is done multiple times, it becomes Repeatability. It is a statistical measure of the consistency of repeated measurements (Elsevier, 2023), and called r. It is the proportion of the variation between measurements that is due to consistent differences between the individuals measured. Reliability coefficient will always range between 0 and 1 and the closer to 1 the better. Generally, experts tend to look for a reliability coefficient in excess of 0.70 (Doverspike, 2017). For instance, a reliability of .5 means that about half of the variance of the observed score is attributable to truth and half is attributable to error. That is, a reliability of .7 means the variability is about 70% true ability and 30% error (Conjointly, 2023).

Repeatability is done to discover the repeatability coefficient that is also known as a measurement of precision (Sullivan, 2023), which denotes the absolute difference between a pair of repeated test results. That is, repeatability and reliability are used to measure difference and a lack of correlation. For instance, if the results differed greatly, it would be probably concluded that the findings were inaccurate, leading, which may lead to further investigations (Labmate, 2023). Repeatability is measured by calculating the standard deviation of the measurements and comparing it to some agreed limit. (Latham, 2023). The closeness of agreement between measurements when measured under the same conditions, the better. When measurement is repeated, it is likely the measurement value is different each time. If the variation is small, the repeatability is high. That is, the smaller the repeatability coefficient (RC), the better.

The mathematics value scale (MaVscale) was developed through the adoption of value theory propounded by Aristotle (384-322 BCE), that, is before common era. Aristotle based his value theory on the idea that, need is what brings about exchange (Fogarty, 1996). Aristotle stated that, value has two major purposes, which are value in use and value in exchange. From this knowledge, it was then discovered that, the mathematics value has two facets: value in mathematics and value for mathematics. Value in mathematics means that mathematics is embedded with numerous skills that learners of mathematics can develop in the process of learning the subject. Likewise, there are values for mathematics (belief or orientation) the learner must inculcate to acquire the values (skills) in mathematics. The two constructs are inseparable. It is the value a learner has for mathematics that will definitely determine the value to acquire in mathematics. Therefore, mathematics value means value in mathematics and value for mathematics. Many learners are ignorant of this. The awareness for mathematics value can be more effective if, there can be means of inducing positive orientation in the learners. Availability of a valid scale like the MaVscale, to diagnose the learners' extent of value for mathematics (orientation) and value in mathematics will go a long way in improving the learners' performance in mathematics. But, the MaVscale level of integrity to perform these functions must be established. That is, the scale must be affirmed to be a reliable and valid instrument that can measure (mathematics value) the construct of interest (DeVellis, 2012).

The processes of making MaVscale reliable and valid have strictly been followed by considering every condition for the scale development. One of the conditions is the repeatability and reproducibility tests of the scale, which is the major purpose for this study. This repeatability study followed every step for the determination and establishment of the MaVscale repeatability capability and the usability of the scale.

Statement of Problem

There is a growing need to understand how students perceive the value of mathematics, both in terms of the skills it develops (value in mathematics) and the personal importance they attach to it (value for mathematics). However, current measurement tools do not adequately capture this dual perspective. While various instruments exist to assess mathematical attitudes and motivation, none are specifically designed to measure both facets of mathematics value comprehensively.

A review of six existing mathematics-related scales revealed key limitations. Tools like the MMQ, ATMI, and the Simple Mathematics Motivation Scale either focus on isolated constructs, are too lengthy, lack a solid theoretical base, or are only validated for specific populations. None of them provides a balanced, generalisable measure of both skill-based value and belief-based orientation toward mathematics, which MaVscale was designed to do.

The Mathematics Value Scale (MaVscale), based on Aristotle's theory of value, addresses these gaps by measuring both "value in" and "value for" mathematics in a concise and reliable format. This study focuses on assessing the repeatability of MaVscale to confirm its reliability and consistency under similar conditions. Proving this will ensure that MaVscale can serve as a valid diagnostic tool for research, teaching, and learning across diverse educational settings.

Research Objectives

1. To establish the repeatability and reliability of the Mathematics Value Scale (MaVscale).
2. To discover whether the results of the repeatability studies provide sufficient evidence of MaVscale's reliability for educational use or not.

