



Article type:
Research Article

The Effects of Teachers' Professional Experience and Academic Background on Their Classroom Assessment and Evaluation Practices

Mohammad Reza Fathi*¹ 

1. *Corresponding author*, Department of Psychology and Counselling, Farhangian University, P.O. Box 14665-889, Tehran, Iran. E-mail: fathi48@cfu.ac.ir

ORCID

Mohammad Reza Fathi  <https://orcid.org/0009-0002-7131-5563>

ABSTRACT

Background and Objectives: Evaluation is core to teaching-learning and plays a strategic role in educational quality. Despite theoretical acceptance of modern approaches, the gap between teachers' positive attitudes and practice remains. Teachers' characteristics, including professional experience and educational background, influence this gap. This study aimed to investigate the impact of teachers' professional experience and educational background on their assessment and evaluation methods in the classroom.

Methods: This research employed a descriptive-survey design. The statistical population comprised teachers working in Tehran province during the 2023-2024 academic year, from which 277 individuals were selected using convenience sampling. Data were collected using a researcher-developed 25-item questionnaire. Construct validity was confirmed through exploratory factor analysis, and reliability was established with a Cronbach's alpha of 0.92 for the overall instrument, with subscale reliabilities ranging from 0.70 to 0.90. Data were analyzed using three-way multivariate analysis of variance.

Findings: The results showed that teaching level had a significant effect on traditional assessment methods ($p = 0.041$) and alternative assessment methods ($p = 0.030$). Field of study demonstrated a significant effect on alternative methods ($p = 0.011$). Middle school teachers exhibited the highest use of traditional methods (4.01), while elementary school teachers performed better in alternative methods (3.60). Teachers with degrees in educational sciences used alternative methods more frequently compared to those in other fields (3.77). Teaching experience had no significant effect on any dimension of assessment practices.

Conclusion: Teaching level and field of study are determining factors in teachers' application of assessment methods. Revisions to teacher education programs and targeted professional development courses are recommended. The findings have important theoretical and practical implications for expanding knowledge and empowering teachers.

Keywords:

Educational evaluation, teachers' professional experience, teachers' educational background, traditional assessment methods, alternative assessment methods

نوع مقاله: پژوهشی

تأثیر تجربه حرفه‌ای و پیشینه تحصیلی معلمان بر شیوه‌های سنجش و ارزشیابی آنان در کلاس درس

محمدرضا فتحی^{۱،*}

۱. نویسنده مسئول، گروه آموزش روان‌شناسی و مشاوره، دانشگاه فرهنگیان، صندوق پستی ۸۸۸-۵۶۴۴۱ تهران، ایران. رایانامه:

fathi48@cfu.ac.ir

چکیده

پیشینه و اهداف: ارزشیابی هسته مرکزی فرآیند یاددهی-یادگیری است و نقشی راهبردی در کیفیت آموزش دارد. با وجود پذیرش نظری رویکردهای نوین، شکاف میان نگرش مثبت معلمان و کاربرد عملی آن مشهود است. ویژگی‌های معلمان از جمله تجربه حرفه‌ای و پیشینه تحصیلی از عوامل مؤثر بر آن هستند. این مطالعه با هدف بررسی تأثیر تجربه حرفه‌ای و پیشینه تحصیلی معلمان بر شیوه‌های سنجش و ارزشیابی آنان در کلاس درس انجام شد.

روش‌ها: پژوهش حاضر توصیفی-پیمایشی است. جامعه آماری شامل معلمان استان تهران در سال تحصیلی ۱۴۰۲-۱۴۰۳ بود که ۲۷۷ نفر به روش نمونه‌گیری در دسترس انتخاب شدند. داده‌ها با پرسشنامه محقق‌ساخته ۲۵ گویه‌ای جمع‌آوری گردید. روایی سازه با تحلیل عاملی اکتشافی تأیید شد و پایایی کل ابزار با آلفای کرونباخ ۰/۹۲ و پایایی خرده‌مقیاس‌ها در بازه ۰/۷۰ تا ۰/۹۰ به دست آمد. داده‌ها با تحلیل واریانس چندمتغیری سهرامه تحلیل شدند.

یافته‌ها: نتایج نشان داد دوره تدریس بر روش‌های سنجش سنتی ($p = ۰/۰۴۱$) و روش‌های نوین سنجش ($p = ۰/۰۳۰$) تأثیر معنادار دارد. رشته تحصیلی بر روش‌های نوین تأثیر معنادار نشان داد ($p = ۰/۰۱۱$). معلمان متوسطه اول بیشترین استفاده را از روش‌های سنتی (۴/۰۱) داشتند، در حالی که معلمان ابتدایی در روش‌های نوین عملکرد بهتری (۳/۶۰) نشان دادند. معلمان علوم تربیتی در مقایسه با سایر رشته‌ها، استفاده بیشتری از روش‌های نوین داشتند (۳/۷۷). سابقه تدریس تأثیر معناداری بر هیچ‌یک از ابعاد ارزشیابی نشان نداد.

نتیجه‌گیری: دوره تدریس و رشته تحصیلی معلمان از عوامل تعیین‌کننده در کاربرد روش‌های سنجش و ارزشیابی هستند. بازنگری در برنامه‌های تربیت معلم و دوره‌های توسعه حرفه‌ای هدفمند توصیه می‌شود. یافته‌ها دلالت‌های نظری و عملی مهمی برای بسط دانش و توانمندسازی معلمان دارد.

واژه‌های کلیدی:

ارزشیابی آموزشی، تجربه حرفه‌ای معلمان، پیشینه تحصیلی معلمان، روش‌های سنتی سنجش، روش‌های جایگزین سنجش

مقدمه

سنجش^۱ و ارزشیابی^۲ آموزشی به عنوان هسته مرکزی فرآیند یاددهی-یادگیری شناخته شده است (Levy-Feldman, 2025). این حوزه نه تنها ابزار سنجش میزان تحقق اهداف آموزشی و قضاوت درباره عملکرد تحصیلی دانش‌آموزان است، بلکه نقشی راهبردی در بهبود کیفیت آموزش، بازنگری برنامه‌های درسی، ارتقای اثربخشی تدریس و تضمین پاسخگویی نظام‌های آموزشی ایفا می‌کند (سیف، ۱۴۰۲). تحول در نظریه‌های یادگیری از رفتارگرایی به شناخت‌گرایی و سپس رویکردهای اجتماعی-فرهنگی، به بازاندیشی اساسی در ماهیت، کارکرد و جایگاه ارزشیابی انجامیده است. این تحولات نشان داده‌اند که یادگیری فرایندی فعال، بافت‌مند^۳ و اجتماعی است و در نتیجه، ارزشیابی نیز باید همسو با این نگاه بازتعریف شود (Willis et al., 2013; Moura et al., 2021). در چنین چشم‌اندازی، ارزشیابی دیگر صرفاً اندازه‌گیری کمی دانش در پایان دوره نیست، بلکه فرآیندی مستمر، تعاملی و کیفی است که در متن آموزش جریان دارد و به هدایت و غنای آن کمک می‌کند (درخشان و غیاثوند، ۲۰۲۲؛ López-Pastor et al., 2013).

در پرتو این تحول پارادایمی، تمایز میان «ارزشیابی از یادگیری»^۴ و «ارزشیابی برای یادگیری»^۵ اهمیت فزاینده‌ای یافته است. ارزشیابی برای یادگیری، بر نقش سازنده و تنظیم‌گر ارزشیابی در جریان آموزش تأکید می‌کند و آن را بخشی جدایی‌ناپذیر از تدریس مؤثر می‌داند (Girgla, et al., 2021; Pastore & Andrade, 2019). در این چارچوب، ارزشیابی مستمر و تکوینی^۶ به عنوان هسته اصلی ارزشیابی برای یادگیری مطرح است؛ زیرا از طریق بازخورد مستمر، شفاف‌سازی انتظارات، مشارکت فعال دانش‌آموزان و اصلاح راهبردهای تدریس، به بهبود کیفیت یادگیری کمک می‌کند (Robert et al., 2024; Veugen et al., 2021). یافته‌های پژوهشی نشان داده‌اند که اجرای مؤثر ارزشیابی تکوینی با افزایش عمق یادگیری، انگیزش تحصیلی، درگیری شناختی و خودکارآمدی دانش‌آموزان همراه است (Yan & Carless, 2021; Bores-García et al., 2020). با این وجود، علی‌رغم پذیرش نظری گسترده این رویکرد در اسناد سیاستی و برنامه‌های درسی، اجرای عملی و پایدار آن در کلاس‌های درس با چالش‌های جدی مواجه است (Serin, 2015; Letina, 2015; Letina et al., 2025).

در مقابل، ارزشیابی سنتی^۷ که عمدتاً بر آزمون‌های مداد-کاغذی، سؤالات عینی، و ارزشیابی تراکمی پایانی استوار است، همچنان در بسیاری از نظام‌های آموزشی جهان سلطه دارد. مطالعات نشان می‌دهد که معلمان، حتی در صورت آشنایی با مبانی نظری ارزشیابی نوین، در عمل به روش‌هایی گرایش دارند که با ساختارهای رسمی، محدودیت‌های زمانی و انتظارات بیرونی سازگارتر است (Ohlsen, 2007; Çalışkan & Kaşıkçı, 2010). در نظام آموزشی ایران نیز شواهد حاکی از وجود شکاف میان نگرش مثبت معلمان به ارزشیابی کیفی-توصیفی و کاربست محدود آن در عمل کلاسی است؛ به گونه‌ای که آزمون‌های کتبی پایانی، پرسش شفاهی و تکالیف تکراری همچنان نقش محوری دارند (علی‌اکبری و همکاران، ۲۰۲۳؛ نراقی‌زاده و همکاران، ۱۴۰۱؛ رضایی و محمدعلیزاده، ۱۴۰۴). افزون بر این، خرم‌آبادی و دلاور (۱۳۸۸) با تأیید اثرات مثبت ارزشیابی جایگزین^۸ بر نگرش و پیشرفت تحصیلی، به موانع ساختاری و اجرایی در مسیر گسترش آن اشاره کرده‌اند. پژوهش‌های اخیر نیز نشان می‌دهند که معلمان با وجود آگاهی از مزایای ارزشیابی جایگزین، همچنان در عمل به روش‌های سنتی پایبند هستند (Arjuna et al., 2024; Suwartono et al., 2024)؛ وحیدی و کریمی، (۱۴۰۴).

