MODESTUM EURASIA Journal of Mathematics, Science and Technology Education, 2025, 21(12), em2748
ISSN:1305-8223 (online)
OPEN ACCESS Research Paper https://doi.org/10.29333/ejmste/17494

Analysis of mathematical literacy and assessment literacy of mathematics preservice teachers

Zhengisbek Ye. Nakypbek [1], Larissa S. Kainbaeva [1*], Gulzat M. Yensebayeva [1], Zinegul A. Yergalauova [1], Saulesh K. Menlikhozhayeva [1], Sherali S. Ibraev [1]

[1] Korkyt Ata Kyzylorda University, KAZAKHSTAN

Received 08 Jun 2025 ▪ Accepted 10 Nov 2025

Abstract

This study aims to examine the levels of mathematical literacy and assessment literacy among prospective mathematics teachers in Kazakhstan, reveal the relationship between these two constructs, and identify profile differences among prospective teachers. A quantitative research design was used in the study, and data were collected from 335 teacher candidates (196 female, 135 male) from seven universities via an online survey. A scale developed based on a literature review and the PISA framework was used for mathematical literacy, while the Assessment Literacy Standards Version 6.0 scale was adapted into Kazakh for assessment literacy. A Bayesian statistical approach was adopted for data analysis, Kendall tau-b correlations were calculated, and latent profile analysis (LPA) was applied. The findings showed that teacher candidates had high perceived competence levels in all seven dimensions of both mathematical literacy and assessment literacy. Overall equality was observed in terms of the gender variable. Moderately strong positive relationships were found between all dimensions of mathematical literacy and assessment literacy. Math anxiety, however, did not show any meaningful relationship with any construct. Latent profile analysis identified three distinct groups: Assessment Experts, Assessment Practitioners, and Assessment Novices. While there were significant differences between the groups in terms of mathematical literacy, math anxiety remained similar. The study offers empirical evidence on the correlation between mathematical literacy and assessment literacy, while also uncovering the diverse composition among teacher candidates. The results show that teacher education programs need to include these two domains of literacy in a way that is connected and provide multiple ways to educate based on individual profiles.

Keywords: mathematical literacy, assessment literacy, teacher candidates, latent profile analysis, Bayesian analysis

INTRODUCTION

One of the fundamental objectives of teacher education programs is to train pre-service teacher to become pedagogically competent professionals. Both knowing the subject matter and being able to evaluate it are very important skills in these domains (Shulman, 1986). For pre-service mathematics teachers, mathematical literacy and assessment literacy constitute two essential foundations of teacher professionalism. Mathematical literacy refers to the ability of humans to construct, utilize, and comprehend mathematical concepts within practical situations (Bolstad, 2020; Haara et al., 2017; Kappassova et al., 2025; Kozakli Ulger et al., 2022; Muhaimin et al., 2024; OECD, 2013). Assessment literacy includes the knowledge and abilities that teachers need to run classroom tests well (DeLuca et al., 2016; Pastore, 2023).

Recent studies have revealed significant deficiencies in pre-service teachers' proficiency in these two areas. Studies conducted in Turkey show that middle school mathematics teacher candidates have insufficient levels of assessment literacy, with correct response rates remaining around 53% (Şahin, 2023). Similarly, Özdil

✉ zhenisbek.nakypbek99@gmail.com ✉ larissa_kain@mail.ru **(*Correspondence)** ✉ gulzat-y83@list.ru ✉ zina_73er@mail.ru ✉ saulesh_menli@mail.ru ✉ ibrayevsheraly@gmail.com

Contribution to the literature

- Empirical evidence on the relationship between mathematical and assessment literacy: This study provides the first quantitative evidence showing that mathematical literacy and assessment literacy are related but distinct constructs, advancing the theoretical understanding of teacher knowledge domains.
- Using latent profile analysis, the study revealed three distinct teacher candidate profiles, demonstrating that teacher development is not linear and highlighting the need for differentiated training strategies based on individual profiles.
- The findings showed that math anxiety does not significantly correlate with other constructs in this context, suggesting that emotional and cognitive domains may function independently in teacher education, challenging deficit-based assumptions.

and Kaçar (2023) found that pre-service mathematics teachers did not have sufficient knowledge of basic concepts such as alternative assessment approaches, validity, and reliability. The situation is no different in terms of mathematical literacy. In their 2020 study in Indonesia, Yustitia et al. (2020) found that elementary school pre-service mathematics teachers could recognize arithmetic knowledge in PISA-focused questions, but they had trouble using such knowledge in real-life circumstances and explaining their math reasoning. Canbazoğlu and Tarım (2020) conducted a study with 66 pre-service teachers in Turkey and discovered that merely 36% of participants were capable of formulating PISA-aligned mathematical literacy questions and accurately identifying mathematical processes.

Within the theoretical framework, the relationship between mathematical literacy and assessment literacy is addressed in the context of pedagogical content knowledge. Ayalon and Wilkie (2020) demonstrated in their study that disciplinary knowledge plays a critical role in establishing assessment criteria for mathematics assignments. This finding suggests that mathematical literacy is a fundamental prerequisite for assessment literacy. Shmigirilova et al. (2024) emphasized the need for an integrative approach in the development of assessment literacy among mathematics teachers, revealing that this process is intertwined with mathematical content knowledge. On the other hand, the reciprocal effect of assessment literacy on mathematical literacy should not be overlooked. Teachers can develop their own mathematical skills while going through assessment processes; for example, when teachers analyze student work and try to understand the reasons for mathematical errors, they can gain a deeper understanding of mathematical topics (Chamberlin et al., 2008).

Current literature contains limited studies that quantitatively examine the relationship between these two literacy areas and, in particular, reveal the profiles of teacher candidates in these areas. Most studies concentrate exclusively on assessment literacy (Özdil & Kaçar, 2023; Şahin, 2023) or exclusively on mathematical literacy (Canbazoğlu & Tarım, 2020; Yustitia et al., 2020). Research that concurrently examine both constructs typically operate at a conceptual level (Shmigirilova et al., 2024); nevertheless, quantitative research experimentally assessing the strength and direction of this link is lacking. Moreover, the connection between affective elements, such as math fear, and these two literacy domains is still ambiguous. While Dirgantoro et al. (2025) indicate that anxiety is an important negative predictor of mathematical literacy, there is insufficient information about the role of this variable in assessment literacy and the interaction between the two literacy domains.

Another important gap is the insufficient examination of heterogeneity among pre-service teachers. Traditional variable-centered approaches assume that all pre-service teachers are similar and differ only along a continuum. However, person-centered analyses can reveal the existence of subgroups with different profiles and help identify the differentiated educational needs of these groups. Hull and Vígh (2025) emphasize that assessment literacy is a multidimensional construct and that it is necessary to adapt to constantly changing technological and pedagogical innovations in this field. Therefore, understanding how teacher candidates cluster within this multidimensional construct and how these clusters relate to mathematical literacy is critical for the design of teacher education programs.

This research aims to examine the levels of mathematical literacy and assessment literacy among prospective mathematics teachers in Kazakhstan and to reveal the relationship between these two structures using a Bayesian statistical approach, in order to address the aforementioned gaps in the literature. Furthermore, it aims to discover the heterogeneous structure among teacher candidates using latent profile analysis and to determine how these profiles differ in terms of mathematical literacy and math anxiety. In this context, the research seeks to answer the following questions:

RQ1 What are the levels of mathematical literacy and assessment literacy among prospective mathematics teachers?

RQ2 Does the gender variable create a significant difference in the levels of mathematical literacy

and assessment literacy among prospective mathematics teachers?

RQ3 What is the relationship between the seven dimensions of mathematical literacy, mathematics anxiety, and assessment literacy?

RQ4 How many different groups are formed according to the assessment literacy profiles of prospective mathematics teachers, and how can these groups be characterized?

RQ5 Are there meaningful differences in mathematical literacy and math anxiety among the groups formed according to assessment literacy profiles?

The findings of this research will provide empirical evidence for the necessity of teacher education programs to address both mathematical literacy and assessment literacy in an integrated manner and will contribute to the development of differentiated teaching strategies for teacher candidates with different competency profiles.

LITERATURE REVIEW

Assessment Literacy

Assessment literacy is defined as the knowledge and skills that teachers need to do assessments well (Pastore, 2023). This concept came about because teachers needed to make tests that helped students learn, and it has been an important part of educational policy and research. In contemporary educational frameworks, assessment literacy is acknowledged as essential to teacher professionalism and pedagogical effectiveness (Darling-Hammond, 2016). Brookhart (2024) rearticulates the requisite educational assessment competencies for teachers as assessment literacy, highlighting the need for adaptation to contemporary classroom dynamics. This notion includes the skills necessary for educators to thrive in technology-enhanced assessment environments, particularly with the increasing prevalence of digital assessment methods (Estaji et al., 2024).

In the realm of the application of the concept of assessment literacy, empirical studies scrutinizing the assessment practices of educators reveal pronounced discrepancies between theoretical understanding in this domain and the actual methodologies employed within classroom settings. Investigations carried out across diverse cultural and educational contexts regarding the enhancement of teachers' assessment literacy substantiate this phenomenon. A research endeavor conducted in Chile demonstrated that instructors of Language and Literature exhibited a moderate level of assessment literacy, while simultaneously lacking comprehensive knowledge regarding authentic assessment methodologies (Alegría Soto et al., 2024). These findings suggest that assessment literacy constitutes a cross-cultural phenomenon, with similar challenges being observed across various educational systems.