Research Questions

1. Is the variation coefficient of the two MaVscale repeatability studies sufficient for its usage?
2. How comparable or consistent is the reliability of MaVscale based on similar studies?

Methodology

The study adopted a Descriptive survey design, because of the strengths of survey research that include generalisability, reliability, and versatility (Mauldin, 2023), which this study proposed to measure in the MaVscale for usage. The study target population comprises all the Senior Secondary 1 (SS1) student in Nigeria. Nine thousand and fifteen (9,015) public schools had a total population of 1,264,237 SS I students 13,423 private schools had a total population of 319,696 SS 1. A four-stage multistage sampling procedure was employed to select participants for the study from two states. At the first stage, a simple random sampling technique was used to select one of Nigeria's six geographical zones, the South-West zone (Wangare et al., 2024). At the second stage, purposive sampling was used to select Ondo State and Osun State from the six states within the South-West zone.

This selection was based on geographical proximity and accessibility. In the third stage, within each of the two selected states, three senatorial districts were identified. From each district, six schools (three public and three private) were randomly selected, making a total of twelve schools. Finally, at the fourth stage, a minimum of 30 SS1 students were randomly selected from each school. In total, 1,100 students were selected from the twelve schools in Ondo and Osun States for the study.

The Mathematics Educator Key Informant Interview Schedule (MEKIIS) was developed by Isaac-Oloniyo (2023) as part of the exploratory (inductive) phase in the construction of the Mathematics Value Scale (MaVscale). It was designed to collect expert input for the generation of the scale items and to explore the key constructs underpinning value orientation and mathematics achievement. MEKIIS was developed with the primary aim of gathering qualitative data from mathematics educators to support the construction of items for the MaVscale. It served as an interview guide to extract expert insights about the meaning, relevance, and dimensions of mathematics value and orientation toward mathematics. Through this tool, the researcher sought to understand the foundational beliefs, perceived benefits, and learning strategies associated with mathematics from the educators' perspectives.

The MEKIIS is divided into two main sections: Section A - Demographic Variables: This section captures basic background information of the respondents, such as: Status, Gender and Occupation. Section B - Interview Questions: This section contains six open-ended questions designed to guide the interviews and collect in-depth qualitative responses. The open-ended items in Section B are grouped conceptually into three key components, each reflecting a central theme relevant to the eventual construction of MaVscale:

Mathematics Orientation/Belief (Items 1 & 2)
– Focuses on respondents' attitudes and orientations toward mathematics.
Value in Mathematics (Items 3 & 4)
– Extracts perceived benefits and purposes of learning mathematics.
Mathematics Learning Processes (Items 5 & 6)
– Gathers insights into how mathematics is learned and how learning goals are achieved.

Number of Items under Each Component are Mathematics Orientation/Belief (2 items), Value

ISSN: 2714-2965 (Print)
ISSN: 2714-3449 (Online)
University of the Gambia supported edition
The African Journal of Behavioural and Scale Development Research
AJB-SDR Vol. 7, No 1, 2025

in Mathematics (2 items, and Mathematics Learning Processes (2 items). This brings the total number of open-ended items in MEKIIS to six. The response format is qualitative and open-ended. Respondents (Mathematics Educators) are expected to answer freely based on their professional experience and views. The researcher guided the interviews using these questions, allowing for flexible but focused discussions. There is no quantitative scoring format for MEKIIS, as it is a qualitative data collection tool. Instead, responses were analyzed thematically. The data collected were combed for recurring themes, phrases, and interpretations, which informed the construction and wording of items for the Mathematics Value Scale (MaVscale). This inductive approach aligns with best practices in exploratory scale development (Carpenter, 2018). Therefore, seven components were discovered from the exploratory stage. The seven components are basic skills, reasoning skills, intelligent skills, emotional intelligent skills, problem-solving and time management skills. Then, a 98-item scale was initially constructed, from which 27-item scale was developed through rigorous tests and analyses.