1 . assessment

2 . evaluation

3 . contextual

4 . evaluation of learning

5 . evaluation for learning

6 . formative evaluation

7 . traditional evaluation

8 . alternative evaluation

این شکاف میان نظر و عمل محدود به یک کشور خاص نیست، بلکه الگویی مشترکی در پژوهش‌های بین‌المللی مشاهده می‌شود. در کشورهایی چون پاکستان، ترکیه، عمان و فیلیپین، استمرار غلبه روش‌های سنتی و استفاده محدود از راهبردهای جایگزین گزارش شده است (Hussain et al., 2019؛ Acar-Erdol & Yıldızlı et al., 2018؛ Alkharusi, 2011؛ Manigbas & De Luna, 2023). در مجارستان، فشار پاسخگویی ناشی از برنامه‌های ملی ارزشیابی، تمرکز بر نمره را تقویت کرده است (Tóth & Csapó, 2022). در عربستان سعودی و اندونزی نیز مقاومت در برابر تغییر و ضعف توانمندی‌های حرفه‌ای از موانع اصلی اجرای ارزشیابی اصیل معرفی شده‌اند (Alansari, 2023؛ Suwartono et al., 2024). حتی در نظام‌های آموزشی پیشرو همچون فنلاند نیز در شرایط خاص، چالش‌هایی در اجرای ارزشیابی کیفی گزارش شده است (Mäkipää et al., 2021؛ Mäkipää, 2023). این مجموعه شواهد نشان می‌دهد که چالش گذار از رویکردهای سنتی به نوین، ماهیتی جهانی دارد. در این راستا، Robert و همکاران (2024) با بررسی ادراک معلمان دوره متوسطه از شیوه‌های ارزشیابی کلاسی نشان دادند که ارزشیابی خود و همتا کمترین کاربرد را دارند و شیوه‌های ارزشیابی همچنان معلم‌محور هستند. همچنین Gyamfi و همکاران (2023) در مطالعه خود در غنا به چالش‌های متعدد معلمان در ارزشیابی کلاسی اشاره کرده و راهبردهایی مانند ارزشیابی عملکردی، پورتفولیو و خودارزشیابی را برای بهبود آن پیشنهاد داده‌اند.

در تبیین این وضعیت، پژوهش‌ها به عوامل زمینه‌ای متعددی اشاره کرده‌اند؛ از جمله تراکم بالای کلاس‌ها، کمبود زمان، فشار والدین، محدودیت منابع، و سیاست‌های آزمون‌محور که فضای عمل معلمان را محدود می‌سازد (Hussain et al., 2019؛ Gyamfi et al., 2023؛ وحیدی و کریمی، ۱۴۰۴؛ صالحی و همکاران، ۱۳۹۴). همچنین اثربخشی برنامه‌های توسعه حرفه‌ای زمانی افزایش می‌یابد که با ساختارهای مدرسه، باورهای معلمان و جو سازمانی همساز باشد (Didier et al., 2025؛ ناصر و محمودی، ۱۴۰۴). بنابراین، بستر نهادی و فرهنگی نقش مهمی در شکل‌گیری شیوه‌های ارزشیابی ایفا می‌کند. Al-Hassan (2024) نیز در پژوهش خود نشان داد که موانع اجرای ارزشیابی جایگزین از دیدگاه معلمان دوره اول، از جمله کمبود انگیزه‌های مادی و معنوی و تراکم بالای کلاس‌ها، بسیار زیاد است.

در کنار عوامل زمینه‌ای، ویژگی‌های فردی معلمان نیز به‌عنوان متغیرهای تعیین‌کننده مورد توجه فزاینده قرار گرفته‌اند. در این میان، «سواد ارزشیابی»^۱ مفهومی محوری است که به مجموعه دانش مفهومی، مهارت‌های عملی و نگرش‌های حرفه‌ای لازم برای طراحی، اجرا، تحلیل و تفسیر ارزشیابی اشاره دارد. Pastore (2023)، Hull and Vigh (2024)، Willis و همکاران (2013) بر ماهیت اجتماعی و بافت‌مند این سواد تأکید کرده‌اند و Pastore and Andrade (2019) آن را در سه بعد مفهومی، عملی و زمینه‌ای تبیین کرده‌اند. پژوهش‌ها نشان می‌دهد که سطح سواد ارزشیابی در بسیاری از بافت‌ها متوسط است و با کیفیت عملکرد ارزشیابی و خودکارآمدی تدریس رابطه مثبت دارد (Khan et al., 2019؛ Wang et al., 2022؛ Ullah, 2023).

با این حال، پرسش اساسی آن است که چه عواملی بر کاربرد اصول و روش‌های سنجش و ارزشیابی آموزشی معلمان اثرگذارند. تجربه حرفه‌ای^۲، دو متغیر سابقه تدریس^۳ و دوره آموزشی (ابتدایی یا متوسطه) را شامل می‌شود که در پژوهش حاضر مورد بررسی قرار گرفتند، از جمله متغیرهایی هستند که نتایج پژوهش درباره آن متناقض بوده است. برخی مطالعات نشان داده‌اند که معلمان باتجربه به دلیل آشنایی بیشتر با مدیریت کلاس و تنوع یادگیرندگان، در به‌کارگیری روش‌های متنوع ارزشیابی موفق‌ترند (Koloi-Keaikitse, 2012؛ Alkharusi, 2011). در مقابل، پژوهش‌هایی دیگر حاکی از آن است که معلمان کم‌سابقه یا معلمان دوره ابتدایی گرایش بیشتری به استفاده از راهبردهای ارزشیابی مستمر و سنجش جایگزین دارند (Korkmaz et al., 2025؛ Letina et al., 2025). همچنین برخی معلمان باتجربه، موانع بیشتری در اجرای ارزشیابی جایگزین ادراک می‌کنند (Al-Hassan, 2024). این یافته‌های ناهمسو، نیازمند بررسی

1 . evaluation literacy
2 . professional experience
3 . teaching experience

نظام‌مندتر است. Korkmaz و همکاران (2025) در پژوهش خود نشان دادند که معلمان با سابقه کمتر از ۵ سال، گرایش بیشتری به استفاده از روش‌های ارزشیابی جایگزین دارند و با افزایش سابقه، این گرایش کاهش می‌یابد. همچنین Didier و همکاران (2025) با بررسی عوامل مؤثر بر اجرای برنامه‌های توسعه حرفه‌ای معلمان نشان دادند که محدودیت‌های زمانی و ساختاری، موانع اصلی اجرای موفق این برنامه‌ها هستند.

پیشینه تحصیلی معلمان، به‌ویژه رشته تحصیلی آنان، سومین متغیر مهم در این پژوهش است. معلمان با رشته‌های گوناگون وارد حرفه می‌شوند و از نظر دانش تربیتی، آشنایی با نظریه‌های یادگیری و مهارت‌های سنجش تفاوت دارند. پژوهش ثامتی و همکاران (۱۴۰۲) نشان داد، نوع رشته تحصیلی دوره متوسطه بر عملکرد حرفه‌ای معلمان تاثیر ندارد. پژوهش دیبایی صابر (۱۳۹۷) نشان داد، بین دانش و معلومات حرفه‌ای معلمان و اثر بخشی آنان در فرایند یاددهی و یادگیری دانش آموزان رابطه وجود دارد. شواهدی از ارتباط حوزه تدریس و آموزش‌های مرتبط با ارزشیابی با عملکرد ارزشیابی و خودکارآمدی وجود دارد (Addow, 2023؛ Matovu & Zubairi, 2014؛ Alkharusi, 2011). اما بررسی تطبیقی اثر رشته تحصیلی بر ابعاد مختلف شیوه‌های ارزشیابی کلاسی همچنان محدود است.

گرچه متغیرهایی چون جنسیت، سطح تحصیلات و شرکت در دوره‌های ضمن خدمت نیز بررسی شده‌اند (Muhson et al., 2025؛ Matovu & Zubairi, 2014؛ Alkharusi, 2011). پژوهشی که به‌صورت همزمان و نظام‌مند تأثیر تجربه حرفه‌ای (سابقه و دوره تدریس) و پیشینه تحصیلی (رشته تحصیلی) را بر طیفی جامع از شیوه‌های ارزشیابی شامل ارزشیابی مستمر، روش‌های سنجش سنتی و روش‌های سنجش نوین و جایگزین بررسی کند، کمتر مشاهده می‌شود. پژوهش حاضر با هدف پاسخ به این خلأ، تأثیر سابقه تدریس، دوره تدریس و رشته تحصیلی را بر سه بعد شیوه‌های ارزشیابی کلاسی شامل ارزشیابی مستمر، روش‌های سنجش سنتی و روش‌های سنجش نوین بررسی می‌کند.

بر اساس مبانی نظری و با توجه به عوامل ساختاری و نهادی حاکم بر نظام آموزشی ایران، انتظار می‌رود معلمان دوره ابتدایی، به دلیل الزامات طرح ارزشیابی کیفی-توصیفی و سرمایه‌گذاری معاونت آموزش ابتدایی، نسبت به معلمان دوره متوسطه از ارزشیابی مستمر و روش‌های نوین بیشتری استفاده کنند و معلمان متوسطه کاربست بیشتری از روش‌های سنتی داشته باشند. همچنین پیش‌بینی می‌شود معلمان با رشته‌های مرتبط با علوم تربیتی در مقایسه با سایر رشته‌ها، گرایش بیشتری به روش‌های نوین و ارزشیابی مستمر نشان دهند. با توجه به نتایج متناقض پژوهش‌های پیشین درباره تأثیر سابقه تدریس بر شیوه‌های ارزشیابی، این متغیر در این پژوهش به صورت اکتشافی بررسی می‌شود. شایان ذکر است که در تمامی دوره‌های تحصیلی، شکاف میان وضع موجود و وضع مطلوب در کاربست روش‌های ارزشیابی پیش‌بینی می‌شود. پژوهش حاضر گامی در جهت تبیین نقش ویژگی‌های فردی معلمان در شکل‌دهی به شیوه‌های ارزشیابی کلاسی برمی‌دارد.

