Educational Studies

ISSN: 0305-5698 (Print) 1465-3400 (Online) Journal homepage: www.tandfonline.com/journals/ceds20

How secondary students make sense of mathematics: psychometric validation and educational implications for teaching practice

Çiğdem Çalişkan Toraman & Osman Aktan

To cite this article: Çiğdem Çalişkan Toraman & Osman Aktan (11 Aug 2026): How secondary students make sense of mathematics: psychometric validation and educational implications for teaching practice, Educational Studies, DOI: 10.1080/03055698.2026.2716066

To link to this article: https://doi.org/10.1080/03055698.2026.2716066

Published online: 11 Aug 2026.

EDUCATIONAL STUDIES
https://doi.org/10.1080/03055698.2026.2716066

Routledge
Taylor & Francis Group

Check for updates

How secondary students make sense of mathematics: psychometric validation and educational implications for teaching practice

Çiğdem Çalişkan Toraman [a] and Osman Aktan [b]

[a]Dr. Ministry of National Education, Edremit Science and Arts Center, Balıkesir, Türkiye; [b]Faculty of Education, Düzce University, Düzce, Türkiye

ABSTRACT

This study developed and validated the Meaning of Mathematics (MofM) scale, a Likert-type instrument assessing the meaning middle and high school students attribute to mathematics, and examined how scores vary across gender, grade, school type, field of study, parental education, and family income. Exploratory factor analyses (EFA) and confirmatory factor analyses (CFA), corroborated by parallel analysis, supported a single-factor, 11-item structure, and item response theory analyses indicated strong discrimination and adequate information across most of the trait continuum. Measurement invariance held across gender and school level; metric invariance did not hold for parental education and family income, and could not be tested for school type and field of study due to inadequate subgroup sizes. Students' scores differed significantly across all demographic variables, though effect sizes were generally small to moderate, and the cross-sectional design does not support causal interpretation of these differences.

ARTICLE HISTORY
Received 18 May 2026
Accepted 4 August 2026

KEYWORDS
Item response theory; meaning of mathematics; measurement invariance; scale development

Introduction

Mathematics is a foundational discipline that supports problem-solving and decision-making in daily life while fostering logical reasoning and critical thinking (Firdaus et al. 2015; Lave 1988; Samo, Kartasasmita, and Kartasasmita 2017; Schoenfeld and Sloane 2016; Tall 1991). Beyond the classroom, mathematics education cultivates lifelong cognitive and analytical skills through instructional approaches spanning all developmental levels, from early childhood to higher education (de Santana, Roazzi, and Nobre 2022; English and Halford 2012; Furinghetti, Matos, and Menghini 2012; Hiebert and Grouws 2007; Schoenfeld and Sloane 2016; Sosa-Gutierrez et al. 2024; Sriraman and English 2010; Van den Heuvel-Panhuizen and Drijvers 2020; Widana et al. 2018; Zakaria and Syamaun 2017), including mathematical literacy that enables the everyday application of these skills (Jablonka 2003). Mathematics education is further linked to individual cognitive and affective development, such as interest, motivation, and systematic work habits (Hui and Mahmud 2023; Ni et al. 2018), and is recognised as contributing to societal and economic progress (Tate and Rousseau 2002). Its effectiveness depends on multiple

CONTACT Osman Aktan karakteregitimi@gmail.com Faculty of Education Düzce University, Düzce, Türkiye

factors, including curriculum, instructional methods, and teacher competence (Hill and Charalambous 2012; Hill, Rowan, and Ball 2005; Llinares 2021; Veldhuis and Van den Heuvel-Panhuizen 2020), with student-centred, inquiry-based, and meaning-focused instruction shown to strengthen students' attitudes and self-efficacy towards the subject (Abrahamson et al. 2020; Harlen 2013; Manaf, Kartowagiran, and Manaf 2021; Usher and Pajares 2009).

The perception of mathematics is a concept encompassing students' attitudes, beliefs, anxiety, emotions, and values regarding mathematics (Di Martino and Zan 2011; Hannula 2002; Szucs and Toffalini 2023). This perception is considered a set of cognitive and affective factors that form the basis of the student's relationship with mathematics (Cargnelutti, Tomasetto, and Passolunghi 2017). According to Dweck's (2008) mindset theory, whether students view mathematical ability as a fixed trait or a potential that can be developed through effort determines their success. Students with a fixed mindset tend to show failure and give up when faced with difficulties (Boaler 2016; Yeager and Dweck 2020), whereas those with a growth mindset perceive these challenges as opportunities for learning. To mitigate mathematics anxiety and support permanent learning, it is crucial for students to develop a positive, growth-oriented perception towards mathematics and their own learning potential.

A positive perception of mathematics increases students' motivation towards the subject, encourages active participation in class, and contributes to the development of self-confidence in mathematics (Appiah et al. 2022; Du et al. 2025). Students' experiences of success in mathematics, positive feedback from their teachers, and the connection of mathematics to daily life strengthen a positive mathematical perception (Edu et al. 2025; Guo, Lau, and Wei 2019; Hascher, Hagenauer, and Sacchi 2024; Zakariya et al. 2019). This, in turn, increases students' interest in mathematics and supports their sense of self-efficacy (Grigg et al. 2018; Özcan and Kültür 2021). Conversely, a negative perception of mathematics leads to emotional responses in students such as anxiety, stress, and fear of failure towards the subject (Chand et al. 2021; Smith, Fotou, and Sharpe 2025; Usher and Pajares 2009). Viewing mathematics as an abstract, complex, and difficult subject to learn causes a loss of motivation for the course (Acharya 2017; Howard and Whitaker 2011). Furthermore, poor quality instructional processes, the lack of student-centred methods, insufficient teacher support, and negative attitudes from school, family, or the environment also contribute to the negative shaping of this perception (Emanet and Kezer 2021; Fufa, Tulu, and Ensene 2023). These conditions negatively affect students' learning processes, making it difficult for them to achieve success in mathematics. One of the primary factors contributing to the deterioration of mathematical perception and the persistence of the failure cycle is incomplete or incorrect understandings regarding the nature and function of mathematics.

Mathematics, beyond merely being about numbers and operations, is a process of thinking and problem-solving aimed at understanding the environment and the universe (Schoenfeld and Sloane 2016). The meaning of mathematics can be assessed as a multidimensional concept shaped by students' conceptual understandings, logical reasoning, and analytical thinking skills. From the students' perspective, the meaning of mathematics is formed based on their experiences, perceptions, and how they relate mathematical concepts to daily life (Van den Heuvel-Panhuizen and Drijvers 2020). This meaning is strengthened by teaching mathematics not just as an abstract subject, but as a functional tool that offers solutions to real-world problems. Mathematical meaning contributes to the development of students' abstract and analytical thinking, problem-solving, and solution-focused analysis skills.

Conceptually, meaning is not treated here as an inherent property of mathematics itself, but as something constituted through the individual's lived experience of it (Husserl and Findlay 1970). Schutz (1967) extended this premise by arguing that people's understanding of everyday and institutional phenomena is not given in advance but is retrospectively constructed as they reflect upon and interpret their own experience. Consistent with this epistemological stance, the present study approached the meaning of mathematics not as a predefined set of beliefs to be measured top-down, but as a construct to be elicited from students' own accounts of their experience with the subject.

This conceptualisation distinguishes the meaning of mathematics from adjacent affective constructs with which it is sometimes conflated. Attitude scales typically capture evaluative affect towards mathematics, specifically in terms of liking or disliking the subject (Di Martino and Zan 2011); self-efficacy captures domain-specific competence beliefs (Pajares 2005); and motivation concerns the energisation and direction of behaviour towards mathematical tasks. Meaning, as conceptualised here, instead concerns students' phenomenological construal of what mathematics is and what it does for them, representing an epistemic and functional understanding of the discipline rather than a felt evaluation of, competence belief about, or drive towards it. This distinction resonates with Skemp's (1976) classic contrast between instrumental understanding ("rules without reasons") and relational understanding ("knowing what to do and why"), recently revisited by Herheim (2023): constructing relational meaning for mathematics requires grasping why the discipline functions as it does for oneself, not merely how to apply it.

At the same time, the meaning students attribute to mathematics is theorised, within Eccles and Wigfield's Situated Expectancy-Value Theory, to overlap substantially with the utility-value component of subjective task value, defined as the perceived usefulness of a domain for one's present and future goals (Eccles 2005; Eccles and Wigfield 2020, 2023). Because the present item pool was generated bottom-up from students' own open-ended accounts of what mathematics means to them, rather than deduced solely from a single a priori theory, it is plausible that the items surviving the psychometric refinement process converged predominantly on this utility-value facet of meaning: several retained items explicitly reference mathematics' role in problem-solving, decision-making, and building confidence. We interpret this convergence not as evidence that the scale merely duplicates existing utility-value measures, but as a substantive empirical finding in its own right, indicating that for the students in this study, meaning-making around mathematics is to a considerable extent anchored in its perceived functional value.

The meaning students ascribe to mathematics is shaped by instructional approaches as well as by cultural and societal context, with teachers playing an important role in its construction (Beswick 2019; Mhlolo, Schafer, and Venkat 2012; Saxe 2015), and mathematics itself valued not only for individual development but also for its contribution to societal and economic progress (Abd Algani 2022; Yadav 2019). This meaning-making process is reciprocally linked to students' broader perception of mathematics: negative experiences of failure make it harder for students to construct meaning and sustain motivation (Aguilar 2021; Hannula 2002; Schoenfeld 1989), whereas finding mathematics meaningful is associated with greater motivation and academic success (Mutodi and

Ngirande 2014). The middle school years represent a particularly critical period in this process, as students' perceptions, attitudes, and beliefs about mathematics take shape during this stage and carry forward into subsequent achievement and motivation (Garcia et al. 2016; Hidayatullah and Csíkos 2024; Moenikia and Zahed-Babelan 2010; Pitsia, Biggart, and Karakolidis 2017), even though many students continue to experience mathematics as abstract, complex, and disconnected from daily life (Martin 2000; Schmitz et al. 2023).

As the meaning students construct around mathematics is closely tied to their academic performance, motivation, and broader cognitive and affective engagement with the subject, understanding how students attribute meaning to mathematics is critical for capturing the multidimensional nature of mathematical learning. Developing a valid and reliable instrument to assess this construct addresses a notable gap in the mathematics education literature, offering both a research tool and practical guidance: the findings are expected to clarify how meaningful engagement with mathematics shapes learning strategies and attitudes, informing curriculum design, instructional practice, and policies aimed at fostering more functional and sustainable student engagement with mathematics.

The primary objective of this research is to develop a measurement instrument aimed at determining the meaning of mathematics from the perspective of middle and high school students and to establish its psychometric properties. Furthermore, among the goals of the study is to compare the data obtained from students based on demographic variables such as gender, grade level, school type, parental education level, and income status.

Method

Participants

The development of the MofM scale involved data collection across three phases. The first phase, conducted in early June 2025, provided data for EFA. The second phase, also in early June, gathered data for CFA. The third phase took place in late September 2025 and supplied data for Item Response Theory (IRT) and Measurement Invariance (MI) analyses. To ensure national representativeness, one province was randomly selected from each of Turkey's 12 Nomenclature of Territorial Units for Statistics (NUTS) regions, including Adana, Ankara, Çanakkale, Denizli, Düzce, Istanbul, Kars, Kırıkkale, Muş, Ordu, Samsun, and Şanlıurfa. Sampling was proportional to population densities (Bakanlar Kurulu Kararı [Council of Ministers Decision] 2002). Data on variables including gender, grade level, school type, parental education level, and family income status were gathered and used for comparative analyses based on the third phase data. Participant characteristics across all phases are presented in Table 1.

Preparation of the MofM draft form

Three main strategies were followed in creating the draft form of the MofM scale. Firstly, theoretical and empirical literature related to education, school, mathematics, the meaning of education and school, and the meaning of mathematics was examined; a potential

Table 1. Distribution of participant students based on certain variables.

Variable		First Round - Data Used for EFA n(%)	Second Round- Data Used for CFA n(%)	Third Round - Data Used for IRT, MI, and Comparative Analyses n(%)
Gender	Female	321 (53.9)	319 (61)	2117 (55.8)
	Male	274 (46.1)	204 (39)	1674 (44.2)
City of Turkey	Adana	58 (9.7)	42 (8)	346 (9.1)
	Ankara	145 (24.4)	99 (18.9)	783 (20.7)
	Çanakkale	14 (2.4)	13 (2.5)	91 (2.4)
	Denizli	29 (4.9)	19 (3.6)	157 (4.1)
	Düzce	7 (1.2)	10 (1.9)	96 (2.5)
	İstanbul	223 (37.5)	221 (42.3)	1539 (40.6)
	Kars	8 (1.3)	6 (1.1)	69 (1.8)
	Kırıkkale	9 (1.5)	13 (2.5)	74 (2)
	Muş	19 (3.2)	8 (1.5)	79 (2.1)
	Ordu	19 (3.2)	24 (4.6)	134 (3.5)
	Samsun	31 (5.2)	39 (7.5)	213 (5.6)
	Urfa	33 (5.5)	29 (5.5)	210 (5.5)
Class	5.Class	45 (7.6)	43 (8.2)	336 (8.9)
	6.Class	80 (13.4)	56 (10.7)	489 (12.9)
	7.Class	94 (15.8)	93 (17.8)	624 (16.5)
	8.Class	103 (17.3)	89 (17)	658 (17.4)
	9.Class	51 (8.6)	35 (6.7)	329 (8.7)
	10.Class	95 (16)	90 (17.2)	620 (16.4)
	11.Class	43 (7.2)	45 (8.6)	252 (6.6)
	12.Class	84 (14.1)	72 (13.8)	483 (12.7)
School Type	Secondary School	290 (48.7)	250 (47.8)	1869 (49.3)
	Religion Secondary School	32 (5.4)	31 (5.9)	238 (6.3)
	Science and Social Science High School	29 (4.9)	32 (6.1)	221 (5.8)
	Anatolian High School	112 (18.8)	94 (18)	713 (18.8)
	Anatolian Religion High School	10 (1.7)	3 (0.6)	43 (1.1)
	Anatolian Vocational and Technical High School	72 (12.1)	76 (14.5)	430 (11.3)
	Others	50 (8.4)	37 (7.1)	277 (7.3)
Field of Study of Students'	Turkish and Mathematics Track	98 (16.5)	94 (18)	667 (17.6)
	Science Track	55 (9.2)	44 (8.4)	316 (8.3)
	Social Sciences Track	8 (1.3)	12 (2.3)	93 (2.5)
	Vocational Training	66 (11.1)	72 (13.8)	430 (11.3)
	No Division/Track	268 (45)	226 (43.2)	1690 (44.6)
	Foreign Language Track	7 (1.2)	11 (2.1)	41 (1.1)
	Others	93 (15.6)	64 (12.2)	554 (14.6)
Education Level of Students' Mother	Primary Education Graduate (Primary School + Secondary School)	270 (45.4)	220 (42.1)	1656 (43.7)
	High School Graduate	160 (26.9)	186 (35.6)	1113 (29.4)
	Bachelor, Master's and Doctorate Degree	165 (27.7)	117 (22.4)	1022 (27)
Education Level of Students' Father	Primary Education Graduate (Primary School + Secondary School)	215 (36.1)	190 (36.3)	1218 (32.1)
	High School Graduate	186 (31.3)	158 (30.2)	1262 (33.3)
	Bachelor, Master's and Doctorate Degree	194 (32.6)	175 (33.5)	1311 (34.6)

(*Continued*)

Table 1. (Continued).

