



Design and Validation of a Measurement Tool for Teacher Skills in the Descriptive Assessment of Mathematics Lessons in Primary School

Leila Moradpoor Kerenkan¹, Yunos Javadi², Maryam Hoseini Asl Nazarlou³

1. Department of Mathematics, Urmia University, Urmia, Iran (**Corresponding author**).
L.moradpoor94@gmail.com

2. Department of statistics, Faculty of Sciences, University of Payame Noor Tehran,
Tehran, Iran. Statman.2010.89@gmail.com

3. Assistant professor of Department of Psychology and Counselling, Farhangian
University, Tehran, Iran. m.hoseini@cfu.ac.ir

Abstract

Background and Objectives: Assessment is an important aspect of educational system improvement and a key factor in enhancing classroom learning. However, one of the main challenges is that teachers lack the necessary skills to perform effective evaluation, particularly in mathematics, whose role in students' educational and vocational achievements is undeniable. Compounding this challenge, the lack of regulation, a specific approved framework, and standard tools to measure teaching skills in descriptive assessment of mathematics has caused serious challenges in Iran's educational system. Given the importance of mathematics in elementary education and the critical role of teacher competencies in effective teaching and assessing it, this study aimed to design and validate a measurement tool for assessing these skills.

Methodology: This study used a mixed-methods design employing an inductive qualitative approach, and was conducted in three phases: a systematic literature review, fuzzy Delphi process with the contribution of 18 experts, and exploratory and confirmatory factor analysis using data from 385 elementary teachers selected across the country through simple random sampling.

Findings: Ultimately, 21 skills were identified and categorized into four overarching domains: Knowledge of descriptive assessment, application of assessment tools, analysis and presentation skills, and mathematical perception. The developed questionnaire demonstrated acceptable validity and reliability indices (CVR > 0.51, CVI > 0.79, Cronbach's alpha > 0.80, CFI= 0.89).

Conclusion: The validated instrument provides a standardized measure of teacher skills in descriptive mathematical assessment and can be effectively applied in managerial, educational, and research contexts.

Keywords: Descriptive assessment, Assessment skills, Skill assessment, Mathematics assessment, Teacher competencies.

Received: 2025/09/15; **Revised:** 2025/12/22; **Accepted:** 2026/01/04; **Published online:** 2026/07/01

Citation: Moradpoor-Kerenkan, L., Javadi, Y., Hoseini-Asl-Nazarlou, M. (2026). Design and validation of a measurement tool for teacher skills in the descriptive assessment of mathematics lessons in primary school. *Research and Innovation in Primary Education*, 8(2), 197-220. <https://doi.org/10.48310/reek.2026.5029>

Publisher: Farhangian University

<https://reek.cfu.ac.ir/>

Article type: Research Article

©2026/authors retain the copyright and full publishing rights



Extended Abstract

Introduction: Teachers are considered the most important factor in curriculum implementation and therefore should possess a wide range of competencies necessary for effective curriculum design and delivery. Teaching is one of the social professions that requires various characteristics, including being a competent assessor (Ebadi et al., 2023). On the other hand, assessment is a core component of the curriculum that lies directly under the teacher's supervision and control (Razi et al., 2017). A curriculum achieves full effectiveness only when there is strong coherence among its main elements: objectives, content, teaching methods, and assessment. Assessment and evaluation have long been complementary parts of the teaching-learning process (James et al., 2002). An effective method to enhance learning during the educational process is to improve assessment procedures. Furthermore, evaluating student performance is one of a teacher's most important tasks (Estaji, 2021). Contemporary educational systems place particular emphasis on descriptive or continuous assessment and feedback to promote student learning (REL Central, 2025; Yan et al., 2023).

Descriptive assessment records students' progress through continuous feedback and precise descriptions of their performance in the classroom, in projects, and in group activities. It typically draws on observation, interviews, and analysis of student work, aiming to provide a comprehensive picture of the learning process (Ozan & Kincal, 2023; Doria et al., 2023).

Ultimately, the lack of regulations, a specific approved framework, and standardized tools for measuring teachers' descriptive assessment skills in mathematics has created serious challenges for Iran's educational system. Considering the lack of a valid instrument in this field, the present study aimed to design and validate a measurement tool for teacher skills in descriptive assessment of mathematics.

Methodology: This mixed-methods exploratory study aimed to design and validate a tool for measuring these skills. The study was conducted in 2025. Initially, an inductive qualitative approach through content analysis was employed to extract and categorize themes (Tabrizi, 2014). Subsequently, fuzzy Delphi analysis was applied, because one of its uses in qualitative research is exploratory (Ahmadi et al., 2009).

Phase one involved searching scientific resources in databases: SID, Magiran, Noormags, and Google Scholar, using relevant Persian and English keywords. Ultimately, 42 articles were identified, and 21 articles meeting the inclusion criteria were selected. The search covered publications from 2001 to 2025. Inclusion criteria required a clear methodology, topical relevance, and complete reporting of results. Phase two included screening, quality evaluation, and inductive qualitative content analysis. Skills were compared and consolidated based on semantic similarities and conceptual distinctions, resulting in 21 teacher skills in descriptive mathematics assessment. Phase three involved implementing the fuzzy Delphi technique through questionnaire development and examination of face and content validity. Questionnaires were distributed online to a sample of 18 experts in mathematics education and assessment in Iran, and their responses were collected. Experts were purposively selected based on the following minimum qualifications: a doctoral degree with at least 5 years of relevant experience, or a master's degree with at least 15 years of relevant experience. Defuzzification was performed using the center of gravity (centroid) method. The stopping criterion was defined as a difference of less than 0.1 between the defuzzified means of two consecutive rounds. This threshold was reached after the second round.

In the quantitative section, for validating the 21-item researcher-developed questionnaire, data were collected from a sample of 385 elementary teachers across the country using simple random sampling. To determine construct validity, exploratory and confirmatory factor analyses were used. Prior to conducting EFA, sampling adequacy was verified using the Kaiser-Meyer-Olkin (KMO) index ($KMO = 0.857$) and Bartlett's test of sphericity ($\chi^2 = 6063.216$, $p < .001$).

Principal component analysis with varimax rotation was applied for factor extraction. The questionnaire's factor structure was evaluated through factor analysis to examine a hypothesized structure regarding the number of factors, items, and loading patterns, testing fit with observed covariances (Calderon, 2011). Cronbach's alpha and composite reliability were used to assess reliability.

Findings: Qualitative content analysis led to the identification and categorization of these skills into four primary domains as follows.

Knowledge of descriptive assessment: (1) classroom time management; (2) mastery of the content of the mathematics textbook; (3) aligning learning objectives with content and students' depth of thinking; (4) understanding the objectives and applications of various assessment methods; (5) understanding and fulfilling legal and ethical responsibilities in assessment; (6) belief in transforming mathematics learning through high-quality assessment methods.

Application of assessment tools: (1) employing a variety of strategies and tools for formative assessment in mathematics; (2) designing engaging classroom assessments aligned with students' culture, identity, experiences, and resources; (3) differentiated assessment for diverse students; (4) developing scoring schemes and enhancing students' motivation; (5) implementing assessments based on out-of-class resources and interpreting the results.

Analysis and presentation skills: (1) analysis of classroom activities and test questions; (2) assisting students in adjusting and improving their learning using assessment information; (3) providing effective and constructive feedback on students' performance, emphasizing strengths and weaknesses; (4) explaining one's own interpretations of assessment results and reasoning about educational decisions; (5) effective interaction and collaboration with others in the school environment (including students).

Mathematical perception: (1) possessing deep and conceptual knowledge in mathematics; (2) ability to identify the type of students' errors in problem-solving and their underlying causes; (3) skill in using mathematics education tools during classroom assessment; (4) logical flexibility in exploring and pursuing different problem-solving methods; (5) ability to rank / sequence / tier tasks and mathematical problems

These domains and their constituent skills were then evaluated and prioritized through the Fuzzy Delphi technique, with expert ratings. The analysis demonstrated consensus among the panel of experts regarding the significance of these skills, reached after just two iterative rounds. Of the 385 participating teachers, 117 (30.4%) were male; mean age was 34.63 years, with an average of 8.85 years of teaching experience. Regarding educational attainment, 200 held a bachelor's degree, 161 held a master's degree, and 24 held a doctorate.

In the quantitative phase, the validity and reliability indices were satisfactory (CVR > 0.51, CVI > 0.79, Cronbach's alpha > 0.80). Confirmatory factor analysis (CFA) showed that the four-factor model exhibited fit indices slightly below conventional thresholds (CFI = 0.89 vs. the ≥ 0.90 benchmark; RMSEA = 0.096 vs. the ≤ 0.08 benchmark). Given the exploratory nature of the study, and the tendency of RMSEA to be inflated in larger samples with multiple indicators (Shi et al., 2019), these values were deemed acceptable. However, this marginal fit is acknowledged as a limitation of the current instrument. The mean scores ranged from 3.02 to 3.67 on a 5-point Likert scale, indicating a moderate level of perceived skill mastery among teachers. All factor loadings were above 0.50, showing adequate associations between the items and their respective constructs. Discriminant validity was also assessed using the Fornell-Larcker criterion, comparing the square root of AVE for each construct with its correlations with other constructs. Results confirmed adequate discriminant validity (Table 6). These findings support the construct validity and reliability of the developed instrument, suggesting that the questionnaire can serve as a reliable and standardized tool for assessing teachers' skills in descriptive mathematics assessment.

Conclusion and Recommendations: Based on statistical results, application of assessment tools showed the highest mean score (mean = 3.67, high factor loadings), and analysis and presentation skills showed the lowest mean score (mean = 3.02, lower factor loadings). Overall skill levels were found to be moderate (ranging from 3.02 to 3.67).

Knowledge of descriptive assessment explained about 20.13% of total variance, reflecting teachers' ability to recognize the objectives, characteristics, and differences from traditional assessment. Application of assessment tools accounted for about 20.03% of the variance and

The strongest dimension was the Application (mean = 3.67), and played a key role in demonstrating teachers' ability to utilize checklists, portfolios, and observation tools. Analysis and presentation skills explained about 17.50% of the variance, being the weakest (mean = 3.02). Mathematical perception accounted for about 14.42% of the variance. This newly identified dimension, comprising five sub-skills (deep mathematical knowledge, detecting student errors, using mathematics teaching tools, flexibility in reviewing solution methods, and task grading ability), may help link pedagogical content knowledge (PCK) with descriptive assessment.

This study had limitations. First, the scarcity of similar studies limited the comparison. Second, simple random sampling was used; future research should employ stratified sampling proportional to provincial teacher populations. Third, the tool's direct impact on student learning was not examined. Although the results demonstrated acceptable validity and reliability, assessing practical effectiveness requires further investigation. It is recommended that future research investigate the relationship between teacher skills (measured by this tool) and student academic progress in mathematics to provide further empirical evidence on its role in enhancing learning.

Ethical Considerations: This study was carried out with full respect for the rights, dignity, and privacy of all participants. Informed consent was obtained prior to data collection, and participation was entirely voluntary. Participants were free to withdraw from the study at any stage without any consequences. All responses were collected anonymously, and confidentiality of participants' information was strictly maintained throughout the research process. This study was deemed exempt from full ethical review by the Farhangian University (Azarbaijan Gharbi Branch) because it relied solely on anonymous questionnaire data and involved no collection of identifiable or sensitive personal information.

Role of Each Author: The first author drafted the manuscript. The second author performed the statistical analysis. The third author provided consultation. All authors reviewed and approved the final version.

Conflict of Interest: The authors declare that there is no conflict of interest.

طراحی و اعتباریابی ابزار سنجش مهارت‌های معلمی در ارزشیابی توصیفی درس ریاضی دوره ابتدایی

لیلا مرادپور کرنگان^۱، یونس جوادی^۲، مریم حسینی اصل نظرلو^۳

۱. کارشناسی ارشد ریاضی متروید، گروه ریاضی، دانشگاه ارومیه، ارومیه، ایران (نویسنده مسئول).

L.moradpoor94@gmail.com

۲. کارشناسی ارشد آمار ریاضی، گروه آمار، دانشکده علوم، دانشگاه پیام نور تهران، تهران، ایران.

Statman.2010.89@gmail.com

۳. استادیار گروه آموزش روان‌شناسی و مشاوره، دانشگاه فرهنگیان، تهران، ایران. m.hoseini@cfu.ac.ir

چکیده

زمینه و هدف: سنجش و ارزشیابی جنبه‌ای مهم در ارتقای نظام آموزشی است. ارزشیابی یکی از عوامل مؤثر در بهبود یادگیری در کلاس است. با این حال، یکی از چالش‌های کلیدی این است که معلمان به‌طور کامل مهارت‌های لازم برای اجرای مؤثر ارزشیابی را دارا نمی‌باشند. نقش عملکرد ریاضی در موفقیت تحصیلی و شغلی دانش‌آموزان غیرقابل‌انکار است. اما فقدان آیین‌نامه و چارچوب مصوب اختصاصی، همراه با ابزار استاندارد برای سنجش مهارت‌های معلمی در ارزشیابی توصیفی درس ریاضی، زیست‌بوم آموزشی ایران را با چالش جدی مواجه کرده است. با توجه به اهمیت درس ریاضی در دوره ابتدایی و نقش کلیدی مهارت‌های معلم در آموزش و ارزشیابی صحیح این درس، پژوهش حاضر با هدف طراحی و اعتباریابی ابزاری استاندارد برای سنجش مهارت‌های معلمی در ارزشیابی توصیفی ریاضی انجام شد.

روش: این مطالعه با رویکرد آمیخته و مبتنی بر روش کیفی استقرائی در سه مرحله اجرا گردید: مرور نظام‌مند متون، اجرای دلفی فازی با مشارکت ۱۸ خبره و تحلیل عاملی اکتشافی و تأییدی بر داده‌های حاصل از ۳۸۵ آموزگار ابتدایی منتخب از سراسر کشور با روش نمونه‌گیری تصادفی ساده. **یافته‌ها:** در نهایت، ۲۱ مهارت در قالب چهار مهارت کلی (آشنایی با ارزشیابی توصیفی، استفاده از ابزارهای ارزشیابی، مهارت تحلیل و ارائه، و ادراک ریاضی) شناسایی شد. پرسش‌نامه طراحی شده نیز از شاخص‌های روایی و پایایی مطلوب برخوردار بود ($CVR > 0.51$ - $CVI > 0.79$ - $CFI = 0.89$ - $\alpha = 0.80$).