The 98 constructed items that have been trial-tested, modified and reviewed twice by the expert reviewers and subjected to all the necessary assumptions' testing were administered on the selected respondents for pilot-testing. As an ordinal data, a non-parametric statistic of Kruskal-Wallis H test was adopted using SPSS version 28 for the exploratory data analysis, for Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity (Table 1).

Table 1: KMO And Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.976
Bartlett's Test of Sphericity	Approx. Chi-Square	38941.613
	Df	4753
	Sig.	.000

KMO = 0.976. Bartlett's Test of Sphericity was significant *p = 0.000*. For a data to be suited for factor analysis, its KMO must be greater than 0.5, while the Bartlett's Test of Sphericity must be less than 0.5. Since KMO = 0.975 and >0.60, while Bartlett's Test of Sphericity p = 0.000 and < 0.05. Therefore, 'psych' and 'GPA rotation' in R Studio software package were adopted for further analyses: Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) The results shown in Figure 1.

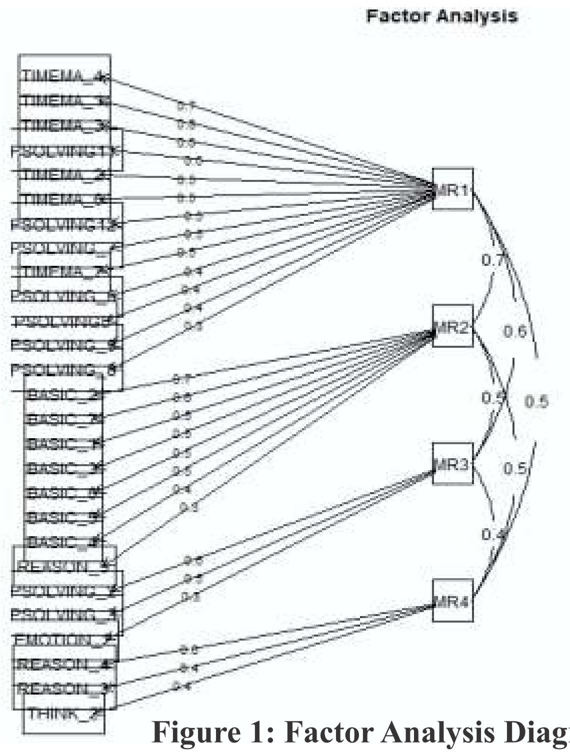

Figure 1: Factor Analysis Diagram of the 27 retained items.

All the items satisfied the benchmarks (cutoff = 0.3) for items and none of the items double-loaded on two factors (Figure 1). Summary of the analysis of the fifth iteration showed that the root mean square of the residuals (RMSR) is 0.02, which is close to zero. Tucker Lewis Index (TLI) factoring reliability = 0.971, which is > 0.9. The root mean square error of approximation (RMSEA) = 0.02 < 0.05. It implies that all the 34 items are valid to be used for measurement. But, factor 5 and factor 6 were expunged for not satisfying the benchmark of at least three items loading on a factor. With the consideration of expert reviews' second round content validity index (CVI = 0.7), 27 items were finally selected. In Figure 1, the twenty-seven retained items were extracted from the modified items (Appendix VI) and converged under four factors. Factor MR1 = Problem-Solving skills has 13 items converged under it. Factor MR2 = Basic skills have 8 items converged under it. Factor MR3 = Intelligence skills and Factor MR4 = Reasoning skills with 3 items converging under each of them. Parallel and factor analyses were performed in RStudio, to test for the validity of the 27 items in MaVscale. The summary showed that the root mean square of the residuals (RMSR) is 0.02, which is close to zero. Tucker Lewis Index (TLI) factoring reliability = 0.977, which is > 0.9 and the root mean square error of approximation (RMSEA) = 0.021 < 0.05. These proved that the twenty-seven items are valid to be used as a measuring tool.

So, data for this study were collected using the 27-item MaVscale. Four months was used for the data collection: The researcher used a month to prepare the mathematics teachers in the 12 selected schools as research assistance and test administrators. In the second month, the first repeatability test was administered to the selected students in all the selected schools. Then, three months later, the second repeatability test was administered to the selected students in the selected schools. Descriptive statistics (Variance and standard deviation) in SPSS were used to analysed the collected data.