همچنین پیش‌بینی می‌شود معلمان با رشته‌های مرتبط با علوم تربیتی در مقایسه با سایر رشته‌ها، گرایش بیشتری به روش‌های نوین و ارزشیابی مستمر نشان دهند. با توجه به نتایج متناقض پژوهش‌های پیشین درباره تأثیر سابقه تدریس بر شیوه‌های ارزشیابی، این متغیر در پژوهش حاضر به صورت اکتشافی بررسی می‌شود. شایان ذکر است که در تمامی دوره‌های تحصیلی، شکاف میان وضع موجود و وضع مطلوب در کاربست روش‌های ارزشیابی پیش‌بینی می‌شود.

روش

پژوهش حاضر با هدف بررسی تأثیر تجربه حرفه‌ای و پیشینه تحصیلی معلمان بر شیوه‌های سنجش و ارزشیابی در کلاس‌های درس، با بهره‌گیری از یک طرح توصیفی-پیمایشی انجام شد. ماهیت توصیفی متغیرها و تمرکز پژوهش بر تحلیل تفاوت‌های گروهی، استفاده از این طرح را برای دستیابی به اهداف پژوهش موجه ساخت.

جامعه آماری پژوهش را کلیه معلمان شاغل در سال تحصیلی ۱۴۰۲-۱۴۰۳ استان تهران تشکیل دادند. حجم نمونه با استفاده از فرمول کوکران، ۲۷۰ نفر برآورد گردید. نمونه پژوهش با استفاده از روش نمونه‌گیری غیرتصادفی در دسترس انتخاب شد. بدین منظور، لینک پرسشنامه از طریق کانال‌های مجازی معلمان در شبکه‌های اجتماعی (شامل گروه‌های تلگرامی و ایتا با تأیید مدیران گروه‌ها) منتشر گردید. معلمان به صورت داوطلبانه در پژوهش شرکت کردند. برای جلوگیری از ورود پاسخ‌های تکراری، امکان ارسال پاسخ تنها یک بار از هر دستگاه فراهم شد. در مجموع، پاسخ‌های ۳۰۰ نفر دریافت شد که پس از پالایش داده‌ها و حذف پرسشنامه‌های ناقص یا نامعتبر، داده‌های معتبر ۲۷۷ برای تحلیل آماری مورد استفاده قرار گرفتند.

ابزار گردآوری داده‌ها پرسشنامه محقق‌ساخته برخط «شیوه‌های سنجش و ارزشیابی کلاسی» بود که در دو بخش طراحی شد. بخش نخست به گردآوری اطلاعات مربوط به تجربه حرفه‌ای و پیشینه تحصیلی معلمان اختصاص داشت و متغیرهای مستقل سابقه تدریس با دو سطح زیر ۲۰ سال، و بالای ۲۰ سال؛ دوره تدریس با سه سطح دوره ابتدایی، متوسطه اول، متوسطه دوم؛ و رشته تحصیلی با سه سطح علوم تربیتی، علوم انسانی و علوم پایه را در بر می‌گرفت. بخش دوم که هسته اصلی ابزار پژوهش را تشکیل می‌داد، شامل ۲۵ گویه بود که شیوه‌های سنجش و ارزشیابی آموزشی را می‌سنجید. این گویه‌ها بر اساس مبانی نظری سنجش کلاسی (Santrock, 2020) تدوین شد. روایی صوری و محتوایی پرسشنامه با اتکا به مبانی نظری معتبر و نظر متخصصان حوزه سنجش و ارزشیابی آموزشی تأیید شد. برای بررسی روایی سازه پرسشنامه، تحلیل عاملی اکتشافی^۱ با روش مؤلفه‌های اصلی و چرخش واریماکس^۲ انجام شد. شاخص KMO برابر با ۰/۹۰۷ و آزمون کرویت بارتلت معنادار بود ($\chi^2 = 3565/058$, $df = 300$, $p < 0/001$). نتایج تحلیل عاملی ساختار سه‌عاملی ارزشیابی مستمر، روش‌های سنجش سنتی و روش‌های سنجش جایگزین (نوبن) را تأیید کرد که ۴۹/۶ درصد از واریانس کل را تبیین می‌نمایند. بارهای عاملی گویه‌ها در دامنه ۰/۴۲ تا ۰/۸۸ قرار داشت که نشان‌دهنده روایی سازه مطلوب ابزار است. پایایی سه خرده‌مقیاس به ترتیب ۰/۷۰، ۰/۸۵ و ۰/۹۰ به دست آمد. پس از طراحی و اعتبارسنجی ابزار، فایل الکترونیکی آن تهیه شد و بعد از دریافت مجوزهای لازم از مدیران شبکه‌های مجازی، لینک پرسشنامه از طریق شبکه‌ها و کانال‌ها و گروه‌های مجازی مرتبط با جامعه معلمان (شامل گروه‌های تلگرامی، واتس‌اپی و ایتای معلمان استان تهران) منتشر شد. مشارکت در پژوهش کاملاً داوطلبانه بود و اصول اخلاق پژوهش، از جمله محرمانه‌ماندن اطلاعات پاسخ‌دهندگان و استفاده پژوهشی از داده‌ها، رعایت شد.

پاسخ‌دهندگان میزان استفاده از هر یک از روش‌های سنجش را در عمل تدریس خود بر اساس یک مقیاس لیکرت پنج‌درجه‌ای از «هرگز» (۱) تا «همیشه» (۵) مشخص کردند. توزیع گویه‌ها در سه خرده‌مقیاس به صورت زیر بود: ارزشیابی مستمر با ۵ گویه (۱، ۲، ۳، ۴ و ۵) که شامل سنجش آغازین و تکوینی می‌شود؛ روش‌های سنجش سنتی با ۹ گویه (۶، ۷، ۸، ۱۰، ۱۱، ۱۲، ۱۳، ۲۲ و ۲۴) که روش‌هایی چون آزمون‌های عینی و تشریحی، پرسش شفاهی و نمره‌دهی کمی و تحلیل آماری نتایج آزمون را دربرمی‌گرفت؛ و روش‌های سنجش جایگزین (نوبن) با ۱۱ گویه (۹، ۱۴، ۱۵، ۱۶، ۱۷، ۱۸، ۱۹، ۲۰، ۲۱، ۲۲ و ۲۵) که روش‌هایی مانند ارزشیابی عملکردی، پورتفولیو، خودارزشیابی و هم‌تارزشیابی، و استفاده از مقیاس کیفی شامل بود. نمره بالاتر در هر بعد نشان‌دهنده میزان کاربرد اصول و روش‌های مربوطه بود.

تجزیه و تحلیل داده‌ها با استفاده از نرم‌افزار SPSS نسخه ۲۶ در دو سطح آمار توصیفی و استنباطی صورت پذیرفت. در سطح آمار توصیفی، شاخص‌هایی نظیر فراوانی، میانگین، انحراف معیار برای توصیف ویژگی‌های نمونه و متغیرهای پژوهش محاسبه شد. در سطح آمار استنباطی، با توجه به وجود سه متغیر وابسته مرتبط و سه متغیر مستقل با سطوح متعدد، از تحلیل واریانس چندمتغیره^۳ سه‌عاملی برای بررسی تفاوت‌های گروهی در کاربرد شیوه‌های سنجش و ارزشیابی

1 . Exploratory factor analysis (EFA)

2 . Varimax rotation

3 . Multivariate Analysis of Variance (MANOVA)

توسط معلمان بر اساس سابقه تدریس، دوره تدریس و رشته تحصیلی معلمان استفاده گردید. پیش از اجرای آزمون، مفروضه‌های آن بررسی شد. با توجه به معناداری آزمون ام باکس^۱ ($p=0/002$) که نشان دهنده نقض مفروضه همگنی ماتریس‌های واریانس-کوواریانس^۲ بود، از آماره مقاوم اثر پیلای^۳ برای تفسیر نتایج چندمتغیره استفاده شد. مفروضه نرمالیتی چندمتغیره برقرار بود. همگنی واریانس‌ها با آزمون لون^۴ بررسی و تأیید شد ($p<0/05$). در مواردی که اثر چندمتغیره معنادار گزارش شد، به‌منظور شناسایی منبع تفاوت‌ها، آزمون‌های تعقیبی تک‌متغیره^۵ شامل آزمون t مستقل و تحلیل واریانس یک‌طرفه (آنووا)^۶ اجرا گردید. برای مقایسه زوجی میانگین‌ها در متغیرهای مستقل با بیش از دو سطح، با توجه به همگنی واریانس‌ها و نابرابری حجم گروه‌ها، از آزمون تعقیبی توکی^۷ استفاده شد (Tabachnick & Fidell, 2019). همچنین برای بررسی اندازه اثر^۸ تفاوت‌های گروهی، ضریب اتا سهمی^۹ محاسبه شد. سطح معناداری آزمون‌های آماری 0/05 در نظر گرفته شد.

یافته‌ها

به منظور بررسی ویژگی‌های جمعیت‌شناختی نمونه و همچنین شاخص‌های مرکزی و پراکندگی متغیرهای وابسته بر حسب سطوح متغیرهای مستقل، از شاخص‌های توصیفی استفاده شد. جدول ۱ فراوانی معلمان شرکت‌کننده در پژوهش به تفکیک متغیرهای مستقل دوره تدریس، سابقه تدریس و رشته تحصیلی را نشان می‌دهد.

جدول ۱ فراوانی و درصد شرکت‌کنندگان پژوهش به تفکیک متغیرهای مستقل

دوره تدریس	رشته تحصیلی			کل
	علوم تربیتی	علوم انسانی	علوم پایه	
دوره اول ابتدایی	زیر ۲۰ سال	۸۰	۱۴	۱۰۸
	بالای ۲۰ سال	۳۸	۶	۴۸
	کل	۱۱۸	۲۰	۱۵۶
متوسطه اول	زیر ۲۰ سال	۴	۸	۲۲
	بالای ۲۰ سال	۴	۱۵	۳۱
	کل	۸	۲۳	۵۳
متوسطه دوم	زیر ۲۰ سال	۵	۹	۳۴
	بالای ۲۰ سال	۲	۱۵	۳۴
	کل	۷	۲۴	۶۸
کل	زیر ۲۰ سال	۸۹	۳۱	۱۶۴
	بالای ۲۰ سال	۴۴	۳۶	۱۱۳
	کل	۱۳۳	۶۷	۲۷۷

- 1 . M-Box test
- 2 . homogeneity of variance-covariance matrices
- 3 . Pillai's Trace
- 4 . Leven's test
- 5 . Univariate post hoc tests
- 6 . One-way analysis of variance (ANOVA)
- 7 . Tukey
- 8 . Effect size
- 9 . Partial eta coefficient

جدول ۲ شاخص‌های توصیفی میانگین، انحراف معیار، کجی و کشیدگی متغیرهای وابسته را به تفکیک متغیرهای مستقل و سطوح آنها نشان می‌دهد. دامنه کجی بین ۲- تا ۲+ و دامنه کشیدگی بین ۷- تا ۷+ نشان دهنده توزیع نرمال داده‌ها است.