نتیجه‌گیری: ابزار استاندارد تدوین شده می‌تواند برای سنجش مهارت‌های معلمی در ارزشیابی توصیفی ریاضی مورد استفاده قرار گیرد و زمینه بهره‌برداری از آن در اهداف مدیریتی، آموزشی و پژوهشی فراهم است.

کلیدواژه‌ها: ارزشیابی توصیفی، مهارت‌های ارزشیابی، سنجش مهارت، ارزشیابی ریاضی، مهارت معلمی.

تاریخ دریافت: ۱۴۰۴/۰۹/۱۵؛ تاریخ اصلاح: ۱۴۰۵/۱۰/۰۱؛ تاریخ پذیرش: ۱۴۰۵/۱۰/۱۴؛ تاریخ انتشار: ۱۴۰۵/۰۴/۱۰

استناد به این مقاله: مرادپور کرنگان، لیلا؛ جوادی، یونس؛ حسینی اصل نظرلو، مریم. (۱۴۰۴). طراحی و اعتباریابی ابزار سنجش مهارت‌های معلمی در ارزشیابی توصیفی درس ریاضی دوره ابتدایی. پژوهش و نوآوری در آموزش ابتدایی، ۸(۲)، ۱۹۷-۲۲۰.

<https://doi.org/10.48310/reek.2026.5029>

نوع مقاله: پژوهشی

<https://reek.cfu.ac.ir/>

ناشر: دانشگاه فرهنگیان

© ۱۴۰۵ / نویسندگان دارنده حق مؤلف مقاله خود بدون محدودیت هستند.



مقدمه

معلم مهم‌ترین شخص در فرایند اجرای برنامه درسی قلمداد می‌شود (Alsubaie, 2016). معلم در جایگاه عنصری کلیدی در طراحی و اجرای برنامه درسی باید از صلاحیت‌های متنوعی برخوردار باشد. معلمی در زمره مشاغل اجتماعی است که فرد برای احراز شرایط ورود به آن به دامنه‌ای وسیع از ویژگی‌ها مانند ارزیاب خوب بودن نیاز دارد (Ebadi et al., 2023). از طرفی، ارزشیابی از جمله مؤلفه‌های برنامه درسی است که مستقیم تحت نظر و کنترل معلم است (Razi et al., 2017). برنامه درسی زمانی کارآمد است که هماهنگی و انسجام بین عناصر اصلی برنامه درسی شامل هدف، محتوا، روش تدریس و ارزشیابی وجود داشته باشد. ارزشیابی از دیرباز بخش مکمل فرایند تدریس-یادگیری بوده است (James et al., 2002). سنجش و ارزشیابی، اهرم اصلاح نظام آموزشی است. هدف آزمون، اطلاع از وضعیت موجود، برنامه‌ریزی آینده و رفع نقایص است (Amini, 2018). ارزشیابی کلاسی نیز یکی از عوامل مؤثر در بهبود یادگیری در کلاس درس محسوب می‌شود (Makvandi, 2021).

یکی از روش‌های مؤثر تقویت یادگیری در دوران تحصیل بهبود رویه‌های ارزشیابی است و ارزشیابی عملکرد دانش‌آموزان یکی از ضروری‌ترین وظایف هر معلم محسوب می‌شود. معلم در این فرایند اطلاعات را جمع‌آوری، تحلیل و تفسیر کرده و میزان تحقق اهداف آموزشی را مورد بررسی قرار می‌دهد.

ارزشیابی برای تعیین و سنجش درک دانش‌آموزان از مفاهیم، اصلاحات و پیشرفت یادگیری استفاده می‌شود. همچنین، به سنجش سطح کسب مهارت‌های یادگیری دانش‌آموزان کمک می‌کند (Dandis, 2013; Simon et al., 2017). ارزشیابی درک دانش‌آموزان دیدگاه بهتری از آنچه دانش‌آموزان می‌دانند ارائه می‌دهد (Unal & Unal, 2019). در نظام‌های آموزشی معاصر تأکید ویژه‌ای بر ارزشیابی توصیفی یا ارزیابی مستمر و بازخوردی برای ارتقای یادگیری دانش‌آموزان مشاهده می‌شود (REL Central,¹ 2025; Yan et al., 2023). ارزشیابی توصیفی عبارت است از روشی که در آن پیشرفت دانش‌آموزان از طریق بازخوردهای مستمر و توصیف‌های دقیق از جنبه‌های مختلف عملکرد آن‌ها در کلاس، پروژه‌ها و فعالیت‌های گروهی ثبت می‌شود. این روش می‌تواند شامل مشاهده، مصاحبه و تحلیل کارهای دانش‌آموز باشد و هدف آن فراهم آوردن تصویری جامع از روند یادگیری است (Ozan & Kincal, 2023; Doria et al., 2023). با این حال، یکی از چالش‌های کلیدی این است که معلمان به‌طور کامل مهارت‌های لازم برای اجرای مؤثر این نوع ارزشیابی را در اختیار ندارند (Ghasemi & Hashemi, 2019; Estaji & Fassihi, 2016). از سوی دیگر، برای حمایت از توسعه حرفه‌ای معلمان، نیاز به ابزاری معتبر و استاندارد برای سنجش «مهارت‌های معلم» در فرایند ارزیابی توصیفی درس احساس می‌شود (Yan et al., 2023; Naghdipour, 2016). در واقع، اگر مهارت‌ها و رفتار معلمان در ارزشیابی توصیفی به‌صورت دقیق اندازه‌گیری نشود، نمی‌توان به وضعیت واقعی نیازها و یا اثربخشی برنامه‌های آموزشی پی برد (Miri & Miri, 2022; Alavi & Dashtestani, 2015).

در ایران ضرورت‌های مختلفی باعث شده است تا ارزشیابی کیفی توصیفی به‌عنوان جایگزین ارزشیابی سنتی پیشرفت تحصیلی از سال تحصیلی ۸۲-۸۱ به‌صورت آزمایشی معرفی و اجرا گردد (Karamalian et al., 2013). تأکید نظام ارزشیابی توصیفی بر تغییر مقیاس کمی (۲۰-۰) به مقیاس کیفی (ارزشیابی توصیفی)، ارزشیابی پایانی به ارزشیابی تکوینی، و استفاده از کارنامه توصیفی و ابزارهایی برای سنجش و ارزشیابی پیشرفت تحصیلی دانش‌آموزان متناسب با مقیاس موردنظر است (Makvandi, 2021).

یکی از مهم‌ترین مشکلات ارزشیابی توصیفی نبود معلمان کیفی و عدم آشنایی معلمان و مدیران با اهداف و ویژگی‌ها و ابزارهای ارزشیابی توصیفی می‌باشد (Karamalian et al., 2013). در ارزشیابی توصیفی انتظار این است که یادگیری بهبود، دوام و پایداری داشته باشد، علاقه به یادگیری ایجاد شود، نسبت به مدرسه نگرش مطلوب ایجاد شود، و مشارکت دانش‌آموزان و والدین در فرایند ارزشیابی تحصیلی افزایش یابد (Shamsolahi & Samie Zafarghandi, 2017). در این نوع ارزشیابی، معلم با هدف یادگیری بهتر

و ماندگارتر زمینه مناسب‌تری را برای جریان یادگیری دانش‌آموزان فراهم می‌آورد؛ بنابراین در اجرای این نوع ارزشیابی نقش معلم تغییر یافته است و معلم، با ارائه بازخوردها و با مهارت‌های حرفه‌ای، دانش‌آموزان را به هدف می‌رساند (Shamsolahi & Samie, 2017). سرکارآرانی و مقدم (به نقل از Shamsolahi & Samie Zafarghandi, 2017) معتقدند تخصص، توانایی، انگیزه و شایستگی‌های معلم مهم‌ترین عناصر بهسازی مداوم آموزش هستند. نقش صلاحیت‌های حرفه‌ای معلمان در پیشرفت تحصیلی دانش‌آموزان به منظور فراهم کردن فرصت‌های بهتر آموزشی اهمیت بسیار دارد (López-Martín et al., 2023). از جمله مهارت‌های حرفه‌ای معلمان می‌توان به «سواد ارزشیابی» و «دانش و مهارت‌های ارزشیابی برای معلمان» اشاره کرد. این دو اصطلاح ایده‌های مشابهی به نظر می‌رسند. با این حال، تفاوت‌هایی بین این دو عبارت وجود دارد. «سواد ارزشیابی یک عمل اجتماعی پویا و وابسته به زمینه است که شامل بیان و مذاکره معلمان در مورد دانش کلاسی و فرهنگی با یکدیگر و با دانش‌آموزان، در شروع، توسعه و تمرین ارزشیابی، برای دستیابی به اهداف یادگیری دانش‌آموزان می‌شود» (Willis et al., 2013).

Pastore و Andrade (2019) سواد ارزشیابی را بر اساس سه بُعد تعریف کردند: دانش (بُعد مفهومی)، مهارت‌ها (بُعد رفتاری) و گرایش‌ها (بُعد اجتماعی-عاطفی). نویسندگان پیشین سواد ارزشیابی را به مدلی از هویت ارزشیابی معلم تبیین نموده‌اند که می‌تواند سلسله‌مراتبی باشد (Xu & Brown, 2016) یا غیرسلسله‌مراتبی (Looney et al., 2018). این مدل شامل دانش معلمان، باورهای آن‌ها در مورد یادگیری و ارزشیابی، احساسات و اعتماد به نفس آن‌ها در مورد ارزشیابی، درک زمینه و سازش در ارزشیابی، و نقش یا هویت خود آن‌ها به عنوان یک ارزشیاب در زمینه آموزشی می‌شود.

قانون «هیچ کودکی جا نماند»^۱ در سال ۲۰۰۱، که به صورت یک قانون مصوب در سال ۲۰۰۲ درآمد، مدارس دولتی را ملزم به اجرای نظام‌های مسئولیت‌پذیری می‌نمود (Vignoles & Meschi, 2010). در این راستا، معلمان از استانداردهای معتبر برای آموزش و ارزشیابی نتایج استفاده می‌کنند.

در واقع، معلم باید بتواند با تغییر در طرز تفکر و روش کار خود در زمینه ارزشیابی تحصیلی، آن را در خدمت بهبود یادگیری دانش‌آموزان قرار دهد (Makvandi, 2021). به کارگیری مناسب این فعالیت‌ها نیازمند دانش و توانایی قوی معلم در سنجش ارزشیابی است که به این توانایی و دانش «سواد ارزشیابی» اطلاق می‌شود: «سواد ارزشیابی به معنای توانایی درک و استفاده مؤثر از داده‌های ارزشیابی، شناسایی معیارها و اهداف مرتبط با ارزشیابی، و استفاده از اطلاعات جمع‌آوری شده برای تصمیم‌گیری آگاهانه در فرایندهای آموزشی است.» (Parsaeian et al., 2023A; Parsaeian et al., 2023B) در واقع مفهوم سواد ارزشیابی گسترده‌تر است و هدف وسیع‌تری را دنبال می‌کند، که معمولاً به عنوان تلاشی برای درک و برآورده کردن نیازهای بالقوه معلمان در حین انجام ارزشیابی توصیف می‌شود و شامل شایستگی‌های ارزشیابی و عوامل شخصی و زمینه‌ای پیرامون آن‌ها می‌شود که بر عملکرد ارزشیابی معلمان تأثیر می‌گذارند (Brookhart, 2024).

انجمن ارزشیابی میشیگان (Michigan Assessment Consortium, 2020) مجموعه‌ای مفصل از دانش و مهارت‌های دوره‌ای را برای معلمان و سایر نقش‌ها (دانش‌آموزان و خانواده‌ها، مدیران مدرسه، مدیران منطقه و سیاست‌گذاران) با هدف توسعه حرفه‌ای تهیه کرده است که می‌تواند منافع را به حداکثر و پیامدهای منفی ارزشیابی دانش‌آموز را به حداقل برساند. اما آنچه معلمان باید در مورد ارزشیابی بدانند به عوامل مختلفی از قبیل نمره، آموزش در حوزه محتواهای مختلف، نظام‌های ارزشیابی مورد استفاده در منطقه و الزامات منطقه‌ای و دولتی بستگی دارد. فهرست معیارهای توانش معلم در ارزشیابی آموزشی دانش‌آموزان (American Federation of Teachers, NCME, & NEA, 1990) اقدام مشترک انجمن آمریکایی معلمان، شورای ملی سنجش آموزش و انجمن ملی آموزش بود. این سند با عنوان «استانداردهای شایستگی مدرس در ارزیابی آموزشی دانش‌آموزان» شامل هفت شایستگی بود (Hoseini et al., 2025). فهرست معیارهای

1. No Child Left Behind (NCLB)

مذکور، که به صورت چارچوبی مفهومی برای معلمان درآمد، شامل موارد ذیل است: انتخاب روش های ارزشیابی متناسب با تصمیمات آموزشی؛ طراحی روش های ارزشیابی مناسب با تصمیمات آموزشی؛ اجرا، نمره دهی و تفسیر نتایج روش های ارزشیابی تدوین شده معلم و بیرونی؛ بهره گیری از نتایج ارزشیابی هنگام تصمیم گیری در مورد دانش آموزان، برنامه ریزی آموزش، طراحی برنامه درسی و ارتقای مدرسه؛ طراحی رویه های نمره دهی معتبر که از ارزشیابی های دانش آموزان بهره می گیرند؛ انتقال نتایج ارزشیابی به دانش آموزان، والدین و سایرین؛ تشخیص روش های ارزشیابی غیر اخلاقی، غیر قانونی و هر نوع ارزشیابی ناصحیح دیگر و استفاده از اطلاعات ارزشیابی (Mertler, 2009).