Recent research indicates that assessment literacy positively influences teacher performance and student academic achievement. A systematic review conducted by Hull and Vígh (2025) demonstrates that teachers' assessment literacy is a multifaceted construct that must evolve in response to ongoing technical and pedagogical advancements. For the theoretical and practical aspects of this field to advance, contemporary assessment literacy models must offer a comprehensive framework that integrates the requisite knowledge, conceptual understanding, and skills for educators to critically evaluate their assessment methodologies and implement strategies that enhance student learning.

The assessment literacy level of pre-service mathematics teachers

Research indicates that the level of assessment literacy among prospective mathematics educators is predominantly deficient. Şahin (2023) examined the extent to which middle school mathematics teachers and teacher candidates in Turkey understood complex issues related to student assessment. The results of this study showed that 53.1% of teacher candidates gave correct answers, while the correct answer rate among in-service teachers was 52.4%. Yamtim and Wongwanich (2014) argue that performance metrics below 60% signify a deficiency in competence that warrants substantial improvement. These results suggest that both pre-sevice teachers and presently in-service teachers demonstrate shortcomings in their measurement and evaluation capabilities. Özdil and Kaçar (2023) ascertained that mathematics teacher candidates demonstrate limited assessment literacy, particularly in relation to alternative assessment techniques. Moreover, the research revealed that teacher candidates lack comprehension of fundamental concepts such as validity and reliability. These results imply that the two-hour theoretical instruction provided during undergraduate training is insufficient in equipping teacher candidates with the necessary knowledge concerning evaluation methodologies and strategies.

Ayalon and Wilkie (2020) investigated the methodology for formulating assessment criteria for a complex mathematical challenge among secondary mathematics teacher candidates. The research indicated that collaborative examination of authentic student replies can illuminate specific assessment criteria for teacher candidates, although fails to render all criteria obvious. This conclusion underscores the significance of teacher candidates' familiarity with assessment techniques through methodology.

The role and development needs of teacher training programs

Research indicates that one of the main reasons for teacher candidates' low levels of assessment literacy is inadequate teacher training programs (DeLuca & Klinger, 2010). Şahin (2023) reported that participants found the assessment training in their undergraduate education to be predominantly theoretical, disconnected from teachers' daily assessment practices, and inconsistent with contemporary educational assessment standards. DeLuca et al. (2019) stated that teachers continue to struggle with assessment practices, particularly when they need to transfer new approaches and theories to the real classroom context. This situation highlights the need for teacher education programs to not only provide theoretical knowledge but also offer opportunities for practice.

Özdil and Kaçar (2023) found that the teaching experiment application had a positive effect on teacher candidates' assessment literacy levels. This finding is consistent with other studies showing that active, collaborative, and classroom-based learning models are perceived as more sustainable by teachers (Neuman & Cunningham, 2009). Research shows that teacher candidates have development needs in various areas. Şahin (2023) found that participants were better at choosing assessment tools, but they were not as good at making assessment tools and coming up with grading methods. The study also revealed that pre-service teachers were inadequate in explaining test results to students, colleagues, and parents. Ayalon and Wilkie (2020) showed that pre-service teachers struggle to develop effective methods for grading open-ended math assignments. This study emphasizes that pre-service teachers need more support in evaluating students' responses and providing constructive feedback.

A systematic review (Pastore, 2023) analyzed ten years of studies in the field of teacher assessment literacy. According to the study's findings, definitions of assessment literacy have changed over time, and socio-cultural perspectives influence how people think about this field. Forty-two articles were reviewed in the study, and teacher candidates emerged as the most frequently studied group. The review additionally indicated that quantitative studies typically employed standardized inventories and scales, whereas qualitative studies utilized smaller samples. The Classroom Assessment Practices Inventory (ACAI), created by DeLuca et al. (2016), is a tool used to compare schools in different nations.

Mathematical Literacy

Mathematical literacy is commonly framed as individuals' capacity to use mathematical ideas, procedures, and reasoning to make sense of and act in real-world contexts, not merely to reproduce school-taught routines. Research acknowledges that this goes beyond basic skills and also includes elements such as the formulation, modeling, interpretation, and critique of problems in a social context (Geiger et al., 2015). This perspective is corroborated by other studies. For example, classroom studies show that even though education emphasizes academic mathematics over practical applications, students acquire mathematical skills more effectively when they work on assignments related to their personal or professional contexts (Bolstad, 2021). These views suggest that mathematical literacy should be taught and assessed as contextual reasoning that integrates concepts with modeling, communication, and reflective judgment, rather than as isolated calculations (Bolstad, 2021; Geiger et al., 2015). Research has identified the characteristics that enhance mathematical literacy and the methods for its evaluation. National-level and programmatic research demonstrates that contextual problem-solving can enhance literacy across many grades when activities directly connect mathematical concepts to real-life scenarios (Kolar & Hodnik, 2021). At the same time, performance varies with learner and school factors; recent analyses using large-scale data and machine learning highlight how resilience-related variables predict literacy outcomes, underscoring the need for targeted support alongside pedagogy (Zhang & Cutumisu, 2024). Classroom interventions that enhance communication and modeling skills demonstrate simultaneous improvements in literacy performance, supporting the notion that instruction should integrate conceptual understanding, representation, and discourse (Ata Baran & Kabael, 2023; Kolar & Hodnik, 2021; Zhang & Cutumisu, 2024).

Mathematical literacy among pre-service teachers

Studies have shown that pre-service teacher lack a lot of the math skills they need. Yustitia et al. (2020) observed that primary school mathematics teacher candidates were able to recognize mathematical knowledge while addressing PISA-oriented questions, although they had difficulties in applying this information to real-world situations and articulating mathematical reasoning. In research of 66 teacher candidates, Canbazoğlu and Tarım (2020) found that only 36% of the participants were able to create PISA-aligned mathematical literacy questions and correctly identify mathematical processes. The most concerning finding is that 18% of the study participants were unable to answer mathematics literacy questions or accurately identify mathematical processes, despite receiving focused training.

Yani et al. (2025) reported that in their comprehensive study with 303 prospective mathematics teachers from nine universities in Indonesia, they found no statistically significant relationship between mathematical literacy and digital literacy. This finding

shows that different literacy areas are independent structures and each requires separate pedagogical attention. Furthermore, it was observed that the gender variable did not create a significant difference in either mathematical literacy or digital literacy. Jacinto and Jakobsen (2020), in their qualitative study of six primary school teacher candidates in Malawi, found that participants developed only an intuitive understanding of the knowledge required for teaching mathematics in theoretical courses. Although the teacher candidates were strong in general content knowledge, they encountered significant difficulties in the areas of horizon knowledge and specialized content knowledge, which are critical for teaching mathematical literacy.

Factors affecting mathematical literacy

Dirgantoro et al. (2025), in their thorough analysis of 13 empirical studies, identified both intrinsic and extrinsic factors that affect the development of mathematical literacy. Intrinsic factors consist of conceptual understanding, self-efficacy, and mathematics-related anxiety, while extrinsic factors encompass instructional methodologies and sociocultural contexts. The study showed that understanding concepts and believing in one's own ability to do arithmetic are both linked to being good at math, while being anxious about math was a strong negative predictor.

Karadağ Yılmaz et al. (2024), in a survey of 562 teacher candidates from seven Turkish universities, identified a modest yet favorable correlation between reading culture and mathematics literacy self-efficacy. Teacher candidates who studied scientific literature exhibited markedly greater mathematical literacy self-efficacy compared to those who engaged with literary texts. This finding indicates that discipline-specific reading techniques might enhance mathematical skill. Frambaugh-Kritzer and Buelow (2022) conducted an ethnographic study tracking 14 elementary teacher candidates in the United States during their two-year preparatory program. They demonstrated that when these candidates observed discipline-specific literacy practices in their methodology courses, they successfully transferred these practices to their internship sites. The notion of productive struggle has developed into a distinct pedagogical approach within mathematics education.

Shortcomings in teacher education programs

Research consistently indicates that the majority of teacher education programs inadequately prioritize the enhancement of mathematics literacy. Dirgantoro et al. (2025) synthesis underscored a significant disparity between PISA frameworks and teacher preparation methodologies. Christiansen and Erixon (2021), in an extensive study of 84 teacher candidates graduating from 13 universities in Sweden, identified substantial deficiencies in students' learning chances for assessing mathematical reasoning, facilitating mathematical discourse, and enhancing mathematical literacy.

A survey of senior mathematics teachers revealed that the majority of questions formulated by six secondary mathematics educators, who had undergone postgraduate mathematics literacy training, were categorized within the 'change and relationships' mathematical subject domain. The majority of inquiries were framed within the "personal" context, whilst the fewest were situated in the "scientific" and "social" settings. These findings indicate that educators continue to struggle with formulating balanced questions across all domains of the PISA framework, despite receiving specialized graduate training.