Variable		First Round - Data Used for EFA n(%)	Second Round- Data Used for CFA n(%)	Third Round - Data Used for IRT, MI, and Comparative Analyses n(%)
Students' Family Income Level	Our income is less than our expenses	284 (47.7)	237 (45.3)	1644 (43.4)
	Our income fully covers our expenses	190 (31.9)	167 (31.9)	1241 (32.7)
	Our income is more than our expenses	121 (20.3)	119 (22.8)	906 (23.9)
Total		595 (100)	523 (100)	3791 (100)

item pool was generated based on this literature. Secondly, an open-ended written application was conducted with a total of 63 students (32 female, 31 male) studying in middle and high school. Students were asked questions such as "What do you feel when the mathematics lesson is mentioned? What are your thoughts about mathematics? Please write your thoughts freely", and they were encouraged to express their responses openly. The obtained student opinions were evaluated through content analysis and used as a resource in the development of the scale items. As a third strategy, the opinions of academics specialising in fields such as measurement and evaluation, mathematics education, and psychological counselling were consulted; scale items were reviewed and edited in line with expert feedback.

In the first two stages of the scale development process, a preliminary item pool of 26 items was compiled into a list and submitted for expert evaluation. For this purpose, the opinions of five academics specialising in mathematics education, educational psychology, measurement and evaluation, and the historical and social foundations of education were obtained. A structured evaluation form was prepared in which each item was accompanied by three response options, namely "the item is appropriate and should be included in the scale", "the item is not appropriate and should be removed from the scale", and "the item may be included in the scale provided it is revised", while experts were also invited to suggest specific wording revisions for each item. Prior to making any revisions, the consistency among the experts' ratings was examined by calculating Krippendorff's alpha coefficient (Krippendorff 2018), which yielded a value of .897, indicating an acceptable level of inter-rater consistency; items were subsequently revised in line with the experts' qualitative feedback. In addition, the content validity ratio (CVR) for each item was calculated using Lawshe's (1975) technique, based on the experts' item-appropriateness ratings. For a panel of five experts, Lawshe's critical CVR value is .99; only items rated as appropriate by all five experts met this threshold and were retained, and the remaining items were revised according to the experts' suggestions. As a result of these evaluations, a preliminary form consisting of 19 items was created, and pilot administrations for psychometric analyses were initiated; all 19 items retained after this stage accordingly obtained a CVR of 1.00.

Likert-type items were chosen as the item format for the scale, and response options were structured using a five-point rating system: "Strongly disagree", "Disagree", "Partially agree", "Agree", and "Strongly agree". The literature emphasises that 5-point Likert scales yield psychometrically reliable results and are comprehensible to respondents (Aybek and Toraman 2022). To enhance the suitability of the scale items' language and expression for the target audience, linguistic editing (redaction) support was obtained from an expert in the field of Turkish Language.

Procedure

The research was completed by following the procedural steps outlined in Figure 1:

Data analysis

The following procedures were performed in data analysis:

- The data obtained from the research were analysed using JAMOVİ and R statistical software.
- Before proceeding with data analysis, missing data control was first performed on the datasets.

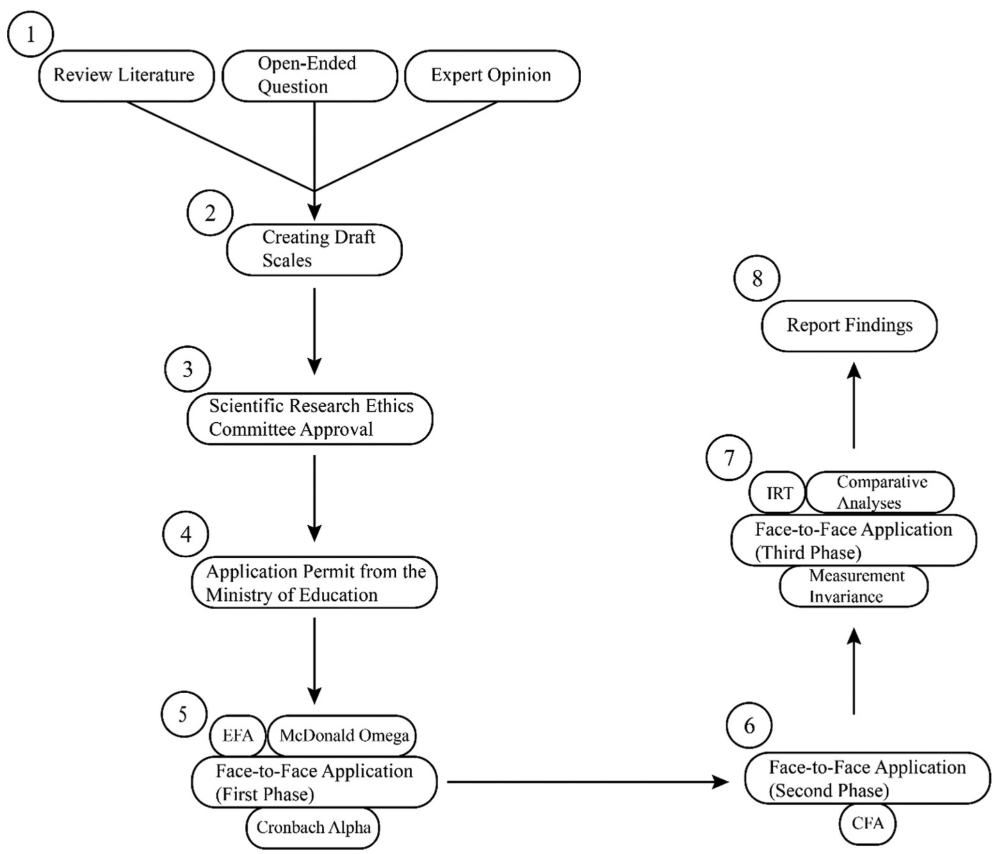

Figure 1. Research procedural steps.

- Subsequently, the multivariate normality of each data file was evaluated using Henze-Zirkler, Mardia, and Doornik-Hansen tests.
- Outlier analysis was conducted on the datasets and considered by taking into account their potential effects. Additionally, multicollinearity among scale items was examined.
- The Kaiser-Meyer-Olkin (KMO) test and Bartlett's Test of Sphericity were applied to determine the suitability of the sample for factor analysis.
- EFA was performed using the Principal Axis Factoring (PAF) method; when deciding on the number of factors, the eigenvalue criterion was considered, with 1 accepted as the reference value. To address the possibility that the eigenvalue-greater-than-one criterion alone may overestimate the number of factors, a parallel analysis (Horn 1965) was additionally conducted.
- The scale's ability to explain the variance related to its target construct (meaning of mathematics) was examined.
- The structure obtained from CFA was tested, and the model's fit level was evaluated through fit indices.
- Cronbach's Alpha, McDonald's Omega, and Marginal Reliability coefficients were calculated for the scale's internal consistency and reliability. Furthermore, IRT analyses were conducted to analyse the discrimination and difficulty levels of each item.
- MI analyses were performed for gender and school level variables to test the comparability of the scale across different groups. Additional metric invariance testing was conducted for parental education and family income, given their adequate subgroup sample sizes, whereas school type and field of study were not tested due to insufficient subgroup sizes for stable multi-group estimation (please see Findings).
- Finally, students' scores obtained from the Meaning of Mathematics Scale were evaluated through comparative analyses in terms of demographic variables.

Ethical committee approval

This research was conducted with the approval of the Duzce University Scientific Research and Publication Ethics Committee, dated 23 January 2025, decision number 2025/1–54, and the permission of the Ministry of National Education, dated 12 June 2025, number MEB.TT.2025.024278.

Findings

First phase

Data file examination

Missing data analysis was conducted across three distinct datasets (for EFA, CFA, and IRT), revealing no missing values. Tests for multivariate normality, including Henze-Zirkler, Mardia, and Doornik-Hansen, indicated that the assumption of multivariate normality was not met ($p < .05$). Outlier detection led to the removal of 18 cases from the EFA dataset, 13 from the CFA dataset, and 125 from the IRT dataset. Multicollinearity was assessed through Variance Inflation Factor (VIF) values, which ranged between 1.52 and

2.35, well below the critical threshold of 10, indicating no multicollinearity concerns (Tabachnick and Fidell 2013). In the EFA dataset, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.927, while Bartlett's Test of Sphericity yielded a significant result ($\chi^2 = 1899$, $p < .05$), confirming the suitability of the data for factor analysis. A KMO value above .90 is considered excellent, and Bartlett's test result suggests that the correlation matrix significantly deviates from an identity matrix, further supporting the appropriateness of the dataset for factor analysis (Field 2018; Kaiser and Rice 1974; Pett, Lackey, and Sullivan 2003).

Deciding on the number of factors

As a result of the EFA performed using the Principal Axis Factoring method, only one factor with an eigenvalue greater than 1 was identified. This factor has an eigenvalue of 5.97, explaining 55% of the total variance. The obtained variance ratios are within the limits accepted in the literature (Hair et al. 2014; Warner 2013). Upon examining the factor loadings, it was determined that items 3, 4, 5, 6, 9, 10, 11, and 12 had loadings below 0.30, and these items were removed from the scale. After item removal, a total of 11 items remained in the scale. The details of the items under the factor are presented in Table 2.

To provide further empirical support for the retained factor structure, a parallel analysis (Horn 1965) was conducted using Jamovi. Random datasets matching the sample size and number of items were generated, and their initial eigenvalues (based on the correlation matrix) were compared against those obtained from the actual data. As illustrated in Figure 2, the eigenvalue of the first factor (8.43) clearly exceeded the corresponding eigenvalue derived from the randomly generated datasets, while the eigenvalues of the remaining factors closely tracked the random-data reference line. To confirm this result numerically rather than rely on visual inspection of the scree plot alone, the number of factors to extract was additionally set to be determined automatically based on the parallel analysis criterion; under this specification, Jamovi's factor extraction algorithm retained exactly one factor. This result corroborates the single-factor solution identified

Table 2. Item loadings after EFA.

Items	Factor	Item Total Correlation
Item 2. Knowing mathematics improves my capacity to solve problems I encounter.	0.861	0.820
Item 7. Mathematics allows me to learn new information.	0.823	0.781
Item 14. Mathematics develops my research skills.	0.817	0.786
Item 19. Mathematics facilitates data-driven decision-making.	0.761	0.728
Item 15. Mathematics enables me to focus on solutions rather than problems.	0.746	0.719
Item 16. Mathematics enables me to use resources efficiently.	0.735	0.711
Item 18. Mathematics allows me to reach correct inferences by starting from assumptions.	0.733	0.703
Item 17. Knowing mathematics positively affects my self-confidence.	0.673	0.636
Item 13. Mathematics is the most useful tool in humanity's hands.	0.645	0.630
Item 8. Mathematics functions like an intelligence game that develops the human mind.	0.643	0.621
Item 1. It is impossible to understand nature without mathematics.	0.623	0.603

Explained Variance = 54.6(%55).
RMSEA = 0.076.
TLI = 0.950.
$X^2 = 197$ (df = 44, $p < .05$).
Bartlett's Test of Sphericity = 3860 (df = 55, $p < .05$).
KMO = 0.949.

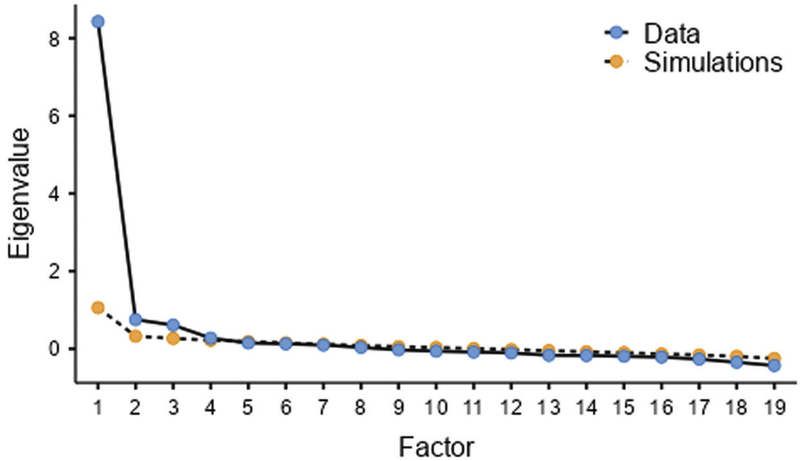

Figure 2. *Scree plot comparing eigenvalues from the actual data and Mean eigenvalues from parallel random data Sets* (jamovi, 500 simulated datasets).

through the Kaiser criterion. It should be noted that this initial eigenvalue reflects the unreduced correlation matrix and is therefore not directly comparable to the eigenvalue obtained after Principal Axis Factoring extraction (5.97, reported above), which is based on communality-adjusted variance; nevertheless, both indices converge on the same single-factor conclusion.