چنین معیارهایی به صراحت حوزه هایی را معرفی می کنند که معلمان در آن ها باید تسلط خیلی خوبی داشته باشند تا بتوانند یادگیری دانش آموزان را ارزشیابی کنند. این معیارها اهمیت آموزش معلم و رشد حرفه ای وی در عرصه ارزشیابی کلاسی را تأیید و تعیین می کنند. این هفت معیار همگی برای طراحی و استفاده از ارزشیابی های کلاسی، اهداف و مقاصد آموزشی توسط معلم به کار می روند (Brookhart, 2001). Brookhart (2024)، ضمن ارائه فهرستی از دانش و مهارت های ارزشیابی معلمان، معتقد است که استفاده مؤثر از هر فهرستی از دانش و مهارت های ارزشیابی، با کمک به معلمان برای ایجاد ارتباط بین دانش و مهارت های ارزشیابی و دانش و مهارت های مربوط به روان شناسی آموزشی، محتوای موضوعی، روش های آموزشی و مدیریت کلاس، منجر به رشد آن ها می شود.

نقش مستقیم و غیرمستقیم عملکرد ریاضی در موفقیت تحصیلی و شغلی دانش آموزان غیرقابل انکار است. از این رو، توجه ویژه به عوامل مؤثر بر آموزش و ارزشیابی این درس حیاتی — که در مقاطع ابتدایی، راهنمایی و آموزش عالی به عنوان پایه ای برای توسعه مهارت های شناختی شناخته می شود — (Rahimi & Sedaghatkhan, 2022; Nabavizadeh et al., 2019) ضروری است. با وجود اهمیت بالای ریاضی و تبیین راهنمای ارزشیابی توصیفی توسط وزارت آموزش و پرورش، مهارت های ارزشیابی توصیفی معلمان به ویژه در این درس تاکنون مورد بررسی علمی قرار نگرفته، هیچ معیار مشخصی برای سنجش آن تعریف نشده و ابزاری معتبر برای ارزیابی این مهارت ها وجود ندارد.

عامل «ادراک ریاضی» در این پژوهش، بازتعریف عملیاتی PCK^۱ (دانش محتوایی-تربیتی؛ Shulman, 1986) در چارچوب ارزشیابی توصیفی است: توانایی معلم در تشخیص و تفسیر درک مفهومی دانش آموز از مفاهیم ریاضی (مانند کسر، هندسه) از طریق ابزارهای توصیفی مانند چک لیست، مشاهده، پوشه کار و تحلیل عملکرد. این بُعد فراتر از PCK تدریس محور است و با مدل سه بعدی سواد ارزشیابی (دانش، مهارت، گرایش؛ Pastore & Andrade, 2019) تلفیق شده است. برخلاف ابزارهای جهانی که PCK را عمدتاً در تدریس بررسی کرده اند (Gallagher et al., 2025)، این بازتعریف اولین تلاش برای عملیاتی سازی PCK در ارزشیابی توصیفی ریاضی است و خلأ موجود در مطالعات ایرانی (van den Heuvel-Panhuizen et al., 2021) و بین المللی (Shaughnessy et al., 2021) را پر می کند.

در حالی که ارزشیابی صحیح یکی از اصلی ترین دغدغه های معلمان (Jazairi, 2017)، و ارزشیابی واقعی تر عملکرد دانش آموزان از نگرانی های کلیدی معلمان ریاضی می باشد (van den Heuvel-Panhuizen et al., 2021)، این خلأ به رفتارهای سلیقه ای و ناهماهنگ در ارزشیابی منجر شده است. در نهایت، فقدان آیین نامه و چارچوب مصوب اختصاصی، همراه با ابزار استاندارد برای سنجش مهارت های معلمی در ارزشیابی توصیفی درس ریاضی، زیست بوم آموزشی ایران را با چالش جدی مواجه کرده است.

تا حدودی ضرورت تدوین معیار برای اجرای موفقیت آمیز ارزشیابی توصیفی درس ریاضی دوره ابتدایی توسط مرتضوی و همکاران (Mortazavi et al., 2020A) تبیین شده است. در این ارتباط، ارزیابی وضعیت کنونی معلمان از منظر مهارت های ارزشیابی توصیفی برای درس ریاضی اجتناب ناپذیر می باشد، که این امر مستلزم به کارگیری ابزاری جهت سنجش این مهارت ها است تا با بررسی نقاط قوت و ضعف و فرصت ها و تهدیدهای موجود، به اصلاح امور جهت بهبود وضعیت و نهادینه سازی آموزش و ارزشیابی بپردازد؛ و، با توجه به عدم وجود ابزاری معتبر (دارای روایی و پایایی) در این زمینه، پژوهش حاضر با هدف طراحی و اعتباریابی ابزار سنجش مهارت های معلمی در ارزشیابی توصیفی درس ریاضی انجام شده است.

روش

این پژوهش آمیخته (کیفی و کمی) از نوع اکتشافی با هدف طراحی و اعتباریابی ابزار سنجش مهارت‌های معلمی در ارزشیابی توصیفی برای درس ریاضی در سال ۱۴۰۴ در دو مرحله (طراحی و اعتباریابی) انجام شده است. در ابتدا با رویکرد کیفی استقرائی با تکیه بر تحلیل محتوا مضامین استخراج و دسته بندی می‌گردد (Tabrizi, 2014) و در ادامه از تحلیل دلفی فازی استفاده می‌شود. یکی از کاربردهای تکنیک دلفی فازی در پژوهش کیفی جنبه اکتشافی آن است. در واقع، پژوهشگران به دنبال شناسایی مهم‌ترین عناصر زیربنایی یک پدیده هستند (احمدی و همکاران، ۱۳۸۷). پژوهش حاضر از نوع کیفی استقرائی بود و در سه مرحله متوالی اجرا شد.

مرحله اول: جستجوی منابع علمی:

بانک‌های اطلاعاتی جهاد دانشگاهی، Google Scholar، Noormags و Magiran جست‌وجو شد. کلیدواژه‌های فارسی شامل «ارزشیابی»، «ارزشیابی کیفی»، «مهارت‌های معلمی»، «مهارت ارزشیابی»، «ارزشیابی ریاضی»، «مهارت ارزشیابی ریاضی»، و کلیدواژه‌های انگلیسی شامل qualitative assessment, descriptive assessment, teacher assessment skills, mathematics assessment بوده است. جست‌وجو در بازه زمانی ۱۳۸۰ تا ۱۴۰۴ و در بین مقالات پژوهشی و گزارش‌های معتبر به زبان فارسی و انگلیسی با تمرکز بر مهارت‌های معلمی در ارزشیابی (مستقیم یا غیرمستقیم) انجام شد و نهایتاً ۴۲ مقاله شناسایی و دانلود گردید.

مرحله دوم: غربالگری، ارزیابی کیفیت و تحلیل محتوای کیفی

گام ۱ - حذف تکراری: مقالات تکراری حذف شد ← ۴۱ مقاله باقی ماند.

گام ۲ - غربال عنوان و چکیده: مقالات غیرمرتبط با مهارت معلمی یا ارزشیابی کیفی ریاضی کنار گذاشته شد ← ۲۵ مقاله باقی ماند.

گام ۳ - مطالعه: متن کامل ۲۵ مقاله توسط دو نویسنده مطالعه شد.

گام ۴ - ارزیابی کیفیت:

در این مرحله، برای ارزیابی کیفیت و تحلیل محتوای کیفی، معیار پذیرش وجود روش‌شناسی واضح، مرتبط بودن با موضوع، و گزارش کامل نتایج.

نتیجه: ۲۱ مقاله معتبر انتخاب شد.

گام ۵ - تحلیل محتوای کیفی استقرائی:

با تکیه بر تحلیل محتوای مضامین، مهارت‌های ذکرشده در ۲۱ مقاله استخراج و دسته‌بندی گردید.

مهارت‌ها بر اساس شباهت معنایی و تفاوت مفهومی مقایسه و ادغام شد.

نتیجه نهایی: ۲۱ مهارت معلمی در ارزشیابی کیفی درس ریاضی به دست آمد.

مرحله سوم: اجرای دلفی فازی

گام ۱ - تدوین پرسش‌نامه و بررسی روایی صوری و محتوایی

گام ۲ - جامعه و نمونه‌گیری: در جامعه متخصصان آموزش ریاضی و ارزشیابی در ایران با نمونه‌گیری هدفمند شامل ۱۸ نفر از متخصصان که حداقل دارای مدرک تحصیلی دکتری با ۵ سال سابقه و فوق‌لیسانس با ۱۵ سال سابقه انجام شد.

گام ۳ - توزیع پرسش‌نامه‌ها به صورت آنلاین و جمع‌آوری داده‌ها

گام ۴ - تحلیل دور اول دلفی فازی

گام ۵ - اجرای دور دوم و نتیجه اجماع کامل

گام ۶ - اجماع نهایی و اولویت‌بندی

معیار توقف: اختلاف میانگین فازی‌زدایی دو دور > 0.1 ← توقف در دور دوم

اولویت‌بندی: بر اساس میانگین فازی‌زدایی دور دوم

در بخش کمی و جهت اعتباریابی پرسش‌نامه ۲۱ سؤالی محقق‌ساخته (مهارت‌های معلمی ارزشیابی درس ریاضی) حاصل از یافته‌های مرحله کیفی، از جمع‌آوری داده از نمونه‌ای شامل ۳۸۵ نفر — که با استفاده از فرمول کوکران به دست آمد — از آموزگاران مقطع ابتدایی سراسر کشور به روش تصادفی ساده استفاده شده است. جهت اطمینان از اعتبار یافته‌های حاصل از مرحله اول، از روایی صوری و محتوایی استفاده گردید. برای ارزیابی روایی صوری به این سؤال پاسخ داده می‌شود که «آیا ظاهر ابزار برای ارزیابی هدف مورد نظر به صورت مناسب طراحی شده است؟» برای تعیین روایی صوری از دو روش کیفی و کمی استفاده شد. در مرحله کیفی روایی صوری پانلی با حضور خبرگان متخصص و باتجربه در حوزه ارزشیابی و مهارت‌های معلمی تشکیل، و سطح دشواری، میزان تناسب و ابهام بررسی گردید و تغییراتی بر اساس نظر آنان در پرسش‌نامه اعمال شد. در گام بعدی جهت کاهش عبارت‌ها، حذف عبارت‌های نامناسب و تعیین اهمیت هر یک از عبارت‌ها انجام شد، و در مرحله کمی از روش تأثیر گویند^۱ استفاده شد. در این روش برای هر یک از گویندگان ابزار، طیف لیکرت پنج قسمتی، شامل «کاملاً مهم است» (امتیاز ۵)، «تا اندازه‌ای مهم است» (امتیاز ۴)، «به‌طور متوسطی مهم است» (امتیاز ۳)، «اندکی مهم است» (امتیاز ۲)، و «اصلاً مهم نیست» (امتیاز ۱)، در نظر گرفته شد. در صورتی که نمره تأثیر مساوی و یا بیشتر از ۱.۵ باشد، عبارت برای تحلیل‌های بعدی مناسب تشخیص داده می‌شود و حفظ می‌گردد (حاجی‌زاده و اصغری، ۱۳۹۰). در پژوهش حاضر شاخص امتیاز تأثیر گویند با استفاده از فرمول زیر محاسبه گردید:

امتیاز روش تأثیر گویند = درصد شرکت‌کنندگانی که آیتم را بررسی کرده اند × متوسط پاسخ کل شرکت‌کنندگان (۱)

برای تعیین روایی محتوا نیز از دو روش کیفی و کمی استفاده شد. در این مطالعه تعیین روایی محتوا مبتنی بر قضاوت افراد متخصص می‌باشد. در روش کیفی تعیین روایی محتوا، گوینده‌ها توسط خبرگان مورد بازنگری قرار گرفتند و اصلاحات ضروری در آن‌ها اعمال شد. برای تعیین روایی محتوا به روش کمی، دو شاخص نسبت روایی محتوا^۲ و شاخص روایی محتوا^۳ مورد استفاده قرار گرفت. برای تعیین نسبت روایی محتوا از خبرگان متخصص درخواست شد تا هر گویند را براساس طیف سه‌قسمتی «ضروری است»، «مفید است ولی ضروری نیست» و «ضرورتی ندارد» بررسی نمایند. بر اساس جدول ۱۹۷۵ لاوشه^۴ حداقل ارزش نسبت روایی محتوا تعیین گردید و نسبت روایی محتوا برای تک‌تک گوینده‌ها با استفاده از فرمول زیر محاسبه شد (Vakili et al., 2012):

$$(2) \quad CVR = \frac{(2 \div \text{تعداد کل شرکت کنندگان}) - \text{تعداد ضروری پاسخ های هر گویند}}{2 \div \text{تعداد کل شرکت کنندگان}}$$

روایی محتوایی بر اساس شاخص روایی محتوایی انجام شد (Sarmad et al., 2019). بدین منظور، پرسش‌نامه طراحی شده در اختیار ۱۴ نفر از خبرگان متخصص قرار داده شد تا میزان مربوط بودن، ساده بودن و واضح بودن هر یک از عبارات موجود در پرسش‌نامه را تعیین نمایند. بدین ترتیب، سه معیار ساده بودن، مربوط بودن و واضح بودن به صورت مجزا در طیف لیکرت چهارقسمتی برای هر یک از گوینده‌ها مورد بررسی قرار گرفت. در این تحقیق، امتیاز شاخص روایی محتوا برای هر گویند به وسیله تقسیم تعداد متخصصان موافق با عبارت دارای رتبه ۳ و ۴ بر تعداد کل متخصصان و به صورت زیر محاسبه شد (Vakili et al., 2012):

$$(3) \quad CVI = \frac{\text{مجموع امتیاز موافق برای هر گویند با رتبه 3 و 4}}{\text{تعداد کل پاسخها}}$$

1. item impact
2. content validity ratio (CVR)
3. content validity index (CVI)
4. Lawshe

پس از اعمال تغییرات لازم، جهت تعیین روایی سازه گویه‌ها، از روش تحلیل عاملی اکتشافی و تأییدی استفاده شد. بدین منظور، پرسش‌نامه‌ها در اختیار نمونه‌ها قرار گرفت. ساختار عاملی پرسش‌نامه از طریق تحلیل عاملی مورد ارزیابی قرار گرفت. هدف پژوهشگر در تحلیل عاملی بررسی ساختار عاملی ویژه‌ای است که درباره تعداد عامل‌ها، تعداد سؤالات و الگوی قرار گرفتن سؤالات در هر عامل به‌طور آشکار فرضیه‌ای بیان شده و برازش ساختار عاملی موردنظر در فرضیه با ساختار کوواریانس‌های اندازه‌گیری شده آزمون می‌گردد (Calderon, 2011). جهت تعیین پایایی پرسش‌نامه از آزمون‌های آلفای کرونباخ و پایایی ترکیبی استفاده گردید. در این پژوهش، بخش کمی به کمک نرم‌افزارهای Smart PLS4 و SPSS27 انجام شده است.

