



Developing a Conceptual Model and Developing a Tool to Measure Teachers' Awareness of Formative Assessment: a Mixed-Method Study

Ronak Sohrabi^{*1}, Asghar Minaei², Hamidreza Maghami³

¹ Masters Student, Psychometrics, Allameh Tabatabaei University, Tehran, Iran

² Associate Professor, Department of Measurement and Evaluation, Allameh Tabatabaei University, Tehran, Iran

³ Associate Professor, Department of Educational Technology, Allameh Tabatabaei University, Tehran, Iran

* Corresponding author: easp2820@gmail.com

Received: 2025-07-13

Accepted: 2025-08-18

Abstract

Purpose: Developing a conceptual model and developing a tool to measure teachers' awareness of formative assessment: A mixed-method study

Objective: Developing a conceptual model 'developing a tool to measure teachers' awareness of formative assessment 'reliability and validation of the tool

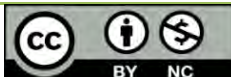
Methods: Exploratory and confirmatory factor analysis methods

Results: New explanation and application of Marzano and Kendall's classification of educational goals for formative assessment 'components of the conceptual model of formative assessment are: knowledge presentation, cognitive system (Retrieval, Comprehension, analysis, Knowledge Utilization), Metacognition, Self-Sytem Thinking, traditional and modern formative evaluation strategies, developing a 74-item questionnaire

Conclusion: Extracting six components and 28 questions from the questionnaire in the exploratory factor analysis method, the Cronbach's alpha of the entire construct is 0.79 and is acceptable. The tool has poor fit in confirmatory factor analysis and its fit was improved by modifying the model.

Keywords: Teachers' Awareness, Formative evaluation, Analysis method, School

© 2023 Journal of Mental Health in School (JMHS)



This work is published under CC BY-NC 4.0 license.

© 2023 The Authors.

How to Cite This Article: Sohrabi, R, et al. (2025). Developing a Conceptual Model and Developing a Tool to Measure Teachers' Awareness of Formative Assessment: a Mixed-Method Study. *JMHS*, 3(2): 34-46.





تدوین مدل مفهومی و ساخت ابزاری برای سنجش آگاهی معلمان از ارزشیابی تکوینی: یک مطالعه آمیخته

روناک سهرابی^{۱*}، اصغر مینائی^۲، حمیدرضا مقامی^۳

^۱ دانشجوی کارشناسی ارشد، رشته روان سنجی، دانشگاه علامه طباطبائی، تهران، ایران

^۲ دانشیار گروه آموزشی سنجش و اندازه گیری، دانشگاه علامه طباطبائی، تهران، ایران

^۳ دانشیار گروه آموزشی تکنولوژی آموزشی، دانشگاه علامه طباطبائی، تهران، ایران

* نویسنده مسئول: easp2820@gmail.com

تاریخ پذیرش مقاله: ۱۴۰۴/۰۵/۲۷

تاریخ دریافت مقاله: ۱۴۰۴/۰۴/۲۲

چکیده

هدف: تدوین مدل مفهومی، ساخت ابزار سنجش آگاهی معلمان از ارزشیابی تکوینی، روایی و اعتبارسنجی ابزار روش پژوهش: روش های تحلیل عاملی اکتشافی و تأییدی یافته ها: تبیین جدید و کاربرد طبقه بندی مارزانو و کندال از اهداف آموزشی برای ارزشیابی تکوینی، مؤلفه های مدل مفهومی ارزشیابی تکوینی عبارتند از: ارائه دانش، سیستم شناختی (بازیابی، درک مطلب، تحلیل، بهره برداری از دانش)، فراشناخت، تفکر نظام شخصی، راهبردهای ارزشیابی تکوینی سنتی و نوین، تدوین پرسش نامه ۷۴ سؤالی نتیجه گیری: استخراج شش مؤلفه و ۲۸ سؤال از پرسش نامه در روش تحلیل عاملی اکتشافی، آلفای کرونباخ کل سازه برابر ۰/۷۹ و قابل قبول است. ابزار در تحلیل عاملی تأییدی دارای برازش ضعیف است و با اصلاح مدل برازش آن بهبود یافت.

واژگان کلیدی: آگاهی معلمان، ارزشیابی تکوینی، روش تحلیل، مدرسه

تمامی حقوق نشر برای فصلنامه سلامت روان در مدرسه محفوظ است.

شیوه استناد به این مقاله: سهرابی، روناک؛ مینائی، اصغر؛ مقامی، حمیدرضا. (۱۴۰۴). تدوین مدل مفهومی و ساخت ابزاری برای سنجش آگاهی معلمان از ارزشیابی تکوینی: یک مطالعه آمیخته. فصلنامه سلامت روان در مدرسه، ۳(۲): ۴۶-۳۴.

مقدمه

ارتقای نتایج آموزشی مطلوب، حل مسئله مشارکتی در محیط های آموزشی، تعدیل آموزش توسط معلم بر اساس شواهد یادگیری دانش آموز، توسعه مهارت های خودتنظیمی و یادگیری مادام العمر، یادگیری به کمک همتایان، از جمله ارزیابی همتا و بازخورد آنها، افزایش تجربه یادگیری، توسعه حرفه ای معلمان،

ارزشیابی تکوینی ابزار ارزشمندی در آموزش است که می تواند تجربه یادگیری را به میزان قابل توجهی افزایش دهد و نتایج مطلوب مختلفی را ارتقا دهد و از رشد مهارت ها و شایستگی های اساسی در دانش آموزان حمایت کند. اهمیت ارزشیابی تکوینی در

معلمان در این مطالعه در سطح التقاطی بودند. درحالی که این معلمان به اهمیت ارزشیابی تکوینی پی بردند، آن را به اندازه کافی در شیوه‌های تدریس خود نمی‌گنجانند. مطالعات متعددی مشاهده کرده‌اند که ارزشیابی تکوینی به‌ندرت در محیط‌های کلاس درس واقعی استفاده می‌شود (Desimone, 2009; ylie & Lyon, 2015)، و زمانی که اجرا می‌شود، کیفیت اغلب کمتر از انتظار است (Yan & Brown, 2021)؛ بنابراین ملاحظات کلیدی برای تدوین ابزارهایی که بر کاربرد عملی مهارت‌ها و دانش‌های عملی ارزشیابی تکوینی تمرکز کنند تا معلمان بتوانند مستقیماً در کلاس‌های درس خود به کار گیرند لازم است. برای ارزیابی جامع، ابزارها باید هم درک نظری و هم کاربرد عملی راهبردهای ارزیابی تکوینی را ارزیابی کنند. ابزارهای جدید باید مبتنی بر نظریه‌های آموزشی معاصر باشد که بر نقش ارزشیابی تکوینی در ارتقای یادگیری دانش‌آموزان تأکید دارند.