جدول ۱. شاخص‌های توصیفی متغیرهای وابسته بر حسب دوره تدریس، سابقه تدریس و رشته تحصیلی

متغیر مستقل	سطح	ارزشیابی مستمر		روش‌های سنجش سنتی		روش‌های سنجش جایگزین	
		میانگین	کجی	میانگین	کجی	میانگین	کجی
		(انحراف معیار)	(کشیدگی)	(انحراف معیار)	(کشیدگی)	(انحراف معیار)	(کشیدگی)
دوره	ابتدایی	۳/۹۱ (۰/۷۸)	۰/۵۰ (-۰/۴۱)	۳/۵۹ (۰/۸۴)	۰/۳۲ (-۰/۲۰)	۳/۶۲ (۰/۸۵)	۰/۲۷ (-۰/۳۴)
تدریس	متوسطه اول	۳/۶۰ (۰/۷۶)	۰/۶۶ (-۰/۴۵)	۳/۹۵ (۰/۸۷)	۱/۱۸ (-۱/۵۱)	۳/۲۸ (۰/۹۹)	۰/۲۸ (-۰/۸۵)
	متوسطه دوم	۳/۵۸ (۰/۷۳)	۰/۴۶ (-۰/۴۶)	۳/۵۵ (۰/۸۳)	۰/۴۶ (-۰/۱۳)	۳/۰۷ (۰/۹۱)	۰/۰۵ (-۰/۳۰)
سابقه	زیر ۲۰ سال	۳/۷۹ (۰/۷۷)	۰/۳۶ (-۰/۴۶)	۳/۵۹ (۰/۸۸)	۰/۵۱ (-۰/۰۴)	۳/۴۴ (۰/۹۵)	۰/۲۳ (-۰/۵۴)
تدریس	بالای ۲۰ سال	۳/۷۴ (۰/۷۹)	۰/۵۹ (-۰/۶۹)	۳/۷۳ (۰/۸۰)	۰/۳۵ (-۰/۴۶)	۳/۳۹ (۰/۸۸)	۰/۰۳ (-۰/۴۹)
رشته	علوم تربیتی	۳/۹۲ (۰/۷۹)	۰/۵۰ (-۰/۴۲)	۳/۶۶ (۰/۸۴)	۰/۲۷ (-۰/۳۳)	۳/۶۷ (۰/۹۰)	۰/۴۱ (-۰/۱۸)
تحصیلی	علوم انسانی	۳/۵۴ (۰/۸۲)	۰/۷۱ (-۰/۸۲)	۳/۶۱ (۰/۸۷)	۰/۶۰ (-۰/۰۵)	۳/۲۲ (۰/۹۳)	۰/۰۹ (-۰/۴۶)
	علوم پایه	۳/۷۱ (۰/۶۸)	۰/۱۱ (-۰/۴۳)	۳/۶۶ (۰/۸۶)	۰/۷۳ (-۰/۴۲)	۳/۱۶ (۰/۸۵)	۰/۲۴ (-۰/۴۰)
کل		۳/۷۷ (۰/۷۸)	۰/۴۶ (-۰/۰۴)	۳/۶۵ (۰/۸۵)	۰/۴۸ (-۰/۰۶)	۳/۴۲ (۰/۹۲)	۰/۱۳ (-۰/۵۴)

به منظور بررسی تأثیر دوره تدریس، سابقه تدریس و رشته تحصیلی بر ترکیب خطی متغیرهای وابسته، از تحلیل واریانس چندمتغیره سه‌عاملی^۱ استفاده شد. پیش از انجام تحلیل اصلی، مفروضه‌های این آزمون بررسی شد. نتایج آزمون ام‌باکس جهت بررسی همگنی ماتریس‌های واریانس-کوواریانس در سطح معنی‌داری کمتر از ۰/۰۵ معنی‌دار بود [$M=177/33$ ، $F(96, 3416/35)=1/465$ ، $p=0/002$] که نشان‌دهنده نقض این مفروضه است. بنابراین، برای تفسیر نتایج چند متغیره از آماره مقاوم اثر پیلای^۲ استفاده شد (Tabachnick & Fidell, 2019). برای بررسی همگنی واریانس‌های خطا از آزمون لون استفاده شد. نتایج این آزمون برای هر سه متغیر وابسته معنادار نبود ($p>0/05$)، بنابراین مفروضه برابری واریانس‌های خطا برقرار است. نتایج آزمون‌های چندمتغیره بر اساس آماره اثر پیلای در جدول ۳ ارائه شده است.

جدول ۳. نتایج تحلیل واریانس چندمتغیره (مانووا) تأثیر ویژگی‌های معلمان بر مؤلفه‌های ارزشیابی

منبع اثر	مقدار اثر پیلای	F	df فرضیه	df خطا	p	مجذور η^2 جزئی
دوره تدریس	۰/۰۸۶	۳/۸۶	۶	۵۱۶	<۰/۰۰۱	۰/۰۴۳
سابقه تدریس	۰/۰۱۰	۰/۸۹	۳	۲۵۷	۰/۴۴۶	۰/۰۱۰
رشته تحصیلی	۰/۰۴۸	۲/۱۰	۶	۵۱۶	۰/۰۵۲	۰/۰۲۴
دوره تدریس × سابقه تدریس	۰/۰۱۲	۰/۵۴	۶	۵۱۶	۰/۷۸۰	۰/۰۰۶
دوره تدریس × رشته تحصیلی	۰/۰۷۷	۱/۷۱	۱۲	۷۷۷	۰/۰۶۰	۰/۰۲۶
سابقه تدریس × رشته تحصیلی	۰/۰۱۷	۰/۷۶	۶	۵۱۶	۰/۶۰۵	۰/۰۰۹
دوره تدریس × سابقه تدریس × رشته تحصیلی	۰/۰۵۳	۱/۱۷	۱۲	۷۷۷	۰/۳۰۰	۰/۰۱۸

$p<0/05$

همانطور که در جدول ۳ مشاهده می‌شود، بین دوره‌های مختلف تدریس از نظر ترکیب متغیرهای وابسته تفاوت معنی‌داری وجود دارد [$Pillai's Trace=0/086$ ، $F(6, 516)=3/86$ ، $p=0/001$ ، $\eta^2 p=0/043$]. اثر اصلی رشته تحصیلی بر ترکیب

1 . Three-Way MANOVA

2 . Pillai's Trace

متغیرهای وابسته در آستانه معناداری قرار داشت [Pillai's Trace=0/048، $F(6, 516) = 2/098$ ، $p=0/024$ ، $\eta^2p=0/052$]. اثر اصلی سابقه تدریس بر ترکیب متغیرهای وابسته معنی‌دار نبود [Pillai's Trace=0/010، $p=0/89$]. همچنین هیچ‌یک از اثرات تعاملی دوگانه و سه‌گانه معنی‌دار نشد ($p>0/05$). از آنجا که اثرات اصلی دوره تدریس و رشته تحصیلی معنادار (یا در آستانه معناداری) بود، برای تعیین دقیق‌تر تأثیر این متغیرها بر هر یک از متغیرهای وابسته، از آزمون‌های تک‌متغیره استفاده شد. نتایج این تحلیل‌ها در جدول ۴ ارائه شده است.

جدول ۴. نتایج آزمون‌های تک‌متغیره (آنووا) برای اثرات اصلی

منبع اثر	متغیر وابسته	مجموع مجزورات	df	میانگین مجزورات	F	p	مجدور η^2 جزئی
دوره تدریس	ارزشیابی مستمر	3/15	2	1/57	2/71	0/069	0/020
	روش‌های سنجش سنتی	4/71	2	2/35	3/25	0/041	0/024
	روش‌های سنجش جایگزین	5/50	2	2/75	3/56	0/030	0/027
سابقه تدریس	ارزشیابی مستمر	0/08	1	0/08	0/13	0/717	0/001
	روش‌های سنجش سنتی	1/05	1	1/05	1/45	0/229	0/006
	روش‌های سنجش جایگزین	0/39	1	0/39	0/51	0/478	0/002
رشته تحصیلی	ارزشیابی مستمر	0/41	2	0/21	0/35	0/703	0/003
	روش‌های سنجش سنتی	1/61	2	0/81	1/12	0/330	0/009
	روش‌های سنجش جایگزین	7/16	2	3/58	4/63	0/011	0/035

$p<0/05$

نتایج جدول ۴ نشان می‌دهد که دوره تدریس بر روش‌های سنتی سنجش ($p=0/041$) و روش‌های جایگزین سنجش ($p=0/030$) تأثیر معنادار دارد. اما بر ارزشیابی مستمر تأثیر معناداری ندارد ($p=0/069$). همچنین رشته تحصیلی بر روش‌های سنجش جایگزین تأثیر معنادار دارد ($p=0/011$)، اما بر ارزشیابی مستمر ($p=0/703$) و روش‌های سنتی ($p=0/330$) تأثیر معناداری ندارد. سابقه تدریس بر هیچ‌یک از سه متغیر وابسته تأثیر معناداری نداشت.

برای شناسایی دقیق منبع تفاوت‌ها، با توجه به همگنی واریانس‌ها، از آزمون تعقیبی توکی برای متغیرهای مستقل با بیش از دو سطح (دوره تدریس و رشته تحصیلی) استفاده شد. نتایج آزمون توکی برای مقایسه زوجی گروه‌های دوره تدریس نشان داد که در مؤلفه «روش‌های سنجش سنتی»، تفاوت میانگین بین معلمان دوره ابتدایی با معلمان متوسطه اول ($p=0/022$ ، $MD = -0/362$) و بین معلمان متوسطه اول با معلمان متوسطه دوم ($p=0/028$ ، $MD = 0/402$) معنادار بود. در مؤلفه «روش‌های سنجش جایگزین»، تفاوت میانگین بین معلمان متوسطه اول ($p=0/045$ ، $MD = 0/336$) و با معلمان متوسطه دوم ($p<0/001$ ، $MD = 0/552$) معنادار بود. همچنین نتایج آزمون تعقیبی توکی برای رشته تحصیلی نشان داد که در مؤلفه «روش‌های سنجش جایگزین»، تفاوت میانگین بین معلمان دارای رشته تحصیلی علوم تربیتی با معلمان با رشته تحصیلی علوم انسانی ($p=0/002$ ، $MD = 0/449$) و با معلمان دارای رشته علوم پایه ($p<0/001$ ، $MD = 0/512$) معنادار بود. با توجه به عدم معناداری اثر اصلی سابقه تدریس و کلیه اثرات تعاملی، از ارائه تحلیل‌های تکمیلی برای این متغیرها خودداری می‌شود.