The Relationship between Mathematical Literacy and Assessment Literacy

Mathematical literacy and assessment literacy are two important areas of professional development for future math instructors that go well together. Mathematical literacy pertains to a person's proficiency in mathematical reasoning and their ability to create, apply, and interpret mathematical concepts across diverse real-world situations (OECD, 2013). Conversely, assessment literacy encompasses the essential knowledge and skills required by educators to administer assessments effectively within the classroom (DeLuca et al., 2016). The connection between these two ideas is quite significant, especially when it comes to teachers' ability to comprehend how students think about arithmetic and help them learn. Shmigirilova et al. (2024) underscore the necessity for an integrative approach in enhancing mathematics instructors' assessment literacy, indicating that this process is interconnected with mathematical content understanding. The theoretical foundation for the connection between the two notions can be examined through the lens of pedagogical subject knowledge. Ayalon and Wilkie (2020) examined the assessment literacy development processes of secondary mathematics teacher candidates, demonstrating the importance of disciplinary knowledge in establishing assessment criteria for mathematical assignments. This conclusion suggests that mathematical literacy is crucial for assessment literacy; educators must possess strong mathematical literacy to effectively analyze students' mathematical reasoning (Veldhuis & van den Heuvel-Panhuizen, 2014). In the same way, teachers need to have a deep understanding of arithmetic topics and be aware of the ways that children absorb these ideas in order to use formative assessment processes efficiently (Schildkamp et al., 2020).

The reciprocal effect of assessment literacy on mathematical literacy should also not be overlooked. Teachers might improve their own arithmetic skills by

going through the process of giving tests. For instance, when teachers look at student work and try to figure out why they made mistakes in math, they can learn more about math topics (Chamberlin et al., 2008). This is especially significant for teacher education since working with real student replies helps teacher candidates improve both their arithmetic skills and their skills in giving tests (Ayalon & Wilkie, 2020). Consequently, a cyclical and reciprocal relationship exists between these two literacy domains; robust mathematical literacy enhances effective assessment practices, while intentional assessment experiences simultaneously augment teachers' mathematical literacy.

As a result, mathematics teacher education programs must incorporate both literacy areas in a cohesive manner. Shmigirilova et al. (2024) propose that this integrative approach should encompass specialized assessment courses, assessment activities within mathematical and pedagogical domains, assessment experiences in educational environments, and possibilities for engagement with seasoned educators. This method will help teacher candidates learn how math and assessment literacy are related, which will help them become good experts in both areas. In particular, the integration of these two literacy areas enables the development of teachers' skills in analyzing student thinking mathematically, selecting appropriate assessment tools, and translating assessment results into instructional decisions.

METHOD

This study employed a quantitative research design to examine the mathematical literacy and assessment literacy levels of preservice mathematics teachers, as well as the relationship between these two constructs. Self-report surveys were used to gather data from preservice teachers in Kazakhstan who were studying to become math teachers at universities. A Bayesian statistical framework was employed to test hypotheses and compare models, providing probabilistic evidence regarding the presence or absence of effects.

Data Collection Tools

Data were collected through an online survey administered via Google Forms, with all instruments presented in the Kazakh language. The survey consisted of four sequential sections. The first section presented an informed consent form, in which participants acknowledged their voluntary participation and granted permission for their data to be used and published in academic research. Participants who declined consent were unable to proceed to the survey items. The second section gathered demographic information, specifically participants' gender. The third section contained the mathematical literacy scale, followed by the assessment literacy scale in the fourth section. No monetary compensation or incentives were provided to participants for their involvement in the study. This procedure ensured informed consent and allowed participants to complete the instruments at their own pace in a familiar digital environment.

Mathematical literacy

The mathematical literacy scale was developed based on literature review and the current PISA assessment framework (see OECD PISA framework). Content validity was established through 8 experts' reviews. The item- and scale-level indices indicated strong agreement: the CVI values ranged from 0.875 to 1.00, the I-CVI was .977, and the CVI/UA was .81. A cognitive pretest with 10 students not included in the main sample confirmed item clarity, after which the final form was administered.

Exploratory factor analysis supported a two-factor solution. Sampling adequacy was excellent (KMO = .91) and Bartlett's test was significant, $\chi^2(496) = 6138$, $p < .001$. Using principal axis factoring with varimax rotation, Factor 1 showed loadings from .539 to .881 across 26 items and was labeled Mathematical Literacy (SS loading = 17.76; explained variance = 52.4%). Factor 2 grouped the reverse-worded items with loadings from .713 to .847 and was labeled General Mathematics Anxiety/Self-Efficacy (SS loading = 3.85; explained variance = 12.0%). After recoding, reverse items remained within the same factor, indicating a stable structure. Higher scores on the reverse-keyed items should be interpreted as lower mathematical fluency.

Confirmatory factor analysis on an independent sample provided acceptable fit: $\chi^2/df = 2.21$ ($\chi^2 = 988$, df = 447), CFI = .920, TLI = .911, SRMR = .048, and RMSEA = .0798, 90% CI [.073, .0866]. Reliability was high for both dimensions (Factor 1: Cronbach's α = .978, McDonald's ω = .983; Factor 2: α = .906, ω = .910). Overall, the scale appears suitable for use with preservice teachers in Kazakhstan.

Assessment literacy

Version 6.0 of the assessment literacy standards developed by the Consortium (2024) was used. The scale was translated into Kazakh by two bilingual experts, reviewed, and finalized. It was then back-translated into English. A third independent expert compared the forward and back translations and confirmed linguistic equivalence. This follows common guidance for translation and adaptation of instruments (Brislin, 1970). The Kazakh form was tested with confirmatory factor analysis (CFA). Model fit indices were: $\chi^2/df = 4.08$ ($\chi^2 = 1327$, df = 325), CFI = .923, TLI = .910, SRMR = .0254, and RMSEA = .0959, 90% CI [.0906, .101]. These results indicate an acceptable, if marginal, fit for research use.

Internal consistency was high across all seven domains. Cronbach's α and McDonald's ω, respectively,

were: Method selection & balanced system (.930, .932), tool development & quality (.931, .933), administration, scoring & interpretation (.932, .933), using results for instruction (.944, .944), grading & standards-based reporting (.940, .941), communication & reporting (.936, .936), and ethics, accessibility & legal responsibility (.950, .950). Taken together, the translation process, CFA outcomes, and reliability estimates support the use of this scale with Kazakh-speaking participants.

Sample

The study was conducted at seven universities in Kazakhstan (Korkyt Ata Kyzylorda University, Altynsarin University, SDU University, Zhanibekov University, Shakarim University, Kazakh National Women's Teacher Training University, and Sherkhan Murtaza University). Initially, 371 preservice mathematics teachers accessed the online survey. Of these, 30 participants did not complete the informed consent form and were therefore excluded from the study. The remaining 341 responses underwent data screening procedures to ensure data quality and completeness. Six participants were removed from the dataset due to incomplete responses, resulting in a final sample of 335 preservice teachers. The sample consisted of 135 males (40.3%), 196 females (58.5%), and 4 participants (1.2%) who preferred not to disclose their gender. All participants were enrolled in mathematics teacher education programs at the participating institutions at the time of data collection.

Data Analysis

All statistical analyses were conducted using Jamovi 2.7.6 (Project, 2025), an open-source statistical software built on the R programming environment. A Bayesian statistical framework was adopted throughout the study to quantify evidence for both the alternative and null hypotheses, providing a more nuanced interpretation of results compared to traditional frequentist approaches (Wagenmakers et al., 2018). Bayes factors (BF_{10}) were interpreted following the classification scheme proposed by Jeffreys (1961) and refined by Lee and Wagenmakers (2014): $BF_{10} > 100$ indicates decisive evidence for the alternative hypothesis, 30-100 very strong evidence, 10-30 strong evidence, 3-10 moderate evidence, and 1-3 anecdotal evidence. Conversely, $BF_{10} < 1$ provides evidence favoring the null hypothesis, with values below 0.33 considered moderate and below 0.10 considered strong evidence for the null.

Descriptive statistics were computed for all study variables, including means, standard deviations, medians, and ranges. To assess whether preservice teachers' mathematical literacy and assessment literacy levels differed significantly from the scale midpoint (3.0 on a 5-point scale), Bayesian one-sample t-tests were performed using default Cauchy priors (scale = 0.707) as recommended by Rouder et al. (2012).

Gender differences in mathematical literacy and assessment literacy dimensions were examined using Bayesian Mann-Whitney U tests due to concerns about distributional assumptions (van Doorn et al., 2021). These non-parametric tests provided Bayes factors quantifying evidence for or against gender-based differences while accounting for ordinal measurement properties of Likert-type scales.

The relationships among mathematical literacy, mathematics anxiety, and the seven assessment literacy dimensions were explored using Kendall's tau-b correlation coefficients. Kendall's tau was preferred over Pearson's r due to its robustness to outliers and appropriateness for ordinal data. Bayesian correlation analyses were conducted using stretched beta priors (width = 1) following the approach of Wetzels and Wagenmakers (2012), yielding Bayes factors that quantified evidence for non-zero correlations.