مبانی نظری و پیشینه پژوهش

مطالعه (McMillan 2017) نشان می‌دهد که تنوع در درک معلمان و اجرای شیوه‌های ارزشیابی تکوینی در زمینه‌های آموزشی مختلف، چالش مهمی برای افزایش یادگیری و پیشرفت دانش‌آموزان است. علی‌رغم شناخت روزافزون اهمیت ارزشیابی تکوینی، مطالعات نشان می‌دهد که بسیاری از معلمان فاقد دانش و مهارت کافی برای به‌کارگیری مؤثر این شیوه‌ها هستند. به‌عنوان مثال، تحقیق Kazinyirako, Ntawiha, Buhigiro & Ndayambaje (2024) نشان می‌دهد که تکنیک‌های ارزشیابی تکوینی به‌ندرت در مدارس متوسطه دولتی در ناحیه کیچوکیرو مورد استفاده قرار می‌گیرد، و معلمان به‌ندرت درگیر برنامه ریزی یا ارائه بازخورد بر اساس ارزیابی‌های تکوینی می‌شوند. اما مطالعه (Entika Fani & Prabowo 2019) نشان داد که محدودیت‌های اصلی معلمان در اجرای ارزیابی‌های تکوینی، دانش ناکافی از ارزشیابی‌های تکوینی، مدیریت زمان، و فقدان راهنمایی رسمی از سوی مقامات است. علاوه بر این، چارچوب‌ها و طبقه‌بندی‌های موجود ارزشیابی تکوینی به‌اندازه کافی پیچیدگی‌های مربوط به آگاهی و کاربرد معلمان از این روش‌ها را بررسی نمی‌کنند. مثلاً (Bennet 2011) پنج راهبرد کلیدی برنامه "یادگیری را در مسیر ادامه دهید" KLT شامل: الف) به‌اشتراک‌گذاری اهداف یادگیری ب) پرسشگری ج) بازخورد د) خودارزیابی ه) ارزیابی همتایان را معرفی می‌کند و مطالعه (Arrafii & Sumarni 2018) به بررسی مفاهیم مختلف مرتبط با ارزشیابی یادگیری از جمله ارزشیابی برای یادگیری، جهت‌گیری هدف یادگیری عملکرد در مقابل تسلط و رویکردهای مختلف به راهبردها و روش‌های ارزشیابی می‌پردازد.

مشارکت و عملکرد آنلاین، بازخورد بلادرنگ، انصاف، شمول، و شیوه‌های آموزشی بلندمدت، یادگیری شبیه‌سازی شده، گفت‌وگو و تعامل اجتماعی، بهبود تعهد دانش‌آموز همراه است، و از سوی دیگر ارزشیابی تکوینی آنلاین با کاهش زمان علامت‌گذاری و صرفه‌جویی در هزینه‌های اداری همراه است و ارزیابی‌های تکوینی آنلاین با اجازه‌دادن به دانش‌آموزان برای تأمل در اشتباه خود و یادگیری از خطاها، درک عمیق‌تری از مطالب ایجاد کنند و به یادگیرندگان خودراهرتر تبدیل شوند (Dixon, 2020; Lee, Chung, Zhang, Abedi & Warschauer, 2020; Haugen & de Lange, 2024)

نقش حیاتی ارزشیابی تکوینی در آموزش به طور گسترده به رسمیت شناخته شده است، با این حال شکاف قابل‌توجهی در دسترسی به ابزارهای نظریه‌محور که به طور خاص برای ارزیابی سواد ارزشیابی تکوینی معلمان طراحی شده‌اند، وجود دارد. این سواد شامل دانش و مهارت‌هایی است که معلمان برای اجرای مؤثر ارزیابی‌های تکوینی در شیوه‌های تدریس خود به آن نیاز دارند، می‌شود. سواد ارزشیابی تکوینی به شایستگی‌هایی اشاره دارد که معلمان باید برای طراحی، اجرا و تفسیر مؤثر ارزشیابی تکوینی داشته باشند. این شامل درک روش‌های مختلف ارزیابی، دانستن نحوه به‌کارگیری مناسب آنها و توانایی تحلیل داده‌های به‌دست‌آمده برای اطلاع از تصمیمات آموزشی است. علی‌رغم اهمیت آن، ادبیات فقدان چارچوب‌ها و ابزارهای جامع متمرکز بر ارزیابی این جنبه خاص از شایستگی معلم را نشان می‌دهد. در رابطه با ابزارهای موجود و محدودیت‌های آنها می‌توان این‌گونه تفسیر کرد؛ درحالی که ابزارهای متعددی برای ارزیابی سواد ارزشیابی معلمان ایجاد شده است، اکثر آنها به‌جای اینکه به طور خاص بر ارزشیابی تکوینی تمرکز کنند، بر روی شیوه‌های ارزیابی گسترده‌تر تمرکز کرده‌اند. به‌عنوان مثال، ابزارهایی مانند مطالعه (Mertler 2004, 2009) که به سنجش سطح سواد ارزشیابی معلمان می‌پردازد، دیگر اینکه Linn & Miller (2005)، به بررسی پنج اصل کلی ارزیابی مؤثر (وضوح اهداف یادگیری، روش ارزیابی مرتبط، روش‌های ارزیابی متنوع، ارزشیابی غیررسمی و همسویی با آموزش) می‌پردازد و سه بعد ارزیابی ("صریح ساختن یادگیری"، "ارتقای استقلال یادگیری" و "جهت‌گیری عملکرد") در مطالعه James & Pedder (2006) بیان شده است که این ابزارهای مختلف دانش و مهارت‌های ارزیابی عمومی را در بین معلمان را اندازه‌گیری می‌کنند. با این حال، این ابزارها به‌اندازه کافی تفاوت‌های ظریف شیوه‌های ارزشیابی تکوینی یا دانش خاص موردنیاز برای اجرای مؤثر در محیط‌های کلاسی متنوع را در بر نمی‌گیرند. در مطالعه Kaya, Atasoy, Candan-Helvaci & Pektaş (2021) آگاهی معلمان علوم از ارزشیابی تکوینی، اکثر

سیستم‌های حمایتی تأکید می‌کند که می‌توانند قابلیت‌های معلمان را در استفاده مؤثر از ارزشیابی تکوینی در کلاس‌هایشان افزایش دهند.