جدول ۵ مقادیر میانگین‌های تعدیل‌یافته^۱ و خطای استاندارد را برای اثرات معنادار دوره تدریس و رشته تحصیلی بر مولفه‌های روش‌های سنجش سنتی و روش‌های سنجش جایگزین نشان می‌دهد.

1 . estimated marginal means

جدول ۵. میانگین‌های تعدیل‌یافته و خطای استاندارد برای اثرات معنادار

متغیر وابسته	متغیر مستقل	سطح	میانگین تعدیل‌یافته	خطای استاندارد
روش‌های سنجش سنتی	دوره تدریس	دوره ابتدایی	۳/۶۳	۰/۱۱
		متوسطه اول	۴/۰۱	۰/۱۳
		متوسطه دوم	۳/۵۸	۰/۱۴
روش‌های سنجش جایگزین	دوره تدریس	دوره ابتدایی	۳/۶۰	۰/۱۴
		متوسطه اول	۴/۴۷	۰/۱۱
		متوسطه دوم	۳/۱۱	۰/۱۴
روش‌های سنجش جایگزین	رشته تحصیلی	علوم تربیتی	۳/۷۷	۰/۱۶
		علوم انسانی	۳/۲۱	۰/۱۱
		علوم پایه	۳/۲۱	۰/۱۲

براساس اطلاعات این جدول، معلمان دوره متوسطه اول در مقایسه با معلمان دوره ابتدایی و متوسطه دوم، میانگین بالاتری را در مؤلفه «روش‌های سنتی» کسب کردند (۴/۰۱۲) در مقابل ۳/۶۲۶ و ۳/۵۸۲). همچنین معلمان دوره ابتدایی نسبت به معلمان متوسطه اول و دوم، میانگین بالاتری را در مؤلفه «روش‌های سنجش جایگزین» داشتند (۳/۶۰) در مقابل ۳/۴۷ و ۳/۱۱). اندازه اثر این تفاوت‌ها در دامنه کوچک قرار داشت (۰/۰۲۴ تا ۰/۰۲۷ $\eta^2p=$). علاوه بر این، رشته تحصیلی بر مؤلفه «روش‌های جایگزین» تأثیر معنی‌دار داشت. معلمان با رشته علوم تربیتی نسبت به معلمان رشته‌های علوم انسانی و علوم پایه، نمرات بالاتری را در روش‌های جایگزین کسب کردند (۳/۷۷) در مقابل ۳/۲۱۴ و ۳/۲۰۹).

بحث و نتیجه‌گیری

هدف پژوهش حاضر بررسی تأثیر تجربه حرفه‌ای (سابقه تدریس و دوره تدریس) و پیشینه تحصیلی (رشته تحصیلی) معلمان بر شیوه‌های سنجش و ارزشیابی آنان در کلاس درس بود. یافته‌های پژوهش نشان داد که دوره تدریس تأثیر معناداری بر ترکیب متغیرهای وابسته دارد و رشته تحصیلی در آستانه معناداری قرار دارد، در حالی که سابقه تدریس کلیه اثرات تعاملی دوگانه و سه‌گانه از نظر آماری معنی‌دار نبودند. تحلیل‌های تک‌متغیره نیز نشان داد که دوره تدریس بر روش‌های سنجش سنتی و روش‌های سنجش جایگزین تأثیر معنادار دارد و رشته تحصیلی بر روش‌های جایگزین تأثیر معنادار نشان داد. در ادامه، هر یک از این یافته‌ها در پرتو پیشینه پژوهشی مورد بحث قرار می‌گیرد.

یافته نخست پژوهش حاکی از آن بود که معلمان دوره متوسطه اول در مقایسه با معلمان دوره ابتدایی و متوسطه دوم، به‌طور معنی‌داری بیشتر از روش‌های سنتی استفاده می‌کنند. روش‌های سنتی در این پژوهش شامل آزمون‌های عینی و تشریحی، پرسش شفاهی، نمره‌دهی کمی و تحلیل آماری نتایج آزمون است. این یافته با پژوهش‌هایی که بر سلطه ارزشیابی سنتی و کمی در دوره متوسطه تأکید کرده‌اند، همخوانی دارد (Hussain et al., 2019؛ Çalışkan, 2010؛ Kaşıkçı, 2015؛ Serin, 2015). Robert و همکاران (2024) نیز در پژوهش خود نشان دادند که در کلاس‌های زبان انگلیسی، خود سنجی و همسال سنجی کمترین کاربرد را دارند و شیوه‌های ارزشیابی همچنان معلم‌محور باقی می‌مانند. در تبیین این یافته باید به ویژگی‌های خاص دوره متوسطه اول توجه کرد. برخلاف متوسطه دوم که فشار ناشی از کنکور سراسری و تأثیر قطعی معدل در آن نقش تعیین‌کننده‌ای دارد، در متوسطه اول، نمرات دانش‌آموزان به طور مستقیم در انتخاب رشته تحصیلی آنان در پایه دهم تأثیر می‌گذارد. همچنین رقابت فشرده برای پذیرش در مدارس خاص، مانند مدارس سمپاد، تیزهوشان و نمونه دولتی که معمولاً از پایه هفتم آغاز می‌شود، معلمان را به سمت استفاده

از روش‌های سنتی و کمی سوق می‌دهد. این عوامل، همراه با انتظارات والدین برای دریافت نمرات عددی و رتبه‌بندی و ساختار متمرکز برنامه‌های درسی، زمینه‌ساز غلبه ارزشیابی سنتی در این دوره است. پژوهش (Tóth and Csapó, 2024) در مجارستان نشان داد که فشار پاسخگویی ناشی از برنامه‌های ملی ارزشیابی، تمرکز بر نمره را تقویت می‌کند. در نظام آموزشی ایران نیز اگرچه تأثیر معدل در کنکور بیشتر متوجه متوسطه دوم است، اما فشار نهادی ناشی از سیستم انتخاب رشته و آزمون‌های ورودی مدارس خاص، فضای ارزشیابی در متوسطه اول را به سمت کمی‌گرایی سوق داده است. از نظر منطقی، ممکن است این پرسش مطرح شود که اگر فشار کنکور علت اصلی استفاده از روش‌های سنتی باشد، چرا معلمان دوره متوسطه دوم که مستقیماً با این فشار مواجه‌اند در یافته‌های این پژوهش بیشترین استفاده را از روش‌های سنتی نشان ندادند؟ تبیین این یافته را باید در تفاوت ماهیت فشارهای ارزشیابی در دو دوره جست. در متوسطه اول، فشار عمدتاً ناشی از نقش نمرات در انتخاب رشته تحصیلی، رقابت برای ورود به مدارس خاص و عدم آشنایی کافی معلمان با روش‌های جایگزین است. در مقابل، معلمان متوسطه دوم با وجود فشار کنکور، به دلیل تجربه حرفه‌ای بیشتر، شرکت در دوره‌های ضمن خدمت تخصصی‌تر و آشنایی نسبی با روش‌های ترکیبی ارزشیابی، مانند ارزشیابی عملکردی در دروس آزمایشگاهی و پروژه‌های پژوهشی، ممکن است تنوع بیشتری در شیوه‌های ارزشیابی خود به کار گیرند و صرفاً به روش‌های سنتی اکتفا نکنند. بنابراین، آنچه معلمان متوسطه اول را بیش از سایر دوره‌ها به سمت روش‌های سنتی سوق می‌دهد، تلفیقی از فشارهای ساختاری ویژه این دوره و محدودیت‌های دانش حرفه‌ای آنان است.

یافته دیگر پژوهش حاکی از آن بود که معلمان دوره ابتدایی در مقایسه با معلمان دوره متوسطه اول و دوم، به‌طور معنی‌داری بیشتر از روش‌های جایگزین استفاده می‌کنند. روش‌های جایگزین در این پژوهش شامل ارزشیابی عملکردی، پورتفولیو، خودارزشیابی، همتارزشیابی، مقیاس کیفی و درجه‌بندی عددی کیفیت دستاوردها است. این یافته با پژوهش‌هایی که بر نقش ارزشیابی اصیل و جایگزین در دوره ابتدایی تأکید کرده‌اند، همخوانی دارد (Yan & Carless, 2021؛ Bores-García et al., 2020؛ Letina, 2015). پژوهش‌های پیشین نشان داده‌اند که اجرای ارزشیابی جایگزین در دوره ابتدایی بیش از آنکه ناشی از آگاهی و نگرش معلمان باشد، ریشه در عوامل ساختاری و سیاست‌های بالادستی دارد (Gyamfi et al., 2023؛ Suwartono et al., 2024). در نظام آموزشی ایران، اجرای الزامی طرح ارزشیابی کیفی-توصیفی در دوره ابتدایی، معلمان را ناگزیر از کاربرت روش‌هایی چون پورتفولیو، خودارزشیابی و همتارزشیابی کرده است. مطالعات داخلی نیز نشان داده‌اند که معاونت آموزش ابتدایی با سرمایه‌گذاری گسترده در برگزاری کارگاه‌های ضمن خدمت و نظارت مستمر، نقش مؤثری در آشنایی معلمان با این روش‌ها داشته است (صالحی و همکاران، ۱۳۹۴؛ ناصری و محمودی، ۱۴۰۴). با این حال، همان پژوهش‌ها تأکید کرده‌اند که علی‌رغم این الزامات، شکاف معناداری میان وضع موجود و وضع مطلوب وجود دارد. به عنوان مثال، علی‌اکبری و همکاران (۲۰۲۳) نشان دادند که آزمون‌های کتبی پایانی و نمره‌دهی عددی در دوره ابتدایی همچنان رواج گسترده‌ای دارد. این یافته با نظریه شکاف نظر و عمل در حوزه سواد ارزشیابی (Pastore & Andrade, 2019) همسو است که نشان می‌دهد عوامل بازدارنده سطح خرد، مانند تراکم کلاس، کمبود زمان و فشار والدین می‌توانند الزامات ساختاری سطح کلان را تضعیف کنند. بنابراین، تفاوت مشاهده‌شده میان دوره ابتدایی و متوسطه در کاربرت روش‌های جایگزین، بیش از آنکه ناشی از آگاهی فردی معلمان ابتدایی باشد، ریشه در الزامات ساختاری تحمیل‌شده از بالا دارد. پژوهش‌های بین‌المللی نیز این الگو را تأیید کرده‌اند؛ به طوری که Letina و همکاران (2025) در مطالعه خود نشان دادند که در نظام‌های آموزشی با سیاست‌های ارزشیابی متمرکز، معلمان ابتدایی تحت فشار نهادی بیشتر از روش‌های جایگزین استفاده می‌کنند، اما همچنان از روش‌های سنتی به طور کامل دست نمی‌کشند. از نظر منطقی و با تکیه بر شواهد ارائه‌شده، می‌توان نتیجه گرفت که در صورت حذف این الزامات نهادی، مانند تغییر سیاست ارزشیابی در دوره ابتدایی، کاربرت روش‌های جایگزین در این دوره به شدت کاهش خواهد یافت و تفاوت معنادار میان دوره ابتدایی و متوسطه احتمالاً از بین خواهد رفت.