Results

Research question 1: Is the variation coefficient of the two MaVscale repeatability studies sufficient for its usage?

The results of the analyses are as shown in Table 3. In Table 2, the repeatability variation coefficients are 0.00, 0.01, 0.02 and 0.03.

Research question 2: How comparable or consistent is the reliability of MaVscale based on similar studies

Answer to Research question 2: To answer this question, the scale for reliability analysis in SPSS was used.

Table 2 shows the results of the MaVscale's reliability analyses of the two repeatability studies.

Table 2 : Reliability statistics of the two repeatability studies.

Repeatability Study	Parallel Alpha	Standardised Alpha	No. of Items
Study 1	0.963	0.963	27
Study 2	0.908	0.908	27

ISSN: 2714-2965 (Print)
ISSN: 2714-3449 (Online)
University of the Gambia supported edition
The African Journal of Behavioural and Scale Development Research
AJB-SDR Vol. 7, No 1, 2025

In Table 2, the reliability coefficients of the first repeatability data = 0.96, while the reliability coefficients of the second repeatability data = 0.91.

Table 3: Comparison of the MaVscale Repeatability test results

	Repeated	N	Sum	Mean	Std. Error	Variance	Std. Dev.	Variation
1. Counting Accurately	1	792	3370	4.26	.04	1.13	1.06	
	2	919	3919	4.26	.04	1.13	1.06	0.00
2. Understanding things like numbers and	1	792	3202	4.04	.04	1.13	1.06	
images that repeat in a logical way.	2	919	3371	4.04	.04	1.15	1.07	0.01
3. Adding things	1	792	3312	4.18	.04	1.07	1.04	
together like '+'	2	919	3859	4.20	.04	1.08	1.04	0.00
4. Subtracting things	1	792	3283	4.15	.04	1.11	1.11	
away from a group using? -? sign	2	919	3829	4.17	.04	1.12	1.12	0.00
5. Multiplying a certain	1	792	3332	4.21	.04	1.06	1.03	
number of objects using ? +? sign	2	919	3861	4.20	.04	1.06	1.03	0.00
6. Dividing a certain number	1	792	3315	4.19	.04	1.08	1.04	
of objects using? /? sign	2	919	3843	4.18	.04	1.11	1.05	0.01
7. Selecting procedures that involve	1	792	3280	4.14	.04	1.10	1.05	
Solving problems that involve addition, subtraction, multiplication and division correctly	2	919	3812	4.15	.04	1.11	1. 05	0.00
8. Applying procedures that involve solving	1	792	3183	4.02	.04	1.19	1.09	
problems that involve addition, subtraction, multiplication, and division correctly	2	919	3714	4.04	.04	1.17	1.08	0.01
9. Solving problems that involve addition,	1	792	3256	4.11	.04	1.14	1.07	
subtraction, multiplication and division correctly.	2	919	3791	4.13	.04	1.12	1.06	0.01
10. Understanding the usefulness of solving	1	792	3176	4.01	.04	1.35	1.16	
problems that involve addition, subtraction, multiplication and division correctly.	2	919	3719	4.05	.04	1.31	1.14	0.02
11. Thinking and believing	1	792	3174	4.01	.04	1.30	1.14	
what to believe about life	2	919	3703	4.03	.04	1.30	1.14	0.00
12. Using old methods in a new method	1	792	3100	3.91	.04	1.42	1.18	
to solve current problems.	2	919	3606	3.92	.04	1.40	1.19	0.01
13. Being able to think orderly to have	1	792	3139	3.96	.04	1.24	1.12	
life-problems' solutions	2	919	3668	3.99	.04	1.23	1.11	
14. Taking part in exploring, discussion	1	792	3180	4.02	.04	1.36	1.17	
or argument to solve problems.	2	919	3692	4.02	.04	1.38	1.17	0.00
15. Expressing solutions to problems	1	792	3268	4.13	.04	1.06	1.03	
clearly and consistently.	2	919	3798	4.13	.04	1.06	1.03	0.00
16. Organising thoughts clearly	1	792	3161	3.99	.04	1.20	1.10	
	2	919	3716	4.04	.04	1.18	1.09	0.01
17. Understanding impressions, feelings,	1	792	3109	3.93	.04	1.26	1.12	
and descriptions that exist in mind	2	919	3678	4.00	.04	1.23	1.11	0.01
18. Using comparison to	1	792	3192	4.03	.04	1.16	1.08	
understanding life-matter.	2	919	3716	4.04	.04	1.20	1.10	0.02
19. Knowing my Mathematical ability	1	792	3202	4.04	.04	1.29	1.13	
and applying it to solve problems.	2	919	3769	4.10	.04	1.28	1.13	0.00
20. Being able to understand other	1	792	3188	4.03	.04	1.17	1.08	
student's problems.	2	919	3741	4.07	.04	1.16	1.08	0.00
21. Handling Mathematics challenge.	1	792	3190	4.03	.04	1.28	1.13	
or problems correctly.	2	919	3761	4.09	.04	1.23	1.11	0.02
22. Gathering information to solve	1	792	3196	4.04	.04	1.20	1.10	
Mathematics problems.	2	919	3755	4.09	.04	1.19	1.09	0.01
23. Developing solutions in solving	1	792	3262	4.12	.04	1.29	1.14	
Mathematics problems.	2	919	3837	4.18	.04	1.25	1.12	0.02
24. Deciding on how to	1	792	3189	4.03	.04	1.22	1.11	
solve Mathematics problems.	2	919	3753	4.08	.04	1.19	1.09	0.01
25. Explaining Mathematics solutions	1	792	3120	3.94	.04	1.36	1.17	
easily and clearly.	2	919	3654	3.98	.04	1.37	1.17	0.00
26. Applying Mathematics knowledge	1	792	3120	3.94	.04	1.45	1.21	
to solve real life problems effectively.	2	919	3690	4.02	.04	1.38	1.18	0.03
27. Raising and sharing opinions with	1	792	3234	4.08	.04	1.41	1.19	
other students when solving Mathematics problems.	2	919	3771	4.10	.04	1.38	1.17	0.02