Following the EFA procedure, factor loading values were found to range between .623 and .861, while item-total correlations ranged from .603 to .820. There are no negatively worded or reverse-coded items among the 11 items included in the scale. The highest possible score on the scale is 55, and the lowest is 11. Higher scores indicate that students attribute a high meaning to mathematics, whereas lower scores suggest they perceive mathematics as meaningless.

Examination of reliability level

The reliability level of the Meaning of Mathematics scale was examined using Cronbach's Alpha, McDonald's Omega, and marginal reliability coefficients. The results are presented in Table 3.

The reliability values obtained for the MofM scale are .70 and above. According to Nunnally and Bernstein (1994), adequate reliability should be at least 0.70 or higher.

Second phase

Confirmation of MofM's factor structure

As a result of the EFA analyses, it was understood that the MofM consists of a total of 11 items and exhibits a single-factor structure. One of the stages of scale development is the

Table 3. MofM Cronbach's Alpha, McDonald's Omega, and Marginal Reliability Level.

Scale	Cronbach Alpha	McDonalds Omega	Marjinal
Meaning of Mathematics Scale	0.927	0.928	0.903

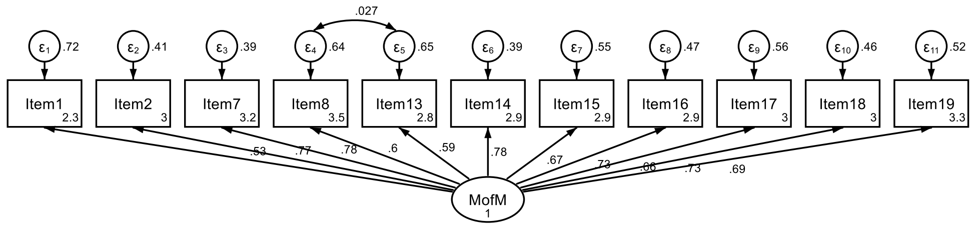

Figure 3. CFA diagram of the MofM single-factor Structure.

confirmation of discovered factor structures with data obtained from another sample from the target audience. The confirmation process is performed using CFA. The diagrams obtained for MofM as a result of CFA can be seen in Figure 3.

As a result of CFA, the fit indices were calculated as follows: CFI = .961, TLI = .950, RMSEA = .069, $X^2 = 148.549$ (df = 43, $p < .05$). These values are above acceptable levels according to scale development literature (Anderson and Gerbing 1984; Bentler 1990; Hooper, Coughlan, and Mullen 2008; Hu and Bentler 1999; Kline 2005; Marsh, Balla, and McDonald 1988; Tabachnick and Fidell 2013; Vieira 2011). The CFA results indicate that the EFA factor structure of the MofM has been confirmed.

Third phase

Providing validity and reliability evidence through IRT analyses

In the third stage, the 11-item MofM scale, determined as a result of EFA and CFA, was administered to 3791 students. Using these data, IRT analyses were conducted. In IRT analyses, the assumptions of unidimensionality and local independence of the scale were evaluated (Zhao 2008). Unidimensionality was confirmed by EFA, proving that MofM has a single-factor structure. The local independence assumption was tested using the Q3 statistics (Yen 1993), and analyses were performed via the "Mirt" package program (version 1.30; Chalmers 2012) within the R program (version 4.1.2). A critical value of 0.20 was adopted, and no item violating local independence was detected in the examinations.

Within IRT, item calibrations were performed using the Generalised Partial Credit Model (GPCM). The calculated statistics are presented in Table 4.

In IRT analyses, RMSEA values were used to evaluate item fit; values of 0.080 and below, accepted as a fit indicator (Stout 1990), were met for MofM items. Additionally, the Scaled Pearson Chi-square Index (S_χ^2) was used to measure the fit of the items to the model, and items were found to be consistent with significance values of $p > .05$ (Orlando and Thissen 2000). Estimations based on the GPCM model (LogLikelihood, $p < .05$) also indicated that item fit was achieved. For the item discrimination parameter ("a"), the acceptable range is between 0.75 and 2.50 (Flannery, Reise, and Widaman 1995), while the item difficulty parameter ("b") is ideally considered between −1.00 and 1.00 (Hambleton 1994). Analysis results showed that the discrimination values of some items exceeded 2.50, and some items' difficulty parameters fell outside the ideal limits. Item information functions and item characteristic curves for MofM items are presented in Figures 4 and 5.

Table 4. MofM item parameters and item fit-Indexes.

	GPCM			Item Difficulty and Discrimination Values				
Items	S_χ2	df	RMSEA	a	b1	b2	b3	b4
Item 1	1270.768	123	0.050	1.249	−1.666	−0.874	0.303	1.168
Item 2	1101.771	97	0.052	3.024	−1.635	−1.124	−0.460	0.353
Item 7	1088.955	102	0.051	2.634	−1.920	−1.272	−0.547	0.236
Item 8	1311.285	127	0.050	1.571	−2.478	−1.889	−0.929	0.192
Item 13	1268.116	115	0.051	1.780	−1.955	−1.121	−0.270	0.954
Item 14	1106.460	100	0.052	2.789	−1.595	−1.050	−0.332	0.481
Item 15	1347.573	111	0.054	2.150	−1.916	−1.265	−0.437	0.385
Item 16	1166.660	113	0.050	2.264	−1.658	−1.132	−0.281	0.595
Item 17	1188.921	123	0.048	1.784	−1.954	−1.394	−0.680	0.254
Item 18	1133.924	110	0.050	2.064	−1.952	−1.423	−0.449	0.589
Item 19	1298.459	114	0.052	1.939	−2.161	−1.520	−0.675	0.368

IRT of Calibration Values for Child Form Factor 1: Log Likelihood = −50, 154.276, $p < .05$ Iteration = 91.
GPCM: Generalised Partial Credit Model.
RMSEA: Root Mean Square Error of Approximation.
S_χ2: Scaled Pearson Chi-square Index.
df: Degree of Freedom.
a: Item Discrimination.
b: Item Difficulty.

The item information function graphically illustrates the range of a trait where the item best discriminates among individuals in the measurement instrument (Edelen and Reeve 2007). Examination of the graphs revealed that the items provided a high level of information regarding the measured trait, the meaning of mathematics. Furthermore, Item 1 and Item 8 were determined to provide lower, but still high, levels of information compared to other items.

The item probability function is a mathematical function that shows the probability of a person correctly answering a specific item or responding with a particular option, given their ability level (θ) (de Ayala 2009). For Likert-type items, the item probability function reveals how the item functions with different choices. In the examined graphs, it was observed that the "disagree" option was less functional than other options in many items. Specifically, items 1, 8, 16, 17, 18, and 19 were determined to function more like four-option items.

Providing evidence for MI

MI is a structural equation modelling (SEM) method used to evaluate the consistency of a measurement instrument's factor structure and whether it measures the same construct across different groups (e.g. gender, age, or cultural groups). The primary goal of this method is to determine whether observed differences between groups stem from true differences in the measured construct or from systematic biases originating from the scale. MI is examined hierarchically through configural, metric, scalar, and strict invariance tests (Meredith 1993; Putnick and Bornstein 2016).

The measurement models in Figure 3 were tested via CFA in separate groups for students based on gender and school level (middle school and high school). The CFA fit indices are as follows:

- For MofM female; $X^2/df = 2.15$, RMSEA = 0.065, CFI = 0.958, and TLI = 0.949,

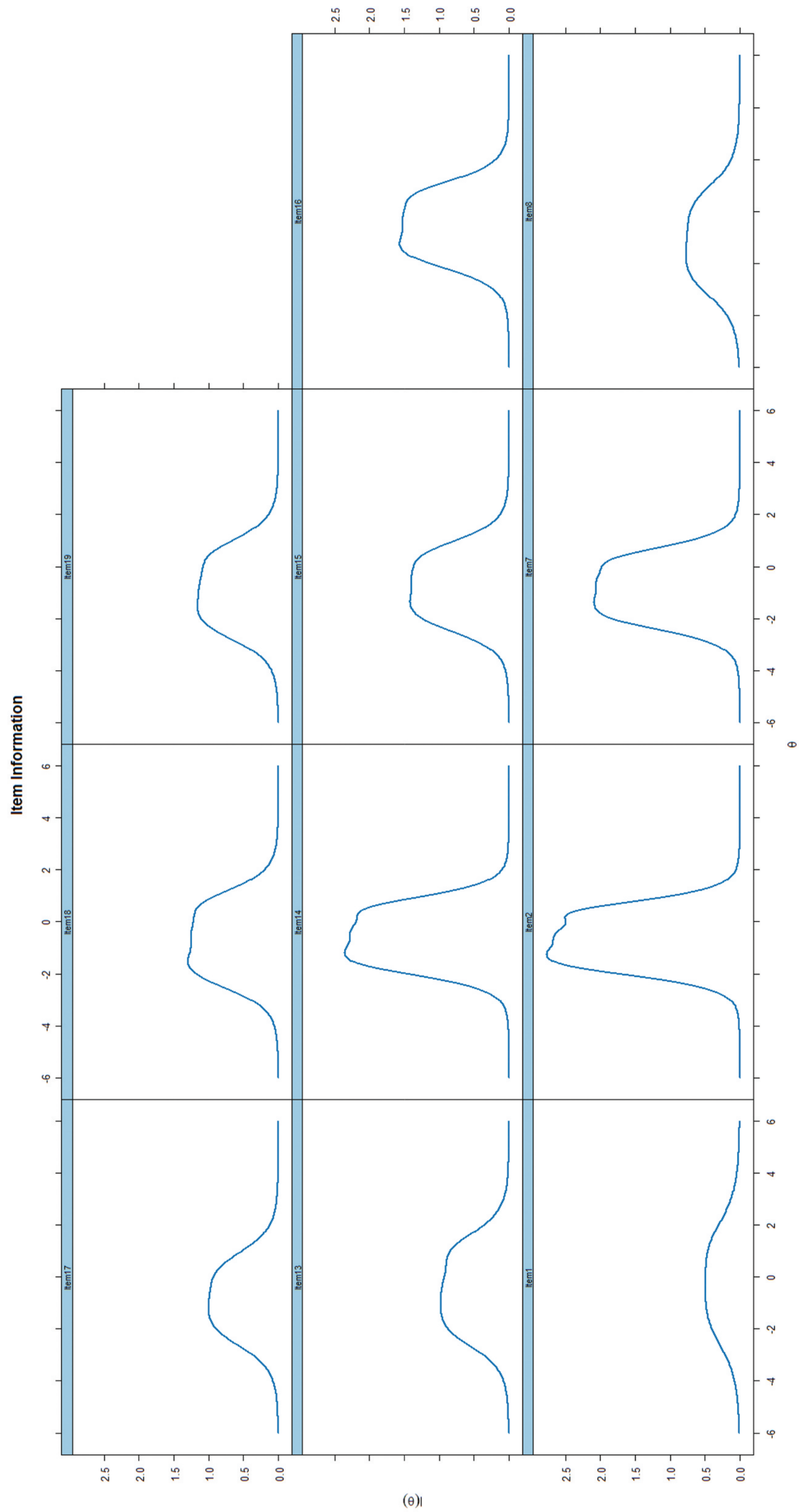

Figure 4. MofM item information function.

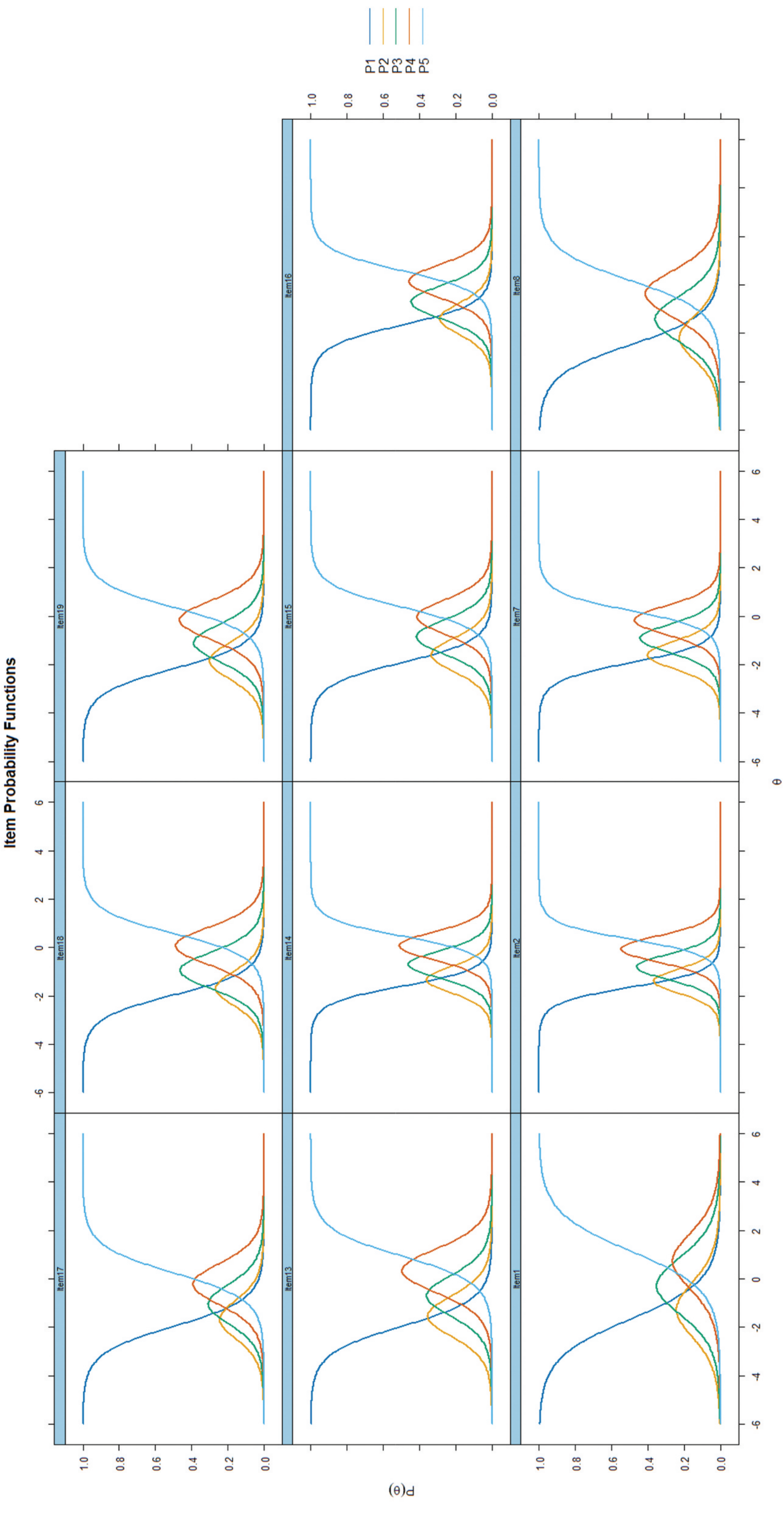

Figure 5. MofM item probability function.