Latent profile analysis (LPA) was performed to identify distinct subgroups of preservice teachers based on their assessment literacy profiles across the seven dimensions. LPA is a person-centered analytical technique that classifies individuals into homogeneous latent classes based on response patterns (Spurk et al., 2020). Models with two through five profiles were estimated and compared using multiple fit indices: Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), sample-size-adjusted BIC (SABIC), entropy, and the Bootstrap Likelihood Ratio Test (BLRT). Lower values of information criteria indicate better fit, while entropy values above 0.80 suggest adequate classification accuracy. BLRT, with 500 bootstrap samples tested whether a k-profile solution fit significantly better than a (k-1)-profile solution (Nylund et al., 2007). The optimal number of profiles was determined by jointly considering statistical fit indices, classification quality, theoretical interpretability, and parsimony.

Following profile identification, Bayesian one-way ANOVA was conducted to test whether mathematical literacy and mathematics anxiety differed across the latent profiles. Default JZS priors (r scale = 0.5) were employed for effect size estimation (Rouder et al., 2012). When the ANOVA provided evidence for profile differences, Bayesian post-hoc comparisons with uniform prior odds across all pairwise contrasts were performed to identify which specific profiles differed from one another. All statistical tests were two-tailed, and results were interpreted within the context of both the magnitude of Bayes factors and the practical significance of observed effects.

Table 1. Mathematical literacy level and Bayesian test

Mathematical literacy dimension	N	Mean	SD	BF_{10}	error %
Mathematical literacy	335	3.89	0.825	$5.89*10^{+54}$	$1.13*10^{-60}$
Mathematics anxiety	335	3.10	1.078	0.273	0.081

Table 2. Assessment literacy level and Bayesian test

Assessment literacy	N	Mean	SD	BF_{10}	error %
MSBS	335	4.04	0.776	$2.24\times10^{+73}$	1.64×10^{-78}
TDQ	335	4.06	0.805	$1.19\times10^{+71}$	1.97×10^{-76}
ASI	335	3.99	0.809	$1.86\times10^{+65}$	2.34×10^{-71}
URI	335	4.08	0.792	$2.68\times10^{+74}$	1.64×10^{-79}
GSBR	335	4.02	0.796	$1.12\times10^{+69}$	1.30×10^{-74}
CR	335	4.05	0.806	$6.76\times10^{+70}$	3.30×10^{-76}
EALR	335	4.06	0.832	$1.75\times10^{+68}$	6.73×10^{-74}

FINDINGS

Descriptive Statistical

The Bayesian one-sample results show a clear contrast between the two constructs (as shown in **Table 1**). Mathematics literacy is high and credibly different from the test value (Mean = 3.89, SD = 0.825, N = 335). The Bayes factor is extreme ($BF_{10} = 5.89\times10^{54}$; error = 1.13×10^{-60}), providing overwhelming evidence that the mean differs from the null value. In contrast, mathematics self-anxiety is close to the test value (Mean = 3.10, SD = 1.078, N = 335), and the Bayes factor favors the null ($BF_{10} = 0.273$; error = 0.081), indicating evidence against a difference. The larger SD for self-efficacy suggests more heterogeneity across candidates. Overall, candidates report strong mathematics literacy but only neutral-to-slightly-above levels of anxiety, with compelling Bayesian evidence for the former difference but not for the latter.

Across all seven Assessment Literacy dimensions (N = 335), Bayesian one-sample tests provide decisive evidence that mean scores differ from the benchmark of 3, and the direction is positive (as shown in **Table 2**). Means cluster around 4.00 with moderate dispersion (SDs ≈ 0.79–0.83). Using Results for Instruction (URI) shows the highest mean (4.08, SD = 0.792; BF_{10} = 2.68E + 74), closely followed by tool development & quality (TDQ, 4.06, SD = 0.805; BF_{10} = 1.19E + 71), ethics, accessibility & legal responsibility (EALR, 4.06, SD = 0.832; BF_{10} = 1.75E + 68), and communication & reporting (CR, 4.05, SD = 0.806; BF_{10} = 6.76E + 70). Method selection & balanced system (MSBS, 4.04; BF_{10} = 2.24E + 73) and grading & standards-based reporting (GSBR, 4.02; BF_{10} = 1.12E + 69) are slightly lower but still clearly above 3. Administration, scoring & interpretation (ASI) has the lowest mean (3.99, SD = 0.809; BF_{10} = 1.86E + 65), yet the bayes factor remains astronomically large. Error percentages are vanishingly small (≈ 10^{-74} to 10^{-79}), indicating stable estimates. Overall, preservice mathematics teachers report high perceived assessment literacy across domains, with the strongest elevation for instructional use of results and test development quality, and comparatively lower (but still high) perceptions for grading/reporting and administration/scoring.

Gender Difference

Across gender, Bayesian comparisons show little evidence of mean differences in mathematics literacy, mathematics anxiety, or most assessment-literacy domains (as shown in **Table 3**). For mathematics literacy, males score slightly higher (3.97 vs. 3.84) but the Bayes factor favors the null (BF_{10} = 0.551), indicating the data do not support a gender gap. Anxety is virtually identical (3.13 vs. 3.10; BF_{10} = 0.126). The assessment-literacy facets likewise favor no difference: MSBS 4.14 vs. 3.98 (BF_{10} = 0.901), TDQ 4.13 vs. 4.01 (BF_{10} = 0.299), URI 4.10 vs. 4.06 (BF_{10} = 0.135), GSBR 4.10 vs. 3.97 (BF_{10} = 0.303), CR 4.12 vs. 4.01 (BF_{10} = 0.270), and EALR 4.11 vs. 4.03 (BF_{10} = 0.158). The only domain with evidence for a gender difference is ASI, where males report higher scores (4.11 vs. 3.92) and the Bayes factor supports the alternative (BF_{10} = 2.43), though the mean gap remains small. Notably, male groups also show slightly lower dispersion on several subscales (e.g., GSBR SD = 0.723 vs. 0.843), suggesting marginally more homogeneous responses. Overall, the pattern indicates broad gender parity, with a modest male advantage emerging only for Assessment Administration, Scoring & Interpretation (ASI).

The correlation pattern shows two clear results (as shown in **Table 4**). First, mathematics literacy relates moderately and consistently to every assessment-literacy dimension: τ ranges from .44 to .52 (e.g., MSBS τ = .519, TDQ τ = .479, ASI τ = .472, URI τ = .461, GSBR τ = .456, CR τ = .443, EALR τ = .447), and the Bayes factors are enormous (e.g., $BF_{10} = 2.40\times10^{42}$ for MSBS), indicating overwhelming evidence for these positive associations. Second, the reverse-scored math self-efficacy shows essentially no relation to any construct (e.g., with Mah_lit τ = −.078, BF_{10} = 0.683; with MSBS τ = .008, BF_{10} = 0.073; with TDQ τ ≈ 0, BF_{10} = 0.072), providing evidence for the null. As expected, assessment-literacy subscales are strongly

Table 3. Gender difference based on mathematics literacy and asssesment literacy

Dimensions	Group	N	Mean	SD	BF_{10}	W
Mathematical literacy	Female	196	3.84	0.794	0.551	11832
	Male	135	3.97	0.854		
Mathematics anxiety	Female	196	3.1	0.998	0.126	12760
	Male	135	3.13	1.184		
MSBS	Female	196	3.98	0.797	0.901	11778
	Male	135	4.14	0.733		
TDQ	Female	196	4.01	0.835	0.299	12185
	Male	135	4.13	0.761		
ASI	Female	196	3.92	0.844	2.43	11665
	Male	135	4.11	0.748		
URI	Female	196	4.06	0.836	0.135	13113
	Male	135	4.1	0.728		
GSBR	Female	196	3.97	0.843	0.303	12162
	Male	135	4.1	0.723		
CR	Female	196	4.01	0.851	0.27	12365
	Male	135	4.12	0.735		
EALR	Female	196	4.03	0.866	0.158	12746
	Male	135	4.11	0.777		

Table 4. Correlation among mathematics literacy and assessment literacy

	BF_{10}	M_Lit	Math Ax	MSBS	TDQ	ASI	URI	GSBR	CR
Math_Ax	Kendall's tau	-0.078	—						
	BF_{10}	0.683	—						
MSBS	Kendall's tau	0.519***	0.008	—					
	BF_{10}	$2.40x10^{+42}$	0.073	—					
TDQ	Kendall's tau	0.479***	$-9.63x10^{-4}$	0.757***	—				
	BF_{10}	$8.34x10^{+35}$	0.071	$1.51x10^{+91}$	—				
ASI	Kendall's tau	0.472***	-0.022	0.697***	0.806***	—			
	BF_{10}	$6.21x10^{+34}$	0.085	$1.81x10^{+77}$	$4.31x10^{+103}$	—			
URI	Kendall's tau	0.461***	0.028	0.701***	0.772***	0.820***	—		
	BF_{10}	$1.20x10^{+33}$	0.096	$1.35x10^{+78}$	$7.73x10^{+94}$	$2.00x10^{+107}$	—		
GSBR	Kendall's tau	0.456***	-0.031	0.651***	0.754***	0.806***	0.820***	—	
	BF_{10}	$2.67x10^{+32}$	0.102	$1.62x10^{+67}$	$2.63x10^{+90}$	$3.77x10^{+103}$	$1.6\ x10^{+107}$	—	
CR	Kendall's tau	0.443***	0.005	0.683***	0.725***	0.767***	0.828***	0.838***	—
	BF_{10}	$3.52x10^{+30}$	0.072	$1.51x10^{+74}$	$5.45x10^{+83}$	$4.91x10^{+93}$	$1.79x10^{+109}$	$1.12x10^{+112}$	—
EALR	Kendall's tau	0.447***	0.003	0.682***	0.722***	0.742***	0.784***	0.770***	0.806***
	BF_{10}	$1.20x10^{+31}$	0.072	$6.20x10^{+73}$	$1.01x10^{+83}$	3.4410^{+87}	$8.52x10^{+97}$	$3.02x10^{+94}$	$3.02x10^{+103}$

Note. * $BF_{10} > 10$, ** $BF_{10} > 30$, *** $BF_{10} > 100$

intercorrelated (many $\tau \geq .75$; e.g., TDQ–MSBS $\tau = .757$, ASI–TDQ $\tau = .806$, URI–CR $\tau = .828$, GSBR–CR $\tau = .838$) with decisive Bayes factors (all $BF_{10} \gg 10^{60}$), supporting a cohesive AL system. Overall, higher mathematics literacy aligns with stronger perceived assessment literacy across its facets, whereas the (reverse) math self-efficacy indicator does not contribute meaningfully to these relationships.