Discussion of Findings

The purpose of this study was to establish the repeatability and reliability of the Mathematics Value Scale (MaVscale), a diagnostic instrument designed to measure students' dual orientation toward mathematics, value in mathematics (skills-oriented) and value for mathematics (belief-oriented). Findings from the repeatability analyses revealed that the variation coefficients across the two test administrations ranged between 0.00 and 0.03, indicating an exceptionally high level of consistency between measurements conducted under similar conditions. According to Bland and Altman (1986), a small variation coefficient reflects minimal measurement error and high stability of results when the same instrument is re-administered to similar samples. Therefore, the MaVscale meets the criteria for repeatability, yielding nearly identical results under the same measurement conditions (Labmate, 2023; Statbooks, 2023).

This result suggests that MaVscale is not sensitive to environmental fluctuations or transient learner factors, a limitation often noted in existing mathematics-related instruments. Previous scales such as the Mathematics Motivation Questionnaire (MMQ) and the Attitude Toward Mathematics Inventory (ATMI) have shown inconsistent performance when replicated across different contexts (Fiorella et al., 2021; Yáñez-Marquina & Villardón, 2017). The exceptionally small variation coefficients in this study confirm that MaVscale demonstrates superior measurement precision and repeatability.

The reliability coefficients obtained from the two repeatability studies both exceeded the minimum reliability threshold of 0.70 recommended for psychometric instruments (Doverspike, 2017). These coefficients indicate excellent internal consistency and confirm that the MaVscale items consistently measure the same underlying constructs of mathematics value. The minor decline from 0.96 to 0.91 between the two administrations aligns with expectations in test–retest reliability studies due to slight changes in respondents' conditions over time (Elsevier, 2023). Overall, both coefficients affirm the instrument's reliability and stability.