- For MofM male; $X^2/df = 2.05$, RMSEA = 0.059, CFI = 0.962, and TLI = 0.950,
- For MofM secondary school; $X^2/df = 2.66$, RMSEA = 0.073, CFI = 0.921, and TLI = 0.911,
- For MofM high school; $X^2/df = 2.56$, RMSEA = 0.071, CFI = 0.912, and TLI = 0.901.

The structure previously determined in EFA analyses and confirmed for the entire group with CFA was found to be consistent with the accepted CFA fit index values in the literature for both male and female students, as well as for middle school and high school students, indicating that it measures the same construct. The MI of the MofM scale according to gender and educational level was examined through configural, metric, scalar, and strict invariance tests, respectively. The obtained fit indices and ΔCFI values were evaluated, and the results are summarised in Table 5.

The MofM scale demonstrates the same factor structure for both genders and for both middle school and high school students. The ΔCFI values obtained in the metric, scalar, and strict MI tests were found to be less than 0.01. The ΔCFI criterion proposed by Cheung and Rensvold (2002) and Chen (2007) was accepted as ΔCFI ≤ 0.01 in this research.

Additional MI testing for parental education and family income

MI was tested only for gender and school level, additional analyses were considered for school type, field of study, parental education, and family income. However, several subgroups within school type and field of study contained as few as 41 to 93 participants, which is well below the sample sizes recommended for stable multi-group CFA estimation (Jackson 2003; MacCallum et al. 1999); invariance testing was therefore not attempted for these two variables. In contrast, the subgroups for parental education and family income each contained at least 906 participants, a sample size adequate for multi-group modelling. Metric invariance (equality of factor loadings) was therefore tested for mother's education, father's education, and family income using a multi-group CFA with Lagrange multiplier (score) tests for each loading constraint, estimated via diagonally weighted least squares given the ordinal nature of the items. The results of these analyses are presented in Table 6.

Table 5. Measurement invariance data.

Child Form					
Measurement invariance		RMSEA	CFI	TLI	ΔCFI
Gender	Configural	0.065	0.951	0.935	
	Metric	0.064	0.952	0.938	0.000
	Scalar	0.061	0.955	0.939	0.002
	Strict	0.059	0.954	0.941	0.006
School Level	Configural	0.075	0.931	0.920	
	Metric	0.072	0.933	0.921	0.000
	Scalar	0.069	0.934	0.923	0.003
	Strict	0.066	0.934	0.925	0.007

Table 6. Multi-group CFA fit and loading invariance tests for parental education and family income.

Variable	Scaled χ^2	df	CFI (scaled)	TLI (scaled)	RMSEA (scaled) [95% CI]
Family Income	1950	150	0.970	0.967	0.097 [0.094, 0.101]
Mother's Education	2121	150	0.965	0.961	0.102 [0.098, 0.106]
Father's Education	1993	150	0.966	0.962	0.099 [0.095, 0.103]

Non-invariant factor loadings (Lagrange multiplier score tests, $p < .05$): Family Income = 17/20; Mother's Education = 14/20; Father's Education = 14/20.

Although the multi-group models showed acceptable overall fit, the majority of factor loadings were found to differ significantly across groups for all three variables (14 to 17 of the 20 tested constraints, $p < .05$), indicating that metric invariance does not hold for mother's education, father's education, or family income. Because scalar (threshold) invariance presupposes metric invariance, and the latter was not supported, scalar invariance testing was not pursued for these three variables; consequently, group comparisons of MofM scores based on these variables cannot be assumed to rest on equivalent measurement. This means that the MofM items do not necessarily relate to the underlying construct in the same way across these subgroups.

Comparison of students' MofM scores across different variables

Within the scope of the third phase, the 11-item MofM scale, determined as a result of EFA and CFA, was administered to 3791 students. In comparative analyses conducted with the obtained data, the Mann-Whitney U Test was used for two-group comparisons, and the Kruskal-Wallis Test for comparisons involving three or more groups. This choice was made because the participants' MofM scores did not show a normal distribution. The significance level was set at $p < .05$, and the results are presented in Table 7.

Comparative analyses revealed significant differences in students' meaning attributed to mathematics based on several demographic and socioeconomic factors ($p < .05$). Male students assigned greater meaning to mathematics than females. The perceived meaning declined as grade level increased, notably after 8^{th} grade, but rose again in 12^{th} grade due to university exam preparation. School type also influenced meaning. Imam Hatip middle schools, science-social sciences high schools, and Anatolian Imam Hatip High Schools attributing higher meaning, while Anatolian High Schools and Anatolian Vocational Technical High Schools showed lower levels. Regarding study fields, students in science, Turkish-mathematics, or undecided fields valued mathematics more than those in foreign language or social sciences. Parental education level impacted meaning: students with mothers or fathers holding only primary education attributed less meaning than those with parents educated at high school level or above. Additionally, students whose family income was below expenses assigned less meaning to mathematics compared to those with income equal to or exceeding expenses, with a gradual increase observed across income levels. Overall, these findings highlight that students' mathematical meaning perceptions are strongly shaped by gender, grade, school type, study field, parental education, and family income. To contextualise the practical significance of these differences given the very large sample size, effect sizes (rank-biserial correlation for two-group comparisons; epsilon-squared for comparisons involving three or more groups) were additionally computed and are reported in Table 7.

Discussion and conclusion

The construct validity of the MoM scale in the research was examined through EFA, CFA, IRT, and MI analyses. The EFA results revealed a reliable single-factor structure composed of 11 items after removing eight items with item-total correlations below 0.30 or containing excessively positive or negative statements. This structure accounted for approximately 55% of the total variance, which is at an acceptable level in the social sciences

Table 7. Comparison of the meaning of mathematics for students based on different variables.

Variable		N	Mean (S. Deviation)	Median (Min.-Max.)	Test (p)	Effect Size	Significant Difference
Gender	1.Female	2117	39.61 (9.99)	41 (12–55)	1356143.500* (<0.001)	0.235	1 < 2
	2.Male	1674	43.24 (9.63)	47 (11–55)			
Class	5.Class	336	45.12 (8.70)	47 (23–55)	382.430** (<0.001)	0.101	5 > 8, 6 > 8, 7 > 8
	6.Class	489	44.48 (8.73)	47 (15–55)			5 > 9, 6 > 9, 7 > 9
	7.Class	624	44.13 (8.36)	47 (23–55)			8 > 9, 5 > 10, 6 >
	8.Class	658	41.59 (10.89)	44 (12–55)			10
	9.Class	329	38.53 (10.28)	36 (14–52)			7 > 10, 8 > 10, 9 >
	10.Class	620	36.69 (9.21)	37 (14–54)			10
	11.Class	252	37.91 (10.85)	39 (11–55)			10 < 12, 5 > 11,
	12.Class	483	40.28 (9.54)	39 (11–55)			6 > 11
							7 > 11, 8 > 11, 5 >
							12
							6 > 12, 7 > 12, 8 >
							12
School Type	1.Secondary School	1869	43.72 (9.32)	46 (12–55)	407.500** (<0.001)	0.108	1 > 4, 1 > 6, 1 > 7
	2.Religion Secondary School	238	42.44 (10.33)	47 (18–55)			2 > 4, 2 > 6, 2 > 7
	3.Science and Social Science High School	221	42.41 (9.79)	45 (17–55)			3 > 4, 3 > 6, 3 > 7
	4.Anatolian High School	713	37.87 (9.71)	38 (14–55)			4 < 5, 4 < 6, 4 > 7
	5.Anatolian Religion High School	43	45.60 (3.76)	46 (41–52)			5 > 6, 5 > 7
	6.Anatolian Vocational and Technical High School	430	38.31 (9.31)	40 (15–53)			6 > 7
	7.Others	277	34.73 (10.16)	33 (11–54)			
Field of Study of Students'	1.Turkish and Mathematics Track	667	40.91 (8.78)	41 (14–55)	181.196** (<0.001)	0.0478	1 < 2, 1 > 3, 1 > 4
	2.Science Track	316	44.34 (8.34)	46.5 (17–55)			1 < 5, 1 > 6
	3.Social Sciences Track	93	36.33 (10.24)	33 (18–51)			2 > 3, 2 > 4
	4.Vocational Training	430	37.86 (10.34)	40 (15–53)			2 > 6, 2 > 7
	5.No Division/Track	1690	42.59 (9.94)	46 (12–55)			3 < 5, 4 < 5
	6.Foreign Language Track	41	34.73 (9.99)	40 (19–46)			5 > 6, 5 > 7
	7.Others	554	39.50 (10.62)	40 (11–55)			6 < 7
Education Level of Students' Mother	1.Primary Education Graduate (Primary School + Secondary School)	1656	39.10 (10.06)	40 (11–55)	155.858** (<0.001)	0.0411	1 < 2 1 < 3
	2.High School Graduate	1113	42.30 (10.20)	45 (12–55)			
	3.Bachelor, Master's and Doctorate Degree	1022	43.45 (8.93)	46 (14–55)			

(*Continued*)

Table 7. (Continued).

Variable		N	Mean (S. Deviation)	Median (Min.-Max.)	Test (p)	Effect Size	Significant Difference
Education Level of Students' Father	1.Primary Education Graduate (Primary School + Secondary School)	1218	38.41 (10.32)	39 (11–55)	163.013** (<0.001)	0.0430	1 < 2 1 < 3
	2.High School Graduate	1262	41.82 (9.92)	44 (12–55)			2 < 3
	3.Bachelor, Master's and Doctorate Degree	1311	43.23 (9.15)	46 (14–55)			
Students' Family Income Level	1.Our income is less than our expenses	1644	40.13 (10.09)	41 (11–55)	48.867** (<0.001)	0.0129	1 < 2
	2.Our income fully covers our expenses.	1241	41.50 (10.18)	44 (14–55)			1 < 3
	3.Our income is more than our expenses	906	42.79 (9.31)	46 (12–55)			2 < 3

*Mann Whitney U Test.
**Kruskal Wallis Test.

(Fabrigar and Wegener 2012; Hair et al. 2014; Kline 2005). Factor loadings for the scale items ranged from .62 to .86, and item-total correlations ranged from .60 to .82, both substantially exceeding the generally accepted minimum thresholds in the social sciences (Field 2018; Hajjar 2018; Williams, Onsman, and Brown 2010).

The accuracy of the single-factor structure derived from the EFA was tested with CFA. The CFA results (CFI = .961, TLI = .950, RMSEA = .069, X2 = 148.549; df = 43, $p < .05$) were found to be above the recommended thresholds in the literature, indicating a good fit of the model to the data (Anderson and Gerbing 1984; Bentler 1990; Hooper, Coughlan, and Mullen 2008; Hu and Bentler 1999; Kline 2005; Marsh, Balla, and McDonald 1988; Tabachnick and Fidell 2013; Vieira 2011). Thus, the CFA confirmed the factor structure determined by the EFA. In the third phase of construct validity, the 11-item MoM scale was administered to a large sample of 3,791 individuals, and IRT analyses were used to examine the assumptions of unidimensionality (DeMars 2010) and local independence (Yen 1993). While factor analysis supported the single-factor structure, no items exceeded the critical value of 0.20 for local independence, indicating that the items functioned independently and the model assumptions were met.

In the IRT analyses, RMSEA values below 0.080 (Hu and Bentler 1999; Stout 1990) demonstrated a high degree of fit between the items and the IRT model used. High significance levels ($p > .05$) in the Scaled Pearson Chi-square analyses further supported the items' fit to the model (Orlando and Thissen 2000). The estimation results from the Generalised Partial Credit Model (GPCM) also confirmed this fit, with Log-Likelihood statistics ($p < .05$) showing that the model fit the data well and successfully reflected variance at the item level (Muraki 1992; Warm 1989). The item discrimination (a) parameter for most items fell within the ideal range of 0.75–2.50, with some exceeding this value, which indicates high sensitivity in differentiating between participants (Gyamfi and Acquaye 2023). Closer inspection of the threshold estimates presented in Table 4, however, shows that the first two threshold parameters (b1 and b2) for most items fell considerably below the conventionally ideal range, ranging from approximately −2.478 to −1.595 for b1, whereas the third and fourth thresholds (b3 and b4) generally remained within the ideal range of −1.00 to 1.00. This pattern indicates that the items are relatively easy to endorse, providing substantial information at low-to-moderate trait levels but comparatively less discrimination among students at the higher end of the meaning-of-mathematics continuum, consistent with the reduced functioning of the lowest response category discussed below. Overall, the IRT findings indicate that the construct validity of the MoM scale is reasonably strong, though measurement precision is less pronounced among students who already attribute high meaning to mathematics.