Latent Profile Analysis

The latent profile analysis fit indices presented in **Table 5** were examined. First, Model 1 yields clean classifications, whereas Model 3 shows pathological solutions for 2–3 classes (Entropy = 0.311 and 0.148; n_min = 0; BLRT p = 1.00), indicating near-degenerate or non-separating profiles. In contrast, Model 1 maintains high classification certainty across K: For K = 3, Entropy = 0.968 and prob_min = 0.978, with no empty classes. Second, within Model 1 the jump from 2→3 classes produces a large improvement in fit (BIC: 4465.41 → 3239.70; AIC: 4381.50 → 3125.27; SABIC: 4395.62 → 3144.53), and the BLRT is significant (p = 0.0099), supporting the additional class. The resulting class sizes are substantively balanced (n_min = 0.167, n_max = 0.501), avoiding dominance or sparsity, and median posterior probabilities are uniformly high (prob_max = 0.989). Third, adding more classes provides diminishing—and potentially misleading—returns. Although some information criteria continue to decrease for K = 4–5, classification quality and practicality deteriorate: for K = 4, Entropy drops (0.938) and the smallest class tightens (n_min = 0.149); for K = 5, despite very low ICs (e.g., BIC = 2398.92), a tiny class emerges

Table 5. Model selection indices in LPA

	Values	Values	Values	Values	Values	Values	Values	Values
Model	1	1	1	1	3	3	3	3
Classes	2	3	4	5	2	3	4	5
LogLik	-2168.748	-1532.636	-1459.411	-1065.734	-961.783	-961.848	-859.5293	-814.153
AIC	4381.497	3125.272	2994.822	2223.467	2009.566	2025.696	1837.059	1762.306
AWE	4657.483	3502.184	3472.82	2802.422	2551.96	2669.441	2580.212	2606.713
BIC	4465.407	3239.696	3139.759	2398.917	2173.574	2220.217	2062.092	2017.853
CAIC	4487.407	3269.696	3177.759	2444.917	2216.574	2271.217	2121.092	2084.853
CLC	4339.332	3067.209	2920.699	2133.412	1924.188	1923.992	1720.972	1629.992
KIC	4406.497	3158.272	3035.822	2272.467	2055.566	2079.696	1899.059	1832.306
SABIC	4395.621	3144.533	3019.22	2253	2037.174	2058.44	1874.938	1805.322
ICL	-4481.956	-3252.13	-3167.227	-2414.11	-2313.684	-2583.643	-2081.474	-2099.684
Entropy	0.91776	0.96822	0.93804	0.97269	0.311	0.148	0.95683	0.84311
prob_min	0.9691	0.97785	0.88507	0.95099	0	0	0.96915	0.76328
prob_max	0.98575	0.98935	0.99471	1	1	1	0.99988	0.99856
n_min	0.40896	0.16716	0.14925	0.02388	0	0	0.04776	0.01493
n_max	0.59104	0.50149	0.43284	0.39403	1	1	0.39403	0.61791
BLRT_val	1276.512	1272.224	146.4497	787.3555	-0.114	-0.13	204.6377	90.75261
BLRT_p	0.0099	0.0099	0.0099	0.0099	1	1	0.0099	0.0099

Table 6. Desciriptive statatistics for profiles

	Class	N	Mean	Median	SD	Minimum	Maximum
MSBS	1	168	3.93	4	0.432	2	5
	2	56	2.92	3	0.632	1	4.25
	3	111	4.78	5	0.358	3.25	5
TDQ	1	168	3.95	4	0.377	2.75	5
	2	56	2.78	3	0.607	1	4
	3	111	4.86	5	0.258	4	5
ASI	1	168	3.85	4	0.397	2.5	5
	2	56	2.76	3	0.534	1	3.5
	3	111	4.84	5	0.297	4	5
URI	1	168	3.97	4	0.370	3	5
	2	56	2.82	3	0.586	1	3.75
	3	111	4.87	5	0.245	4	5
GSBR	1	168	3.89	4	0.415	2.5	5
	2	56	2.81	3	0.558	1	3.75
	3	111	4.83	5	0.284	4	5
CR	1	168	3.96	4	0.393	2.5	5
	2	56	2.77	3	0.574	1	4
	3	111	4.85	5	0.277	4	5
EALR	1	168	3.95	4	0.436	2	5
	2	56	2.75	3	0.615	1	4
	3	111	4.88	5	0.231	4	5

(n_min = 0.024), which risks over-extraction and unstable, non-generalizable profiles. As a result, Model 1 (K = 3) offers the best balance of fit, classification certainty, and interpretable, well-sized classes.

Across all seven assessment-literacy dimensions, the three-class solution is highly ordered and consistent (as shown in **Table 6**). Class 3 shows near-ceiling mastery (e.g., TDQ = 4.86, URI = 4.87, EALR = 4.88; medians = 5 throughout) with very small dispersion (SDs ≈ 0.23–0.36). Class 1 centers around solid but non-ceiling proficiency (e.g., MSBS = 3.93, URI = 3.97, CR = 3.96; SDs ≈ 0.37–0.44), and Class 2 is uniformly low (e.g., MSBS = 2.92, TDQ = 2.78, EALR = 2.75) with the largest variability (SDs up to 0.63). Min–max ranges confirm this structure: Class 3 rarely dips below 4.0 (mins ≥ 3.25–4.0), whereas Class 2 includes the full lower spectrum (mins often = 1.0), and Class 1 spans moderate ranges (mins ~2.5–3.0).

The tight SDs in Class 3 suggest a homogeneous group confidently applying assessment practices across Method Selection & Balanced System, Tool Development & Quality, Administration, Scoring & Interpretation, Using Results for Instruction, Grading & Standards-Based Reporting, Communication &

Table 7. Matematical literacy in assessment profile

Math literacy dimension	Class	N	Mean	Median	SD
Math literacy	Assessment practitioners	168	3.74	3.85	0.665
	Assessment novices	56	3.21	3.13	0.854
	Assessment experts	111	4.47	4.65	0.648
Math anxiety	Assessment practitioners	168	3.14	3.17	0.815
	Assessment novices	56	3.08	3.17	0.913
	Assessment experts	111	3.05	3	1.448

Table 8. Bayesian ANOVA based on profiles

Dimensions	Models	P(M)	P(M | data)	BFM	BF10	error %
Math lit	Null model	0.5	1.65E-23	$1.65x10^{-23}$	1	
	Class	0.5	1	$6.05x10^{+22}$	$6.05x10^{+22}$	$1.42x10^{-4}$
Math anxiety	Null model	0.5	0.9543	20.8875	1	
	Class	0.5	0.0457	0.0479	0.0479	0.0246

Table 9. Bayesian post-hoc test results

		Prior odds	Posterior odds	$BF_{10, U}$	error %
Assessment practitioners	Assessment novices	0.587	3336	5679	$2.95x10^{-10}$
	Assessment experts	0.587	$1.30x10^{+14}$	$2.21x10^{+14}$	$1.79x10^{-21}$
Assessment novices	Assessment experts	0.587	$1.22x10^{+17}$	$2.08x10^{+17}$	$1.37x10^{-19}$

Reporting, and Ethics, Accessibility & Legal Responsibility. In contrast, the wider SDs in Class 2 signal uneven skill development and likely gaps in core procedures and ethics/accessibility; Class 1 appears stable but with clear headroom below the ceiling (means ~3.85–3.97).

When examining the average standard deviations of the classes, Class 1 represents 50.1% of the group. They are named Assessment Practitioners. They have mid-to-high means (~3.9) with moderate variability, showing established, working proficiency. Class 2 represents 16.7% of the group. They are named Assessment Novices. The consistently low means (≈ 2.75–2.92) and broad spread indicate early-stage competence. Class 3 represents 33.1% and is named Assessment Experts. They near-ceiling means (≥ 4.78) and minimal variability denote comprehensive, consistent mastery across all facets.