یافته دیگر پژوهش نشان داد که معلمان با رشته علوم تربیتی در مقایسه با معلمان رشته‌های علوم انسانی و علوم

پایه، صرف نظر از مقطع تدریس، به طور معناداری بیشتر از روش های جایگزین استفاده می کنند. این یافته با پژوهش هایی که بر نقش دانش تربیتی و آشنایی با نظریه های یادگیری در کیفیت ارزشیابی تأکید کرده اند، همخوانی دارد (Pastore, 2023؛ Hull & Vigh, 2024؛ Pastore & Andrade, 2019). در تبیین این یافته باید به تمایز میان دو مسیر ورود به حرفه معلمی توجه کرد. معلمان فارغ التحصیل از دانشگاه فرهنگیان - صرف نظر از رشته تخصصی آنان (علوم تربیتی، علوم پایه یا انسانی) - معمولاً دروس پایه روانشناسی، روش های تدریس و سنجش را در دوره کارشناسی خود می گذرانند. با این حال، تفاوت مشاهده شده در این پژوهش احتمالاً ناشی از آن است که بخشی از معلمان رشته های غیر علوم تربیتی (به ویژه در دوره متوسطه) از مسیرهای غیردانشگاه فرهنگیان، مانند استخدام های ماده ۲۸، حق التدریسی های بلندمدت، یا تغییر رشته از سایر حوزه ها وارد حرفه شده اند و آموزش های نظام مند تربیت معلم را دریافت نکرده اند. در مقابل، معلمان دارای مدرک علوم تربیتی، حتی اگر از مسیر غیردانشگاه فرهنگیان وارد شده باشند، به دلیل ماهیت رشته خود، واحدهای تخصصی بیشتری را در زمینه سنجش و ارزشیابی گذرانده اند. بنابراین، تفاوت در کاربست روش های جایگزین را باید بیش از آنکه ناشی از عنوان رشته باشد، در نوع و میزان آموزش های نظام مند تربیت معلم و تعداد واحدهای تخصصی ارزشیابی جستجو کرد. پژوهش Addow (2023) نیز نشان داد که آموزش های مرتبط با ارزشیابی با عملکرد ارزشیابی و خودکارآمدی معلمان رابطه مثبت دارد. از نظر منطقی، می توان استدلال کرد که در صورتی که معلمان رشته های غیر علوم تربیتی نیز از همان ابتدا آموزش های تخصصی یکسانی را دریافت کرده باشند، تفاوت معنادار میان گروه ها کاهش می یابد.

دروسی مانند روان شناسی یادگیری، سنجش و اندازه گیری، و روش های تدریس را به طور تخصصی می گذرانند و این آموزش ها می تواند به درک عمیق تر آنان از کارکردهای ارزشیابی و آشنایی با روش های جایگزین منجر شود. Willis و همکاران (2013) نیز بر ماهیت اجتماعی و بافت مند سواد ارزشیابی تأکید کرده و نقش آموزش های دانشگاهی را در شکل گیری آن برجسته ساخته اند. پژوهش Addow (2023) نیز نشان داد که آموزش های مرتبط با ارزشیابی با عملکرد ارزشیابی و خودکارآمدی معلمان رابطه مثبت دارد. در مقابل، معلمان رشته های علوم انسانی و علوم پایه که دانش تخصصی آنان معطوف به محتوای درسی است، ممکن است فرصت کمتری برای یادگیری مبانی نظری و عملی ارزشیابی آموزشی داشته باشند و در نتیجه، گرایش آنان به استفاده از روش های سنتی و کمی بیشتر باشد. این یافته با پژوهش Alkharusi (2011) که تفاوت در عملکرد ارزشیابی معلمان را بر اساس حوزه تدریس نشان داد، همسو است. همچنین با پژوهش Matovu and Zubairi (2014) که به ارتباط رشته تحصیلی با خودکارآمدی در ارزشیابی اشاره کردند، همخوانی دارد. از نظر منطقی، می توان استدلال کرد که معلمان رشته علوم تربیتی به دلیل برخورداری از دانش نظری عمیق تر در زمینه یادگیری و سنجش، توانایی بیشتری در طراحی و اجرای روش های جایگزین دارند.

نکته قابل تأمل در یافته های پژوهش، عدم معنی داری اثر اصلی سابقه تدریس بر شیوه های ارزشیابی است. این یافته با پژوهش هایی که تفاوت معنی داری بین معلمان کم سابقه و باتجربه در استفاده از روش های ارزشیابی گزارش نکرده اند، همخوانی دارد (Muhson et al., 2025؛ Alkharusi, 2011). با این حال، با برخی پژوهش ها که معلمان باتجربه را در به کارگیری روش های متنوع ارزشیابی موفق تر معرفی کرده اند (Koloi-Keaikitse, 2012) و نیز با پژوهش هایی که معلمان کم سابقه را گرایش مندتر به راهبردهای تکوینی نشان داده اند (Korkmaz et al., 2025؛ Letina et al., 2025)، ناهمسو است. Al-Hassan (2024) نیز نشان داد که معلمان باتجربه موانع بیشتری در اجرای ارزشیابی جایگزین ادراک می کنند. Gyamfi و همکاران (2023) در پژوهش خود به چالش های متعدد معلمان در ارزشیابی کلاسی اشاره کرده و راهبردهایی مانند ارزشیابی عملکردی، پورتفولیو و خودارزشیابی را برای بهبود آن پیشنهاد داده اند. این ناهمسویی را می توان از چند منظر تبیین کرد. نخست، سابقه تدریس به تنهایی نمی تواند شاخص مناسبی برای کیفیت تجربه حرفه ای باشد. آنچه اهمیت دارد، ماهیت تجربه، بازخوردهای دریافتی، مشارکت در برنامه های توسعه حرفه ای و تأمل در عمل است (ناصری و محمودی، ۱۴۰۴؛ Didier et al., 2025). دوم، ممکن است معلمان کم سابقه به دلیل آموزش های اولیه

و آشنایی با رویکردهای نوین، نگرش مثبت‌تری داشته باشند، اما در عمل به دلیل کم‌تجربگی و محدودیت‌های ساختاری، نتوانند این نگرش را به عمل تبدیل کنند. از نظر منطقی، می‌توان استدلال کرد که سابقه تدریس بدون همراهی با بازاریابی و مشارکت در برنامه‌های توسعه حرفه‌ای، به تنهایی نمی‌تواند پیش‌بین‌کننده قوی برای شیوه‌های ارزشیابی باشد.

عدم معناداری اثرات تعاملی دوگانه و سه‌گانه نیز قابل تأمل است. این یافته نشان می‌دهد که تأثیر دوره تدریس و رشته تحصیلی بر شیوه‌های ارزشیابی، مستقل از یکدیگر و مستقل از سابقه تدریس است. به عبارت دیگر، مزیت معلمان دوره ابتدایی در استفاده از روش‌های جایگزین، در تمام سطوح سابقه تدریس و در تمام رشته‌های تحصیلی به یک اندازه وجود دارد. همچنین تأثیر رشته تحصیلی بر کاربرد روش‌های جایگزین، در تمام دوره‌های تدریس و سطوح سابقه مشابه است. این یافته می‌تواند دلالت‌های مهمی برای برنامه‌ریزی توسعه حرفه‌ای داشته باشد.

در نتیجه، پژوهش حاضر نشان داد که تجربه حرفه‌ای (سابقه تدریس و دوره تدریس) و پیشینه تحصیلی (رشته تحصیلی) معلمان نقش مهمی در شکل‌دهی به شیوه‌های سنجش و ارزشیابی آنان ایفا می‌کنند. از منظر نظری، این یافته‌ها با تأکید بر نقش دوره تدریس و رشته تحصیلی در تبیین تفاوت‌های فردی معلمان، به بسط دانش موجود درباره سواد ارزشیابی و تعیین‌کننده‌های آن کمک می‌کند و چارچوبی مفهومی برای مطالعات آتی در زمینه شکل‌گیری شیوه‌های ارزشیابی فراهم می‌آورد. در سطح سیاست‌گذاری، یافته‌ها بر ضرورت بازنگری در برنامه‌های تربیت معلم و توسعه حرفه‌ای مبتنی بر نیازهای خاص هر دوره تحصیلی و هر رشته تأکید می‌کند. ارتقای سواد ارزشیابی معلمان به ویژه در دوره متوسطه و در میان معلمان با رشته‌های غیرمرتبط با علوم تربیتی، باید در اولویت سیاست‌گذاران آموزشی قرار گیرد. در سطح عملی، این یافته‌ها می‌تواند مدیران مدارس را در تصمیم‌گیری‌های مرتبط با انتصاب، حمایت و توانمندسازی معلمان یاری رساند. همچنین تقویت برنامه‌های درسی رشته‌های غیرعلوم تربیتی با دروس مرتبط با سنجش و ارزشیابی، و طراحی دوره‌های ضمن خدمت هدفمند برای معلمان این رشته‌ها، می‌تواند به کاهش شکاف میان نظر و عمل در ارزشیابی کلاسی کمک کند.