یافته‌ها

یافته‌های کیفی

فراوانی خبرگانی که در بررسی پرسش‌نامه دلفی فازی شرکت نمودند ۴ نفر مرد و ۱۱ نفر زن بود، شامل ۱۲ نفر متأهل و ۳ نفر مجرد. متوسط سن ایشان ۴۰ سال و متوسط سابقه خدمت ۱۶ سال بوده است. سطح تحصیلات دکتری و بالاتر دارای فراوانی ۶ نفر و فوق‌لیسانس دارای ۹ نفر بوده است.

یافته‌های حاصل از بخش کیفی با ترکیب پژوهش مرادپور کرنکان (۱۴۰۳) و کدهای استخراج‌شده از Sahito et al., 2024; Hill et al., 2008; Schilling & Hill, 2007; An & Wu, 2012 توسعه یافته، و پرسش‌نامه نهایی شامل ۲۱ مهارت کلی برای معلمان در ارزشیابی درس ریاضی در قالب ۴ بُعد به دست آمد. نتایج اجرای روش دلفی فازی در دو مرحله در جدول ۱ ارائه شده است. فرایند دفازی‌سازی با روش مرکز ثقل^۱ انجام شد.

$$(۴) \quad S_j = \frac{L_j + 2 \times M_j + U_j}{4}$$

در دلفی فازی پایایی بین ارزیابان الزامی نیست، زیرا خبرگان ناشناس و مستقل هستند و هدف کاهش اختلاف میانگین‌های دور اول و دوم و به تبع آن کاهش واریانس می‌باشد که نشانه همسویی نظر خبرگان و اجماع نظر ایشان می‌باشد (Hsu & Chen, 1996; Okoli & Pawlowski, 2004).

جدول ۱. اولویت‌بندی مهارت‌های معلمی ارزشیابی درس ریاضی و ابعاد آن

ردیف	بُعد	مهارت	میانگین فازی‌زدایی‌شده	رتبه
۱	آشنایی با ارزشیابی توصیفی	مدیریت زمان کلاسی	۰.۶۰۵	۱۶
۲		یادگیری محتوای کتاب ریاضی	۰.۷۱۲۵	۸
۳		انطباق اهداف یادگیری با محتوا و عمق تفکر دانش‌آموزان	۰.۶۹	۱۰
۴		درک اهداف و کاربردهای روش‌های مختلف ارزشیابی	۰.۷۲۷۵	۶
۵		درک و انجام مسئولیت‌های قانونی و اخلاقی خود در ارزشیابی	۰.۷۲۷۵	۶
۶		باور به ایجاد تغییر در یادگیری ریاضی با تکیه بر شیوه‌های ارزشیابی با کیفیت	۰.۷۵	۳

1. center of gravity

۷	استفاده از مجموعه‌ای از راهبردها و ابزارها برای ارزشیابی تکوینی درس ریاضی	۰.۷۲۵	۷
۸	ایجاد ارزشیابی‌های کلاسی جذاب مرتبط با فرهنگ، هویت، تجارب و منابع دانش‌آموزان	۰.۷۴۷۵	۸
۹	تمایز ارزشیابی برای دانش‌آموزان مختلف	۰.۷۵	۹
۱۰	ساخت طرح‌های امتیازدهی و تقویت انگیزش دانش‌آموزان	۰.۶۶	۱۰
۱۱	اجرای ارزشیابی‌های مبتنی بر منابع خارج از کلاس و تفسیر نتایج	۰.۶۷۷۵	۱۱
۱۲	تجزیه و تحلیل سؤالات کلاسی و سؤالات آزمون	۰.۶۴۷۵	۱۲
۱۳	کمک به دانش‌آموزان برای تنظیم یادگیری با استفاده از اطلاعات ارزشیابی	۰.۷۷۵	۱۳
۱۴	ارائه بازخورد مؤثر و مفید در مورد کار دانش‌آموز با تأکید بر نقاط ضعف و قوت	۰.۶۲	۱۴
۱۵	تبیین تفاسیر خود از نتایج ارزشیابی و استدلال در مورد تصمیمات آموزشی	۰.۷۵	۱۵
۱۶	تعامل و کار مؤثر با دیگران در مدرسه (از جمله دانش‌آموزان)	۰.۷۳۲۵	۱۶
۱۷	داشتن دانش عمیق در ریاضی	۰.۶۰۵	۱۷
۱۸	توانایی تشخیص نوع خطای دانش‌آموزان در حل مسائل و دلایل آن	۰.۷۰۲۵	۱۸
۱۹	مهارت استفاده از ابزارهای آموزش ریاضی در ارزشیابی کلاسی	۰.۸۰۲۵	۱۹
۲۰	انعطاف‌پذیری منطقی در بررسی روش‌های مختلف حل مسئله	۰.۶۱۵	۲۰
۲۱	توانایی درجه‌بندی تکالیف و مسائل ریاضی	۰.۷۳۲۵	۲۱

با توجه به اینکه اختلاف میانگین فازی‌زدایی شده نظر خبرگان در دو مرحله کمتر از ۰.۱ می‌باشد خبرگان در مورد ۲۱ مهارت‌های معلمی جهت ارزشیابی درس ریاضی به اجماع رسیده‌اند و نظرسنجی در این مرحله متوقف گردید. این بدان معنی است که خبرگان به مهارت‌های معلمی ارزشیابی درس ریاضی شناسایی شده در پژوهش نگاه تقریباً یکسانی داشته‌اند.

یافته های کمی

این گویه‌ها به صورت پرسش‌نامه‌ای در ۴ بُعد و ۲۱ سؤال طراحی گردید. جهت اطمینان از اعتبار یافته‌های پژوهش، ابتدا از روایی صوری و محتوایی استفاده می‌شود. فراوانی خبرگانی که در بررسی روایی صوری و محتوایی پرسش‌نامه شرکت نمودند ۳ نفر مرد و ۱۱ نفر زن بود، شامل ۱۰ نفر متأهل و ۴ نفر مجرد. متوسط سن ایشان ۴۱ سال و متوسط سابقه خدمت ۱۶ سال بوده است. سطح تحصیلات دکتری و بالاتر دارای فروانی ۳ نفر و فوق‌لیسانس دارای ۱۱ نفر بوده است.

تأثیر گویه: در این مرحله پس از محاسبه شاخص امتیاز تأثیر گویه، تمام گویه‌ها تأیید شدند.

نسبت روایی محتوا: نتایج حاصل از محاسبه این نسبت نشان داد که مقادیر نسبت روایی محتوا بر اساس جدول لاوشه برای ۱۴ خبره سؤالات کمتر از ۰.۵۱ باید حذف شوند. در این بخش مقادیر نسبت روایی محتوا همه سؤالات بالاتر از ۰.۵۱ بود و همه سؤالات پرسش‌نامه باقی ماندند. شاخص روایی محتوا: بر اساس نتایج حاصل از محاسبه این شاخص، همه سؤالات با مقادیر شاخص روایی محتوای بالاتر از ۰.۷۹ بوده‌اند و همه سؤالات پرسش‌نامه باقی ماندند.

تعداد نمونه ۳۸۵ نفر از جامعه همکاران فرهنگی مقطع ابتدایی در کل کشور در بخش نهایی همکاری داشته‌اند که ۱۱۷ نفر از ایشان مرد بودند. میانگین سن همکاران ۳۴.۶۳ سال و سابقه ایشان ۸.۸۵ سال می‌باشد که از این میان ۲۰۰ نفر فارغ‌التحصیل مقطع کارشناسی، ۱۶۱ نفر کارشناسی ارشد و ۲۴ نفر دکتری بوده‌اند.

برای بررسی روایی سازه‌ای، از تحلیل عاملی اکتشافی استفاده شد. جهت رعایت پیش‌فرض‌ها، شاخص کایزر-مایر-اولکین (KMO) به عنوان شاخص کیفیت نمونه‌برداری محاسبه شد و ۰.۸۵۷ به دست آمد. همچنین، آزمون کرویت بارلت اجرا شد که در آن مجذور کای به دست آمده (برابر با ۶۰۶۳.۲۱۶) در سطح کمتر از ۰.۰۰۰ کاملاً معنی‌دار بود. این نتایج حاکی از تحقق پیش‌فرض‌های تحلیل عاملی اکتشافی بود. بر اساس نتایج ابتدایی تحلیل عاملی اکتشافی، عوامل استخراجی و چرخش واریماکس ۴ عامل شناسایی گردید.

جدول ۲. مقادیر ویژه و واریانس کل تبیینی عوامل استخراج‌شده پرسش‌نامه

عوامل	مقادیر ویژه اولیه			مقادیر عناصر استخراجی پس از چرخش واریماکس		
	کل	درصد واریانس	درصد تراکمی	کل	درصد واریانس	درصد تراکمی
۱	۵.۷۴۳	۲۷.۳۴۵	۲۷.۳۴۵	۴.۲۲۷	۲۰.۱۲۷	۲۰.۱۲۷
۲	۴.۹۶۲	۲۳.۶۲۷	۵۰.۹۷۳	۴.۲۰۶	۲۰.۰۲۸	۴۰.۱۵۵
۳	۲.۶۵۷	۱۲.۶۵۲	۶۳.۶۲۵	۳.۶۷۴	۱۷.۴۹۷	۵۷.۶۵۲
۴	۱.۷۷۳	۸.۴۴۴	۷۲.۰۶۹	۳.۰۲۸	۱۴.۴۱۷	۷۲.۰۶۹

با توجه به نتایج جدول ۲ و پس از چرخش واریماکس، ۴ عامل با مقادیر ویژه ۷۲.۰۶۹ درصد از واریانس کل پرسش‌نامه (۲۷.۳۴۵) درصد به وسیله عامل اول، ۲۳.۶۲۷ درصد به وسیله عامل دوم، ۱۲.۶۵۲ درصد به وسیله عامل سوم و ۸.۴۴۴ درصد به وسیله عامل چهارم) را تبیین می‌کنند. بارهای عاملی نشان‌دهنده میزان همبستگی هر گویه با عامل است. بر اساس ۴ عامل اصلی استخراج‌شده پس از چرخش واریماکس، بار عاملی هر کدام از گویه‌ها مشخص شد که در جدول ۳ قابل مشاهده است.

جدول ۳. بارهای عاملی گویه‌ها پس از چرخش واریماکس

عوامل				گویه‌ها
۴	۳	۲	۱	
		۰.۷۹۸		مدیریت زمان کلاس درس یکی از مهارت‌های مؤثر در ارزشیابی درس ریاضی می‌باشد.
		۰.۸۴		یادگیری محتوای کتاب ریاضی و آگاهی از ارتباط طولی و عرضی مباحث درسی، لازمه ارتقای سطح ارزشیابی این درس است.
		۰.۸۷۸		توانایی انطباق اهداف یادگیری و عمق تفکر دانش‌آموزان برای یک ارزشیابی خوب لازم است.
		۰.۷۰۱		معلم باید درک بالایی از کاربرد روش‌های مختلف ارزشیابی در ارتباط با اهداف آموزشی مقطع و درس ریاضی داشته باشد.
		۰.۸۰۸		معلم بایستی مسئولیت‌های قانونی و اخلاقی خود در ارزشیابی را درک و به آن‌ها عمل کند.
		۰.۸۰۹		آشنایی با شیوه‌های ارزشیابی باکیفیت بر باور ایجاد تغییر در یادگیری ریاضی می‌تواند مؤثر باشد.
		۰.۸۵۵		استفاده از مجموعه‌ای از راهبردها و ابزارها در ارزشیابی تکوینی درس ریاضی مهم است.
		۰.۸۶۸		معلم باید بتواند ارزشیابی کلاسی جذاب مرتبط با فرهنگ، هویت، تجارب و منابع دانش‌آموزان انجام دهد.
		۰.۹۲۳		معلم باید بتواند ارزشیابی دانش‌آموزان را از هم متمایز نماید.
		۰.۹۲۱		ساخت طرح‌های امتیازدهی و تقویت انگیزش دانش‌آموزان جزو ابزارهای مهم ارزشیابی است.
		۰.۸۷۷		معلم باید بتواند ارزشیابی‌های مبتنی بر منابع خارج از کلاس و تفسیر نتایج را اجرا نماید.
۰.۷۲۷				معلم باید توانایی تجزیه و تحلیل سؤالات کلاسی و سؤالات آزمون را داشته باشد.
۰.۷۸۵				معلم باید بتواند به یادگیری دانش‌آموزان با استفاده از اطلاعات ارزشیابی کمک کند.
۰.۸۳۴				معلم بایستی بتواند بازخورد مؤثر و مفید در مورد کار دانش‌آموز با تأکید بر نقاط ضعف و قوت ایشان ارائه نماید.
۰.۸۰۴				معلم بایستی بتواند تفسیر خود از نتایج ارزشیابی و استدلال در مورد تصمیمات آموزشی را بیان نماید.