Item Information Function analyses demonstrated that the items in the MoM scale measure the "meaning of mathematics" at a high level (Hambleton and Swaminathan 2013). In the Item Characteristic Curve plots, the "disagree" option was observed to have lower discrimination compared to other response categories (Embretson and Reise 2013). This finding does not contradict the rationale for adopting a five-point format at the scale-development stage, given that five-point response formats have generally been shown to offer advantages over three-point formats in terms of reliability and test information (Aybek and Toraman 2022). Nevertheless, the present IRT results suggest that, specifically for this instrument, the lowest response category may be underperforming for several items. Since re-administering the scale with a different number of response categories was beyond the scope of the present study, we recommend that future research administer three- and four-point versions of the MofM scale and empirically compare their psychometric performance against the present five-point format; such comparative evidence would help determine the optimal response format for this specific construct and would further contribute to the broader literature on response-category functioning in Likert-type scales. The MI of the scale was tested based on gender (female-male) and education level (middle school-high school). The structure determined by EFA and confirmed by CFA in the general sample was found to be valid in the subgroups; fit indices in the confirmatory factor analyses for the gender and school level groups were at acceptable levels. These findings indicate that the MoM scale measures the same construct across different groups and that MI is established (Millsap 2011). In conclusion, the MoM scale is determined to be a valid and reliable instrument for measuring the meaning students in middle and high school ascribe to mathematics.

Comparative analyses have shown a significant difference in the meaning of mathematics based on the gender variable. The level of meaning ascribed to mathematics was found to be higher among male students than female students. Similar findings in the literature from studies on variables such as mathematics achievement (Hyde and Mertz 2009; Hyde, Fennema, and Lamon 1990; Perez Mejias et al. 2021), attitudes towards mathematics (Brandell and Staberg 2008; Dowker and Sheridan 2022; Hyde, Fennema, and Lamon 1990; Reilly, Neumann, and Andrews 2019), mathematics self-efficacy (Hackett and Betz 1989; Huang 2013; Louis and Mistele 2012; Pajares 2005), and mathematics learning motivation (Gasco, Goñi, and Villarroel 2014; Skaalvik and Skaalvik 2004; Xiao and Sun 2021) generally indicate this difference favouring male students. This finding suggests that male students develop a more positive perception of mathematics, establish a stronger personal connection with the subject, and view it as a more functional and valuable discipline. Gender roles, family attitudes, teacher behaviours, and implicit messages conveyed in the instructional environment are considered to be influential in the emergence of this difference. Furthermore, the lower perception of the meaning of mathematics among female students may be related to their mathematical achievement, self-efficacy beliefs, motivation levels, and attitudes, which aligns with previous research. However, the magnitude of this difference was modest (rank-biserial correlation = 0.235), corresponding to a small-to-moderate effect according to conventional benchmarks (Cohen 1988); although statistically significant given the large sample size, the practical difference between male and female students was not substantial. It should be noted, however, that these explanations remain plausible interpretations rather than empirically confirmed causal mechanisms, as the cross-sectional design of this study does not allow for causal inference regarding the role of gender roles, family attitudes, or teacher behaviour.

Analysis indicated a significant inverse relationship between grade level and the meaning students ascribe to mathematics, showing a diminished sense of meaning as students advance. While middle schoolers reported higher meaning, a notable decline began in the 8^{th} grade. This trend was reversed in the 12^{th} grade, where meaning re-increased, coinciding with the proximity of the university entrance examination. This suggests students' perception of mathematics shifts from intrinsic meaning towards association with examination success and academic necessity as they progress through their education. Students, who exhibit a more open, curiosity-based, and meaning-seeking approach to mathematics particularly in the early years of middle school, may become disengaged from the subject as their grade level increases, a pattern that may be related to examination pressure, achievement expectations, and instructional styles. This trend is also supported by research conducted with different age groups. Niemivirta et al. (2024) indicated that students' perceptions of mathematical competence and intrinsic value decrease over time. Campbell, Vela, and Powell (2025) reported that students' interest and self-confidence in mathematics decline as the grade level advances. Similarly, a study conducted in Turkey observed significant decreases in students' attitude scores towards mathematics as the grade level increased (Gulburnu and Yildirim 2021). These findings suggest that changes in the meaning and perception of mathematics are influenced not only by cognitive factors but also by affective and motivational factors. However, according to the research findings, the meaning of mathematics generally decreases as the grade level increases, but this meaning rises again in the 12th grade. This increase may be associated with the perception of mathematics as a more valuable and functional subject for exam success, coinciding with the Higher Education Institutions Exam (HEIE) administered in the 12th grade in Turkey. The value students attribute to mathematics tends to increase during examination periods (Köller, Baumert, and Schnabel 2001). This phenomenon supports the view, in accordance with Eccles' Expectancy-Value Theory (Eccles 2005; Eccles and Wigfield 2020), that increased expectations of success lead students to perceive mathematics as a more meaningful and important domain. The overall findings reveal that the meaning students ascribe to mathematics changes periodically, and this change coincides with examination timing and academic necessities. The corresponding effect size ($\varepsilon^2 = 0.101$) indicated a moderate effect (Tomczak and Tomczak 2014), suggesting that grade level accounts for a non-trivial, though not large, proportion of the variance in the meaning attributed to mathematics. It is important to note that these grade-level differences were examined using a cross-sectional design; therefore, they may reflect cohort differences between students of different ages and school-entry cohorts rather than genuine developmental change within individuals over time, and the proposed explanation involving university entrance examination preparation, while plausible, remains a post hoc interpretation that was not directly tested in this study.

The research determined that the type of school students attend creates significant differences in the meaning of mathematics. The students who attributed the least meaning to mathematics were those from Anatolian High Schools and Anatolian Vocational and Technical High Schools. Conversely, the students who assigned the highest meaning to mathematics were those from middle schools, Imam Hatip middle schools (religious vocational schools), Science-Social Sciences High Schools, and Anatolian Imam Hatip High Schools. This finding is consistent with previous studies in Turkey that examined the effect of school type on students' mathematical perception (Cansız-Aktaş and Aktaş 2012; Çelik and Ceylan 2009; Yasar 2016). Furthermore, PISA 2003 Turkey data indicated that students in high-achieving schools, such as science high schools and Anatolian high schools, had stronger mathematical self-efficacy and mathematical self-concept, and clear differences existed between vocational high schools and general high schools (World Bank 2005). PISA 2018 results for Turkey also reported that the mathematics performance of schools admitting students by examination was significantly higher (OECD 2019).

The meaning students attribute to mathematics appears to be associated not only with individual beliefs but also with institutional factors such as school type, student profile, and academic expectations. Students in high-achieving, exam-admitting schools, including Science-Social Sciences High Schools and Anatolian Imam Hatip High Schools, report elevated meaning, a finding linked to their selective admission based on high mathematical achievement and intrinsic motivation. Conversely, lower meaning in vocational and Anatolian high schools may be associated with differences in student demographics, resource availability, and motivation.

Furthermore, field of study significantly differentiates the ascribed meaning. Students in Science and Turkish-Mathematics fields, as well as those who have not yet chosen a field, attribute the highest meaning to mathematics. This is explained by the subject's direct relevance and heavy curricular weighting. In contrast, students in Foreign Language and Social Sciences fields assign the lowest meaning, which is attributed to the more indirect connection with mathematics, course content divergence, and subsequent decline in attitude and motivation. These outcomes affirm that students' field choices and associated achievement expectations are significantly associated with their connection with mathematics, a finding consistent with Eccles' Expectancy-Value Theory (Eccles 2005; Eccles and Wigfield 2020). The effect size associated with school type ($\varepsilon^2 = 0.108$) indicated a moderate effect, whereas the effect associated with field of study ($\varepsilon^2 = 0.048$) was comparatively smaller, though still within the moderate range (Tomczak and Tomczak 2014), indicating that both institutional factors meaningfully, but not overwhelmingly, differentiate students' ascribed meaning. As with the other demographic comparisons, these institutional and motivational explanations are plausible interpretations consistent with the observed pattern of scores rather than causal conclusions that can be drawn from the present cross-sectional data.

The research also determined that the educational level of the students' mothers created a significant difference in the meaning of mathematics. The meaning ascribed to mathematics by students whose mothers had a primary education level was lower compared to students whose mothers had high school, university, and higher levels of education. This indicates that the educational level within the family is associated with students' academic attitudes, perceptions, and motivations towards mathematics and with their perception of meaning. Support from highly educated mothers enables the student to better grasp the importance of mathematics, while this support and awareness may be limited in families with lower educational levels (Davis-Kean 2005; Fan and Chen 2001; Hidayatullah and Csíkos 2024; Marks 2008; Susperreguy et al. 2022; Tomul, Önder, and Taslidere 2021). Furthermore, different studies have also found that children whose mothers have a higher educational level exhibit higher mathematics achievement and more positive attitudes towards mathematics (Çanakçı and Özdemir 2015; Deniz and Cıtdır 2020). This finding suggests that maternal education level is a significant factor, extending beyond its role as a socioeconomic indicator, by being closely associated with students' perceptions, values, academic goals, and self-efficacy in mathematics. High maternal education facilitates higher achievement expectations in the home, increased access to extracurricular resources, and a more academically motivating environment, all of which are associated with students perceiving mathematics as more meaningful and developing a positive attitude towards the subject. Concurrently, this result highlights educational inequalities, as students whose mothers have lower educational attainment may be disadvantaged in their attitudes and achievement due to a lack of sufficient resources and support. This difference corresponded to a moderate effect size ($\varepsilon^2 = 0.041$; Tomczak and Tomczak 2014).

Similarly, it was determined that the educational level of the students' fathers also created significant differences in the meaning of mathematics. The meaning ascribed to mathematics by students whose fathers had a primary education level was lower compared to students whose fathers had high school, university, and higher levels of education. Furthermore, the meaning ascribed to mathematics by students whose fathers had a high school education was lower than that of students whose fathers had a university education or higher. The tendency for students to attribute a higher meaning to mathematics as the father's educational level increases may be related to academic parental modelling, expectation levels, and the value placed on learning within the family. Educated fathers generally become more involved in their children's education, demonstrate higher expectations for achievement, and provide an environment that supports academic goals (Cruz, Barbosa-Ducharne, and Canário 2020; Del Río et al. 2017; Putnick et al. 2020; Spera 2005). This pattern is consistent with the possibility that the father's educational level is indirectly associated with the meaning ascribed to mathematics. This difference likewise corresponded to a moderate effect size ($\varepsilon^2 = 0.043$; Tomczak and Tomczak 2014). As with maternal education, metric invariance was not supported for this variable, so this comparison should likewise be interpreted with caution.

The research determined that the income level of the students' families created a significant difference in the meaning of mathematics. Students whose family income was less than their expenditures ascribed less meaning to mathematics than those whose family income was equal to or exceeded their expenditures. Furthermore, the meaning ascribed to mathematics by students whose family income was equal to their expenditures was lower than that of students whose family income exceeded their expenditures. This pattern suggests that socioeconomic status is associated not only with access to physical resources but also with how students make sense of mathematics. Davis-Kean (2005) stated that there is a positive relationship between family income level and children's academic achievement, and this relationship coincides with parental expectations and the learning environment provided at home. Similarly, Sirin's (2005) comprehensive meta-analysis revealed that students from lower socioeconomic backgrounds have lower academic achievement and motivation for learning. Children with higher-income families may perceive mathematics as a more functional, valuable, and meaningful discipline due to having greater learning opportunities, extracurricular resources, and a supportive environment. The OECD's (2019) PISA 2018 results also indicate that economic resources are associated with not only academic achievement but also how students make sense of the learned knowledge. Additionally, various studies support that the socioeconomic status of families is significantly associated with mathematics achievement (Chiu and Chow 2015; Choi et al. 2015; Østbø and Zachrisson 2022; Papanastasiou 2002). These findings suggest that the meaning students ascribe to mathematics is associated not only with individual interest and effort but also with external factors such as the family's socioeconomic status and social environment. Although statistically significant, this difference corresponded to a weak effect size ($\varepsilon^2 = 0.013$; Tomczak and Tomczak 2014), suggesting that, given the very large sample size, family income accounted for only a small proportion of the variance in students' ascribed meaning of mathematics.

Recommendations

To strengthen the validity and reliability of the scale, it is recommended to re-administer it to samples varying in age, region, and socioeconomic groups. The factors influencing the perception of the meaning of mathematics (gender, grade level, school type, field of study, parental education, and socioeconomic status) should be investigated in depth through qualitative research. To enhance students' meaning-making of mathematics, engaging, interdisciplinary learning experiences linked to daily life and careers should be offered; content should be enriched, particularly in fields where mathematics is perceived as less valuable. Awareness and guidance services should be provided to parents, and teachers should develop motivational methods targeting diverse student profiles. Policymakers should ensure equitable resource distribution to reduce the disparities between school types and socioeconomic differences. Furthermore, the impact of the examination-focused structure should be considered, and in-class academic performance should also be taken into account. In designing such teacher-facing interventions, the classification system of autonomy-, competence-, and relatedness-supportive motivational behaviours proposed by Ahmadi et al. (2023) offers a concrete, empirically grounded framework that teachers and teacher educators could draw upon to foster more meaningful engagement with mathematics.