When compared with other mathematics-related scales, MaVscale demonstrates competitive or superior reliability. For instance, the Simple Mathematics Motivation Scale reported a reliability coefficient of approximately .87 (Arellano-Garcia et al., 2022), while the Elementary Mathematics Motivation Inventory achieved $\alpha = .91$ (Smart & Linder, 2018). The Visualised Scale of Attitude Toward Mathematics (Kılıç & Bölükbaş, 2025) reported $\alpha = .94$, focusing mainly on attitudes. However, MaVscale extends beyond measuring attitudes alone by capturing both the skill-based and belief-based dimensions of mathematics value, thereby offering broader construct coverage and stronger psychometric evidence.

The factor analyses conducted during MaVscale development confirmed that the 27 items load meaningfully under four factors: problem-solving skills, basic skills, intelligence skills, and reasoning skills. All the fit indices obtained met accepted psychometric benchmarks, thereby confirming the construct validity of the instrument. These findings align with DeVellis (2012), who emphasized that valid measurement tools must reflect the theoretical constructs they are intended to measure. More also, the theoretical foundation of MaVscale is rooted in Aristotle's theory of value, which distinguishes between value in use and value in exchange (Fogarty, 1996). By applying this philosophical model to mathematics education, MaVscale conceptualizes value in mathematics (skills development) and value for mathematics (belief or orientation) as interdependent components of learners' engagement with mathematics. This conceptual integration differentiates MaVscale from prior instruments based mainly on expectancy-value or self-efficacy frameworks (Fiorella et al., 2021).

The findings of this study present several important implications for both educational research and classroom practice. First, the high repeatability and reliability of the Mathematics Value Scale (MaVscale) affirm its suitability as a dependable diagnostic tool for assessing students' perceptions of mathematics across diverse educational contexts. Its consistency of measurement ensures that educators and researchers can confidently use the instrument to evaluate how learners perceive and value mathematics, thereby providing a more accurate understanding of students' orientations toward the subject.

Second, the MaVscale's validation across multiple regions establishes its strength in terms of cross-context application. Unlike other mathematics-related instruments that were developed and validated within narrow or specific populations, such as rural communities or elementary school learners. MaVscale demonstrates broader generalisability and contextual flexibility. This means it can be effectively used in varied educational settings, encompassing differences in school type, location, and student background, without compromising the accuracy or meaning of the results.

Third, the distinction the MaVscale makes between value in mathematics (skills orientation) and value for mathematics (belief orientation) offers valuable instructional insights. Teachers can use this dual construct to design more targeted interventions that address both cognitive and affective dimensions of mathematics learning. By fostering an awareness of the skills embedded in mathematics as well as the beliefs that support engagement with the subject, educators can cultivate deeper learner motivation, confidence, and persistence. This approach moves beyond traditional methods that focus solely on skill acquisition, highlighting instead the interplay between students' attitudes and their competence development.

Lastly, the robust psychometric properties of MaVscale position it as a valuable instrument for advancing educational research. Its reliability, validity, and theoretical grounding make it a powerful tool for future studies exploring the complex relationships among mathematical values, motivation, and academic achievement. Researchers can employ MaVscale to generate data that inform curriculum design, teacher education, and policy formulation aimed at improving mathematics learning outcomes. In essence, the MaVscale not only fills a critical measurement gap in mathematics education but also provides practical and theoretical foundations for enhancing teaching, learning, and scholarly inquiry within the discipline.

Comparison with existing instruments revealed that, the Mathematics Motivation Questionnaire (MMQ) and the Attitude Toward Mathematics Inventory (ATMI) are well-established instruments for assessing students' motivation and attitudes toward mathematics; however, both are relatively lengthy and exhibit limited theoretical cohesion (Fiorella et al., 2021; Yáñez-Marquina & Villardón, 2017). Similarly, the Simple Mathematics Motivation Scale (SMMS) and the Elementary Mathematics Motivation Inventory (EMMI) offer concise and practical alternatives but were validated within narrow population groups, which restricts their applicability and generalisability across diverse educational contexts (Arellano-Garcia et al., 2022; Smart & Linder, 2018).