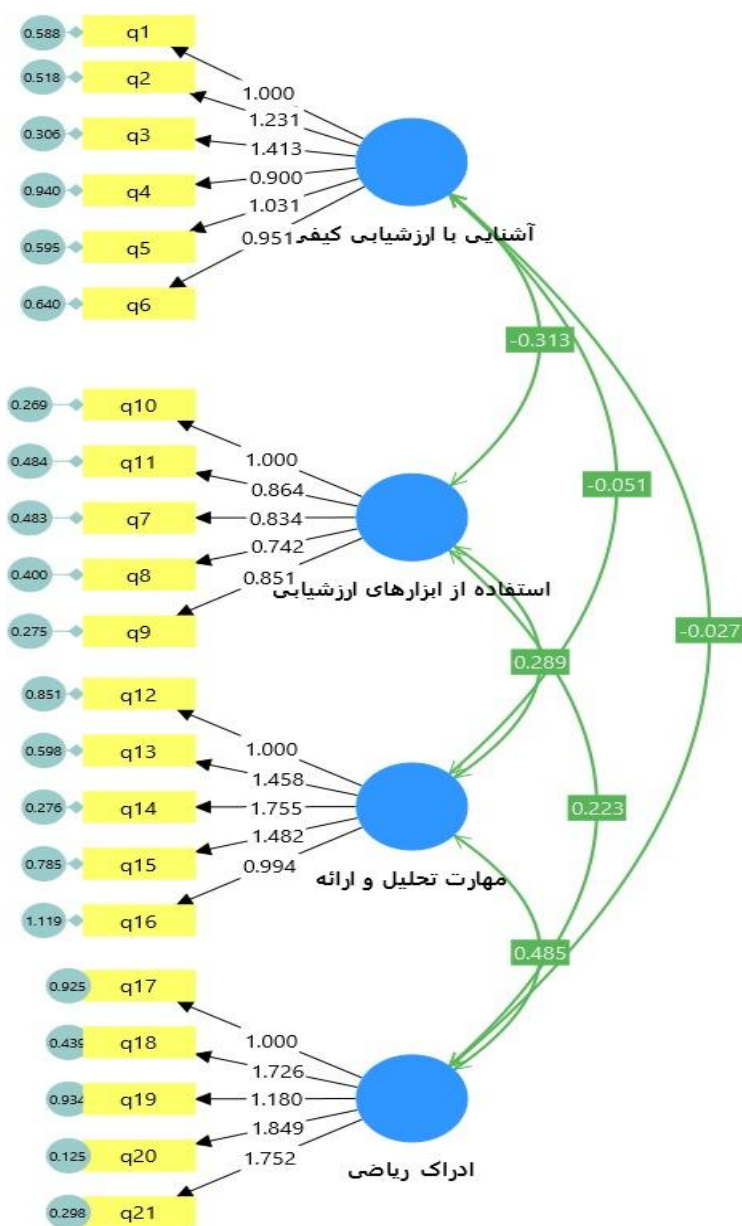
۰.۵۶۵				داشتن تعامل و کار مؤثر با دیگران در مدرسه (از جمله دانش‌آموزان) در ارزشیابی نقش دارد.
	۰.۶۷۳			معلم بایستی در علم ریاضی دانش عمیقی داشته باشد.
	۰.۸۹۲			معلم باید در تشخیص نوع خطای دانش‌آموزان در حل مسائل و دلایل آن توانا باشد.
	۰.۷۱۷			معلم باید از ابزارهای آموزش ریاضی در ارزشیابی کلاسی استفاده نماید.
	۰.۹۰۹			معلم باید در بررسی روش‌های مختلف حل مسئله منعطف باشد.
	۰.۸۸۵			معلم بایستی توانایی درجه‌بندی تکالیف و مسائل ریاضی را داشته باشد.

در جدول ۴ بارهای عاملی هر کدام از گویه‌های پرسش‌نامه سنجش مهارت‌های معلمی ارزشیابی درس ریاضی بر روی ۴ عامل استخراج شده مشخص شده است. بارهای عاملی ۰.۵ و بالاتر ملاک اختصاص هر گویه به عامل‌ها در نظر گرفته شد. با توجه به اینکه همه گویه‌ها بار عاملی بالاتر از متوسط با عامل‌های خود داشتند، در این مرحله هیچ کدام از آن‌ها حذف نشدند. همچنین، در این مطالعه روایی سازه‌ای پرسش‌نامه با استفاده از تحلیل عاملی تأییدی نیز بررسی شد. شکل ۱ مدل تحلیل عاملی تأییدی را نشان می‌دهد.

جدول ۴. بار عاملی، مقدار آماره t ، مقدار سطح معنی‌داری، میانگین، انحراف معیار، کمینه و بیشینه گویه‌ها

گویه‌ها	بار عاملی	T values	P values	میانگین	انحراف معیار	کمینه	بیشینه
q1	۰.۷۵۹	n/a	n/a	۳.۳۳۵	۱.۳۸۶	۱	۵
q2	۰.۸۳۷	۱۷.۴۴	۰	۳.۲۴۷	۱.۷۲۹	۱	۵
q3	۰.۹۱۶	۱۸.۵۸۷	۰	۳.۱۲۲	۱.۸۹۹	۱	۵
q4	۰.۶۳۸	۱۲.۶۲۴	۰	۲.۹۲۵	۱.۵۸۷	۱	۵
q5	۰.۷۶۷	۱۵.۴۳۵	۰	۳.۴۲۸	۱.۴۵۱	۱	۵
q6	۰.۷۲۸	۱۴.۵۰۳	۰	۳.۵۲	۱.۳۶	۱	۵
q7	۰.۸۴۹	۲۵.۷۰۴	۰	۳.۵۱۴	۱.۷۳۶	۱	۵
q8	۰.۸۴۴	۲۴.۴۰۹	۰	۳.۸۲۶	۱.۳۹	۱	۵
q9	۰.۹۰۹	۳۰.۱۰۶	۰	۳.۸۷۵	۱.۵۸۳	۱	۵
q10	۰.۹۳۳	n/a	n/a	۳.۵۰۴	۲.۰۹۳	۱	۵
q11	۰.۸۵۸	۲۶.۶۸	۰	۳.۶۲۹	۱.۸۲۸	۱	۵
q12	۰.۵۸۸	n/a	n/a	۲.۹۰۶	۱.۳	۱	۵
q13	۰.۷۸۴	۱۱.۶۱۹	۰	۳.۱۷۴	۱.۵۵۲	۱	۵
q14	۰.۹۱۳	۱۲.۲۹۶	۰	۳.۱۴۸	۱.۶۵۹	۱	۵
q15	۰.۷۴۶	۱۱.۲۵۹	۰	۲.۸۴۷	۱.۷۷۱	۱	۵
q16	۰.۵۳۳	۸.۸۰۳	۰	۳.۰۶	۱.۵۶۳	۱	۵
q17	۰.۵۷	n/a	n/a	۲.۹۶۴	۱.۳۷	۱	۵
q18	۰.۸۶۷	۱۲.۳۵۵	۰	۳.۳۶۶	۱.۷۶۵	۱	۵
q19	۰.۶۳۲	۱۰.۱۸۲	۰	۳.۰۹۹	۱.۵۵۴	۱	۵
q20	۰.۹۶۱	۱۲.۸۵۴	۰	۳.۴۷۳	۱.۶۴۷	۱	۵
q21	۰.۹۰۶	۱۲.۵۶	۰	۳.۴۰۸	۱.۶۶۵	۱	۵

یافته‌ها نشان می‌دهد که مقدار بار عاملی تمامی سؤالات پرسش‌نامه بالاتر از ۰.۵ بود، و با توجه به اینکه مقدار T-value به‌دست‌آمده برای همه سؤالات مربوط به ۴ بُعد شناسایی شده در حد مناسب بود، در نتیجه مجموع سؤالات اعتبار لازم برای حضور در پرسش‌نامه را دارند.



شکل ۱. تحلیل عاملی تأییدی پرسش‌نامه سنجش مهارت‌های معلمی ارزشیابی درس ریاضی

در جدول ۴ و شکل ۱ بارهای عاملی هر کدام از سؤالات پرسش‌نامه سنجش مهارت‌های معلمی ارزشیابی درس ریاضی بر روی ۴ عامل استخراج شده مشخص شده است. سؤالات ۱ تا ۶ بیشترین بار عاملی را با عامل اول داشتند. عامل دوم بار عاملی سؤالات ۷ تا ۱۱ را به خود اختصاص داد. سؤالات ۱۲ تا ۱۶ بیشترین بار عاملی را با عامل سوم داشتند. و عامل چهارم بار عاملی سؤالات ۱۷ تا ۲۱ را نصیب خود ساخته است. در نهایت، با در نظر گرفتن محتوای سؤالات هر عامل و با توجه به یافته‌های بخش کیفی، عامل اول با ۶ سؤال آشنایی با ارزشیابی توصیفی، عامل دوم با ۵ سؤال استفاده از ابزارهای ارزشیابی، عامل سوم با ۵ سؤال مهارت تحلیل و ارائه، و عامل چهارم با ۵ سؤال ادراک ریاضی نام گرفت.

جدول ۵. شاخص‌های روایی و پایایی تحلیل عاملی تأییدی

آشنایی با ارزشیابی توصیفی	آلفای کرونباخ	پایایی ترکیبی	میانگین واریانس استخراج شده (AVE)
آشنایی با ارزشیابی توصیفی	۰.۹	۰.۹۰۵	۰.۶۰۷
ادراک ریاضی	۰.۸۹۵	۰.۹۰۲	۰.۶۴۴
استفاده از ابزارهای ارزشیابی	۰.۹۴۴	۰.۹۴۵	۰.۷۷۳
مهارت تحلیل و ارائه	۰.۸۳۸	۰.۸۴۷	۰.۵۲۷

پایایی هر دو ملاک (آلفای کرونباخ، پایایی ترکیبی) همه ابعاد بالای ۰.۸، و میزان میانگین واریانس استخراج شده برای تمام عامل‌ها بیشتر از ۰.۵ است.

جدول ۶. مقایسه ریشه دوم میانگین واریانس استخراج شده هر سازه با مقادیر ضرایب همبستگی بین سازه‌ها

آشنایی با ارزشیابی توصیفی	ادراک ریاضی	استفاده از ابزارهای ارزشیابی	مهارت تحلیل و ارائه
آشنایی با ارزشیابی توصیفی	۰.۷۷۹		
ادراک ریاضی	-۰.۰۲۷	۰.۸۰۲	
استفاده از ابزارهای ارزشیابی	-۰.۳۱۳	۰.۲۲۳	۰.۸۷۹
مهارت تحلیل و ارائه	-۰.۰۵۱	۰.۴۸۵	۰.۲۸۹

همچنین، مجذور همبستگی هر یک از عامل‌ها با عوامل دیگر کوچک‌تر از میزان میانگین واریانس استخراج شده آن عامل می‌باشد، که نشان‌دهنده وجود روایی تشخیصی می‌باشد (جدول ۶). بنابراین، ابزار اندازه‌گیری از نقطه‌نظر روایی همگرا و روایی تشخیصی (واگرا) مورد پذیرش قرار می‌گیرد و در نهایت سازه موردنظر با ۴ بُعد و ۲۱ سؤال مورد تأیید قرار گرفت. بررسی شاخص‌های برازش مدل نشان داد که شاخص‌ها با اختلاف کم، حد نصاب لازم را برای برازش مدل را دارا می‌باشند. در واقع این مقادیر در محدوده قابل قبول برای نمونه‌های بزرگ قرار دارد (Shi et al., 2019). با وجود حجم نمونه $n=385$ و $CFI=0.89$ و $RMSEA=0.096$ نشان‌دهنده برازش مناسب است (Hu & Bentler, 1999). شاخص‌های برازش مدل در جدول ۷ آمده است.

جدول ۷. شاخص‌های برازش مدل

شاخص	مقدار	معیار قابل قبول (Shi et al., 2019)
χ^2/df	۴.۵۴۷	$5 >$
CFI	۰.۸۹۲	$0.85 \leq$
TLI	۰.۸۷۷	$0.85 \leq$
RMSEA	۰.۰۹۶	$0.1 \geq$
SRMR	۰.۰۹۴	$0.1 \geq$
p-value	۰.۰۰۰	-

بحث و نتیجه‌گیری

تحقیقات در حوزه مهارت و دانش معلمی در آموزش ریاضی گزاره‌های قوی و قابل‌قبولی را در مورد چگونگی اهمیت مهارت و دانش برای تدریس مطرح کرده‌اند (Hill et al., 2008; Kaya, 2018; Rahimi & Sedaghatkhah, 2022; Sahito et al., 2024; Brookhart, 2024; Hosseini et al., 2024)، اما در این تحقیق، با جمع‌بندی مطالعات پیشین و نظرسنجی از خبرگان، مقوله‌های مهمی در مهارت‌های معلمی در ارزشیابی توصیفی درس ریاضی آشکار شد و پرسش‌نامه مربوطه با تکیه بر نتایج آماری مورد تأیید قرار گرفت. بدین منظور از روش آمیخته (کیفی و کمی) اکتشافی با روش تحلیل عاملی استفاده شد. در بخش کیفی و در تکمیل یافته‌های مرادپور کرنکان (Moradpoor Kerenkan, 2024)، ۲۱ مهارت معلمی در ارزشیابی درس ریاضی به دست آمد. پس از تأیید سه شاخص تأثیر گویند، نسبت روایی محتوا و شاخص روایی محتوا در بخش یافته‌های کمی، نتایج تحلیل عاملی نیز مؤید ۲۱ گویند در قالب ۴ عامل اصلی بوده است.