Black & Wiliam (1998) ارزشیابی تکوینی را این‌گونه تعریف کردند: «ارزشیابی تکوینی شامل تمام آن فعالیت‌هایی است که توسط معلمان و/یا دانش‌آموزان انجام می‌شود، که اطلاعاتی را برای استفاده به‌عنوان بازخورد برای اصلاح فعالیت‌های آموزشی و یادگیری که در آن درگیر هستند ارائه می‌دهد» برای این منظور این مطالعه کاربرد و تعریف طبقه‌بندی مارزانو از اهداف آموزشی را برای ارزشیابی تکوینی گسترش داده است و معلمان و دانش‌آموزان از آنها در روند کلاس استفاده می‌کنند در حقیقت نظریه مارزانو جنبه عملیاتی مؤلفه‌های اصلی ارزشیابی تکوینی در طی کلاس است. زیرا به‌عنوان معلم، درک فرایندهای بهره‌برداری از دانش برای کمک به دانش‌آموزان برای به‌کارگیری مؤثر دانش و مهارت‌های خود در موقعیت‌های عملی بسیار مهم است. معلمان بر اهمیت فرایندهای استفاده از دانش تکلیف محور که بر به‌کارگیری دانش کسب شده برای دستیابی به اهداف خاص متمرکز است، تأکید می‌کنند. در فعالیت‌های کلاس تکالیفی را طراحی می‌کنند که دانش‌آموزان را به چالش می‌کشد تا از درک و بینش خود برای رسیدگی به چالش‌ها، حل مسائل، تصمیم‌گیری یا تکمیل کارها به طور مؤثر استفاده کنند. معلمان دانش‌آموزان را در اجرای عملی دانش در سناریوهای دنیای واقعی راهنمایی می‌کنند و به آنها کمک می‌کنند مفاهیم نظری را به کاربردهای عملی متصل کنند این مسئله مطابق بهره‌برداری از دانش در طبقه‌بندی مارزانو و کندال است. طبقه‌بندی مارزانو و کندال، کنترل و تنظیم فرایندهای شناختی را به‌عنوان اصل سازمان‌دهنده، در نظر می‌گیرد و فرایندهای آموزشی دانش‌آموزان و معلمان را در کلاس کنترل می‌کند. این مدل چهار مؤلفه را ارائه می‌دهد که عبارت‌اند از: دانش یعنی محتوای اطلاعاتی که به سه سیستم دیگر وارد می‌شود. بعد از دانش، سیستم نظام شخصی قرار دارد؛ یعنی مسئول تصمیم‌گیری در مورد شرکت در یک کار جدید است. دیگری سیستم شناختی است که پس از تصمیم‌گیری برای تعامل، مسئول پردازش اطلاعات است. در نهایت سیستم فراشناختی است که در نظارت و تنظیم فرایندهای شناختی نقش دارد (Marzano & Kendall, 2007)؛ درحالی‌که روش‌های ارزشیابی تکوینی سستی می‌توانند برای اندازه‌گیری دانش در یک مقطع زمانی خاص کارآمد باشند، اما اغلب در ارائه بازخورد و انعطاف‌پذیری لازم که شیوه‌های آموزشی مدرن نیاز دارند، کوتاهی می‌کنند. از آنجایی‌که معلمان به دنبال افزایش مشارکت دانش‌آموز و نتایج یادگیری هستند، ادغام راهبردهای ارزیابی پویاتر و پاسخگوتر ممکن است سودمند باشد. اگر معلمان در کلاس

(2018) Khodabakhshzadeh, Kafi & Hosseinnia با بررسی مفاهیم و سواد معلمان زبان انگلیسی از ارزشیابی تکوینی به تدوین پرسش‌نامه‌ای با دو عامل دانش ضمنی و عملی ارزشیابی تکوینی پرداختند. در این مطالعه از رویکرد ساختاریافته نظری و راهبردهای مختلف ارزشیابی تکوینی استفاده نشده است و در تدوین سؤال‌های پرسش‌نامه به ارزیابی نگرش معلمان از اجرای ارزشیابی تکوینی پرداخته است. مثلاً سؤال یک (به چه چیزی در ارزشیابی تکوینی اعتقاد دارید؟) یا سؤال دو (به نظر شما ارزشیابی تکوینی تا چه اندازه می‌تواند اجرا شود؟) و ... از سوی دیگر Yan & Pastore (2022) به ابعاد مفهومی، عملی، و اجتماعی - عاطفی پرداخته‌اند علی‌رغم این طبقه‌بندی‌های دقیق، مطالعه نشان می‌دهد که چارچوب‌های موجود به طور کامل به پیچیدگی‌های اجرای دانش ارزشیابی تکوینی توجه نکرده‌اند و سؤال‌های تدوین شده پرسش‌نامه مؤلفه‌ها را به خوبی نمی‌سنجد. مثلاً در بعد مفهومی سؤال یک (من می‌توانم منطق ارزشیابی تکوینی را توضیح دهم.) یا سؤال هفت (من روش‌های ارزشیابی متنوعی را می‌شناسم که به دانش‌آموزان اجازه می‌دهد تا یادگیری خود را نشان دهند) و ... از سوی دیگر با قرار دادن مقیاس لیکرت (هرگز، به ندرت، گاهی اوقات، اغلب، و بسیار مکرر) برای همه سؤال‌ها به ارزشیابی کیفی عملکرد معلمان در میزان استفاده از ارزشیابی تکوینی در کلاس می‌پردازد و در واقع سواد معلمان را از ارزشیابی تکوینی را نمی‌سنجد. محققین می‌توانستند از مقیاس لیکرت برای بعد عملی معلمان در ارزشیابی تکوینی استفاده کنند و برای بعد مفهومی از گزینه‌های (صحیح، غلط) در سنجش سواد معلمان استفاده کنند. در یک مطالعه مرتبط توسط Kaya et al (2021)، آگاهی معلمان از ارزشیابی تکوینی را بر اساس به‌کارگیری دانش ارزشیابی تکوینی در سه سطح بی‌تجربه، التقاتی و آگاه طبقه‌بندی شد. با این حال، این طبقه‌بندی چارچوب ظریف‌تر را در نظر نمی‌گیرد و مطالعه بر اساس مؤلفه‌های اساسی ارزشیابی تکوینی صورت نگرفته است و به معرفی مؤلفه‌های مورد اندازه‌گیری نپرداخته است. در مطالعه یادگارزاده (۲۰۰۵) برای بررسی میزان آگاهی معلمان از روش‌های ارزشیابی تکوینی و میزان کاربرد آن در کلاس درس مؤلفه‌های دانش پایه (مفاهیم، تعاریف، فرایندها)، روش‌ها و مراحل اجرایی ارزشیابی تکوینی را در نظر می‌گیرد باتوجه‌به اینکه ما در عصر فناوری هستیم این مطالعه و پژوهش‌های دیگر تا کنون راهبردهای ارزشیابی تکوینی نوین را در معلمان بررسی نکرده‌اند. در نتیجه، نیاز مبرمی به توسعه ابزارهای تحقیق جامع، مانند پرسشنامه‌های جدید، برای ارزیابی بهتر آگاهی معلمان و اجرای روش‌های ارزشیابی تکوینی وجود دارد. این شکاف در تحقیقات بر ضرورت توسعه حرفه‌ای هدفمند و

جامعه مورد بررسی در این پژوهش معلمان شهر تهران است. در این مطالعه با استفاده از نمونه‌گیری تصادفی طبقه‌ای برای اطمینان از حضور مناطق مختلفی که جمعیت بر اساس کیفیت زندگی و سطح برخورداری به اقشاری تقسیم می‌شود که به‌عنوان توسعه‌یافته، متوسط، محروم و نیازمند مداخله طبقه‌بندی می‌شوند (Ahmadi & Jahangard, 2020). برای تعیین حجم نمونه لازم از هر قشری، محققان نمونه را با توجه به نسبت معلمان در هر دسته کیفیت زندگی در مناطق مختلف محاسبه می‌کنند. پس از آن، معلمان به طور تصادفی از هر قشری با استفاده از روش نمونه‌گیری تصادفی طبقه‌ای انتخاب می‌شوند تا اطمینان حاصل شود که نمونه منعکس‌کننده تنوع مدارس موجود در کیفیت زندگی و سطوح برخورداری در هر منطقه است که در این پژوهش مناطق (۵، ۶، ۱۵، ۱۸) شهر تهران به‌صورت تصادفی انتخاب شده‌اند.