Limitations

The first limitation of the study is that it only covers the middle school and high school levels, thus excluding an investigation into the meaning-perception of mathematics among primary school students. This creates a gap in knowledge regarding the development of mathematical meaning during the primary school period. Secondly, the exclusive use of quantitative measurement instruments constitutes another limitation. The developed scale may have restricted the in-depth and contextual understanding of the meaning students ascribe to mathematics; therefore, future studies supported by qualitative methods are recommended. Another limitation of the research is that the developed scale is restricted to the Turkish language and culture. This limits the scale's validity in different languages and cultures. Moreover, the fact that the research was conducted solely with a Turkish sample indicates that the findings and conclusions should only be interpreted within this context. Adapting and administering the scale in different cultural and linguistic settings is important for increasing generalisability. An additional limitation concerns the cross-sectional design of the study. As data were collected at a single time point across different grade levels, the observed differences between grades may reflect cohort effects rather than genuine within-student developmental change. Longitudinal designs would be needed to draw firmer conclusions about developmental trajectories in the meaning students attribute to mathematics.

Although Turkish-language instruments exist to measure closely related but conceptually distinct constructs, such as attitudes towards mathematics and engagement in mathematics lessons, no validated Turkish instrument currently exists that measures the meaning students attribute to mathematics, or a sufficiently equivalent construct, against which the MofM scale's convergent, discriminant, or criterion-related validity could be meaningfully assessed. Consequently, the validity evidence presented in this study is limited to internal structure (EFA, CFA, IRT, and MI) and does not extend to external validity criteria. Future research should examine the relationship between the MofM scale and validated measures of mathematics achievement, attitudes, self-efficacy, or motivation, either as comparable Turkish instruments become available or through cross-cultural adaptations of related international measures, in order to further substantiate the nomological network of the meaning-of-mathematics construct. A related limitation concerns MI. Because several subgroups within school type and field of study were too small ($n = 41–93$) for stable multi-group CFA estimation (Jackson 2003; MacCallum et al. 1999), invariance testing could not be extended to these variables. For parental education and family income, subgroup sizes were adequate, and metric invariance testing was conducted; however, the majority of factor loadings were found to be non-invariant across groups for all three variables. Consequently, the corresponding group comparisons reported in this study should be interpreted with caution, and future research should further investigate the sources of this non-invariance, for example through differential item functioning analyses at the item level.

Acknowledgements

All authors have contributed equally to this study.

Author contributions

CRediT: **Çiğdem Çalişkan Toraman:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing; **Osman Aktan:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Çiğdem Çalişkan Toraman has a bachelor's degree in Primary Mathematics Education. She has a master's degree and then a doctorate in Mathematics Education. Çalışkan Toraman began teaching mathematics while still a university student and has been teaching mathematics for many years. Having taught middle and high school mathematics, she currently serves as a mathematics teacher at the Science and Art Center (a type of school for gifted students in Türkiye). She has worked on numerous national and international projects and provided consulting services. Her articles include: "Metaphorical perceptions of middle school students towards math," "The impact of cognitive flexibility and classroom engagement on high school students' creative problem-solving skills," "Cognitive flexibility: A bibliometric analysis of research and insights for the future," "What do we know about underachievement in mathematics? A review based on SCOPUS and WOS," "I have a letter to my math teachers: A qualitative study with gifted students," and "Are Virtual Manipulatives a Disappointment for Middle School Mathematics Teachers?" and "Teachers' tendency to use alternative assessment methods".

Osman Aktan PhD, is is a Associate Professor in Department of Special Education in Düzce University, Faculty of Education. His research interests include teacher education, teachers' professional development, curriculum and instruction, inclusive education. He has published several book chapters, articles and conference papers on teacher education, professional development, curriculum and instruction, inclusive education, research methods in education. Among his publications are “Investigation of Primary School Mathematics Curriculum Lesson Acquisitions According to the Renewed Bloom’s Taxonomy”, “Analysis of the Relationships between Mathematics Achievement, Reflective Thinking in Problem Solving, and Metacognitive Awareness”, “Evaluation of Teachers’ Opinions About Using the Team-Assisted Individualization Technique in Mathematics Courses within Inclusive Education Practices”, and “Investigation of Primary School Students’ Attitudes towards Mathematics in Terms of Various Variables.” These articles have received a high number of citations and are regarded as significant studies in the field of mathematics education

ORCID

Çiğdem Çalişkan Toraman http://orcid.org/0000-0002-5862-628X
Osman Aktan http://orcid.org/0000-0001-6583-3765

Data availability statement

The data used in the current study are available from the corresponding author upon reasonable request.

Ethical approval

This research was conducted with the approval of the Duzce University Scientific Research and Publication Ethics Committee (Decision No: 2025/1-54, Date: 23.01.2025) and the permission of the Ministry of National Education (Permission No: MEB.TT.2025.024278, Date: 12.06.2025).

The procedures used in this study adhere to the tenets of the Declaration of Helsinki.

Informed consent

All participants, or their legal guardian, provided informed written consent prior to study enrolment.

All scientific, ethical and legal responsibilities of our article belong to the authors whose names and signatures are below.

References

Abd Algani, Y. M. 2022. “Role, Need and Benefits of Mathematics in the Development of Society.” *Journal for the Mathematics Education and Teaching Practices* 3 (1): 23–29.

Abrahamson, D., M. J. Nathan, C. Williams-Pierce, C. Walkington, E. R. Ottmar, H. Soto, and M. W. Alibali. 2020. “The Future of Embodied Design for Mathematics Teaching and Learning.” *Frontiers in Education* 5:147. https://doi.org/10.3389/feduc.2020.00147.

Acharya, B. R. 2017. “Factors Affecting Difficulties in Learning Mathematics by Mathematics Learners.” *International Journal of Elementary Education* 6 (2): 8–15. https://doi.org/10.11648/j.ijeedu.20170602.11.

Aguilar, J. J. 2021. “High School Students’ Reasons for Disliking Mathematics: The Intersection Between Teacher’s Role and Student’s Emotions, Belief and Self-Efficacy.” *International Electronic Journal of Mathematics Education* 16 (3): 1–11. https://doi.org/10.29333/iejme/11294.

Ahmadi, A., M. Noetel, P. Parker, R. M. Ryan, N. Ntoumanis, J. Reeve, M. Beauchamp, et al. 2023. “A Classification System for teachers’ Motivational Behaviors Recommended in Self-

Determination Theory Interventions." *Journal of Educational Psychology* 115 (8): 1158–1176. https://doi.org/10.1037/edu0000783.

Anderson, J. C., and D. W. Gerbing. 1984. "The Effect of Sampling Error on Convergence, Improper Solutions, and Goodness of Fit Indices for Maximum Likelihood Confirmatory Factor Analysis." *Psychometrika* 49 (2): 155–173. https://doi.org/10.1007/BF02294170.

Appiah, J. B., S. Korkor, Y. D. Arthur, and B. A. Obeng. 2022. "Mathematics Achievement in High Schools, the Role of the Teacher-Student Relationship, Students' Self-Efficacy, and Students' Perception of Mathematics." *International Electronic Journal of Mathematics Education* 17 (3): 1–11. https://doi.org/10.29333/iejme/12056.

Aybek, E. C., and Ç. Toraman. 2022. "How Many Response Categories are Sufficient for Likert Type Scales? An Empirical Study Based on the Item Response Theory." *International Journal of Assessment Tools in Education* 9 (2): 534–547. https://doi.org/10.21449/ijate.1132931.

Bakanlar Kurulu Kararı [Council of Ministers Decision]. 2002. "İstatistiki bölge birimleri sınıflandırması [Classification of istatistical territorial units]." August Tarihli ve 2002/4720 Sayılı Kararname 28. https://www.resmigazete.gov.tr/eskiler/2002/09/20020922.htm.

Bentler, P. M. 1990. "Comparative Fit Indexes in Structural Models." *Psychological Bulletin* 107 (2): 238–246. https://doi.org/10.1037/0033-2909.107.2.238.

Beswick, K. 2019. "The Role of Knowledge and Beliefs in Helping Learners to Progress Their Mathematical Understanding." *Journal of Mathematics Teacher Education* 22 (2): 125–128. https://doi.org/10.1007/s10857-019-09432-5.

Boaler, J. 2016. *Mathematical Mindsets: Unleashing students' Potential Through Creative Math, Inspiring Messages and Innovative Teaching*. San Francisco, CA: Jossey-Bass.

Brandell, G., and E. M. Staberg. 2008. "Mathematics: A Female, Male or Gender-Neutral Domain? A Study of Attitudes Among Students at Secondary Level." *Gender and Education* 20 (5): 495–509. https://doi.org/10.1080/09540250701805771.

Campbell, T. G., K. N. Vela, and T. Powell. 2025. "Interest, Enjoyment, and Confidence in Mathematics in the United States: Exploring Patterns Across Age, Gender, Race, and Time." *Large-Scale Assessments in Education* 13 (1): 1–23. https://doi.org/10.1186/s40536-025-00258-7.

Çanakçı, O., and A. Ş. Özdemir. 2015. "Matematik Başarısı ve Anne-Baba Eğitim Düzeyi." *İstanbul Aydın Üniversitesi Dergisi* 7 (25): 19–36. https://doi.org/10.17932/IAU.IAUD.m.13091352.2015.7/25.19-36.

Cansız-Aktaş, M., and D. Y. Aktaş. 2012. "Investigating high school students' attitudes towards geometry according to different variables: Ordu sample." *Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi* 18:156–167. https://izlik.org/JA76BF39WD.

Cargnelutti, E., C. Tomasetto, and M. C. Passolunghi. 2017. "The Interplay Between Affective and Cognitive Factors in Shaping Early Proficiency in Mathematics." *Trends in Neuroscience and Education* 8-9:28–36. https://doi.org/10.1016/j.tine.2017.10.002.

Çelik, H. C., and H. Ceylan. 2009. "Lise Öğrencilerinin matematik ve bilgisayar tutumlarının çeşitli değişkenler açısından karşılaştırılması." *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi* 26 (26): 92–101. https://izlik.org/JA58JP33FK .

Chalmers, R. P. 2012. "Mirt: A Multidimensional Item Response Theory Package for the R Environment." *Journal of Statistical Software* 48 (6): 1–29. https://doi.org/10.18637/jss.v048.i06.

Chand, S., K. Chaudhary, A. Prasad, and V. Chand. 2021. "Perceived Causes of students' Poor Performance in Mathematics: A Case Study at Ba and Tavua Secondary Schools." *Frontiers in Applied Mathematics and Statistics* 7:1–13. https://doi.org/10.3389/fams.2021.614408.

Chen, F. F. 2007. "Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance." *Structural Equation Modeling: A Multidisciplinary Journal* 14 (3): 464–504. https://doi.org/10.1080/10705510701301834.

Cheung, G. W., and R. B. Rensvold. 2002. "Evaluating Goodness-Of-Fit Indexes for Testing Measurement Invariance." *Structural Equation Modeling* 9 (2): 233–255. https://doi.org/10.1207/S15328007SEM0902_5.

Chiu, M. M., and B. W. Y. Chow. 2015. "Classmate Characteristics and Student Achievement in 33 Countries: Classmates' Past Achievement, Family Socioeconomic Status, Educational Resources, and Attitudes Toward Reading." *Journal of Educational Psychology* 107 (1): 152–169. https://doi.org/10.1037/a0036897.

Choi, N., M. Chang, S. Kim, and T. G. Reio Jr. 2015. "A Structural Model of Parent Involvement with Demographic and Academic Variables." *Psychology in the Schools* 52 (2): 154–167. https://doi.org/10.1002/pits.21813.

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cruz, O., M. Barbosa-Ducharne, and C. Canário. 2020. "Do adolescents' Perceptions of mothers' And fathers' Parenting Behaviors Predict Academic Achievement and Social Skills?" *Child Indicators Research* 13 (1): 237–254. https://doi.org/10.1007/s12187-019-09687-7.

Davis-Kean, P. E. 2005. "The Influence of Parent Education and Family Income on Child Achievement: The Indirect Role of Parental Expectations and the Home Environment." *Journal of Family Psychology* 19 (2): 294–304. https://doi.org/10.1037/0893-3200.19.2.294.

de Ayala, R. J. 2009. *The Theory and Practice of Item Response Theory*. New York, NY: Guilford Press.

Del Río, M. F., M. I. Susperreguy, K. Strasser, and V. Salinas. 2017. "Distinct Influences of Mothers and Fathers on kindergartners' Numeracy Performance: The Role of Math Anxiety, Home Numeracy Practices, and Numeracy Expectations." *Early Education & Development* 28 (8): 939–955. https://doi.org/10.1080/10409289.2017.1331662.

DeMars, C. 2010. *Item Response Theory*. New York: Oxford University Press.

Deniz, L., and N. Cıtdır. 2020. "Ortaokul Öğrencilerinin Matematiğe Yönelik Tutumlarının İncelenmesi." *Anadolu University Journal of Education Faculty* 4 (3): 294–322. https://doi.org/10.34056/aujef.704071.

de Santana, A. N., A. Roazzi, and A. P. M. C. Nobre. 2022. "The Relationship Between Cognitive Flexibility and Mathematical Performance in Children: A Meta-Analysis." *Trends in Neuroscience and Education* 28 (2022): 1–9. https://doi.org/10.1016/j.tine.2022.100179.

Di Martino, P., and R. Zan. 2011. "Attitude Towards Mathematics: A Bridge Between Beliefs and Emotions." *ZDM Mathematics Education* 43 (4): 471–482. https://doi.org/10.1007/s11858-011-0309-6.

Dowker, A., and H. Sheridan. 2022. "Relationships Between Mathematics Performance and Attitude to Mathematics: Influences of Gender, Test Anxiety, and Working Memory." *Frontiers in Psychology* 13:1–11. https://doi.org/10.3389/fpsyg.2022.814992.

Du, W., T. T. Wijaya, Y. Cao, and X. Li. 2025. "What Factors Influence Secondary students' Active Participation in Mathematics Classrooms: A Theory of Planned Behavior Perspective." *Acta Psychologica* 253:1–14. https://doi.org/10.1016/j.actpsy.2025.104694.

Dweck, C. S. 2008. *Mindsets and Math/Science Achievement*. New York: Carnegie Corporation. http://www.growthmindsetmaths.com/uploads/2/3/7/7/23776169/mindset_and_math_science_achievement_-_nov_2013.pdf .