جدول ۸. نتایج تحلیل عاملی و آمار توصیفی مهارت‌های معلمی در ارزشیابی توصیفی ریاضی

میانگین (طیف لیکرت ۱-۵)	بار عاملی (دامنه)	مهارت کلی
۳.۲۵	۰.۶۷-۰.۹۲	آشنایی با ارزشیابی کیفی
۳.۶۷	۰.۹۳-۰.۸۴	استفاده از ابزارهای ارزشیابی
۳.۰۲	۰.۵۳-۰.۹۱	مهارت تحلیل و ارائه
۳.۲	۰.۵۷-۰.۹۶	ادراک ریاضی

مطابق نتایج آماری، قوی‌ترین بُعد استفاده از ابزارهای ارزشیابی (میانگین=۳.۶۷، بارهای عاملی بالا) و ضعیف‌ترین بُعد مهارت تحلیل و ارائه (میانگین=۳.۰۲، بارهای عاملی پایین‌تر) می‌باشد. سطح کلی مهارت‌ها در حد متوسط (۳.۰۲ تا ۳.۶۷) گزارش شده است. تحلیل زیرمهارت‌ها به شرح ذیل می‌باشد:

مهارت اول؛ آشنایی با ارزشیابی توصیفی: نتایج تحلیل عاملی اکتشافی بر روی عامل اول نشان داد که مهارت آشنایی با ارزشیابی توصیفی به‌تنهایی ۲۰/۱۲۷ درصد از واریانس کل پرسش‌نامه را تبیین می‌کند. این عامل دارای ۶ زیرمهارت می‌باشد. زیرمهارت اول، مدیریت زمان کلاسی: حسینی و همکاران (2024) از جمله چالش‌های اجرای ارزشیابی توصیفی از دیدگاه معلمان در ارتباط با مدیریت زمان کلاس را شامل کمبود زمان برای انجام تمامی اشکال ارزشیابی توصیفی، زمان‌بر بودن تکمیل برگه‌های مربوط به ارزشیابی توصیفی و کمبود زمان مخصوصاً در کلاس‌های چندپایه می‌دانند. زیرمهارت دوم، یادگیری محتوای کتاب ریاضی: معلمان باید یادگیری را در حوزه محتوایی که تدریس می‌کنند درک کنند (Brookhart, 2024). معلمی که قصد ارزیابی یک تکلیف را در هر حوزه‌ای دارد نیازمند شناخت ماهیت آن حوزه و دانش و تسلط لازم برای قضاوت راه‌حل‌ها و پاسخ‌های باز و منعطف دانش‌آموزان در آن حوزه است (Mortazavi et al., 2020A). معلمان باید یادگیری را در حوزه محتوایی که تدریس می‌کنند درک کنند (Brookhart, 2024). زیرمهارت سوم، انطباق اهداف یادگیری با محتوا و عمق تفکر دانش‌آموزان: معلمان باید بتوانند مقاصد یادگیری روشنی را بیان کنند که هم با محتوا و هم با عمق تفکری که توسط استانداردها و اهداف برنامه درسی بیان می‌شود، و هم با هویت‌های فرهنگی دانش‌آموزان، به‌گونه‌ای که قابل‌دستیابی و ارزشیابی باشند، مطابقت داشته باشند (Brookhart, 2024). زیرمهارت چهارم، درک اهداف و کاربردهای روش‌های مختلف ارزشیابی: مدرسان باید قادر باشند شیوه‌های ارزشیابی گوناگونی همچون پاسخ انتخابی، پاسخ مکتوب، ارزشیابی عملکرد و ارتباط شخصی مستقیم را انجام دهند (Estaji, 2021). نمره‌های آزمون‌های استاندارد باید قسمتی از اطلاعات لازم برای ارزشیابی دانش‌آموزان تلقی شوند که تنها همراه با اطلاعات دیگر، مانند بازخوردهای توصیفی و مشاهدات کلاسی، معنا و اعتبار پیدا می‌کنند (Yan et al., 2023). دانش ناکافی

معلمان درباره هدف‌ها، رسالت و رویه‌های ارزشیابی توصیفی جهت ارتقای جریان یاددهی-یادگیری از دلایل عدم موفقیت اجرای این طرح شناخته شد (Mortazavi et al., 2020B). زیرمهارت پنجم، درک و انجام مسئولیت‌های قانونی و اخلاقی خود در ارزشیابی: معلمان باید مسئولیت‌های قانونی و اخلاقی خود را در ارزشیابی در حین انجام کار درک کرده و انجام دهند (Brookhart, 2024). زیرمهارت ششم، باور به ایجاد تغییر در یادگیری ریاضی با تکیه بر شیوه‌های ارزشیابی باکیفیت: معلمان باید باور داشته باشند که شیوه‌های ارزشیابی با کیفیت بالا برای یادگیری کودکان تفاوت ایجاد می‌کند و آن‌ها (معلمان) می‌توانند در هنگام استفاده از آن‌ها تفاوت ایجاد کنند (Brookhart, 2024). این مهارت با مدل سه‌بعدی سواد ارزشیابی (دانش، مهارت، گرایش؛ Pastore & Andrade, 2019) هم‌خوانی دارد. Brookhart (2001) تأکید می‌کند که معلمان باید مسئولیت‌های قانونی/اخلاقی و باور به تغییر با ارزشیابی باکیفیت را داشته باشند. این دو زیرمهارت در ابزار ما اولویت‌دار شدند.

مهارت دوم: استفاده از ابزارهای ارزشیابی: نتایج آماری نشان داد که این مهارت ۲۰/۰۲۸ درصد از واریانس کل پرسش‌نامه را تبیین می‌کند. این عامل دارای ۵ زیرمهارت می‌باشد. زیرمهارت اول، استفاده از مجموعه‌ای از راهبردها و ابزارها برای ارزشیابی تکوینی درس ریاضی: معلمان باید مجموعه‌ای از راهبردها را برای ارزشیابی تکوینی کلاس درس داشته باشند و از آن‌ها استفاده کنند و بتوانند کار دانش‌آموزان را از نظر تفکرشان تفسیر کنند، نه فقط از نظر درستی آن‌ها (Brookhart, 2024). معلمان از ابزارهای دیگری مانند مشاهده دانش‌آموزان در شرایط واقعی، ثبت اطلاعات لازم، کارپوشه، چک‌لیست، آزمون‌های عملکردی، خودارزشیابی، ارزشیابی همسال و ارائه گزارش‌های پیشرفت دانش‌آموزان استفاده می‌کنند. این ابزارها تصویر دقیق‌تری از فرایند یادگیری و نتایج تلاش‌های دانش‌آموزان برای معلم فراهم می‌کنند (Hassani & Ahmadi, 2012). زیرمهارت دوم، ایجاد ارزشیابی‌های کلاسی جذاب مرتبط با فرهنگ، هویت، تجربیات و منابع دانش‌آموزان: معلمان باید بتوانند ارزشیابی‌های کلاسی جذابی را انتخاب یا ایجاد کنند که با فرهنگ‌ها، هویت‌ها، تجربیات و منابع دانش‌آموزانشان مرتبط باشد (Brookhart, 2024). زیرمهارت سوم، تمایز ارزشیابی برای دانش‌آموزان مختلف: معلمان باید بتوانند ارزشیابی را برای دانش‌آموزان مختلف متمایز کنند (Brookhart, 2024). برای ارائه تصویری کامل از وضعیت آموزشی دانش‌آموزان، باید ویژگی‌ها، رفتارها، و توانایی‌های ارتباطی نوشتاری یا شفاهی آنان مورد ارزشیابی چندجانبه قرار گیرد (Ozan & Kincal, 2023). در بُعد تمیز، به آسیب‌هایی مانند تصور اشتباه در مورد نبودن نمره و امتحان، دادن نمره به دانش‌آموزان در حین سال تحصیلی و استفاده از آزمون‌های مدادکاغذی با حجم زیاد، توسط دانش‌آموزان مصاحبه‌شونده اشاره شده است (Hosseini et al., 2024). زیرمهارت چهارم، ساخت طرح‌های امتیازدهی و تقویت انگیزش دانش‌آموزان: معلمان باید بتوانند طرح‌های امتیازدهی‌ای بسازند که به عملکرد دانش‌آموز در ارزشیابی‌های کلاسی به صورت اطلاعات مفیدی برای تصمیم‌گیری در مورد دانش‌آموزان، کلاس‌های درس، مدارس و مناطق، کمیت می‌دهد. این تصمیمات باید به بهبود یادگیری، رشد یا توسعه دانش‌آموزان منجر شود (Brookhart, 2024). زیرمهارت پنجم، اجرای ارزشیابی‌های مبتنی بر منابع خارج از کلاس و تفسیر نتایج: معلمان باید بتوانند ارزشیابی‌های بیرونی را اجرا کنند و نتایج آن‌ها را برای تصمیم‌گیری در مورد دانش‌آموزان، کلاس‌های درس، مدارس و مناطق تفسیر کنند (Brookhart, 2024). این مهارت با چارچوب IUA¹ (Black & Wiliam, 1998) و معیارهای ۷گانه صلاحیت معلم (AFT, NCME, NEA, 1990) هم‌خوانی دارد. Brookhart (2024) بر تمایز ارزشیابی و طرح‌های امتیازدهی انگیزشی تأکید دارد. ابزار این پژوهش این مهارت‌ها را عملیاتی کرد.

مهارت سوم: مهارت تحلیل و ارائه: نتایج تحلیل عامل بر روی عامل سوم نشان داد که مهارت تحلیل و ارائه ۱۷/۴۹۷ درصد از واریانس کل پرسش‌نامه را تبیین می‌کند. این مهارت دارای ۵ بُعد می‌باشد. زیرمهارت اول، تجزیه و تحلیل سؤالات کلاسی و سؤالات آزمون: معلمان باید مهارت تجزیه و تحلیل سؤالات کلاسی، سؤالات آزمون، و وظایف ارزشیابی عملکرد را داشته باشند تا دانش و مهارت‌های فکری خاص

1. instrumental, unobtrusive, actionable

مورد نیاز دانش‌آموزان را برای انجام آن‌ها مشخص کنند (Brookhart, 2024). زیرمهارت دوم، کمک به دانش‌آموزان برای تنظیم یادگیری با استفاده از اطلاعات ارزشیابی: معلمان باید بتوانند به دانش‌آموزان کمک کنند تا از اطلاعات ارزشیابی برای تنظیم یادگیری خود استفاده کنند (Brookhart, 2024). زیرمهارت سوم، ارائه بازخورد مؤثر و مفید در مورد کار دانش‌آموز با تأکید بر نقاط ضعف و قوت: معلمان برای برقراری ارتباط مناسب با دیگران در مسائل ارزشیابی یادگیرندگان باید بتوانند به‌درستی از اصطلاحات تخصصی ارزشیابی استفاده و، معنی، محدودیت‌ها و کاربردهای نتایج آن‌ها را بیان کنند (Estaji, 2021). سمیعی و شمس الهی (Shamsolahi & Samie Zafarghandi, 2017) توانایی ارائه بازخورد متناسب با عملکرد دانش‌آموزان را جزو مهارت‌های معلمی که دارای همبستگی با ارزشیابی توصیفی هستند می‌دانند. زیرمهارت چهارم، تبیین تفاسیر خود از نتایج ارزشیابی و استدلال در مورد تصمیمات آموزشی: یکی از حوزه‌های مورد نیاز، بیان و انتقال نتایج ارزشیابی و اتخاذ تصمیمات آموزشی بر اساس آن است (Estaji, 2021). معلمان باید بتوانند تفاسیر خود را از نتایج ارزشیابی و استدلال خود در مورد تصمیمات آموزشی بر اساس نتایج ارزشیابی را برای جمعیت‌های آموزشی که به آن‌ها خدمت می‌کنند (دانش‌آموز و خانواده آن‌ها، کلاس، مدرسه و جامعه) بیان کنند (Brookhart, 2024). زیرمهارت پنجم، تعامل و کار مؤثر با دیگران در مدرسه (از جمله دانش‌آموزان): معلمان دارای سواد ارزشیابی شیوه‌های درست ارزشیابی، سنجش و برقراری ارتباط را تشخیص می‌دهند (Estaji, 2021). معلمان باید قادر به تعامل و کار مؤثر با دیگران در مدرسه (از جمله دانش‌آموزان) و در جامعه باشند، زیرا دانش و مهارت‌های ارزشیابی را در این زمینه به کار می‌برند (Brookhart, 2024). این مهارت با استانداردهای ارتباط و تصمیم‌گیری ارزشیابی (Estaji, 2021) و مدل سه‌بعدی سواد ارزشیابی (Pastore & Andrade, 2019) هم‌خوانی دارد. Gallagher و همکاران (2025) نشان دادند که کمتر از ۱۰٪ ابزارهای جهانی تحلیل بازخورد را پوشش می‌دهند. ابزار حاصل از این پژوهش این خلأ را پر کرد.

مهارت چهارم؛ ادراک ریاضی: تبیین ۱۴/۴۱۷ درصد از واریانس کل پرسش‌نامه توسط این مهارت بر اساس یافته‌های تحلیل عاملی نشان داده می‌شود. این عامل دارای ۵ زیرمهارت می‌باشد. زیرمهارت اول، داشتن دانش عمیق در ریاضی: ما با بررسی رابطه بین دانش ریاضی معلم برای تدریس و کیفیت ریاضی آموزش، پویایی استفاده از دانش در تدریس را روشن کردیم (Hill et al., 2008). در ریاضیات وقتی دانش‌آموزان درگیر پرسش‌های متفکرانه می‌شوند درک عمیق‌تری از مفاهیم پیدا می‌کنند. پرسش همچنین به معلمان کمک می‌کند تا توجه دانش‌آموزان را جلب کرده و آن‌ها را در یادگیری‌شان نگه دارند (Sahito et al., 2024). زیرمهارت دوم، توانایی تشخیص نوع خطای دانش‌آموزان در حل مسائل و دلایل آن: فرایند تجزیه و تحلیل خطا می‌تواند به پیشرفت در توسعه استراتژی‌های آموزشی کارآمد کمک کند، اگر زمانی صرف بررسی چگونگی و چرایی اشتباهات دانش‌آموزان و انجام اصلاحات فوری برای خطاها شود (Pajares, 2002). با تکرار کمتر خطاها، دانش‌آموزان مفاهیم و مهارت‌ها را در سطح مهارت به‌طور کامل فرا خواهند گرفت. علاوه بر این، فرایند تحلیل خطای آینده‌ای شفاف فراهم می‌کند که از طریق آن معلمان می‌توانند اثربخشی تدریس خود را منعکس و ارزشیابی کنند، و این انعکاس به معلمان اجازه می‌دهد تا نقاط قوت و ضعف خود را در تدریس حوزه‌های محتوایی خاص شناسایی کرده و تجربه و اعتمادبه‌نفس خود را در تدریس افزایش دهند (An & Wu, 2012). زیرمهارت سوم، مهارت استفاده از ابزارهای آموزش ریاضی در ارزشیابی کلاسی: برخی از مؤسسات آموزشی با اتخاذ انواع ابزارهای مدرن ارزشیابی، که برای ارتقای یادگیری دانش‌آموزان طراحی شده‌اند، به اهمیت روش‌های ارزشیابی که از یادگیری دانش‌آموزان پشتیبانی می‌کنند پی‌برده‌اند (Sahito et al., 2024). زیرمهارت چهارم، انعطاف‌پذیری منطقی در بررسی روش‌های مختلف حل مسئله: راهبردها و فرایندهای استفاده‌شده توسط دانش‌آموزان باید به اندازه پاسخ‌های نهایی آنان مورد توجه و تحلیل قرار گیرند (van den Heuvel-Panhuizen et al., 2021). زیرمهارت پنجم، توانایی درجه‌بندی تکالیف و مسائل ریاضی: مضمون آزمون‌ها باید به‌گونه‌ای متعادل و با توجه به استانداردهای برنامه درسی و ارزشیابی تنظیم شوند که تنها تکرار مکرر محاسبات ریاضی نباشند، بلکه تفکر، استدلال و حل مسئله را نیز پوشش دهند (REL Central, 2025). این مهارت بازتعریف عملیاتی PCK در ارزشیابی توصیفی است (Shulman, 1986) و تحقیق Shaughnessy و همکاران (2021) مهارت‌های هدایت بحث‌های ریاضی را گزارش کرده، اما هیچ ابزاری PCK را در ارزشیابی توصیفی عملیاتی نکرده است. ابزار حاصل از این پژوهش اولین در ایران و جهان است.