پژوهش حاضر با محدودیت‌هایی همراه بود که در تعمیم‌پذیری یافته‌ها باید مد نظر قرار گیرد. نخست، استفاده از روش نمونه‌گیری در دسترس و جمع‌آوری داده‌ها به صورت برخط، ممکن است منجر به مشارکت معلمانی شده باشد که دسترسی بیشتری به فناوری دارند یا انگیزه بیشتری برای شرکت در پژوهش‌های آموزشی نشان می‌دهند. دوم، داده‌های پژوهش مبتنی بر خودگزارشی معلمان بود و ممکن است بین آنچه معلمان گزارش می‌دهند و آنچه در عمل در کلاس انجام می‌دهند، تفاوت وجود داشته باشد. سوم، ماهیت مقطعی پژوهش امکان بررسی تغییرات شیوه‌های ارزشیابی در طول زمان و شناسایی روابط علی را فراهم نمی‌سازد. چهارم، نقض مفروضه همگنی ماتریس‌های واریانس-کوواریانس (آزمون باکس) اقتضا کرد که از آماره مقاوم پیلای برای تفسیر نتایج چندمتغیره استفاده شود و در تفسیر یافته‌ها احتیاط بیشتری به عمل آید.

بر اساس یافته‌ها و محدودیت‌های پژوهش، پیشنهادهای زیر به تفکیک برای پژوهش‌های آتی و اقدامات عملی ارائه می‌شود: الف) با توجه به اینکه معلمان متوسطه اول بیشترین استفاده را از روش‌های سنتی دارند و این پدیده با عواملی چون تأثیر نمرات در انتخاب رشته و رقابت برای مدارس خاص مرتبط است، بازنگری در نظام ارزشیابی این دوره، کاهش فشار نمره‌محوری در پایه‌های هفتم تا نهم، طراحی جایگزین‌های ارزشیابی برای فرآیند انتخاب رشته، و توانمندسازی معلمان در کاربری روش‌های جایگزین توصیه می‌شود. ب) با توجه به اینکه معلمان ابتدایی به دلیل الزامات نهادی و سرمایه‌گذاری معاونت آموزش ابتدایی بیشتر از روش‌های جایگزین استفاده می‌کنند اما همچنان شکاف قابل توجهی میان وضع موجود و مطلوب وجود دارد، انجام پژوهش‌های طولی و مشاهده‌ای برای شناسایی موانع اجرای کامل ارزشیابی کیفی-توصیفی و بررسی نقش متغیرهایی مانند جو مدرسه، باورهای معلمان و فرهنگ سازمانی پیشنهاد می‌شود. ج) با توجه به اینکه معلمان رشته علوم تربیتی (و به طور کلی معلمانی که آموزش‌های نظام‌مند تربیت معلم را دریافت کرده‌اند)

بیشتر از روش‌های جایگزین استفاده می‌کنند، گنجاندن دروس تخصصی‌تر سنجش و ارزشیابی در برنامه درسی تربیت معلم برای رشته‌های غیرعلوم تربیتی، به ویژه برای معلمان مسیرهای غیرفرهنگیان، و طراحی دوره‌های ضمن خدمت هدفمند برای این معلمان توصیه می‌گردد. (د) با توجه به عدم تأثیر سابقه تدریس بر شیوه‌های ارزشیابی، تمرکز پژوهش‌های آتی بر کیفیت تجربه حرفه‌ای، میزان مشارکت در دوره‌های توسعه حرفه‌ای، و تأمل‌پذیری معلمان به جای کمیت سابقه تدریس ضروری است. (ه) در نهایت، فراهم‌آوردن بسترهای نهادی و ساختاری مناسب شامل کاهش فشار آزمون‌های بیرونی (به ویژه در متوسطه اول)، تأمین منابع و امکانات لازم برای اجرای روش‌های جایگزین در تمامی دوره‌ها، و ایجاد فرهنگ سازمانی حامی ارزشیابی کیفی، برای تسهیل تغییر در شیوه‌های ارزشیابی معلمان ضروری به نظر می‌رسد.

مشارکت نویسندگان

مقاله یک نویسنده دارد.

تشکر و قدردانی

بدین وسیله از تمامی معلمان گرامی که با مشارکت صادقانه خود در این پژوهش، امکان گردآوری داده‌ها را فراهم ساختند، و از آقای علی قره داغی که در فرایند گردآوری داده‌ها کمک نمودند، صمیمانه قدردانی می‌گردد.

تعارض منافع

«هیچ‌گونه تعارض منافع توسط نویسندگان بیان نشده است»

منابع

- ثامتی، ناهید؛ امامقلی‌وند، فاطمه؛ موسوی، سید محسن؛ رستمی، مریم؛ و محمد اسمعیلی، فائزه. (۱۴۰۲). نقش رشته تحصیلی دوره متوسطه با عملکرد حرفه‌ای فارغ‌التحصیلان رشته آموزش ابتدایی دانشگاه فرهنگیان. نظریه و عمل در تربیت معلمان، ۹(۱۵)، ۴۹-۶۴. <https://dor.isc.ac/dor/20.1001.1.26457156.1402.9.15.3.1>
- دیبايي صابر، محسن. (۱۳۹۷). تحلیل رابطه توسعه شایستگی حرفه ای معلمان دوره متوسطه با اثربخشی آنان در فرایند یاددهی-یادگیری دانش آموزان. نظریه و عمل در تربیت معلمان، ۴(۶)، ۲۳-۱. <https://dor.isc.ac/dor/20.1001.1.26457156.1397.4.6.1.6>
- رضائی، عباس و محمد علی‌زاده، اعظم. (۱۴۰۴). بررسی روش‌های سنتی سنجش یادگیری و روش‌های سنجش جایگزین، دوازدهمین کنفرانس بین‌المللی مطالعات میان رشته ای علوم بهداشتی، روانشناسی، مدیریت و علوم تربیتی. <https://civilica.com/doc/2393830>
- سانتراک، جان دبلیو. (۲۰۲۰). روان‌شناسی تربیتی (ترجمه شاهده سعیدی، مهشید عراقچی و حسین دانشفر). تهران: رسا.
- سیف، علی اکبر. (۱۴۰۲). اندازه‌گیری، سنجش و ارزشیابی آموزشی. تهران: دوران.
- صالحی، کیوان؛ بازرگان، عباس؛ صادقی، ناهید؛ و شکوهی یکتا، محسن. (۱۳۹۴). تحلیل پدیدارشناسانه ادراک و تجربه زیسته معلمان مدارس ابتدایی از نقاط ضعف و قوت برنامه ارزشیابی توصیفی. پژوهش در نظام‌های آموزشی، ۹(۳۱)، ۶۸-۱۹. <https://dor.isc.ac/dor/20.1001.1.23831324.1394.9.31.2.0>
- ناصری، فاطمه، و محمودی، فیروز. (۱۴۰۴). تحلیل عوامل مؤثر بر فرآیندهای توسعه حرفه‌ای معلمان در تهران. آموزش، تعلیم و توسعه پایدار، ۱۱(۱)، ۱۵-۱.
- نراقی‌زاده، مصطفی، عزیزملایی، فرامرز، و خلجی، حمید رضا. (۱۴۰۱). ادراکات و عملکرد معلمان زبان انگلیسی از استراتژی‌های ارزیابی جایگزین و رابطه آنها با تامل معلم. نشریه آموزش زبان انگلیسی و عمل، ۱۵(۳۰)، ۲۶۶-۲۹۰. <https://doi.org/10.30495/jal.2022.697012>

وحیدی، زهرا و کریمی، اعظم (۱۴۰۴). شیوه های سنجش دانش آموزان ابتدایی روایتی از معلمان و مدیران ابتدایی، هشتمین کنفرانس بین المللی پژوهش های نوین در حوزه مشاوره، علوم تربیتی و روانشناسی ایران، تهران.
<https://civilica.com/doc/2419097>

References

- Acar-Erdol, T., & Yıldızlı, H. (2018). Classroom assessment practices of teachers in Turkey. *International Journal of Instruction*, 11(3), 587-602.
<https://doi.org/10.12973/iji.2018.11340a>
- Addow, A. (2023). Factors affecting lecturers' assessment practices in Somalia higher education. *European Journal of Education Studies*, 10(5). <https://doi.org/10.46827/ejes.v10i5.4820>
- Alansari, W. M. (2023). The application of the alternative evaluation strategy in social studies in general education in KSA. *Journal of Educational and Psychological Researches*, 19(73), 88-49. <https://jperc.uobaghdad.edu.iq/index.php/jperc/article/view/1526>
- Al-Hassan, A. A. (2024). Obstacles of applying alternative evaluation strategies from the point of view of the first cycle teachers in the schools of Hama city and ways of solving them. *Alustath Journal for Human and Social Sciences*, 63(2), 68-89.
<https://doi.org/10.36473/x3md1066>
- Aliakbari, M., Yasini, A., & Sadeghi, S. (2023). Iranian EFL teachers' classroom assessment practices: Discrepancy between theoretical and practical aspects of classroom assessment practices. <https://civilica.com/doc/1774437>
- Alkharusi, H. (2011). Teachers' classroom assessment skills: Influence of gender, subject area, grade level, teaching experience and in-service assessment training. *Journal of Turkish Science Education*, 8(2), 39-48. <https://www.tused.org/index.php/tused/article/view/358>
- Arjuna, R., Hikmat, M., & Candraningrum, D. (2024). Teachers' perception about authentic assessment of English learning based on Merdeka Curriculum: A case in Papua. *AL-ISHLAH: Jurnal Pendidikan*, 16(3), 1-12. <https://doi.org/10.35445/alishlah.v16i3.4628>
- Bores-García, D., Hortigüela-Alcalá, D., González-Calvo, G., & Barba-Martín, R. (2020). Peer assessment in physical education: A systematic review of the last five years. *Sustainability*, 12(21), 9233. <https://doi.org/10.3390/su12219233>
- Çalışkan, H., & Kaşıkçı, Y. (2010). The application of traditional and alternative assessment and evaluation tools by teachers in social studies. *Procedia - Social and Behavioral Sciences*, 2, 4152-4156. <https://doi.org/10.1016/j.sbspro.2010.03.656>
- Derakhshan, A., & Ghiasvand, F. (2022). Demystifying Iranian EFL teachers' perceptions and practices of learning-oriented assessment (LOA): Challenges and prospects in focus. *Language Testing in Asia*, 12(1). <https://doi.org/10.1186/s40468-022-00204-2>
- Dibaei Saber, M. (2018). Analysis of the Relationship between Professional Competence Development of High School Teachers with Their Effectiveness in Teaching-Student Learning Process. *Theory and Practice in Teachers Education*, 4(6), 1-23. [In Persian].
<https://dor.isc.ac/dor/20.1001.1.26457156.1397.4.6.1.6>
- Didier, L., Schildkamp, K., Visscher, A., & Bosker, R. (2025). Factors influencing the implementation of a teacher professional development program to improve teaching quality. *Frontiers in Education*. <https://doi.org/10.3389/feduc.2025.1546448>
- Girgla, A., Good, L., Krstic, S., McGinley, B., Richardson, S., Sneider-Gregory, S., & Star, J. (2021). Developing a teachers' assessment literacy and design competence framework. Australian Council for Educational Research.
<https://www.ibo.org/globalassets/publications/if-research/assessment-literacy-final-report-en.pdf>