Mathematics literacy levels align clearly with the assessment-literacy classes (as shown in **Table 7**). Assessment Experts show the highest mathematics literacy (Mean = 4.47, Median = 4.65, SD = 0.648), well above Assessment Practitioners (Mean = 3.74, Median = 3.85, SD = 0.665) and Assessment Novices (Mean = 3.21, Median = 3.13, SD = 0.854). The ordered medians (4.65 > 3.85 > 3.13) and the relatively small dispersion for Experts indicate a strong, consistent advantage in mathematics literacy among those with advanced assessment literacy.

In contrast, math anxiety differs little across classes (Means ≈ 3.14, 3.08, 3.05; Medians = 3.17, 3.17, 3.00), and variability is largest within Experts (SD = 1.448). This suggests that the observed gains in mathematics literacy are not mirrored by parallel increases in anxiety, and that confidence levels among Experts are heterogeneous. Overall, higher assessment-literacy profiles are associated with substantially higher mathematics literacy, while anxiety about mathematics remains comparatively flat and dispersed across groups.

Bayesian ANOVA strongly supports class-based differences in mathematics literacy (as shown in **Table 8**). The class model outperforms the null with BF_{10} = 6.05×10^{22} and P(M | data) = 1.00, indicating decisive evidence that mathematics literacy varies by assessment-literacy class. Post hoc tests confirm an ordered pattern: Experts > Practitioners > Novices (as shown in **Table 9**). Evidence is extreme for Experts vs Practitioners (BF_{10} = 2.21×10^{14}) and Experts vs Novices (BF_{10} = 2.08×10^{17}), and very strong for Practitioners vs Novices (BF_{10} = 5,679). These results align with the descriptive statistics (Experts Mean = 4.47, Practitioners = 3.74, Novices = 3.21) and the ordered medians (4.65 > 3.85 > 3.13). Error rates are negligible, supporting the stability of these inferences.

For math anxiety, the model comparison favors the null: P(M | data) = 0.954, BF_{10} = 0.048 for the class model. This indicates no reliable class differences in anxiety despite the clear gaps in mathematics literacy. Descriptives are consistent with this: Means cluster near 3.05–3.14 with overlapping medians, and the Experts group shows the largest dispersion (SD = 1.448), suggesting heterogeneity in confidence even among those with high performance.

Overall, advanced assessment literacy profiles correspond to substantially higher mathematics literacy but do not translate into higher math anxiety. Instructional implications are twofold: Leverage the strengths of the Expert profile to scaffold mathematical

performance in other groups, and design separate, targeted interventions to build anxiety, as confidence appears decoupled from assessment competence and mathematical achievement in this sample.

DISCUSSION

The current study indicated that preservice mathematics instructors in Kazakhstan exhibited elevated levels of both mathematical literacy and assessment literacy. These findings differ from previous research in other contexts. Research conducted in Turkey revealed that preservice teachers exhibited insufficient assessment literacy, achieving correct response rates of approximately 53% (Şahin, 2023). In the same way, Özdil and Kaçar (2023) found that Turkish math teacher applicants had low levels of evaluation literacy. Studies conducted in Indonesia have similarly recorded challenges in mathematics literacy among preservice teachers (Yustitia et al., 2020). The elevated literacy levels identified in the present study may indicate many influencing variables. First, the use of self-report measures may inflate perceived competence compared to performance-based assessments. Şahin (2023) utilized a knowledge assessment, whereas the present study applied Likert-scale inventories. Second, the cultural and educational environments vary from one country to another. Kazakhstan's teacher education system might focus on these skills in a different way than those in Turkey or Indonesia. Third, social desirability bias in online surveys may lead participants to report higher proficiency (van de Mortel, 2008).

The strong positive relationship between mathematical literacy and assessment literacy supports theoretical predictions. This finding is consistent with educational content knowledge frameworks that highlight the interrelation of subject matter expertise and assessment methodologies (Shulman, 1986). Ayalon and Wilkie (2020) showed that having understanding of the subject is important for establishing useful math evaluation criteria.The current correlations (τ = .44–.52) provide empirical support for this connection. Teachers cannot effectively assess mathematical thinking without understanding the mathematical content themselves (Veldhuis & van den Heuvel-Panhuizen, 2014). Conversely, engaging in assessment activities may deepen teachers' own mathematical understanding (Chamberlin et al., 2008). This reciprocal relationship has important implications for teacher education. Programs should integrate mathematical literacy and assessment literacy development rather than treating them as separate competencies (Shmigirilova et al., 2024).

The absence of mathematics anxiety's relationship with other constructs was unexpected. Previous research identified anxiety as a significant negative predictor of mathematical literacy (Dirgantoro et al., 2025). However, Yani et al. (2025) similarly found no relationship between mathematical literacy and digital literacy among Indonesian preservice teachers. This suggests that different literacy domains may function independently. The neutral anxiety levels (M = 3.10) and lack of correlation may indicate that preservice teachers in this sample view assessment and mathematical tasks separately from their emotional responses to mathematics. Alternatively, the reverse-coded nature of these items may have introduced measurement issues that weakened observable relationships.

Gender parity in both literacy domains contradicts some past findings but agrees with recent developments. Yani et al. (2025) found no variations in mathematics literacy between male and female preservice teachers. The only difference in this study is that men scored a little better on administration, scoring, and interpretation (BF_{10} = 2.43). This is a tiny effect. This area deals with the technical parts of running tests and understanding the results. The Bayes factor, on the other hand, simply gives anecdotal evidence for this distinction. In general, the findings indicate that disparities in mathematical and evaluative skills between genders may be decreasing in modern teacher education settings. This finding corroborates parity in professional preparation.

The latent profile analysis constitutes the study's most innovative contribution. Three unique profiles developed with a clear hierarchical framework. Assessment experts (33.1%) showed almost perfect competence in all areas and were much better at math (M = 4.47) than the other groups. Assessment practitioners (50.1%) demonstrated substantial competence in both areas. Assessment novices (16.7%) demonstrated inferior performance across various metrics. Bayesian ANOVA validated significant disparities in mathematical literacy among profiles (BF_{10} = 6.05×10^{22}).Yet mathematics anxiety remained equivalent across groups. This pattern indicates that advanced assessment literacy co-occurs with strong mathematical understanding but not with reduced anxiety. Previous research has not examined such profile structures among preservice mathematics teachers. The profiles suggest that teacher education programs serve students with varying readiness levels. The substantial expert group (one-third of the sample) may benefit from advanced coursework or peer-teaching opportunities. The novice group requires targeted foundational support. Most importantly, the profiles demonstrate that assessment literacy is not uniformly distributed. Program designers should consider differentiated instruction based on entry-level competencies.

This study adds to the body of knowledge about how to teach math teachers. First, it offers empirical evidence for the correlation between mathematical literacy and evaluation literacy. Prior studies examined this relationship conceptually (Shmigirilova et al., 2024) but were deficient in substantial quantitative data. The

moderate-to-strong correlations (τ = .44–.52) indicate that these constructs are similar but not identical. This discovery expands the theory of instructional content knowledge (Shulman, 1986). Shulman argued that effective teaching requires both subject matter knowledge and knowledge of how to teach that subject. The current results suggest that assessment knowledge should be added as a third essential component. To help students learn well, teachers need to know math, how to teach, and how to grade.

Second, the latent profile structure provides a person-centered perspective that enhances variable-centered methodologies. Most previous studies investigated the mean correlations among variables (Özdil & Kaçar, 2023; Şahin, 2023). This method presumes that all preservice teachers are alike, differing only in their placement on a continuum. The profile analysis shows that there are different subgroups with distinctive trends. Experts in assessment are quite good at both areas. Assessment Novices struggle with both. Assessment Practitioners fall in between. This heterogeneity has theoretical implications. It suggests that teacher development does not follow a single linear path. Different preservice teachers may need different types of support based on their profile membership. This aligns with differentiation theories in education (Tomlinson, 2014) but extends them to teacher preparation contexts.

Third, the study contributes to understanding literacy relationships. Yani et al. (2025) found that mathematical literacy and digital literacy were independent among preservice teachers. They concluded that different literacies require separate attention. The current findings reveal a distinct trend for mathematical literacy and evaluation literacy. There is a significant link between these two ideas. This indicates that certain literacies possess shared foundations, whereas others do not. Subsequent research ought to investigate the factors that influence the interrelation of literacies. Some possible reasons are content overlap, cognitive demands, and real-world uses. Comprehending these interrelations can assist teacher educators in formulating more effective programs that cultivate various talents concurrently.

Fourth, the null finding with mathematics anxiety enhances theoretical discourse. Dirgantoro et al. (2025) recognized anxiety as a significant internal element influencing mathematical literacy. Nonetheless, the present investigation identified no correlation between anxiety and either literacy domain. This indicates that emotive and cognitive dimensions may operate independently within teacher education situations. Preservice teachers can develop strong mathematical and assessment skills regardless of their anxiety levels. Alternatively, teacher candidates may regulate anxiety differently than school students. They know that teaching math means getting over their own uneasiness. This finding contradicts deficit models that posit worry inherently diminishes professional performance (Ramirez et al., 2018). It indicates that teacher education programs ought to treat anxiety and skill enhancement as distinct objectives, rather than presuming that alleviating anxiety inherently enhances competence.