با توجه به بررسی‌های انجام‌شده و در راستای هدف پژوهش و طراحی پرسش‌نامه‌ای برای سنجش مهارت‌های معلمی در ارزشیابی توصیفی درس ریاضی، مؤلفه‌های به‌دست‌آمده از نظرات خبرگان و ذی‌نفعان و بر پایه مستندات پژوهشی و در امتداد پژوهش مرادپور کرنکان (۱۴۰۳) و تکمیل مهارت‌های معلمی موردنیاز در ارزشیابی توصیفی درس ریاضی به دست آمد. مهم‌ترین نکته اضافه شدن مهارت چهارم ادراک ریاضی و ۵ زیرمهارت آن شامل: داشتن دانش عمیق در ریاضی، توانایی تشخیص نوع خطای دانش‌آموزان در حل مسائل و دلایل آن، مهارت استفاده از ابزارهای آموزش ریاضی در ارزشیابی کلاسی، انعطاف‌پذیری منطقی در بررسی روش‌های مختلف حل مسئله و توانایی درجه‌بندی تکالیف و مسائل ریاضی، است که در آن بر توانایی معلم در حوزه علوم ریاضی و آموزش ریاضی تأکید شده است.

با تکیه بر اندازه شاخص‌های نسبت روایی محتوا، شاخص روایی محتوای آلفای کرونباخ، پایایی ترکیبی و شاخص‌های برازش مدل، در یک نتیجه‌گیری کلی می‌توان گفت که پرسش‌نامه سنجش مهارت‌های معلمی در ارزشیابی توصیفی درس ریاضی، با ۲۱ گویه و ۴ مؤلفه آشنایی با ارزشیابی توصیفی، استفاده از ابزارهای ارزشیابی، مهارت تحلیل، و ارائه و ادراک ریاضی، از اعتبار و روایی مطلوبی جهت سنجش مهارت‌های مورنیاز در معلمان برای ارزشیابی توصیفی درس ریاضی برخوردار می‌باشد.

هر چند چارچوب‌های سنجش کیفیت تدریس ریاضی (Gallagher et al., 2025) و مهارت‌های هدایت بحث‌های ریاضی (Shaughnessy et al., 2021) وجود دارند، ابزار حاضر اولین تلاش در ایران برای سنجش مهارت‌های ارزشیابی توصیفی ریاضی با استفاده از روش ترکیبی دلفی فازی و تحلیل عاملی تأییدی (CFA) است. ابعاد چهارگانه با چارچوب IUA (Black & Wiliam, 1998) هم‌خوانی دارد. اندازه شاخص‌های روایی و پایایی، نیکویی برازش و تناسب مدل شواهد تجربی قوی‌ای را ارائه می‌دهد. لذا این روش‌شناسی قابلیت تعمیم به کشورهای دارای ارزشیابی تکوینی (مانند فنلاند و مالزی) را دارد و نوآوری روش‌شناختی قابل توجه است.

این پژوهش با برخی محدودیت‌ها همراه بوده است: نخست آنکه کمبود مطالعات مشابه در حوزه نظام ارزشیابی امکان مقایسه و تطبیق نتایج را محدود ساخته است. دومین محدودیت به روش نمونه‌گیری تصادفی ساده از سراسر کشور مربوط می‌شود؛ هر چند تلاش شد پوشش جغرافیایی مناسبی حاصل شود، اما احتمال سوگیری انتخابی همچنان وجود دارد و این امر می‌تواند تعمیم‌پذیری یافته‌ها به مناطق خاص، به‌ویژه مناطق محروم، را با محدودیت مواجه سازد. بنابراین، پیشنهاد می‌شود در پژوهش‌های آتی از نمونه‌گیری طبقه‌ای متناسب با جمعیت معلمان هر استان استفاده شود. محدودیت دیگر این مطالعه آن است که تأثیر مستقیم ابزار طراحی‌شده بر یادگیری واقعی دانش‌آموزان مورد بررسی قرار نگرفته است. با وجود آنکه نتایج نشان داد ابزار از روایی و پایایی مطلوبی برخوردار است، ارزیابی اثربخشی آن در عمل نیازمند مطالعات تکمیلی است. به‌ویژه پیشنهاد می‌شود در تحقیقات آینده رابطه میان مهارت‌های معلمان (اندازه‌گیری‌شده با این ابزار) و پیشرفت تحصیلی دانش‌آموزان در درس ریاضی بررسی شود تا شواهد تجربی دقیق‌تری در خصوص نقش این ابزار در ارتقای یادگیری دانش‌آموزان فراهم گردد.

به لحاظ کاربردی پیشنهاد می‌گردد مسئولان آموزش و پرورش، مربیان و معلمان با استفاده از این پرسش‌نامه مهارت‌های معلمی در ارزشیابی توصیفی درس ریاضی را بسنجند و در صورت نیاز نسبت به طراحی و اجرا و مشارکت در دوره‌های ارتقای مهارت اقدام نمایند. و اما سه پیشنهاد اجرایی نیز ارائه می‌گردد: ۱. خودارزیابی سالانه معلمان در سامانه همگام؛ ۲. طراحی دوره‌های هدفمند ارتقای مهارت، مانند «تشخیص خطای ریاضی و بازخورد اصلاحی»؛ ۳. ارزیابی عملکرد مدارس و مناطق: رتبه‌بندی سالانه بر اساس نمره میانگین مهارت معلمان هر مدرسه.

مشارکت نویسندگان

میزان مشارکت نویسندگان در مقاله به‌صورت نسبی بوده است. نویسنده اول نگارش مقاله، نویسنده دوم انجام تحلیل آماری و نویسنده سوم مشاوره را بر عهده داشته است.

تعارض منافع

هیچ‌گونه تعارض منافع توسط نویسندگان بیان نشده است.

منابع

- احمدی، فضل‌اله؛ نصیریانی، خدیجه؛ ابادری، پروانه. (۱۳۸۷). تکنیک دلفی: ابزار برای تحقیق. *مجله ایرانی آموزش در علوم پزشکی*، ۸(۱۸) (پیاپی ۱۹)، ۱۷۵-۱۸۵. <https://sid.ir/paper/441616/fa>.
- استاجی، معصومه. (۱۴۰۰). بررسی سواد ارزشیابی معلمان در کلاس‌های آموزش زبان انگلیسی به‌عنوان یک زبان خارجی: تأثیر پیشینه تحصیلی. *اندازه‌گیری تربیتی*، ۱۲(۴۵)، ۷-۳۸. [10.22054/jem.2022.59105.2153](https://doi.org/10.22054/jem.2022.59105.2153).
- امینی، زهره. (۱۳۹۷). بهبود کیفیت ارزشیابی درس ریاضی با استفاده از فناوری. دومین کنفرانس آموزش و کاربرد ریاضیات، کرمانشاه. <https://civilica.com/doc/824693>.
- تبریزی، منصوره. (۱۳۹۳). تحلیل محتوای کیفی از منظر رویکردهای قیاسی و استقرائی. *علوم اجتماعی*، ۲۱(۶۴)، ۱۰۵-۱۳۸. <https://doi.org/10.22054/qjss.2014.344>.
- جزایری، اسماعیل. (۱۳۹۶). بررسی تأثیر ارزشیابی فرایندی در پیشرفت تحصیلی درس ریاضی دانش‌آموزان مقطع متوسطه شهرستان گتوند. پایان‌نامه کارشناسی ارشد رشته آموزش ریاضی، دانشگاه فرهنگیان.
- حاجی‌زاده، ابراهیم؛ اصغری، محمد. (۱۳۹۰). روش‌ها و تحلیل‌های آماری با نگاه به روش تحقیق در علوم زیستی و بهداشتی. تهران: جهاد دانشگاهی. چاپ اول.
- حسینی، محمد؛ احمدی، غلامعلی. (۱۳۸۶). زمینه‌یابی اجرای ارزشیابی کیفی-توصیفی در مدارس ابتدایی شهر تهران. *نوآوری‌های آموزشی*، ۶(۲۳)، ۸۵-۱۲۲. [SID. https://sid.ir/paper/75548/fa](https://sid.ir/paper/75548/fa).
- حسینی، زهرا سادات؛ شعبانی، عنایت‌اله. (۱۴۰۴). بررسی سواد ارزیابی مدرسان زبان فارسی به غیرفارسی‌زبانان. *پژوهش‌نامه آموزش زبان فارسی به غیرفارسی‌زبانان*، ۱۴(۱)، ۹۵-۱۲۷. <https://doi.org/10.30479/jtpsol.2025.21944.1707/>.
- حسینی، مسعود؛ تهرانیان، ابوالفضل؛ بهزادی، محمدحسن؛ علم‌الهدایی، سیدحسن. (۱۴۰۳). تجارب زیسته معلمان و دانش‌آموزان پایه ششم از ارزشیابی کیفی درس ریاضی، طراحی مقیاس مهارت‌های معلمی در ارزشیابی کیفی درس ریاضی. *علوم روانشناختی*، ۲۳(۱۳۷)، ۲۴۷-۲۶۵. <http://psychologicalscience.ir/article-۲۱۶۲-fa.html>.
- رزی، جمال؛ امام جمعه، سید محمدرضا؛ احمدی، غلامعلی. (۱۳۹۶). استخراج ابعاد معلمان اثربخش به منظور ارائه چهارچوب مفهومی آموزش و تربیت‌معلم اثربخش و اعتباربخشی آن. *مطالعات سیاست‌گذاری تربیت‌معلم (پژوهش در تربیت‌معلم)*، ۱۱(۱)، ۱۳-۴۶. [SID. https://sid.ir/paper/269462/fa](https://sid.ir/paper/269462/fa).
- سرمذ، زهره؛ بازرگان، عباس؛ حجازی، الهه. (۱۳۹۸). روش‌های تحقیق در علوم رفتاری. تهران: نشر آگه.
- شمس‌اللهی، مریم؛ سمیعی زفرقندی، مرتضی. (۱۳۹۶). سهم مهارت‌های حرفه‌ای معلمان دوره ابتدایی در اجرای ارزشیابی توصیفی. *مطالعات پیش‌دبستان و دبستان*، ۳(۹)، ۲۳-۴۷. <https://doi.org/10.22054/soece.2020.37429.1189>.
- عبادی، موسی؛ سراجی، فرهاد؛ بختیاری، ابوالفضل. (۱۴۰۱). صلاحیت‌های تدریس مبتنی بر فاوا برای آموزش ریاضی: مقایسه دانش‌جو معلمان با مهارت‌آموزان ماده ۲۸. *فصلنامه علمی-پژوهشی تعلیم و تربیت*، ۳۸(۴)، ۲۹-۴۸. <http://qjoe.ir/article-۲۳۷۰-fa.html>.
- کرمعلیان، حسن؛ هرندی، رضا؛ عبادی، میرحسین. (۱۳۹۲). بررسی مشکلات اجرای ارزشیابی کیفی توصیفی از دیدگاه معلمان و مدیران مدارس ابتدایی. *رویکردهای نوین آموزشی*، ۸(۲)، ۷۳-۹۲. [SID. https://sid.ir/paper/504320/fa](https://sid.ir/paper/504320/fa).
- مرادپور کرنکان، لیلان. (۱۴۰۳). شناسایی و اولویت‌بندی مهارت‌های معلمی در ارزشیابی کیفی درس ریاضی. اولین کنفرانس بین‌المللی مطالعات کاربردی در فرایندهای تعلیم و تربیت، بندرعباس. <https://civilica.com/doc/2246854>.
- مرتضوی، مرتضی؛ گویا، زهرا؛ ملکی، حسن؛ غلام آزاد، سهیلا. (۱۳۹۹الف). ضرورت تدوین معیارهای روبریک کیفی برای اجرای موفقیت‌آمیز ارزشیابی توصیفی درس ریاضی دوره ابتدایی. *مطالعات برنامه درسی ایران*، ۱۵(۵۷)، ۵-۴۰.
- مرتضوی، مرتضی؛ ملکی، حسن؛ گویا، زهرا؛ غلام آزاد، سهیلا؛ حسن، محمد. (۱۳۹۹ب). چالش‌های اجرای ارزشیابی توصیفی درس ریاضی در دوره ابتدایی. *نظریه و عمل در برنامه درسی*، ۸(۱۵)، ۱۹۳-۲۲۸. [SID. https://sid.ir/paper/952615/fa](https://sid.ir/paper/952615/fa).
- مکوندی، محمد. (۱۴۰۰). بررسی دیدگاه معلمان در رابطه با مشکلات اجرایی طرح ارزشیابی کیفی-توصیفی در مقطع ابتدایی و ارائه راهکارهایی برای برطرف کردن آن‌ها. *پژوهش در مسائل تعلیم و تربیت اسلامی*، ۲۹(۵۱)، ۱۸۱-۲۰۴. <https://sid.ir/paper/1072818/fa>.
- وکیلی، محمدمسعود؛ حیدرنیا، علیرضا؛ نیکنامی، شمس‌الدین. (۱۳۹۱). طراحی و روان‌سنجی ابزار سنجش مهارت‌های ارتباطی (ا. س. م. ا). میان‌فردی در جمعیت رابطان سلامت شهر زنجان. *حیات*، ۱۸(۱)، ۵-۱۹. <http://hayat.tums.ac.ir/article-1-34-fa.html>.