Eccles, J. S. 2005. "Subjective Task Values and the Eccles et al. Model of Achievement-Related Choices." In *Handbook of Competence and Motivation*, edited by A. J. Elliot and C. S. Dweck, 105–121. Guilford.

Eccles, J. S., and A. Wigfield. 2020. "From Expectancy-Value Theory to Situated Expectancy-Value Theory: A Developmental, Social Cognitive, and Sociocultural Perspective on Motivation." *Contemporary Educational Psychology* 61 (2020): 1–13. https://doi.org/10.1016/j.cedpsych.2020.101859.

Eccles, J. S., and A. Wigfield. 2023. "Expectancy-Value Theory to Situated Expectancy-Value Theory: Reflections on the Legacy of 40+ Years of Working Together." *Motivation Science* 9 (1): 1–12. https://doi.org/10.1037/mot0000275.

Edelen, M. O., and B. B. Reeve. 2007. "Applying Item Response Theory (IRT) Modeling to Questionnaire Development, Evaluation, and Refinement." *Quality of Life Research* 16 (1): 5–18. https://doi.org/10.1007/s11136-007-9198-0.

Edu, N. B., X. J. Minquez, S. C. Dabalos, and E. Taja-On. 2025. "Untangling the 'Why' Of Math: Examining Student Perceptions on the Value of Mathematics in the Modern World." *Journal of Tertiary Education and Learning* 3 (1): 36–43. https://doi.org/10.54536/jtel.v3i1.3399.

Emanet, E. A., and F. Kezer. 2021. "The Effects of Student-Centered Teaching Methods Used in Mathematics Courses on Mathematics Achievement, Attitude, and Anxiety: A Meta-Analysis Study." *Participatory Educational Research* 8 (2): 240–259. https://doi.org/10.17275/per.21.38.8.2.

Embretson, S. E., and S. P. Reise. 2013. *Item Response Theory for Psychologists*. New York: Psychology Press.

English, L. D., and G. S. Halford. 2012. *Mathematics Education: Models and Processes*. New York: Routledge.

Fabrigar, L. R., and D. T. Wegener. 2012. *Exploratory Factor Analysis*. New York and Oxford: Oxford University Press.

Fan, X., and M. Chen. 2001. "Parental Involvement and students' Academic Achievement: A Meta-Analysis." *Educational Psychology Review* 13 (1): 1–22. https://doi.org/10.1023/A:1009048817385.

Field, A. 2018. *Discovering Statistics Using IBM SPSS Statistics*. 5th ed. London: SAGE Publications.

Firdaus, F., I. Kailani, M. N. B. Bakar, and B. Bakry. 2015. "Developing Critical Thinking Skills of Students in Mathematics Learning." *Journal of Education and Learning (EduLearn)* 9 (3): 226–236. https://doi.org/10.11591/edulearn.v9i3.1830.

Flannery, W. P., S. P. Reise, and K. F. Widaman. 1995. "An Item Response Theory Analysis of the General and Academic Scales of the Self-Description Questionnaire II." *Journal of Research in Personality* 29 (2): 168–188. https://doi.org/10.1006/jrpe.1995.1010.

Fufa, F. S., A. H. Tulu, and K. A. Ensene. 2023. "Examining the Challenges of Using Student-Centred Teaching Strategies in Secondary Schools: A Qualitative Approach." *Journal of Pedagogical Sociology and Psychology* 5 (3): 61–72. https://doi.org/10.33902/jpsp.202323181.

Furinghetti, F., J. M. Matos, and M. Menghini. 2012. "From Mathematics and Education, to Mathematics Education." In *Third International Handbook of Mathematics Education*, edited by M. A. Clements, A. J. Bishop, C. Keitel, J. Kilpatrick, and F. K. S. Leung, 273–302. New York, NY: Springer New York.

Garcia, T., C. Rodriguez, L. Betts, D. Areces, and P. Gonzalez-Castro. 2016. "How Affective-Motivational Variables and Approaches to Learning Predict Mathematics Achievement in Upper Elementary Levels." *Learning & Individual Differences* 49:25–31. https://doi.org/10.1016/j.lindif.2016.05.021.

Gasco, J., A. Goñi, and J. D. Villarroel. 2014. "Sex Differences in Mathematics Motivation in 8th and 9th Grade." *Procedia-Social and Behavioral Sciences* 116:1026–1031. https://doi.org/10.1016/j.sbspro.2014.01.340.

Grigg, S., H. N. Perera, P. McIlveen, and Z. Svetleff. 2018. "Relations Among Math Self Efficacy, Interest, Intentions, and Achievement: A Social Cognitive Perspective." *Contemporary Educational Psychology* 53:73–86. https://doi.org/10.1016/j.cedpsych.2018.01.007.

Gulburnu, M., and K. Yildirim. 2021. "Development and Implementation of Mathematics Attitudes Scale for the Primary and Secondary Students." *Pegem Journal of Education and Instruction* 11 (4): 177–184. https://doi.org/10.47750/pegegog.11.04.17.

Guo, W., K. L. Lau, and J. Wei. 2019. "Teacher Feedback and students' Self-Regulated Learning in Mathematics: A Comparison Between a High-Achieving and a Low-Achieving Secondary Schools." *Studies in Educational Evaluation* 63:48–58. https://doi.org/10.1016/j.stueduc.2019.07.001.

Gyamfi, A., and R. Acquaye. 2023. "Parameters and Models of Item Response Theory (IRT): A Review of Literature." *Acta Educationis Generalis* 13 (3): 68–78. https://doi.org/10.2478/atd-2023-0022.

Hackett, G., and N. E. Betz. 1989. "An Exploration of the Mathematics Self-Efficacy/mathematics Performance Correspondence." *Journal for Research in Mathematics Education* 20 (3): 261–273. https://doi.org/10.2307/749515.

Hair, J. F., W. C. Black, B. J. Babin, and R. E. Anderson. 2014. *Multivariate Data Analysis*. 7th ed. Harlow, Essex: Pearson Education Limited.

Hajjar, S. T. 2018. "Statistical Analysis: Internal-Consistency Reliability and Construct Validity." *International Journal of Quantitative and Qualitative Research Methods* 6 (1): 27–38.

Hambleton, R. K. 1994. "Guidelines for Adapting Educational and Psychological Test: A Progress Report." *European Journal of Psychological Assessment* 10 (3): 229–244.

Hambleton, R. K., and H. Swaminathan. 2013. *Item Response Theory: Principles and Applications*. New York: Springer Science & Business Media.
Hannula, M. S. 2002. "Attitude Towards Mathematics: Emotions, Expectations and Values." *Educational Studies in Mathematics* 49 (1): 25–46. https://doi.org/10.1023/A:1016048823497.
Harlen, W. 2013. "Inquiry-Based Learning in Science and Mathematics." *Review of Science, Mathematics and ICT Education* 7 (2): 9–33.
Hascher, T., G. Hagenauer, and S. Sacchi. 2024. "Secondary students' Emotions, Perceptions of Instructional Quality, and Achievement in Mathematics: A Representative Study with Swiss Students Towards the End of Compulsory Education." *Contemporary Educational Psychology* 78:1–16. https://doi.org/10.1016/j.cedpsych.2024.102295.
Herheim, R. 2023. "On the Origin, Characteristics, and Usefulness of Instrumental and Relational Understanding." *Educational Studies in Mathematics* 113 (3): 389–404. https://doi.org/10.1007/s10649-023-10225-0.
Hidayatullah, A., and C. Csíkos. 2024. "The Role of students' Beliefs, parents' Educational Level, and the Mediating Role of Attitude and Motivation in students' Mathematics Achievement." *The Asia-Pacific Education Researcher* 33 (2): 253–262. https://doi.org/10.1007/s40299-023-00724-2.
Hiebert, J., and D. A. Grouws. 2007. "The Effects of Classroom Mathematics Teaching on students' Learning." In *Second Handbook of Research on Mathematics Teaching and Learning*, edited by F. K. Lester, 371–404. Vol. 1. Charlotte, NC: Information Age.
Hill, H. C., and C. Y. Charalambous. 2012. "Teacher Knowledge, Curriculum Materials, and Quality of Instruction: Lessons Learned and Open Issues." *Journal of Curriculum Studies* 44 (4): 559–576. https://doi.org/10.1080/00220272.2012.716978.
Hill, H. C., B. Rowan, and D. L. Ball. 2005. "Effects of teachers' Mathematical Knowledge for Teaching on Student Achievement." *American Educational Research Journal* 42 (2): 371–406. https://doi.org/10.3102/00028312042002371.
Hooper, D., J. Coughlan, and M. R. Mullen. 2008. "Structural Equation Modelling: Guidelines for Determining Model Fit." *The Electronic Journal of Business Research Methods* 6 (1): 53–60.
Horn, J. L. 1965. "A Rationale and Test for the Number of Factors in Factor Analysis." *Psychometrika* 30 (2): 179–185. https://doi.org/10.1007/BF02289447.
Howard, L., and M. Whitaker. 2011. "Unsuccessful and Successful Mathematics Learning: Developmental students' Perceptions." *Journal of Developmental Education* 35 (2): 2–16.
Hu, L. T., and P. M. Bentler. 1999. "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives." *Structural Equation Modeling: A Multidisciplinary Journal* 6 (1): 1–55. https://doi.org/10.1080/10705519909540118.
Huang, C. 2013. "Gender Differences in Academic Self-Efficacy: A Meta-Analysis." *European Journal of Psychology of Education* 28 (1): 1–35. https://doi.org/10.1007/s10212-011-0097-y.
Hui, H. B., and M. S. Mahmud. 2023. "Influence of Game-Based Learning in Mathematics Education on the students' Cognitive and Affective Domain: A Systematic Review." *Frontiers in Psychology* 14:1–15. https://doi.org/10.3389/fpsyg.2023.1105806.
Husserl, E.1970. *Logical Investigations*. Translated by J. N. Findlay. New York: Humanities Press.
Hyde, J. S., E. Fennema, and S. J. Lamon. 1990. "Gender Differences in Mathematics Performance: A Meta-Analysis." *Psychological Bulletin* 107 (2): 139–155. https://doi.org/10.1037/0033-2909.107.2.139.
Hyde, J. S., and J. E. Mertz. 2009. "Gender, Culture, and Mathematics Performance." *Proceedings of the National Academy of Sciences* 106 (22): 8801–8807. https://doi.org/10.1073/pnas.0901265106.
Jablonka, E. 2003. "Mathematical Literacy." In *Second International Handbook of Mathematics Education*, edited by, A. Bishop, M. A. Clements, C. Keitel, J. Kilpatrick, and F. K. S. Leung, 75–102. Dordrecht: Kluwer Academic Publishers.
Jackson, D. L. 2003. "Revisiting Sample Size and the Number of Parameter Estimates: Some Support for the N: Q Hypothesis." *Structural Equation Modeling* 10 (1): 128–141. https://doi.org/10.1207/s15328007sem1001_6.
Kaiser, H. F., and J. Rice. 1974. "Little Jiffy, Mark IV." *Educational and Psychological Measurement* 34 (1): 111–117. https://doi.org/10.1177/001316447403400115.
Kline, T. J. B. 2005. *Psychological Testing: A Practical Approach to Design and Evaluation*. Thousand Oaks, CA: SAGE Publications.

Köller, O., J. Baumert, and K. Schnabel. 2001. "Does Interest Matter? The Relationship Between Academic Interest and Achievement in Mathematics." *Journal for Research in Mathematics Education* 32 (5): 448–470. https://doi.org/10.2307/749801.

Krippendorff, K. 2018. *Content Analysis: An Introduction to Its Methodology*. Thousand Oaks, CA: Sage Publications.

Lave, J. 1988. *Cognition in Practice: Mind, Mathematics and Culture in Everyday Life*. Cambridge: Cambridge University Press.

Lawshe, C. H. 1975. "A Quantitative Approach To Content Validity 1." *Personnel Psychology* 28 (4): 563–575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x.

Llinares, S. 2021. "Instructional Quality of Mathematics Teaching and Mathematics Teacher Education." *Journal of Mathematics Teacher Education* 24 (1): 1–3. https://doi.org/10.1007/s10857-021-09488-2.

Louis, R. A., and J. M. Mistele. 2012. "The Differences in Scores and Self-Efficacy by Student Gender in Mathematics and Science." *International Journal of Science & Mathematics Education* 10 (5): 1163–1190. https://doi.org/10.1007/s10763-011-9325-9.

MacCallum, R. C., K. F. Widaman, S. Zhang, and S. Hong. 1999. "Sample Size in Factor Analysis." *Psychological Methods* 4 (1): 84–99. https://doi.org/10.1037/1082-989X.4.1.84.

Manaf, B., A. Kartowagiran, and A. Manaf. 2021. "Student Attitude and Mathematics Learning Success: A Meta-Analysis." *International Journal of Instruction* 14 (4): 209–222. https://doi.org/10.29333/iji.2021.14413a.

Marks, G. N. 2008. "Are father's or mother's Socioeconomic Characteristics More Important Influences on Student Performance? Recent International Evidence." *Social Indicators Research* 85 (2): 293–309. https://doi.org/10.1007/s11205-007-9132-4.

Marsh, H. W., J. R. Balla, and R. P. McDonald. 1988. "Goodness-Of-Fit Indexes in Confirmatory Factor Analysis: The Effect of Sample Size." *Psychological Bulletin* 103 (3): 391–410. https://doi.org/10.1037/0033-2909.103.3.391.

Martin, D. B. 2000. *Mathematics Success and Failure Among African-American Youth: The Roles of Sociohistorical Context, Community Forces, School Influence, and Individual Agency*. Mahwah, NJ: Lawrence Erlbaum Associates.

Meredith, W. 1993. "Measurement Invariance, Factor Analysis and Factorial Invariance." *Psychometrika* 58 (4): 525–543. https://doi.org/10.1007/BF02294825.