The Visualised Scale of Attitude Toward Mathematics (VSATM), though visually engaging and psychometrically strong, primarily measures affective components of mathematics learning, such as enjoyment and self-efficacy without addressing the dual constructs of value in and value for mathematics (Kılıç & Bölükbaş, 2025). In contrast, the Mathematics Value Scale (MaVscale) integrates theoretical rigor with conciseness and robust psychometric properties, offering a comprehensive, generalisable, and contextually adaptable instrument for assessing students' mathematical value orientations.

The limitations identified in these existing instruments provided a strong justification for the development of the Mathematics Value Scale (MaVscale). While prior tools effectively measured motivation, attitude, or self-efficacy, none comprehensively captured the interrelated dimensions of value in mathematics, the acquisition of cognitive and problem-solving skills and value for mathematics are the beliefs, orientations, and personal importance learners attach to the subject. The MaVscale was therefore conceptualised to bridge this measurement gap by offering a theoretically grounded, psychometrically robust, and contextually flexible instrument capable of diagnosing students' holistic valuation of mathematics. This development responds to the increasing need for valid and reliable tools that can inform instructional strategies, promote positive learner dispositions, and support empirical research on mathematics education across diverse learning environments.

Conclusion

This study aimed to establish the repeatability and reliability of the Mathematics Value Scale (MaVscale) as a valid instrument for assessing students' perceptions of mathematics through two interrelated constructs: value in mathematics (skills orientation) and value for mathematics (belief orientation). Grounded in Aristotle's theory of value, MaVscale was developed to address gaps in existing instruments by integrating both the cognitive and affective dimensions of learners' engagement with mathematics.

Analyses of repeatability and reliability provided strong evidence of the scale's psychometric robustness. Extremely low variation coefficients (0.00 - 0.03) confirmed measurement stability, while high reliability indices (0.91 - 0.96) indicated strong internal consistency. Factor analyses further supported the scale's construct validity and theoretical coherence, demonstrating its suitability for both research and educational practice.

Compared with instruments such as the Mathematics Motivation Questionnaire (MMQ), Attitude Toward Mathematics Inventory (ATMI), Simple Mathematics Motivation Scale (SMMS), and Elementary Mathematics Motivation Inventory (EMMI), MaVscale offers a more concise, theoretically grounded, and contextually flexible tool that captures both intrinsic and instrumental dimensions of mathematical value. These findings highlight MaVscale as a significant contribution to mathematics education, providing a nuanced framework for understanding students' motivation, values, and orientation toward mathematics.

Recommendations

Based on the findings and conclusions of this study, the following recommendations are proposed:

1. Educational Practice and Teacher Training: Adopt MaVscale as a diagnostic tool in schools and integrate its use into preservice and in-service teacher training programs to guide targeted interventions and pedagogical strategies that align with students' mathematical values and motivation.
2. Policy, Curriculum, and Cross-Disciplinary Application: Utilize insights from MaVscale to inform curriculum design, value-based educational programs, and broader policy initiatives, while exploring adaptation of the scale for other disciplines to examine the impact of value-based constructs on learning and performance.
3. Research and Validation: Conduct further studies to validate MaVscale across diverse populations, educational levels, and regions, including longitudinal research to track changes in students' mathematical values and their influence on achievement and STEM-related career pathways.

References

American Educational Research Association (AERA), 2018.AERA, APA, & NCME. (2018). *Standards for Educational and Psychological Testing*. American Educational Research Association.

Arellano-Garcia, Y., Vargas-De-Leon, C., & Reyes-Carreto, R. (2022). A Simple Mathematics Motivation Scale and Study of Validation in Mexican Adolescents. https://doi.org/10.1177/21582440221085264

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet, 327*(8476), 307–310.

Byrne, B. M. (2023). *Structural Equation Modeling with Mplus: Basic Concepts, Applications, and Programming* (3rd ed.). Routledge.

Carpenter, S. (2018). Ten steps in scale development and reporting: A guide for researchers. *Communication Methods and Measures, 12*(1), 25–44.

Conjointly, 2023. Theory of Reliability. https://conjointly.com/kb/theory-of-reliability/

Costello, A. B., & Osborne, J. (2021). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment,* *Research, and Evaluation, 10*(1).