References

- Ahmadi, F., Nasiriani, K., & Abazari, P. (2009). Delphi technique: A tool in research. *Iranian Journal of Medical Education*, 8(1), 175–185. <https://sid.ir/paper/441616/fa> [In Persian]
- Alavi, S. M., & Dashtestani, R. (2015). Obstacles to the implementation of formative assessment in English for academic purposes instruction in Iran. *Journal of Language Teaching and Research*, 6(3), 499–508. <https://doi.org/10.17507/jltr.0603.05>
- Alsubaie, M. A. (2016). Curriculum development: Teacher involvement in curriculum development. *Education and Practice*, 7(9), 105–107. <https://files.eric.ed.gov/fulltext/EJ1095725.pdf>
- American Federation of Teachers, National Council on Measurement in Education, & National Education Association. (1990). *Standards for teacher competence in the educational assessment of students*. Washington, DC.
- Amini, Z. (2018). Improving the quality of mathematics lesson evaluation using technology. The Second Conference on Mathematics Education and Application, Kermanshah. <https://civilica.com/doc/824693/> [In Persian]
- An, S., & Wu, Z. (2012). Enhancing mathematics teachers' knowledge of students' thinking from assessing and analyzing misconceptions in homework. *International Journal of Science and Mathematics Education*, 10, 717–753. <https://doi.org/10.1007/s10763-011-9324-x>
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Brookhart, S. M. (2001). *The standards and classroom assessment research*. Paper presented at the annual meeting of the American Association of Colleges for Teacher Education, Dallas, TX.
- Brookhart, S. M. (2024). Educational assessment knowledge and skills for teachers revisited. *Education Sciences*, 14(7), 751. <https://doi.org/10.3390/educsci14070751>
- Calderon, A. (2011). Editor challenges and paradigms for institutional research in a globalized higher education system. Keynote address, Fourth Conference of UK and Ireland Institutional Research.
- Dandis, P. (2013). *Assessment of student learning: Approaches and tools*. Tehran University Press.
- Doria, M. G., Todesco, F., & Bagattini, C. (2023). Formative assessment in higher education: An exploratory study within programs for professionals in education. *Frontiers in Education*, 9, 1366215. <https://doi.org/10.3389/educ.2024.1366215>
- Ebadi, M., Seraji, F., & Bakhtiari, A. (2023). Competencies in teaching math with an IT-based approach: A comparison between student-teachers and article 28 trainees. *QJOE*, 38(4), 29–48. <http://qjoe.ir/article-1-2370-fa.html> [In Persian]
- Estaji, M. (2021). Examining teachers' assessment literacy in EFL classrooms: The effect of educational background. *Quarterly of Educational Measurement*, 12(45), 7–38. <https://doi.org/10.22054/jem.2022.59105.2153> [In Persian]
- Estaji, M., & Fassihi, S. (2016). On the relationship between the implementation of formative assessment strategies and Iranian EFL teachers' self-efficacy: Do gender and experience make a difference? *Innovation in Language Learning and Teaching*, 10(3), 235–251. <https://doi.org/10.1080/17501229.2015.1080692>
- Gallagher, M. A., Folger, T. D., Walkowiak, T. A., Wilhelm, A. G., & Zelkowski, J. (2025). Measuring mathematics teaching quality: The state of the field and a call for the future. *Education Sciences*, 15(9), 1158. <https://doi.org/10.3390/educsci15091158>
- Ghasemi, F., & Hashemi, M. (2019). The language assessment literacy needs of Iranian EFL teachers with a focus on reformed assessment policies. *Language Testing in Asia*, 9(1), 3. <https://doi.org/10.1186/s40468-019-0078-7>
- Hajizadeh, E. Asghari, M. (2012). *Methods and statistical analysis by looking at research method in biology and health sciences* (first edition). Tehran: Jahad Daneshgahi. [In Persian]
- Hassani, M., & Ahmadi, G. A. (2012). Groundwork for implementing qualitative descriptive assessment in elementary schools in Tehran City. *Journal of Educational Innovations*, 23(6), 85–123. <SID:https://sid.ir/paper/75548/fa> [In Persian]

- Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction*, 26(4), 430–511. <https://doi.org/10.1080/07370000802177235>
- Hoseini, Z., & Shabani, E. A. (2025). Examining the assessment literacy of teachers of Persian to speakers of other languages. *Journal of Teaching Persian to Speakers of Other Languages (JTPSOL)*, 14(1), 95–127. <https://doi.org/10.30479/jtpsol.2025.21944.1707>. [In Persian]
- Hosseini, M., Tehranian, A., Behzadi, M. H., & Alam Al-Hadaei, S. H. (2024). Lived experiences of sixth grade teachers and students from the qualitative assessment of math lessons, designing a scale of teacher skills in the qualitative evaluation of math lessons. *Journal of Psychological Science*, 23(137), 247–265. <https://doi.org/10.52547/JPS.23.137.1243> [In Persian]
- Hsu, T. H., & Chen, Y. C. (1996). A fuzzy Delphi approach. *Chinese Fuzzy Systems Association*.
- Hu, L., & Bentler, P. M. (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>
- James, R., McInnis, C., & Devlin, M. (2002). *Assessing learning at Australian universities*. Melbourne: Center for the Study of Higher Education.
- Jazairi, I. (2017). *Investigating the effect of process evaluation on the academic progress of mathematics subjects of secondary school students in Gotvand city*. Master's thesis, Farhangian University. [In Persian]
- Karamalian, H., Jafari Harandi, R., & Ebadi, H. (2013). A study of the qualitative descriptive evaluation problems from primary school teachers' and managers' point of view. *New Educational Approaches*, 8(2), 73–92. [SID. https://sid.ir/paper/504320/fa](https://sid.ir/paper/504320/fa) [In Persian]
- Kaya, İ. (2018). Examination of preschool teachers' opinion on alternative assessment. *Universal Journal of Educational Research*, 6(10), 2294–2299. <https://DOI:10.13189/ujer.2018.061028>
- Looney, A., Cumming, J., van Der Kleij, F., & Harris, K. (2018). Reconceptualising the role of teachers as assessors: Teacher assessment identity. *Assessment in Education: Principles, Policy & Practice*, 25, 442–467.
- López-Martín, E., Gutiérrez-de-Rozas, B., González-Benito, A. M., & Expósito-Casas, E. (2023). Why do teachers matter? A meta-analytic review of how teacher characteristics and competencies affect students' academic achievement. *International Journal of Educational Research*, 120, 102229. <https://doi.org/10.1016/j.ijer.2023.102229>
- Makvandi, M. (2021). An investigation of the teachers' views regarding the implementation problems of the qualitative-descriptive evaluation plan in the primary schools, and proposing solutions. *Islamic Education*, 29(51), 181–204. <https://sid.ir/paper/1072818/fa> [In Persian]
- Mertler, C. A. (2009). Teachers' assessment knowledge and their perceptions of the impact of classroom assessment professional development. *Improving Schools*, 12(1), 101–113. <https://doi.org/10.1177/1365480209105575>
- Michigan Assessment Consortium. (2020). *Assessment literacy standards*. <https://www.michiganassessmentconsortium.org/assessment-literacy-standards/>
- Miri, M. A., & Miri, M. A. (2022). Afghan English teachers' and students' perceptions of formative assessment: A comparative analysis. *Cogent Education*, 9(1), 2107297. <https://doi.org/10.1080/2331186X.2022.2107297>
- Moradpoor Kerenkan, L. (2024). Identifying and prioritizing teacher skills in qualitative evaluation of mathematics lessons. Fourth National Conference and First International Conference on Applied Studies in Educational Processes, Hormozgan. <https://civilica.com/doc/2246854> [In Persian]
- Mortazavi, M., Gooya, Z., Maleki, H., & Gholamazad, S. (2020B). The challenges of implementing descriptive evaluation for mathematics at the elementary school from teachers' perspective. *Theory and Practice in the Curriculum*, 8(15), 193–228. <https://sid.ir/paper/952615/fa> [In Persian]
- Mortazavi, M., Maleki, H., Gooya, Z., Gholamazad, S., & Hasani, M. (2020A). The necessity of developing a qualitative rubric to implement the descriptive evaluation of mathematics at the elementary school in Iran. *Journal of Curriculum Studies*, 15(57), 5–40. <https://dor.isc.ac/dor/20.1001.1.17354986.1399.15.57.1.9> [In Persian]

- Nabavizadeh, H., Gholami, M., & Karimi, A. (2019). The role of mathematics performance in students' academic achievement: An educational perspective. *Journal of Mathematics Education*, 14(3), 45–62.
- Naghdipour, B. (2016). *Writing assessment literacy: Exploring the beliefs and practices of Iranian EFL teachers*. Doctoral dissertation, University of Ottawa. <https://ruor.uottawa.ca/handle/10393/36348>
- Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: An example, design considerations and applications. *Information & Management*, 42(1), 15–29. <https://doi.org/10.1016/j.im.2003.11.002>
- Ozan, C., & Kincal, R. Y. (2023). The effects of formative assessment on academic achievement, attitudes toward the lesson, and self-regulation skills of middle school students. *International Electronic Journal of Elementary Education*, 15(4), 1–15. <https://doi.org/10.26822/iejee.2023.045>
- Pajares, F. (2002). Gender and perceived self-efficacy in self-regulated learning. *Theory Into Practice*, 41(2), 116–125. https://doi/abs/10.1207/s15430421tip4102_8
- Parsaeian, A., Rezaei, M., & Ahmadi, P. (2023A). Developing teachers' assessment literacy in the Iranian educational system: A qualitative study. *Educational Innovations*, 22(86), 51–74. <https://doi.org/10.22054/jti.2023.73145.3562>
- Parsaian, M., Rafieyan, A., & Khaleghi, F. (2023B). Assessment literacy: A new approach to effective teaching and learning. *Educational Research*, 47(3), 234–245. <https://doi.org/10.1016/j.jedures.2023.02.015>
- Pastore, S., & Andrade, H. L. (2019). Teacher assessment literacy: A three-dimensional model. *Teaching and Teacher Education*, 84, 128–138. <https://doi.org/10.1016/j.tate.2019.05.003>
- Rahimi, T., & Sedaghatkhan, A. M. (2022). The relationship between teacher skills in the qualitative evaluation of math lessons with math anxiety, academic procrastination and thinking style. *Iranian Journal of Educational Research*, 1(2), 56–72. <https://doi.org/10.52547/ijer.1.2.56>
- Razi, J., Emam Jomea, M. R., & Ahmadi, G. (2017). The extracting the dimensions of effective teachers with regard to provide a conceptual framework of education and teacher training. *Research in Teacher Education*, 1(1), 13–46. [SID. https://sid.ir/paper/269462/fa](https://sid.ir/paper/269462/fa) [In Persian]
- REL Central. (2025). Formative assessment and elementary school student academic achievement: A review of the evidence. *Regional Educational Laboratory Central, U.S. Department of Education*. https://ies.ed.gov/ncee/edlabs/regions/central/pdf/REL_2025001.pdf
- Sahito, Z., Özer, Ö., Abro, G. A., & Junejo, K. A. (2024). Perception of the elementary mathematics teachers about assessment for learning: A case study of Sukkur IBA community colleges, Sindh, Pakistan. *Frontiers in Education*, 9, 1430318. <https://doi.org/10.3389/feduc.2024.1430318>
- Sarmad, Z., Bazargan, A., & Hejazi, E. (2019). *Research methods in behavioral sciences*. Tehran: Ageh Publish. [In Persian]
- Schilling, S. G., & Hill, H. C. (2007). Assessing measures of mathematical knowledge for teaching: A validity argument approach. *Measurement*, 5(2–3), 70–80.
- Shamsolahi, M., & Samie Zafarghandi, M. (2017). The contribution of professional skills of elementary school teachers to descriptive evaluation. *Preschool and Elementary School Studies*, 3(9), 23–47. <https://doi.org/10.22054/soece.2020.37429.1189> [In Persian]
- Shaughnessy, M., Garcia, N. M., O'Neill, M. K., Selling, S. K., Willis, A. T., Wilkes, C. E., II, Salazar, S. B., & Ball, D. L. (2021). Formatively assessing prospective teachers' skills in leading mathematics discussions. *Educational Studies in Mathematics*, 108(3), 451–472. <https://doi.org/10.1007/s10649-021-10070-z>
- Shi, D., Lee, T., & Maydeu-Olivares, A. (2019). Understanding the model size effect on SEM fit indices. *Educational and Psychological Measurement*, 79(2), 310–334. <https://doi.org/10.1177/0013164418783530>
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14. <https://doi.org/10.3102/0013189X015002004>
- Simon, A., Smith, J., & Johnson, K. (2017). Assessment of student learning outcomes in higher education. *Educational Evaluation*, 42(3), 215–230. <https://doi.org/10.1007/s11092-017-9256-4>
- Tabrizi, M. (2014). Qualitative content analysis from the perspective of deductive and inductive approaches. *Social Sciences*, 21(64), 105–138. <https://doi.org/10.22054/qjss.2014.344> [In Persian]

- Unal, S., & Unal, E. (2019). Understanding student comprehension through formative assessment. *International Assessment Tools in Education*, 6(2), 145–158. <https://doi.org/10.21449/ijate.567890>
- Vakili, M. M., Hidarnia, A. R., & Niknami, S. (2012). Development and psychometrics of an interpersonal communication skills scale (A.S.M.A) among Zanjan health volunteers. *Hayat*, 18(1), 5–19. <http://hayat.tums.ac.ir/article-۱-۳۴-fa.html> [In Persian]
- Van den Heuvel-Panhuizen, M., Veldhuis, M., & Sangari, A. A. (2021). Teachers' use of descriptive assessment in primary school mathematics education in Iran. *Education Sciences*, 11(3), 100. <https://doi.org/10.3390/educsci11030100>
- Vignoles, A. & Meschi, E. (2010). The determinants of non-cognitive and cognitive schooling outcomes [Report to the Department of Children, Schools and Families. CEE Special Report 004]. *London Center for the Economics of Education, London School of Economics*.
- Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualising teachers' assessment literacies in an era of curriculum and assessment reform. *The Australian Educational Researcher*, 40, 241–256. <https://doi.org/10.1007/s13384-013-0109-8>
- Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149–162. <https://doi.org/10.1016/j.tate.2016.05.010>
- Yan, Z., Li, Z., Panadero, E., Yang, M., Yang, L., & Lao, J. (2023). Are teachers literate in formative assessment? The development and validation of the Teacher Formative Assessment Literacy Scale. *Studies in Educational Evaluation*, 77, 101261. <https://doi.org/10.1016/j.stueduc.2022.101261>