در این مطالعه رضایت آگاهانه^۵ از همه شرکت‌کنندگان گرفته شد و محرمانه بودن و ناشناس بودن آنها تضمین شد. طرح تحقیق مورد تأیید کمیته اخلاق مربوطه قرار خواهد گرفت و در جمع‌آوری، تحلیل و گزارش‌دهی داده‌ها، کلیه دستورالعمل‌های اخلاقی رعایت شد. شرکت‌کنندگان ۸۱۱ معلم از کلیه مقاطع تحصیلی در مدارس دولتی چهار منطقه (۵، ۶، ۱۵، ۱۸) بودند. از این تعداد، همه آنها خانم بودند. معلمان منطقه ۱۵ با ۱۳/۹۲ درصد بیشترین مشارکت را در این پژوهش داشتند و منطقه ۶ با ۱۲/۳۵ درصد کمترین مشارکت را در این پژوهش داشتند. در نهایت کل مشارکت معلمان در این مناطق ۱۲/۴۶ درصد بود. شرکت‌کنندگان در این پژوهش از ۲۳ تا ۵۹ سال بودند که بیشترین مشارکت (۳۳/۴۲) مختص طبقه سنی در بازه ۴۱-۵۰ سال است و سطح تحصیلات اکثر شرکت‌کنندگان (۵۲/۷۲) در سطح کارشناسی است و سابقه کار بیشتر آنها (۴۰/۹۷) در بازه ۱-۱۰ سال است. باین‌حال، توزیع بر اساس سن و سطح تحصیلات و سابقه کار همگن نبود. این نمونه به طور تصادفی به دو نمونه فرعی تقریباً ۵۰٪ تقسیم شد. نمونه فرعی اول برای تحلیل عاملی اکتشافی و دومی برای تحلیل عاملی تأییدی استفاده شد. در نهایت داده‌های کمی با ابزارهای SPSS، R و EXCEL محاسبه و تحلیل شدند.

یافته‌ها

پس از پایان آزمون و گردآوری ۸۱۱ پاسخ‌نامه از افراد، در طی عملیات پیش‌پردازش داده‌ها، پاسخ‌نامه افراد با کلید پاسخ‌نامه پرسش‌نامه، کدگذاری مجدد شد، ۳ پاسخ‌نامه به دلیل تعداد سؤال‌های بدون پاسخ زیاد یا داده‌های گمشده حذف شد و ۸۰۸

بیش از حد بر روی روال‌های خودکار و تجارب شخصی تکیه کنند بدون اینکه آنها را برای موقعیت‌های جدید تنظیم کنند باعث عدم انعطاف‌پذیری و ناسازگاری با موقعیت‌های جدید و استفاده از اطلاعات قدیمی یا نامربوط می‌شود؛ لذا پرورش فرهنگ اقدام هوشمندانه در آموزش برای بهبود عملکرد، تقویت نوآوری و ایجاد موفقیت پایدار اهمیت دارد که در اینجا لزوم سنجش آگاهی معلمان از راهبردهای ارزشیابی تکوینی نوین را مورد تأکید قرار می‌دهد. در نتیجه برای سنجش سواد ارزشیابی تکوینی معلمان، توسعه ابزارهای نظریه‌محور که به طور خاص به این حوزه می‌پردازد، ضروری است. چنین ابزارهایی نه تنها به ارزیابی شایستگی‌های فعلی کمک می‌کنند، بلکه به تلاش‌های هدفمند توسعه حرفه‌ای کمک می‌کنند و در نهایت منجر به بهبود شیوه‌های آموزشی و نتایج بهتر دانش‌آموز می‌شوند.

روش پژوهش

ساخت ابزار پرسش‌نامه‌ای: طرح تحقیق شامل تدوین ابزارهای اندازه‌گیری قابل‌اعتماد و معتبر است، شامل یک رویکرد هفت‌مرحله‌ای سیستماتیک برای اطمینان از ایجاد مقیاس‌های قابل‌استفاده برای اهداف تحقیقاتی است شامل: (۱) شناسایی نیاز به مقیاس جدید (۲) تعریف ساختار (۳) تولید سؤال (۴) اصلاح سؤال (۵) تست مقدماتی (۷) اعتبارسنجی (Clark & Watson, 1995; Miller - Carpenter, Serena, 2018). در این مطالعه تدوین ابزار سنجش آگاهی معلمان از ارزشیابی تکوینی از نوع خودگردان، ساختاریافته و متشکل از ۹۸ سؤال در مقیاس دوازده‌گانه (صحیح، غلط) که ابعاد سؤال‌های آن عبارت‌اند از: (۱) ارائه دانش (۲)، سیستم شناختی (بازیابی، درک مطلب، تحلیل، بهره‌برداری از دانش) (۳) فراشناخت (۴) تفکر نظام شخصی (۵) راهبرد ارزشیابی تکوینی سنتی (۶) راهبرد ارزشیابی تکوینی نوین است.

سپس شاخص روایی محتوای سطح سؤال (I-CVI) پرسش‌نامه توسط ۷ نفر از متخصصان، هر سؤال بر اساس مقیاس لیکرت (غیرمرتبط، نیاز به بازبینی اساسی، مرتبط اما نیاز به بازبینی، کاملاً مرتبط) بررسی شد تا ارتباط آن را با ساختار مورد اندازه‌گیری تعیین کنند. با محاسبه I-CVI از ۹۸ سؤال، ۲۴ سؤال حذف و ۱۵ سؤال بازننگری شد و بعد از تأیید سؤال‌های بازننگری شده توسط متخصص، پرسش‌نامه ۷۴ سؤالی پس از اخذ مجوز توسط دانشگاه، سازمان آموزش و پرورش شهر تهران و در انتها با اجازه مدیران مدارس لینک پرسش‌نامه در پیام رسان ایتا و به برای معلمان به اشتراک گذاشته شد سپس در سایت پرس لاین سؤال‌های پرسش‌نامه پاسخ داده شد. مشارکت آنها، داوطلبانه و با رضایت آنها بود.

۰/۸۸ است و این یعنی همسانی درونی سؤال‌ها بعد از حذف سؤال‌هایی که همبستگی منفی داشتند افزایش یافته است؛ بنابراین پرسش‌نامه ۳۹ سؤالی برای تحلیل عاملی اکتشافی بکار می‌رود.

آزمون کروییت بارتلت^۶

یک آزمون آماری است که در تحلیل عاملی برای تعیین همبستگی متغیرهای مشاهده شده استفاده می‌شود که نشان می‌دهد ممکن است برای تشخیص ساختار مناسب باشند. این آزمون ارزیابی می‌کند که آیا ماتریس همبستگی یک ماتریس همانی^۷ است یا خیر، به این معنی که متغیرها همبستگی ندارند. (Tobias & Carlson, 1969). در این مطالعه آزمون کروییت بارتلت برای پرسش‌نامه ۳۹ سؤالی ($df=741$, $p=0$)، $\chi^2=16272.821$ معنادار است و p از ۰/۰۵ کوچکتر است و ماتریس همبستگی دارای اطلاعات معناداری است بنابراین می‌توان تحلیل عاملی اکتشافی را انجام داد.