- Gyamfi, A., Loganathan, S., & Acquaye, R. (2023). Improving teachers' classroom assessment practices: Perceptions of teachers in the Ellembele District of Ghana. *e-mentor*, 4(101), 56-62. <https://doi.org/10.15219/em101.1625>
- Hussain, S., Shaheen, N., Ahmad, N., & Islam, S. U. (2019). Teachers' classroom assessment practices: Challenges and opportunities to classroom teachers in Pakistan. <https://www.researchgate.net/publication/331669191>
- Khan, F. N., Hussain, S., & Imad, M. (2019). Classroom assessment, literacy and practices of teacher educators in Pakistan. *Global Social Sciences Review*, 4(4), 45-51. [http://dx.doi.org/10.31703/gssr.2019\(IV-IV\).07](http://dx.doi.org/10.31703/gssr.2019(IV-IV).07)
- Khorramabadi, Y., & Delavar, A. (2009). Investigating the effects of alternative assessment methods (performance and portfolio) on pre-university students' attitude and academic achievement in Khorramabad city. *Educational Psychology*, 5(15), 88-109. https://jep.atu.ac.ir/article_2226.html
- Koloi-Keaikitse, S. (2012). Classroom assessment practices: A survey of Botswana primary and secondary school teachers (Doctoral dissertation, Ball State University). <https://eric.ed.gov/?id=ED543182>
- Korkmaz, G., Özdemir, H. F., Çalışkan Toraman, Ç., & Toraman, Ç. (2025). Teachers' tendency to use alternative assessment methods. *Pegem Journal of Education and Instruction*, 15(3), 60-77. <https://doi.org/10.47750/pegegog.15.03.08>
- Letina, A. (2015). Application of traditional and alternative assessment in science and social studies teaching. *Croatian Journal of Education*, 17(0). <https://doi.org/10.15516/CJE.V17I0.1496>
- Letina, A., Škugor, A., & Tomaš, S. (2025). Teachers' perceptions and practices of assessment in primary schools. *European Journal of Educational Research*, 14(4), 1105-1121. <https://doi.org/10.12973/eu-jer.14.4.1105>
- Levy-Feldman, I. (2025). The role of assessment in improving education and promoting educational equity. *Education Sciences*, 15(2), 224. <https://doi.org/10.3390/educsci15020224>
- López-Pastor, V. M., Kirk, D., Lorente-Catalán, E., MacPhail, A., & Macdonald, D. (2013). Alternative assessment in physical education: A review of international literature. *Sport, Education and Society*, 18(1), 57-76. <https://doi.org/10.1080/13573322.2012.713860>
- Mäkipää, T. (2023). Students' perceptions of teacher assessment practices in foreign language emergency remote teaching during Covid-19 in Finland. *Studies in Language Assessment*, 1-25. <http://hdl.handle.net/10138/569070>
- Mäkipää, T., Hahl, K., & Luodonpää-Manni, M. (2021). Teachers' perceptions of assessment and feedback practices in Finland's foreign language classes during the Covid-19 pandemic. *Center for Educational Policy Studies Journal*, 11(Special Issue), 153-173. <https://doi.org/10.26529/cepsj.1108>
- Manigbas, J. P., & De Luna, Y. V. V. (2023). Factors influencing the assessment practices of senior high school teachers in Goa district, Philippines. *Psychology and Education: A Multidisciplinary Journal*, 14(10), 998-1005. <https://ssrn.com/abstract=4615160>
- Matovu, M., & Zubairi, A. (2014). Factors influencing assessment practices among university academic staff: A multiple regression analysis. *Mevlana International Journal of Education*, 4(1), 176-188. <https://doi.org/10.13054/mije.13.57.4.1>
- Moura, A., Graça, A., MacPhail, A., & Batista, P. (2021). Aligning the principles of assessment for learning to learning in physical education: A review of literature. *Physical Education and Sport Pedagogy*, 26(4), 388-401. <https://doi.org/10.1080/17408989.2020.1834528>

- Muhson, A., Mwakapemba, J., Siswanto, S., Prilanita, Y., & Susilowati, N. (2025). Teachers' assessment literacy: The impact of gender, education, training, and work experience. *Journal of Education and Learning (EduLearn)*, 19(3), 1410-1417. <https://doi.org/10.11591/edulearn.v19i3.22040>
- Naseri, F., & Mahmoudi, F. (2025). Analysis of factors affecting teachers' professional development processes in Tehran. *Training, Education, and Sustainable Development*, 1(1), 1-15. [In Persian]. <https://www.journaltesd.com/index.php/tesd/article/view/108>
- Naraghizadeh, M., Azizmalayeri, F., & Khalaji, H. R. (2022). Teachers' perceptions and performance of alternative assessment strategies and their relationship with teacher reflection. *Journal of English Language Pedagogy and Practice*, 15(30), 266-290. [In Persian]. <https://doi.org/10.30495/jal.2022.697012>
- Ohlsen, M. (2007). Classroom assessment practices of secondary school members of NCTM. *American Secondary Education*. <https://www.researchgate.net/publication/234652265>
- Pastore, S. (2023). Teacher assessment literacy: A systematic review. *Frontiers in Education*, 8, Article 1217167. <https://doi.org/10.3389/feduc.2023.1217167>
- Pastore, S., & Andrade, H. L. (2019). Teacher assessment literacy: A three-dimensional model. *Teaching and Teacher Education*, 84, 1-12. <https://doi.org/10.1016/j.tate.2019.05.003>
- Rezaei, A., & Mohammad Alizadeh, A. (2025). A comparison of traditional and alternative assessment methods in learning. In Proceedings of the 12th International Conference on Interdisciplinary Studies in Health Sciences, Psychology, Management and Educational Sciences. [In Persian] <https://civilica.com/doc/2393830>
- Robert, S. C., Yunus, M. M., & Said, N. E. M. (2024). Secondary school teachers' perception of classroom assessment practices in ESL classrooms. *International Journal of Academic Research in Progressive Education and Development*, 13(2), 319-336. <http://dx.doi.org/10.6007/IJARPED/v13-i3/21716>
- Sameti, N., Imamgholivand, F., Mousavi, S. M., Rostami, M. and Mohammad Esmaili, F. (2023). The Role of Secondary School Course with The Professional Performance of Graduates of Primary Education Field of Farhangian University. *Theory and Practice in Teachers Education*, 9(15), 64-49. [In Persian]. <https://dor.isc.ac/dor/20.1001.1.26457156.1402.9.15.3.1>
- Seif, A. A. (2023). Educational measurement, assessment and evaluation. Tehran: Dowran. (In Persian)
- Salehi, K., Bazargan, A., Sadeghi, N., & Shokouhi Yekta, M. (2015). Phenomenological analysis of elementary school teachers lived experience of strengths and weaknesses of descriptive evaluation program. *Journal of Research in Educational Systems*, 9(31), 19-68. [In Persian]. <https://dor.isc.ac/dor/20.1001.1.23831324.1394.9.31.2.0>
- Santrock, J. W. (2020). Educational psychology. Tehran: Rasa Publications. [In Persian]
- Serin, G. (2015). Alternative assessment practices of a classroom teacher: Alignment with reform-based science curriculum. *Eurasia Journal of Mathematics, Science and Technology Education*, 11(2), 277-297. <https://doi.org/10.12973/eurasia.2015.1330a>
- Suwartono, T., Zulehah, R., & Dada, E. (2024). EFL teacher voices of authentic assessment in the current curriculum in Indonesia. *Journal on English as a Foreign Language*. <https://www.researchgate.net/publication/388195444>
- Tabachnick, B. G., & Fidell, L. S. (2019). Using Multivariate Statistics (6th ed.). Boston, MA: Pearson.
- Tóth, E., & Csapó, B. (2022). Teachers' beliefs about assessment and accountability. *Educational Assessment, Evaluation and Accountability*, 34, 459-481. <https://doi.org/10.1007/s11092-022-09396-w>

- Ullah, N. (2023). Assessment practices and assessment skills of teachers at secondary school level: Factorability and associative perspective. *Journal of Professional & Applied Research*, 7(1), 101–116. <https://doi.org/10.36968/JPDC-V07-I01-10>
- Vahidi, Z., & Karimi, A. (2025). Elementary students' assessment practices: A narrative of elementary teachers and principals. In Proceedings of the 8th International Conference on New Research in Counseling, Educational Sciences and Psychology. [In Persian] <https://civilica.com/doc/2419097>
- Veugen, M. J., Gulikers, J. T. M., & den Brok, P. (2021). We agree on what we see: Teacher and student perceptions of formative assessment practice. *Studies in Educational Evaluation*, 70, Article 101027. <https://doi.org/10.1016/j.stueduc.2021.101027>
- Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualising teachers' assessment literacies in an era of curriculum and assessment reform. *The Australian Educational Researcher*, 40(2), 241-256. <https://doi.org/10.1007/s13384-013-0089-9>
- Yan, Z., & Carless, D. (2021). Self-assessment is about more than self: The enabling role of feedback literacy. *Assessment & Evaluation in Higher Education*, 47(7), 1116-1128. <https://doi.org/10.1080/02602938.2021.2001431>