CONCLUSIONS

This study examined the levels of mathematical literacy and assessment literacy among prospective mathematics teachers in Kazakhstan and revealed the relationship between these two constructs. The findings showed that prospective teachers had high perceived competence levels in both mathematical literacy and assessment literacy. In terms of the gender variable, a general equality was observed between female and male teacher candidates, with only a small difference in favor of males in the assessment management dimension. Moderately strong positive relationships were found between all dimensions of mathematical literacy and assessment literacy. This finding reveals that the two literacy domains are related but not identical constructs. It is noteworthy that math anxiety did not show a meaningful relationship with any construct.

Latent profile analysis identified three distinct groups among teacher candidates. Assessment experts demonstrated high proficiency in all dimensions, assessment practitioners showed moderate proficiency, and assessment novices demonstrated low performance in all dimensions. There were substantial disparities in how well each group understood arithmetic, but the degrees of math anxiety were the same for all groups. The assessment experts had a far better math skills than the other groups. These results indicate that teacher candidates exhibiting advanced assessment literacy also demonstrate robust mathematics literacy.

The study offers three significant contributions to the domain of mathematics teacher education. First, it provides strong empirical evidence for the relationship between mathematical literacy and assessment literacy. Second, it reveals the heterogeneous structure among teacher candidates using a person-centered analysis approach and identifies the existence of groups with different competency profiles. Third, it has been found that math anxiety functions independently of these two literacy domains. These findings suggest that teacher education programs need to address the processes of developing both content knowledge and assessment skills in an integrated manner.

LIMITATIONS

The study has some limitations. First, data collection is based on self-report scales. Participants' assessment of their own competencies may lead to reporting higher levels than their actual performance, particularly due to social desirability bias. Second, the study has a cross-sectional design. Because data was collected at a single

point in time, definitive conclusions about causal relationships cannot be drawn. Longitudinal studies are necessary to comprehend the evolution of mathematics and assessment literacy within teacher education programs and the factors that affect this progression. Third, the sample was obtained from merely two universities in Kazakhstan. The findings have limited generalizability to various school systems due to varying cultural and educational contexts among countries. Cross-validation studies in different cultural contexts and education systems are necessary to test the universality of these findings. Fourth, the measurement of math anxiety involves methodological complexities. The fact that reverse-coded items form a separate factor and the requirement to interpret high scores as low mathematical fluency may have led to uncertainty in the relationships between anxiety and other constructs. Subsequent research ought to evaluate math anxiousness employing directly constructed and psychometrically superior assessments. Lastly, the research did not examine the actual pedagogical practices of teacher candidates. It is uncertain if teacher candidates who indicated a strong sense of competence successfully utilized this information in actual teaching contexts.

RECOMMENDATIONS

The findings offer important practical recommendations for teacher education programs. First, given the strong relationship between mathematical literacy and assessment literacy, these two areas should be taught in an integrated manner rather than as separate courses. The integrative approach proposed by Shmigirilova et al. (2024) should be adopted. In mathematics content courses, teacher candidates should both solve mathematical problems and develop assessment criteria for these problems. In mathematics pedagogy courses, they should analyze real student work and apply formative assessment strategies. This integration will enable teacher candidates to simultaneously develop both mathematical thinking and the processes for assessing this thinking.

Second, differentiated teaching strategies should be developed based on the three distinct profiles identified. assessment experts can benefit from advanced courses, research projects, or peer teaching opportunities. This group can take on assistant teacher or mentor roles and provide support to other candidates. Assessment novices should have extra classes, one-on-one mentorship programs, or helpful workshops set up for them. This group needs a lot of help with basic ideas. To get better at their jobs, assessment practitioners require structured practice opportunities and ongoing feedback.

Teacher education programs ought to implement diagnostic evaluations aligned with entry-level competencies and provide customized learning routes to meet students' specific requirements. Third, performance-based assessment components should be integrated into teacher education programs. Going beyond self-report measures, teacher candidates' performance on real practice tasks should be evaluated.

Tasks such as developing rubrics for assessment literacy, designing assessment tools, scoring real student work, and translating test results into educational decisions should be assigned. For mathematical literacy, tasks such as solving PISA-style contextual problems, completing mathematical modeling tasks, and explaining mathematical reasoning both in writing and orally should be included. Working with real student responses, as suggested by Ayalon and Wilkie (2020), can be effective in developing both the assessment and mathematical skills of teacher candidates.

Fourth, math anxiety and proficiency development should be addressed as separate educational goals. Findings indicate that anxiety functions independently of assessment and mathematical proficiency. It would be incorrect to assume that reducing anxiety will automatically increase competence. Structured practice, regular feedback, and a focus on gradual skill progression should be prioritized for skill development. For affective support, self-efficacy building, strategies for coping with math anxiety, and access to psychological counseling services when needed should be provided.

DeLuca and Klinger (2010) highlight the deficiencies in teacher candidates' assessment literacy, indicating that these deficiencies arise from both insufficient knowledge and inadequate practical experience. Fifth, subsequent research ought to employ longitudinal designs to monitor the evolution of these literacies during teacher preparation programs and the initial years of professional practice.

Research tracking teacher candidates from program inception to graduation and their initial teaching years will elucidate the most successful program attributes and the application of acquired skills in practice. Furthermore, by using performance-based measures and self-report measures together, the relationship between perceived and actual competencies can be better understood. Dirgantoro et al. (2025), while examining the factors that influence mathematical literacy, emphasize the importance of considering both internal factors (conceptual understanding, self-efficacy, anxiety) and external factors (pedagogical instruction, sociocultural conditions).

Finally, the generalizability of these findings across different education systems should be tested through multicultural and multinational comparative studies. Hull and Vígh (2025) state that assessment literacy is a multidimensional construct and that this field must adapt to constantly changing technological and pedagogical innovations. Conducting similar studies in

different countries will help to understand which findings are universal and which are context-specific. In particular, the impact of factors such as the structure of teacher education programs, assessment culture, and the importance given to mathematics education on these literacies should be examined comparatively.

Author contributions: ZYN: conceptualization, writing – original draft; **LSK**: methodology, investigation; **GMY**: data curation, formal analysis; **ZAY & SKM**: investigation, visualization; **SSI**: supervision, writing – review & editing. All authors have read and approved the final manuscript.

Funding: No funding source is reported for this study.

Ethical statement: The study protocol was reviewed and approved by the Ethics Committee of Korkyt Ata University (Approval No.: 4304, Date: 2024).

AI statement: The authors declare that no artificial intelligence tools were used for production, data analysis, data collection, or literature review for this article.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Alegría Soto, J. G., Carrasco Díaz, C. E., Silva Barbini, R. O., & Núñez-Valdés, K. (2024). Alfabetización evaluativa: Acompañamiento docente para promover la evaluación auténtica [Evaluative literacy: Teacher guidance for promoting authentic assessment]. *Fides Et Ratio, 28*(28). https://doi.org/10.55739/fer.v28i28.157

Ata Baran, A., & Kabael, T. (2023). An investigation of eighth-graders' mathematical communication competency and mathematical literacy performance. *The Journal of Educational Research, 116*(4), 216-229. https://doi.org/10.1080/00220671.2023.2250752

Ayalon, M., & Wilkie, K. J. (2020). Developing assessment literacy through approximations of practice: Exploring secondary mathematics pre-service teachers developing criteria for a rich quadratics task. *Teaching and Teacher Education, 89*, Article 103011. https://doi.org/10.1016/j.tate.2019.103011

Bolstad, O. H. (2020). Secondary teachers' operationalisation of mathematical literacy. *European Journal of Science and Mathematics Education, 8*(3), 115-135. https://doi.org/10.30935/scimath/9551

Bolstad, O. H. (2021). Lower secondary students' encounters with mathematical literacy. *Mathematics Education Research Journal, 35*(1), 237-253. https://doi.org/10.1007/s13394-021-00386-7

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology, 1*(3), 185-216. https://doi.org/10.1177/135910457000100301

Brookhart, S. M. (2024). Educational assessment knowledge and skills for teachers revisited. *Education Sciences, 14*(7). https://doi.org/10.3390/educsci14070751

Canbazoğlu, H. B., & Tarım, K. (2020). An activity-based practice for improving mathematical literacy and awareness of elementary school teacher candidates. *Pegem Eğitim ve Öğretim Dergisi, 10*(4), 1183-1218. https://doi.org/10.14527/pegegog.2020.036

Chamberlin, M. T., Powers, R. A., & Novak, J. D. (2008). Teachers' perceptions of mathematics content knowledge assessments in professional development courses. *International Electronic Journal of Mathematics Education, 3*(3), 155-178. https://doi.org/10.29333/iejme/224

Christiansen, I. M., & Erixon, E.-L. (2021). Opportunities to learn mathematics pedagogy and learning to teach mathematics in Swedish mathematics teacher education: A survey of student experiences. *European Journal of Teacher Education, 47*(1), 159-177. https://doi.org/10.1080/02619768.2021.2019216

Consortium, M. A. (2024). *Assessment literacy standards: A national imperative*. MAC. https://www.michiganassessmentconsortium.org/wp-content/uploads/29-Assessment-Literacy-Standards.pdf

Darling-Hammond, L. (2016). Research on teaching and teacher education and its influences on policy and practice. *Educational Researcher, 45*(2), 83-91. https://doi.org/10.3102/0013189x16639597

DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice, 17*(4), 419-438. https://doi.org/10.1080/0969594x.2010.516643

DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016). Approaches to classroom assessment inventory: A new instrument to support teacher assessment literacy. *Educational Assessment, 21*(4), 248-266. https://doi.org/10.1080/10627197.2016.1236677

DeLuca, C., Schneider, C., Coombs, A., Pozas, M., & Rasooli, A. (2019). A cross-cultural comparison of German and Canadian student teachers' assessment competence. *Assessment in Education: Principles, Policy & Practice, 27*(1), 26-45. https://doi.org/10.1080/0969594x.2019.1703171

Dirgantoro, K. P. S., Kusumah, Y. S., Jupri, A., & Melissa, M. M. (2025). Mathematical literacy of prospective teachers: A systematic review of factors, readiness, and teacher education preparation. *Jurnal Pendidikan MIPA, 26*(3), 1355-1376. https://doi.org/10.23960/jpmipa.v26i3.pp1355-1376

Estaji, M., Banitalebi, Z., & Brown, G. T. L. (2024). The key competencies and components of teacher

assessment literacy in digital environments: A scoping review. *Teaching and Teacher Education, 141*, Article 104497. https://doi.org/10.1016/j.tate.2024.104497

Frambaugh-Kritzer, C., & Buelow, S. (2022). Problem solving like a mathematician: Disciplinary literacy instruction in elementary mathematics. *Reading Horizons: A Journal of Literacy and Language Arts, 61*(2), 31-64.