اندازه‌گیری کفایت نمونه‌گیری کایزر-میر-الکین^۸ (KMO)

یک آزمون آماری است که در تحلیل عاملی برای ارزیابی مناسب بودن داده‌ها برای تحلیل عاملی استفاده می‌شود. آزمون KMO نسبت واریانس بین متغیرهایی را که ممکن است واریانس مشترک^۹ باشند، ارزیابی می‌کند. (Nkansah, 2018). در این پژوهش آزمون کفایت نمونه‌گیری کایزر مایر اولکین (KMO) پرسش‌نامه ۳۹ سؤالی تقریباً برابر ۰/۳۷ است نتیجه این آزمون نشان می‌دهد که داده‌ها همبستگی ضعیفی با هم دارند؛ لذا برای تحلیل ایده‌آل نیستند.

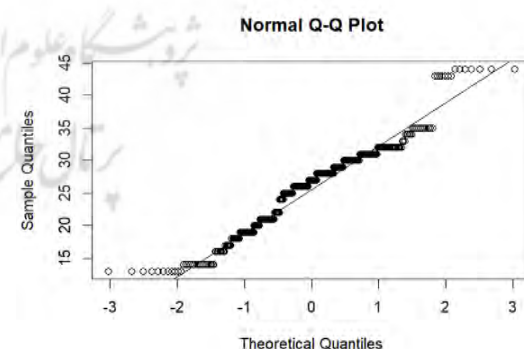
تحلیل مؤلفه اصلی (PCA)^{۱۰}

یک تکنیک آماری است که برای کاهش ابعاد یک مجموعه داده با حفظ واریانس تا حد ممکن استفاده می‌شود. در PCA، متغیرهای اصلی به مجموعه جدیدی از متغیرهای غیرهمبسته به نام مؤلفه اصلی تبدیل می‌شوند. این مؤلفه‌ها ترکیبی خطی از متغیرهای اصلی هستند و بر اساس میزان واریانس که در داده‌ها توضیح می‌دهند مرتب می‌شوند. بردار ویژه با بالاترین ارزش ویژه اولین مؤلفه اصلی را نشان می‌دهد، دومین ارزش ویژه مربوط به مؤلفه اصلی دوم است و غیره. زیرمجموعه‌ای از مؤلفه‌های اصلی انتخاب می‌شود که مقدار قابل توجهی از واریانس در داده‌ها را ثبت کند و در عین حال ابعاد آن را کاهش دهد (Jaques, Depoian, Murrell, Xie, ۲۰۱۸).

پاسخ‌نامه باقی ماند؛ یعنی نسبت مقادیر از دست رفته ۰/۳۷٪ است. پاسخ‌نامه‌های با بیش از ۸٪ مقادیر گم‌شده حذف شدند. برای استانداردسازی داده‌ها و حذف نقاط پرت با محاسبه نمره Z، افرادی که نمره‌های خارج از (۳-، ۳) داشتند حذف شدند در نهایت ۷۸۸ پاسخ‌نامه باقی ماند و سؤال‌هایی که قدرت افتراقی نداشتند؛ یعنی همه افراد به آنها پاسخ یکسان داده بودند و نمره Z آنها یکسان بود سؤال‌های (۱، ۲، ۳، ۴، ۱۱، ۱۲، ۱۵، ۲۳، ۲۵، ۳۵، ۴۴، ۴۵، ۴۸، ۵۰، ۵۴، ۵۶، ۶۱، ۷۱) حذف شدند. از حذف سؤالات، افراد حذف شده (به جز ۳ نفر) برگردانده شدند دوباره هنجاریابی شدند این بار سؤال‌هایی که دامنه کجی خارج از دامنه (۲-، ۲) و کشیدگی خارج از دامنه (۳-، ۳) داشتند سؤال‌های (۱۶، ۲۴، ۲۶، ۲۸، ۳۲، ۳۷، ۴۷) حذف شدند. با هنجاریابی مجدد سؤال‌ها، سؤال (۴۹) حذف شد. در نهایت ۴۸ سؤال برای تحلیل عاملی اکتشافی باقی ماند.

دترمینان ماتریس همبستگی در این پژوهش دترمینان (3.254996×10^{-3}) است. یعنی کمتر از ۰.۰۰۰۰۱ است. واریانس ماتریس همبستگی برای محاسبه اشتراک و بارهای عاملی استفاده می‌شود، بنابراین اگر نتوان آن را معکوس کرد، نمی‌توان تحلیل عاملی را انجام داد. برای حل این مشکل، ممکن است لازم باشد یک یا چند متغیر را از تحلیل حذف شود یا از روش‌های جایگزین مانند تحلیل مؤلفه‌های اصلی یا مدل‌سازی معادلات ساختاری استفاده شود.

بررسی خطی بودن: برای اطمینان از اینکه روابط بین متغیرها خطی است، نمودار (۱) نمودار QQ رسم شد زیرا تحلیل عاملی روابط خطی را فرض می‌کند.



شکل ۱: نمودار Q-Q

با بررسی همبستگی سؤال با مجموع سؤال‌ها، نتایج نشان می‌دهد که سؤال‌های (۵، ۱۳، ۲۲، ۳۱، ۳۶، ۳۸، ۴۲، ۵۳، ۵۷) دارای همبستگی منفی هستند؛ بنابراین از مجموع سؤال‌ها حذف می‌شوند. پرسش‌نامه ۴۸ سؤالی ضریب همسانی درونی ($\alpha=0.81$) خوبی دارد. اما آلفای کرونباخ ۳۹ سؤال باقی‌مانده

نمودار اسکری (۲) و نمودار تحلیل موازی (۳) تصویر گرافیکی مقدار ویژه برای هر عامل را استخراج کرده‌اند. مقدار واریانس تبیین شده (ارزش ویژه) به ترتیب با استخراج عامل‌های بعدی شیب آن به سرعت کاهش می‌یابد. با شمارش نقاط نمودارهای فوق و به طور دقیق‌تر محاسبات مقادیر ویژه اولیه، تعیین‌کننده مقادیر ویژه ماتریس همبستگی است و عامل‌های باقی‌مانده در تحلیل را مشخص می‌کند. دوازده عامل دارای مقدار ویژه بالاتر از یک است که در این تحلیل عامل اول بارزش ویژه (۸/۴۴) بالاترین و عامل دوازدهم بارزش ویژه (۱/۰۱۲) کمترین مقدار را در ارزش ویژه دارند. زیرا ارزش ویژه برای هر عامل برابر تعداد سؤال‌های آن است. سپس با چرخش اکوامکس^{۱۷}، که تعداد عامل‌های در حال چرخش را خودتنظیم می‌کند و بیشتر ترکیبی از واریمکس و کوارتیمکس است. در تحلیل ۶ مؤلفه با مقدار ویژه بزرگ‌تر از یک باقی‌مانده‌اند که تقریباً ۵۹/۲ درصد از واریانس کل را تبیین می‌کنند.