DeVellis, R.F. 2012. *Scale development: theory and application,* SAGE Publications Inc., pp10-30.

DeVellis, R. F., & Thorpe, C. T. (2021). *Scale Development: Theory and Applications* (5th ed.). Sage Publications.

Doverspike, D. 2017. *What are the characteristics of a good test?* http://ar.ipma-hr.org/2017/04/what-are-the-characteristics-of-a-good-test/

Elsevier, B.V. 2023. Advances in Agronomy: Temporal and spatial repeatability. https://www.sciencedirect.com/bookseries/

Fiorella, L., Yoon, S. Y., Atit, K., Power, J. R., Panther, G., Sorby, S., Uttal, D. H., & Veurink,

N. (2021). Validation of the Mathematics Motivation Questionnaire (MMQ) for secondary school students. *International Journal of STEM Education, 8*(52). https://stemeducationjournal.springeropen.com/articles/ Fogarty, M. 1996. *A history of value theory*. https://www.tcd.ie/Economics/assets/pdf/SER/1996/

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2021). *Multivariate Data Analysis* (8th ed.). Cengage.

Isaac-Oloniyo, F. O. (2023). Development of Mathematics Value Scale and its predictive estimate on Mathematics achievement among Senior Secondary I Students in Nigeria *[Unpublished doctoral dissertation, University of Ibadan, Ibadan, Nigeria]*.

Kılıç, R. & Bölükbaş, F. A. (2025). Visualised Scale of Attitude Towards Mathematics: A Validity and Reliability Study. https://onlinelibrary.wiley.com/doi/full/10.1111/ejed.70072

Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling* (4th ed.). Guilford Press.

Labmate O., 2023. What is the Difference Between Repeatability and Reproducibility? https://www.labmate-online.com/news/

Latham, A. (2023). How do I calculate repeatability? https://www.ehow.co.uk/

Mauldin, R.L. 2023. *Foundations of Social Work Research*. https://uta.pressbooks.pub/foundationsofsocialworkresearch/chapter/7-2-assessing- survey-research/

Petersen, I. T. (2025). Principles of psychological assessment: Reliability as the proportion of true-score variance. In *Principles of psychological assessment: With applied examples in R* (Ch. 4.3). https://isaactpetersen.github.io/Principles-Psychological-Assessment/reliability

Ryan, R. M., & Deci, E. L. (2020). *Intrinsic and Extrinsic Motivation from a Self-Determination Theory Perspective: Definitions, Theory, Practices, and Future Directions*. *Contemporary Educational Psychology, 61*, 101860.

Robertson, G. 2017. *Developing valid and reliable survey scales*. www.i2iacility.org

Smart, J. & Linder, S. M. (2018). Development, validation, and implementation of the

elementary mathematics motivation inventory (EMMI): examining motivational constructs in elementary mathematics. *Fields Mathematics Education Journal, 3*(2) https://fieldsmathed.springeropen.com/articles/10.1186/Statbooks, 2023. http://pcool.dyndns.org:8080/statsbook/?page_id=835statistics.com, (2025). https://www.statistics.com/glossary/repeatability/

Sullivan B., 2023. The Importance of Repeatability, Reliability, Consistency and Accuracy. https://www.valin.com/

Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.

Wangare, J., Sinwa, A., & Ugbodaga, M. (March 06, 2024). Geopolitical zones in Nigeria and their states: detailed Updated Wednesday, https://www.legit.ng/1094595-geopolitical-zones-nigeria-states.html

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist, 34*(6), 806–838.

Yáñez-Marquina, L. & Villardón, L. (2017). Attitudes towards mathematics at secondary

level: Development and structural validation of the Scale for Assessing Attitudes towards Mathematics in Secondary Education (SATMAS). *Electronic Journal of Research in Educational Psychology, 14*(40). 10.25115/ejrep.40.15163

Zakariya, Y. F. & Massimiliano, B. (2021). Development of mathematics motivation scale: A preliminary exploratory study with a focus on secondary school students.

International Journal of Progressive Education, 17(1) 314-324. https://doi.org/10.29329/ijpe.2021.329.20