Geiger, V., Goos, M., & Forgasz, H. (2015). A rich interpretation of numeracy for the 21st century: A survey of the state of the field. *Zdm, 47*(4), 531-548. https://doi.org/10.1007/s11858-015-0708-1

Haara, F. O., Bolstad, O. H., & Jenssen, E. S. (2017). Research on mathematical literacy in schools - Aim, approach and attention. *European Journal of Science and Mathematics Education, 5*(3), 285-313. https://doi.org/10.30935/scimath/9512

Hull, P., & Vígh, T. (2025). Teachers' assessment literacy: A descriptive literature review. *Hungarian Educational Research Journal, 15*(2), 210-225. https://doi.org/10.1556/063.2024.00317

Jacinto, E. L., & Jakobsen, A. (2020). Mathematical knowledge for teaching: How do primary pre-service teachers in Malawi understand it? *African Journal of Research in Mathematics, Science and Technology Education, 24*(1), 31-40. https://doi.org/10.1080/18117295.2020.1735673

Jeffreys, H. (1961). *The theory of probability* (3erd edition ed.). Oxford.

Kappassova, S., Abylkassymova, A., Bulut, U., Zykrina, S., Zhumagulova, Z., & Balta, N. (2025). Mathematical literacy and its influencing factors: A decade of research findings (2015-2024). *Eurasia Journal of Mathematics, Science and Technology Education, 21*(7), Article em2671. https://doi.org/10.29333/ejmste/16615

Karadağ Yılmaz, R., Horzum, T., & Koyuncu, İ. (2024). An examination of the relationship between reading culture and mathematical literacy self-efficacy of pre-service teachers. *International Journal of Progressive Education, 20*(4), 18-41. https://doi.org/10.29329/ijpe.2024.1055.2

Kolar, V. M., & Hodnik, T. (2021). Mathematical literacy from the perspective of solving contextual problems. *European Journal of Educational Research, 10*(1), 467-483. https://doi.org/10.12973/eu-jer.10.1.467

Kozakli Ulger, T., Bozkurt, I., & Altun, M. (2022). Analyzing in-service teachers' process of mathematical literacy problem posing. *International Electronic Journal of Mathematics Education, 17*(3), Article em0687. https://doi.org/10.29333/iejme/11985

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge university press. https://doi.org/10.1017/CBO9781139087759

Muhaimin, L. H., Sholikhakh, R. A., Yulianti, S., Ardani, Hendriyanto, A., & Sahara, S. (2024). Unlocking the secrets of students' mathematical literacy to solve mathematical problems: A systematic literature review. *Eurasia Journal of Mathematics, Science and Technology Education, 20*(4), Article em2428. https://doi.org/10.29333/ejmste/14404

Neuman, S. B., & Cunningham, L. (2009). The impact of professional development and coaching on early language and literacy instructional practices. *American Educational Research Journal, 46*(2), 532-566. https://doi.org/10.3102/0002831208328088

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(4), 535-569. https://doi.org/10.1080/10705510701575396

OECD. (2013). PISA 2012 assessment and analytical framework. https://doi.org/10.1787/9789264190511-en

Özdil, G., & Kaçar, A. (2023). An investigation of the effect of a teaching experiment on the assessment literacy of pre-service mathematics teachers according to teacher competence standards. *Turkish Journal of Mathematics Education, 4*(2), 27-58.

Pastore, S. (2023). Teacher assessment literacy: A systematic review. *Frontiers in Education, 8*. https://doi.org/10.3389/feduc.2023.1217167

Project, T. J. (2025). Jamovi (Version 2.7.6). https://www.jamovi.org

Ramirez, G., Shaw, S. T., & Maloney, E. A. (2018). Math anxiety: Past research, promising interventions, and a new interpretation framework. *Educational Psychologist, 53*(3), 145-164. https://doi.org/10.1080/00461520.2018.1447384

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default bayes factors for ANOVA designs. *Journal of Mathematical Psychology, 56*(5), 356-374. https://doi.org/10.1016/j.jmp.2012.08.001

Şahin, B. (2023). *Investigating pre-service and in-service mathematics teachers' assessment literacy: Suggestions for an effective teacher education program* [Master's thesis, Middle East Technical University].

Schildkamp, K., van der Kleij, F. M., Heitink, M. C., Kippers, W. B., & Veldkamp, B. P. (2020). Formative assessment: A systematic review of critical teacher prerequisites for classroom practice. *International Journal of Educational Research, 103*, Article 101602. https://doi.org/10.1016/j.ijer.2020.101602

Shmigirilova, I. B., Rvanova, A. S., Tadzhigitov, A. A., & Kopnova, O. L. (2024). Razvitiye otsenochnoy gramotnosti budushchikh uchiteley matematiki: Kompleksnyy podkhod [Developing assessment literacy of future mathematics teachers: An integrative approach]. *The Education and Science Journal, 26*(3), 91-122. https://doi.org/10.17853/1994-5639-2024-3-91-122

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14. https://doi.org/10.3102/0013189X015002004

Spurk, D., Hirschi, A., Wang, M., Valero, D., & Kauffeld, S. (2020). Latent profile analysis: A review and "how to" guide of its application within vocational behavior research. *Journal of Vocational Behavior, 120*, Article 103445. https://doi.org/10.1016/j.jvb.2020.103445

Tomlinson, C. A. (2014). *The differentiated classroom: Responding to the needs of all learners*. Ascd.

van de Mortel, T. F. (2008). Faking it: Social desirability response bias in self-report research. *The Australian Journal of Advanced Nursing, 25*(4), 40-48.

van Doorn, J., van den Bergh, D., Bohm, U., Dablander, F., Derks, K., Draws, T., Etz, A., Evans, N. J., Gronau, Q. F., Haaf, J. M., Hinne, M., Kucharsky, S., Ly, A., Marsman, M., Matzke, D., Gupta, A., Sarafoglou, A., Stefan, A., Voelkel, J. G., & Wagenmakers, E. J. (2021). The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin & Review, 28*(3), 813-826. https://doi.org/10.3758/s13423-020-01798-5

Veldhuis, M., & van den Heuvel-Panhuizen, M. (2014). Primary school teachers' assessment profiles in mathematics education. *PLoS ONE, 9*(1), Article e86817. https://doi.org/10.1371/journal.pone.0086817

Wagenmakers, E. J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Smira, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2018). Bayesian inference for psychology (Part I: Theoretical advantages and practical ramifications). *Psychonomic Bulletin & Review, 25*(1), 35-57. https://doi.org/10.3758/s13423-017-1343-3

Wetzels, R., & Wagenmakers, E.-J. (2012). A default Bayesian hypothesis test for correlations and partial correlations. *Psychonomic Bulletin & Review, 19*(6), 1057-1064. https://doi.org/10.3758/s13423-012-0295-x

Yamtim, V., & Wongwanich, S. (2014). A study of classroom assessment literacy of primary school teachers. *Procedia - Social and Behavioral Sciences, 116*, 2998-3004. https://doi.org/10.1016/j.sbspro.2014.01.696

Yani T, A., Buyung, B., Rosmaiyadi, R., Marhayani, D. A., Nirawati, R., Triani, S. N., & Soeharto, S. (2025). Examining the relationship between mathematical literacy and digital literacy among pre-service mathematics teachers. *Eurasia Journal of Mathematics, Science and Technology Education, 21*(5), Article em2635. https://doi.org/10.29333/ejmste/16392

Yustitia, V., Amin, S. M., & Abadi. (2020). Mathematical literacy in pre-service elementary school teacher: A case study. Journal of Physics: Conference Series, 1613(1). https://doi.org/10.1088/17426596/1613/1/012054

Zhang, Y., & Cutumisu, M. (2024). Predicting the mathematics literacy of resilient students from high-performing economies: A machine learning approach. *Studies in Educational Evaluation, 83*, Article 101412. https://doi.org/10.1016/j.stueduc.2024.101412