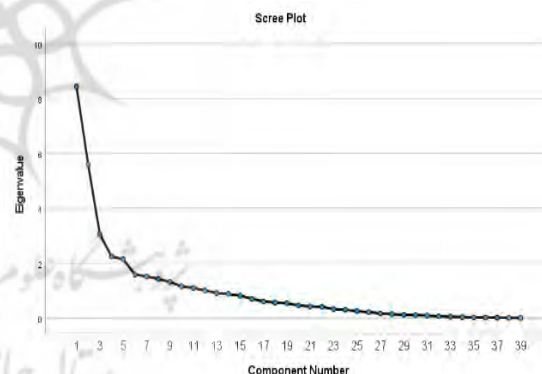
ارزیابی بار عاملی سؤال‌ها

ارزیابی بار عاملی شامل تجزیه و تحلیل همبستگی بین متغیرها و عوامل برای تعیین قدرت و جهت آنها است. بارهای عاملی بالا (نزدیک به ۱ یا ۱-۱) نشان دهنده روابط قوی بین متغیرها و عوامل خاص است و این متغیرها را کلیدی برای تعریف عامل می‌کند. بارگذاری کم (نزدیک به ۰) نشان دهنده رابطه ضعیف یا عدم وجود رابطه با هیچ عاملی است. بارگذاری متقاطع، که در آن متغیرها به طور متوسط بر روی چندین عامل بارگذاری می‌شوند، ممکن است نشان دهنده ابهام یا چند بعدی بودن باشد. بر اساس الگوهای بارگذاری، محققان پرسش‌نامه‌ها را با حفظ سؤالات پربار، اصلاح یا حذف سؤال‌های کم‌بار یا متقاطع، و اطمینان از ساختار عاملی واضح و قابل تفسیر، اصلاح می‌کنند. این فرآیند اعتبار و وضوح ابزار اندازه‌گیری را با همسوسازی متغیرها با عوامل متمایز افزایش می‌دهد (Widaman, 1990; Winter & Dodou, 2012). با استفاده از چرخش اکوامکس و سؤال‌هایی که بار عاملی بالاتر از ۰/۳ و بر روی یک عامل بارگذاری شده بودند یا سؤال‌هایی که بیشترین بار عاملی را بر عامل خاصی داشتند استخراج شدند. در بار عاملی که به ۱ نزدیک‌تر باشد بهتر است. در نتیجه برای عامل اول ۹ سؤال، عامل دوم ۳ سؤال، عامل سوم ۶ سؤال، عامل چهارم ۳ سؤال، عامل پنجم ۴ سؤال، عامل ششم ۳ سؤال استخراج شد (جدول ۵).

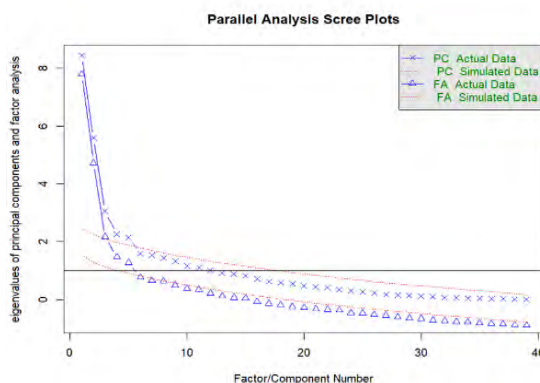
آلفای کرونباخ^{۱۸} یک معیار آماری است که برای ارزیابی همسانی درونی یا پایایی یک مقیاس یا آزمون استفاده می‌شود. آلفای کرونباخ به عنوان یک حد پایینی از پایایی عمل می‌کند،

تعیین تعداد عوامل

رویکردهای رایج استفاده از نمودار اسکری^{۱۱} که یک روش گرافیکی است که ارزش ویژه^{۱۲} عوامل را به ترتیب نزولی نمایش می‌دهد. محققان به دنبال نقطه‌ای در نمودار می‌گردند که در آن ارزش ویژه مسطح می‌شوند و تعداد عواملی که باید حفظ شوند را نشان می‌دهد. ملاک کایزر^{۱۳} حفظ عوامل بارزش ویژه بیشتر از ۱ را پیشنهاد می‌کند. با درصد واریانس تبیین^{۱۴} شده محققان عواملی را حفظ می‌کنند که درصد مشخصی از واریانس را در داده‌ها توضیح می‌دهد، مانند عواملی که حداقل ۷۰ یا ۸۰ درصد از واریانس کل را تبیین می‌کند. تحلیل موازی^{۱۵} ارزش ویژه به دست آمده از داده‌های واقعی را بارزش ویژه از داده‌های تصادفی مقایسه می‌کند. عامل‌هایی بارزش ویژه بالاتر از عوامل حاصل از داده‌های تصادفی حفظ می‌شوند. آزمون جزئی حداقل میانگین ویلیر^{۱۶} (MAP) میانگین مجذور ضرایب همبستگی جزئی را برای تعیین تعداد بهینه عوامل بررسی می‌کند (Cattell, 1988; de Mesquita & Kosteljk, 2021; Jöreskog, Olsson & Wallentin, 2016).



شکل ۲: نمودار اسکری



شکل ۳: تحلیل موازی

عامل ذکر شده است که در آن عامل پنجم ($\alpha = 0.81$) بیشترین پایایی را دارد و عامل دوم با آلفای کرونباخ 0.61 کمترین پایایی را دارد.

که نشان‌دهنده حداقل سطح همسانی درونی است که می‌توان از یک مقیاس انتظار داشت (Kılıç, 2016). در این مطالعه آلفای کرونباخ کل سازه تقریباً 0.79 برآورد شده است؛ بنابراین مقیاس، پایایی قابل‌قبولی دارد و در جدول (۲) آلفای کرونباخ هر

جدول (۲) آلفای کرونباخ عامل‌های سازه

عامل	عامل ۱	عامل ۲	عامل ۳	عامل ۴	عامل ۵	عامل ۶	کل سازه
آلفای کرونباخ	0.77	0.61	0.78	0.70	0.81	0.78	0.79
تفسیر	قابل قبول	متوسط	قابل قبول	قابل قبول	مناسب	قابل قبول	قابل قبول

مناطق که ممکن است مدل برای همسویی بهتر با داده‌ها نیاز به اصلاح داشته باشد، استفاده می‌شود (Okan & Ekşi, 2020; Nie, 2020; Mohd Matore, Zainal, Mohd Noh, Khairani & Abd Razak, 2021). لذا تخمین شاخص‌های برازش در جدول (۵) شرح داده شده است.

ارزیابی برازش مدل: در تحلیل عاملی تأییدی (CFA)، ارزیابی برازش مدل با مقایسه ماتریس کوواریانس پیش‌بینی‌شده از مدل نظری مشخص شده با ماتریس کوواریانس واقعی داده‌های مشاهده شده محاسبه می‌شود. این ارزیابی از چندین شاخص برازش موجود در جدول (۳) استفاده می‌کند. ترکیبی از شاخص‌های برازش برای ارزیابی برازش کلی مدل و شناسایی

جدول ۳: شاخص‌های برازش برای مدل تدوین شده و مدل اصلاحی

NPar	RMSEA	CI90%		SRMR	CFI	TLI	مدل‌ها	
		LO	UP				تدوین	اصلاح
۷۱	0.176	0.172	0.181	0.135	0.410	0.335	تدوین	اصلاح
۶۷	0.166	0.161	0.171	0.126	0.464	0.377	تدوین	اصلاح
۶۶	0.185	0.181	0.190	0.212	0.241	0.267	تدوین	اصلاح
۶۲	0.179	0.171	0.184	0.193	0.367	0.277	تدوین	اصلاح

مدل مقدار برازش آنها افزایش یافته است. از سوی دیگر مقادیر آنها در مدل مرتبه اول بیشتر است. مقادیر ریشه میانگین مربعات خطای برآورد (RMSEA) و ریشه دوم میانگین مجزورات باقی‌مانده (SRMR) در مدل‌های تدوینی و اصلاحی، بالاتر از 0.1 است و دارای برازش ضعیف هستند. در نتیجه مقدار مدل مرتبه اول اصلاحی به برازش نزدیک‌تر است.