Mhlolo, M. K., M. Schafer, and H. Venkat. 2012. "The Nature and Quality of the Mathematical Connections Teachers Make." *Pythagoras* 33 (1): 1–9. https://doi.org/10.4102/pythagoras.v33i1.22.

Millsap, R. E. 2011. *Statistical Approaches to Measurement Invariance*. New York: Routledge.

Moenikia, M., and A. Zahed-Babelan. 2010. "A Study of Simple and Multiple Relations Between Mathematics Attitude, Academic Motivation and Intelligence Quotient with Mathematics Achievement." *Procedia-Social and Behavioral Sciences* 2 (2): 1537–1542. https://doi.org/10.1016/j.sbspro.2010.03.231.

Muraki, E. 1992. "A Generalized Partial Credit Model: Application of an EM Algorithm." *ETS Research Report Series* 1992 (1): i–30. https://doi.org/10.1002/j.2333-8504.1992.tb01436.x.

Mutodi, P., and H. Ngirande. 2014. "The Influence of Studentsperceptions on Mathematics Performance. A Case of a Selected High School in South Africa." *Mediterranean Journal of Social Sciences* 5 (3): 431–445. https://doi.org/10.5901/mjss.2014.v5n3p431.

Ni, Y., D. H. R. Zhou, J. Cai, X. Li, Q. Li, and I. X. Sun. 2018. "Improving Cognitive and Affective Learning Outcomes of Students Through Mathematics Instructional Tasks of High Cognitive Demand." *Journal of Educational Research* 111 (6): 704–719. https://doi.org/10.1080/00220671.2017.1402748.

Niemivirta, M., A. Tapola, H. Tuominen, and J. Viljaranta. 2024. "Developmental Trajectories of School-beginners' Ability Self-Concept, Intrinsic Value and Performance in Mathematics." *British Journal of Educational Psychology* 94 (2): 441–459. https://doi.org/10.1111/bjep.12655.

Nunnally, J. C., and I. H. Bernstein. 1994. *Psychometric Theory*. 3rd ed. New York: McGraw-Hill.

OECD. 2019. *PISA 2018 Results (Volume I): What Students Know and Can Do*. Paris: PISA, OECD Publishing. https://doi.org/10.1787/5f07c754-en.

Orlando, M., and D. Thissen. 2000. "Likelihood-Based Item-Fit Indices for Dichotomous Item Response Theory Models." *Applied Psychological Measurement* 24 (1): 50–64. https://doi.org/10.1177/01466216000241003.

Østbø, I. U., and H. D. Zachrisson. 2022. "Student Motivation and Parental Attitude as Mediators for SES Effects on Mathematics Achievement: Evidence from Norway in TIMSS 2015." *Scandinavian Journal of Educational Research* 66 (5): 808–823. https://doi.org/10.1080/00313831.2021.1939138.

Özcan, B., and Y. Z. Kültür. 2021. "The Relationship Between Sources of Mathematics Self-Efficacy and Mathematics Test and Course Achievement in High School Seniors." *Sage Open* 11 (3). https://doi.org/10.1177/21582440211040124.

Pajares, F. 2005. "Gender Differences in Mathematics Self-Efficacy Beliefs." In *Gender Differences in Mathematics: An Integrative Psychological Approach*, edited by A. M. Gallagher and J. C. Kaufman, 294–315. New York: Cambridge University Press.

Papanastasiou, C. 2002. "Effects of Background and School Factors on the Mathematics Achievement." *Educational Research and Evaluation* 8 (1): 55–70. https://doi.org/10.1076/edre.8.1.55.6916.

Perez Mejias, P., D. E. McAllister, K. G. Diaz, and J. Ravest. 2021. "A Longitudinal Study of the Gender Gap in Mathematics Achievement: Evidence from Chile." *Educational Studies in Mathematics* 107 (3): 583–605. https://doi.org/10.1007/s10649-021-10052-1.

Pett, M. A., N. R. Lackey, and J. J. Sullivan. 2003. *Making Sense of Factor Analysis: The Use of Factor Analysis for Instrument Development in Health Care Research*. Thousand Oaks, CA: SAGE Publications. https://doi.org/10.4135/9781412984898.

Pitsia, V., A. Biggart, and A. Karakolidis. 2017. "The Role of students' Self-Beliefs, Motivation and Attitudes in Predicting Mathematics Achievement: A Multilevel Analysis of the Programme for International Student Assessment Data." *Learning & Individual Differences* 55:163–173. https://doi.org/10.1016/j.lindif.2017.03.014.

Putnick, D. L., and M. H. Bornstein. 2016. "Measurement Invariance Conventions and Reporting: The State of the Art and Future Directions for Psychological Research." *Developmental Review* 41:71–90. https://doi.org/10.1016/j.dr.2016.06.004.

Putnick, D. L., C. S. Hahn, C. Hendricks, J. T. Suwalsky, and M. H. Bornstein. 2020. "Child, Mother, Father, and Teacher Beliefs About Child Academic Competence: Predicting Math and Reading Performance in European American Adolescents." *Journal of Research on Adolescence* 30 (2): 298–314. https://doi.org/10.1111/jora.12477.

Reilly, D., D. L. Neumann, and G. Andrews. 2019. "Investigating Gender Differences in Mathematics and Science: Results from the 2011 Trends in Mathematics and Science Survey." *Research in Science Education* 49 (1): 25–50. https://doi.org/10.1007/s11165-017-9630-6.

Samo, D. D., B. Kartasasmita, and B. Kartasasmita. 2017. "Developing Contextual Mathematical Thinking Learning Model to Enhance Higher-Order Thinking Ability for Middle School Students." *International Education Studies* 10 (12): 17–29. https://doi.org/10.5539/ies.v10n12p17.

Saxe, G. B. 2015. *Culture and Cognitive Development: Studies in Mathematical Understanding*. New York: Psychology Press.

Schmitz, E. A., B. R. Jansen, R. W. Wiers, and E. Salemink. 2023. "Math-Failure Associations, Attentional Biases, and Avoidance Bias: The Relationship with Math Anxiety and Behaviour in Adolescents." *Cognitive Therapy and Research* 47 (5): 788–801. https://doi.org/10.1007/s10608-023-10390-9.

Schoenfeld, A. H. 1989. "Explorations of students' Mathematical Beliefs and Behavior." *Journal for Research in Mathematics Education* 20 (4): 338–355. https://doi.org/10.2307/749440.

Schoenfeld, A. H., and A. H. Sloane. 2016. *Mathematical Thinking and Problem Solving*. New York: Routledge.

Schutz, A. 1967. *The Phenomenology of the Social World*. Translated by G. Walsh & F. Lehnert. Evanston, IL: Northwestern University Press.

Sirin, S. R. 2005. "Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research." *Review of Educational Research* 75 (3): 417–453. https://doi.org/10.3102/00346543075003417.

Skaalvik, S., and E. M. Skaalvik. 2004. "Gender Differences in Math and Verbal Self-Concept, Performance Expectations, and Motivation." *Sex Roles* 50 (3): 241–252. https://doi.org/10.1023/B:SERS.0000015555.40976.e6.

Skemp, R. R. 1976. "Relational Understanding and Instrumental Understanding." *Mathematics Teaching* 77 (1): 20–26.

Smith, J., N. Fotou, and R. Sharpe. 2025. "Changes in Mathematics Anxiety and Mathematics Confidence." *International Journal of Mathematical Education in Science and Technology*: 1–19. https://doi.org/10.1080/0020739X.2025.2475928.

Sosa-Gutierrez, F., H. M. V. Apaza, S. V. Valdivia-Yábar, and W. W. Condori-Castillo. 2024. "Critical Thinking and Teaching Mathematics: An Analysis from Education." *International Journal of Religion* 5 (9): 958–976. https://doi.org/10.61707/94v23344.

Spera, C. 2005. "A Review of the Relationship Among Parenting Practices, Parenting Styles, and Adolescent School Achievement." *Educational Psychology Review* 17 (2): 125–146. https://doi.org/10.1007/s10648-005-3950-1.

Sriraman, B., and L. English, Eds. 2010. *Theories of Mathematics Education: Seeking New Frontiers*, 309–331. New York: Springer.

Stout, W. F. 1990. "A New Item Response Theory Modeling Approach with Applications to Unidimensionality Assessment and Ability Estimation." *Psychometrika* 55 (2): 293–325. https://doi.org/10.1007/BF02295289.

Susperreguy, M. I., S. D. L. Burr, H. Douglas, C. Xu, J. A. LeFevre, M. F. Del Río, and V. Salinas. 2022. "Home Mathematics Environment and Math Performance of Chilean Students in Kindergarten and Grades 1 to 3." *Early Childhood Research Quarterly* 59:84–95. https://doi.org/10.1016/j.ecresq.2021.11.004.

Szucs, D., and E. Toffalini. 2023. "Maths Anxiety and Subjective Perception of Control, Value and Success Expectancy in Mathematics." *Royal Society Open Science* 10 (11): 1–17. https://doi.org/10.1098/rsos.231000.

Tabachnick, B. G., and L. S. Fidell. 2013. *Using Multivariate Statistics*. 6th ed. Boston, MA: Pearson Education.

Tall, D., ed. 1991. *Advanced Mathematical Thinking*. New York: Kluwer Academic.

Tate, W. F., and C. Rousseau. 2002. "Access and Opportunity: The Political and Social Context of Mathematics Education." In *H*andbook of International Research in Mathematics Education, edited by L. English, 271–299. New York: Routledge.

Tomczak, M., and E. Tomczak. 2014. "The Need to Report Effect Size Estimates Revisited. An Overview of Some Recommended Measures of Effect Size." *Trends in Sport Sciences* 1 (21): 19–25.

Tomul, E., E. Önder, and E. Taslidere. 2021. "The Relative Effect of Student, Family and School-Related Factors on Math Achievement by Location of the School." *Large-Scale Assessments in Education* 9 (1): 1–29. https://doi.org/10.1186/s40536-021-00117-1.

Usher, E. L., and F. Pajares. 2009. "Sources of Self-Efficacy in Mathematics: A Validation Study." *Contemporary Educational Psychology* 34 (1): 89–101. https://doi.org/10.1016/j.cedpsych.2008.09.002.

Van den Heuvel-Panhuizen, M., and P. Drijvers. 2020. "Realistic Mathematics Education." In *Encyclopedia of Mathematics Education*, edited by S. Lerman, 713–717. Cham: Springer International Publishing.

Veldhuis, M., and M. Van den Heuvel-Panhuizen. 2020. "Supporting Primary School teachers' Classroom Assessment in Mathematics Education: Effects on Student Achievement." *Mathematics Education Research Journal* 32 (3): 449–471. https://doi.org/10.1007/s13394-019-00270-5.

Vieira, A. L. 2011. *Interactive LISREL in Practice: Getting Started with a SIMPLIS Approach*. Heidelberg: Springer. https://doi.org/10.1007/978-3-642-18044-6.

Warm, T. A. 1989. "Weighted Likelihood Estimation of Ability in Item Response Theory." *Psychometrika* 54 (3): 427–450. https://doi.org/10.1007/BF02294627.

Warner, R. M. 2013. *Applied Statistics: From Bivariate Through Multivariate Techniques*. 2nd ed. Thousand Oaks, CA: SAGE Publications.

Widana, I. W., I. M. Y. Parwata, N. N. Parmithi, I. G. A. T. Jayantika, K. Sukendra, and I. W. Sumandya. 2018. "Higher Order Thinking Skills Assessment Towards Critical Thinking on Mathematics Lesson." *International Journal of Social Sciences and Humanities* 2 (1): 24–32. https://doi.org/10.29332/ijssh.v2n1.74.

Williams, B., A. Onsman, and T. Brown. 2010. "Exploratory Factor Analysis: A Five-Step Guide for Novices." *Australasian Journal of Paramedicine* 8:1–13. https://doi.org/10.33151/ajp.8.3.93.

World Bank. 2005. "Turkey-Education Sector Study: Sustainable Pathways to an Effective, Equitable and Efficient Education System for Preschool Through Secondary School education (Report No. 32450)." Washington, DC: World Bank.

Xiao, F., and L. Sun. 2021. "Students' Motivation and Affection Profiles and Their Relation to Mathematics Achievement, Persistence, and Behaviors." *Frontiers in Psychology* 11:1–15. https://doi.org/10.3389/fpsyg.2020.533593.

Yadav, S. 2019. "Role of Mathematics in the Development of Society." *International Journal of Research and Analytical Reviews* 6 (4): 295–298.

Yasar, M. 2016. "High School students' Attitudes Towards Mathematics." *Eurasia Journal of Mathematics, Science and Technology Education* 12 (4): 931–945. https://doi.org/10.12973/eurasia.2016.1571a.

Yeager, D. S., and C. S. Dweck. 2020. "What Can Be Learned from Growth Mindset Controversies?" *The American Psychologist* 75 (9): 1269–1284. https://doi.org/10.1037/amp0000794.

Yen, W. M. 1993. "Scaling Performance Assessments: Strategies for Managing Local Item Dependence." *Journal of Educational Measurement* 30 (3): 187–213. https://doi.org/10.1111/j.1745-3984.1993.tb00423.x.

Zakaria, E., and M. Syamaun. 2017. "The Effect of Realistic Mathematics Education Approach on students' Achievement and Attitudes Towards Mathematics." *Mathematics Education Trends and Research* 1 (1): 32–40. https://doi.org/10.5899/2017/metr-00093.

Zakariya, Y. F., S. Goodchild, K. Bjørkestøl, and H. K. Nilsen. 2019. "Calculus Self-Efficacy Inventory: Its Development and Relationship with Approaches to Learning." *Education Sciences* 9 (3): 1–14. https://doi.org/10.3390/educsci9030170.

Zhao, Y. 2008. "Approaches for Addressing the Fit of Item Response Theory Models to Educational Test Data." [Doctoral dissertation, University of Massachusetts Amherst], ProQuest Dissertations and Theses Global.