اصل مقتصد در تحلیل عاملی بر انتخاب ساده‌ترین و مختصرترین مدلی که داده‌ها را بدون پیچیدگی غیرضروری به‌اندازه کافی توضیح می‌دهد، تأکید دارد و تضمین می‌کند که مدل‌ها هم از نظر آماری صحیح و هم از نظر مفهومی معنادار هستند. هنگام ارزیابی سؤالات در تحلیل عاملی، مقتصد در کنار معیارهای مناسب برای تعیین تعداد بهینه عوامل و تفسیر بارهای عاملی استفاده می‌شود که منجر به تحلیل قوی و روشنگری می‌شود (Schmitt & Sass, 2011; Beauducel, 2015).

برای رسیدگی به محدودیت‌های CFA سنتی، محققان اغلب از رویکرد اصلاح مدل تعقیبی^{۱۹} (PMM) استفاده می‌کنند. این روش امکان تنظیم مدل را برای بهبود برازش بدون بازتعریف کامل آن فراهم می‌کند. هدف اصلی PMM افزایش برازش مدل با آزادکردن برخی از مفروضات دقیق در مورد کوواریانس‌های باقیمانده است. در این مطالعه اصلاحیه‌هایی مانند بارگذاری سؤال (۱) بر عامل سوم، بارگذاری سؤال (۲) بر عامل اول، حذف سؤال‌های (۹، ۷) از عامل اول و در انتها حذف سؤال (۲۵) از عامل پنجم صورت گرفته است و نتایج اصلاح‌های مختلف در جداول (۴، ۳) ارائه شده است. هرچه شاخص برازش مدل به نقاط مورد قبول نزدیک‌تر باشد احتمال ارتقا مدل افزایش می‌یابد. در جدول (۵)، مدل‌های تدوینی و اصلاح به دلیل اینکه شاخص‌های برازش تطبیقی (CFI) و شاخص توکر لوئیس (TLI) کمتر از 0.9 هستند برازش ندارند، ولی با اصلاح

جدول (۴) شاخص‌های مقتصد برای مقایسه مدل تدوین شده و مدل اصلاح شده

مدل	BIC	AIC
تدوین	۱۰۴۰۵.۴۷۳	۱۰۱۲۱.۳۷۳
اصلاح	۹۱۶۳.۱۹۱	۸۸۹۵.۰۹۶
تدوین	۱۰۸۷۶.۹۱۰	۱۰۶۱۲.۸۱۷
اصلاح	۹۶۵۸.۷۶۷	۹۴۱۰.۶۷۹

اصلاح ترکیبی مرتبه دوم کمتر است بنابراین مدل بهتر و مقتصدتری است.

نام‌گذاری اکنون می‌توان عوامل را باتوجه به مجموعه سؤال‌ها، در جدول (۵) نام‌گذاری کرد: عامل اول سیستم شناختی، عامل دوم تفکر نظام شخصی، عامل سوم راهبرد ارزیابی تکوینی نوین، عامل چهارم راهبرد ارزیابی تکوینی سنتی، عامل پنجم ارائه دانش و عامل ششم فراشناخت است.

در جدول (۴) اصلاح مدل‌های اصلاح در شاخص‌های مقتصد AIC و BIC مقدار کوچکتری نسبت به مدل‌های دیگر دارند زیرا با حذف سؤال پیچیدگی مدل کاهش یافته است. از سوی دیگر مقدار مدل‌های مرتبه دوم در شاخص‌های مقتصد بیشتر از مرتبه اول است، زیرا پیچیدگی آن افزایش می‌یابد. از سوی دیگر مقدار شاخص‌های مقتصد مدل اصلاح ترکیبی مرتبه اول از مدل

جدول (۵) سؤال‌های مدل تدوین

ردیف	نام‌گذاری عامل	سؤال	کلید پاسخ	گزینه پاسخ	
				غلط	صحیح
۱.	سیستم شناختی	من در ارزشیابی تکوینی تصورات نادرست دانش آموز را شناسایی و برطرف می‌کنم.	۱		
۲.		من به دانش‌آموزان کمک می‌کنم تا مهارت‌های خود را برای ایجاد راه‌حل‌ها و راهبردهای بالقوه برای ایجاد موانع توسعه دهند.	۰		
۳.		من با گنجاندن تحلیل خطا در ارزیابی، دانش‌آموزان را به تشخیص خطا و تصحیح آن، ترغیب می‌کنم.	۱		
۴.		من در ارزشیابی تکوینی، دانش‌آموزان را به تصمیم‌گیری بر اساس معیارهای مرتبط مانند امکان‌سنجی، اثربخشی، هزینه و تأثیر، راهنمایی می‌کنم.	۱		
۵.		من برای مدیریت بار شناختی اطلاعات ساده، برای دانش‌آموزان داربست مفهومی ارائه می‌دهم.	۰		
۶.		من در ارزشیابی تکوینی، با فعالیتهای گروهی، فرصتهایی را برای دانش‌آموزان فراهم می‌کنم تا مسائل پیچیده را حل کنند.	۱		
۷.		من برای درک مطلب دانش‌آموزان، مانع ارتباط اطلاعات جدید با اطلاعات قبلی می‌شوم.	۰		
۸.		من در ارزشیابی تکوینی استدلال منطقی دانش‌آموزان را ترغیب می‌کنم.	۱		
۹.		من در ارزشیابی تکوینی، به ارتباط مفاهیم با تجارب غیرشخصی دانش‌آموزان می‌پردازم.	۰		
۱۰.	فرایند تفکر	من با آموزش تکنیک‌های مدیریت هیجان مثبت و پرورش احساس منفی در دانش‌آموزان، محیطی امن برای یادگیری مؤثر ایجاد می‌کنم.	۰		
۱۱.		من در ارزشیابی پایانی، به پیگیری تغییر در باورهای کارآمدی دانش‌آموزان در طول زمان توجه می‌کنم.	۰		
۱۲.		من دانش‌آموزان را در ایجاد راهبردهایی برای تنظیم احساسشان راهنمایی می‌کنم، به‌ویژه آنهایی که ممکن است مانع یادگیری و عملکرد شود.	۱		
۱۳.	تجربه	من می‌دانم که با پلت فرم ارزیابی همتایان لرن ایول دانش‌آموزان کار همسالان خود را بدون معیارهای مشخص بررسی می‌کنند.	۰		
۱۴.		من می‌دانم که نیروی ترکیب عناصر تعاملی مختلف، مانند آزمون‌ها، نظرسنجی‌ها و محتوای چندرسانه‌ای را امکان‌پذیر می‌کند.	۱		

۱۵	من می‌دانم که سیستم قضاوت تمرین آردوینو با نظارت انسانی کار عملی دانش‌آموزان را ارزیابی می‌کند.	۰	
۱۶	در برنامه شاد، نظرسنجی از دانش‌آموزان امکان‌پذیر نیست.	۰	
۱۷	من می‌دانم که متاتیتور می‌تواند داربست‌های تطبیقی را برای هدایت دانش‌آموزان در مسیر یادگیری فراهم کند.	۱	
۱۸	من می‌دانم که اجزای کلیدی ارزیابی آنلاین مبتنی بر شایستگی (OCBA) عبارت‌اند از: (الف) مدل شایستگی، (ب) مدل ارزیابی، (ج) مدل یادگیرنده، (د) سیستم تألیف.	۱	
۱۹	من در ارزشیابی تکوینی، از دانش‌آموزان درباره رویدادهای روزمره سؤال می‌پرسم	۰	
۲۰	من می‌دانم از گوگل فرم نمی‌توان برای ارزیابی استفاده کرد.	۰	
۲۱	من در ارزشیابی تکوینی، از دانش‌آموزان می‌خواهم تا سطح درک خود از مفاهیم آموزشی را با یکی از رنگ‌های چراغ‌راه‌نما (سبز، زرد، قرمز) نشان دهند.	۱	
۲۲	من دانش‌آموزان را به اصلاح بی‌درنگ برای درک مطلب ترغیب می‌کنم.	۰	
۲۳	من در مرحله تداعی، به اصلاح خطاهای درک مطلب دانش‌آموزان می‌پردازم.	۱	
۲۴	من در کلاس از یک راهبرد ارزشیابی تکوینی استفاده می‌کنم.	۰	
۲۵	من با آموزش قوانین به دانش‌آموزان، پاسخ واضح و مستقیم در یک موقعیت خاص ارائه می‌دهم.	۱	
۲۶	در ارزشیابی تکوینی، نیازی به ارزیابی توانایی دانش‌آموزان برای نظارت بر وضوح تفکرشان نیست.	۰	
۲۷	من با ارزشیابی تکوینی می‌توانم، بر توانایی دانش‌آموزان برای نظارت بر صحت کارشان تمرکز کنم.	۱	
۲۸	من راهبردهای فراشناختی (برنامه‌ریزی، نظارت و ارزیابی) را به دانش‌آموزان آموزش می‌دهم.	۱	

بحث

این مطالعه چارچوب جامعی را برای تحویل دانش و ارزیابی تکوینی ترسیم می‌کند و بر همسویی آن با سطوح مختلف شناختی برای ایجاد تجربیات یادگیری معنادار تأکید می‌کند. نتایج کلیدی عبارت‌اند از: ۱) رویکرد تلفیقی: ارزیابی‌های تکوینی باید بر انواع مختلف دانش متمرکز شوند - درک حقایق، توسعه تفکر انتقادی و به‌کارگیری مهارت‌های عملی. این رویکرد اثربخشی تدریس را افزایش می‌دهد و نتایج یادگیری دانش‌آموزان را بهبود می‌بخشد. ۲) سطوح شناختی: الف) سطح ۱ (بازیابی): بر ارزیابی توانایی دانش‌آموزان و معلمان برای بازیابی و به‌کارگیری اطلاعات، با استفاده از روش‌هایی مانند یادآوری و استدلال استنباطی تأکید می‌کند. ب) سطح ۲ (درک مطلب): تمرکز بر نحوه ذخیره دانش معلمان به طور مؤثر، شامل تکنیک‌های یکپارچه سازی و نمادسازی برای افزایش درک و حفظ است. ج) سطح ۳ (تحلیل): چرخه‌های تکراری تجزیه و تحلیل را برای عمیق‌تر کردن درک، با استفاده از فرآیندهایی مانند طبقه‌بندی و تحلیل خطا تشویق می‌کند. د) سطح ۴ (بهره برداری از دانش): اهمیت استفاده معلمان از دانش خود برای طراحی راهبردهای آموزشی مؤثر از جمله

تصمیم‌گیری و حل مسئله را برجسته می‌کند. ۳) سطح ۵ (فراشناخت): شامل تعیین اهداف یادگیری روشن، نظارت بر پیشرفت، و تأمل در شیوه‌های آموزشی برای اطمینان از وضوح و دقت است. ۴) سطح ۶ (تفکر نظام شخصی): تأثیر سیستم‌های شخصی را بر تعامل و انگیزه تشخیص می‌دهد و بر نیاز به آگاهی از نگرش‌ها و پاسخ‌های عاطفی در یادگیری تأکید می‌کند. ۵) نقش معلم: معلمان تشویق می‌شوند تا از دانش خود به‌صورت راهبردی برای تقویت موفقیت دانش‌آموزان از طریق طراحی آموزشی مؤثر و شیوه‌های ارزیابی تکوینی استفاده کنند. به‌طور کلی، این مطالعه از رویکردی ساختاریافته برای ارزشیابی تکوینی حمایت می‌کند که نه تنها یادگیری دانش‌آموزان را افزایش می‌دهد، بلکه از معلمان در نقش‌های آموزشی آنها نیز حمایت می‌کند. این راهبرد چند وجهی برای توسعه مهارت‌های تفکر انتقادی و پرورش درک عمیق‌تر از موضوع در بین دانش‌آموزان ضروری است. این مطالعه نقش مکمل راهبردهای ارزشیابی تکوینی سنتی و مدرن را در تقویت یادگیری دانش‌آموز و اطلاع‌رسانی به شیوه‌های آموزشی برجسته می‌کند. درحالی‌که هر دو ارزیابی تکوینی سنتی و مدرن به تقویت یادگیری دانش‌آموز کمک می‌کنند، راهبردهای مدرن رویکرد تعاملی‌تر و

فرایند محوری را ارائه می‌دهند که تعامل و درک بیشتر را تقویت می‌کند.

نتیجه‌گیری

در اینجا تکامل و تنوع تحقیقات در مورد آگاهی معلمان از ارزشیابی تکوینی مورد بحث قرار گرفت و این مطالعه نشان می‌دهد که چگونه مطالعات مختلف از زوایای مختلف و با سطوح مختلف عمق به موضوع نزدیک شده‌اند. مطالعه فعلی تدوین ابزاری جامع و چندبعدی برای سنجش آگاهی معلمان از ارزشیابی تکوینی با شش مؤلفه (ارائه دانش، سیستم شناختی، سیستم فراشناختی، تفکر نظام شخصی، راهبرد ارزشیابی تکوینی سنتی، راهبرد ارزشیابی تکوینی نوین) است که به جنبه‌های نظری و عملی آگاهی ارزیابی تکوینی می‌پردازد. این مطالعه با تمایز بین راهبردهای ارزیابی تکوینی سنتی و مدرن، ماهیت در حال تحول شیوه‌های ارزیابی را تصدیق می‌کند. این مطالعه پیشرفت قابل توجهی را در زمینه تحقیقات ارزشیابی تکوینی نشان می‌دهد. زیرا کاربرد طبقه‌بندی مارزانو و کندال را در قالب ارزشیابی تکوینی گسترش داد و برای مفاهیم آن توضیحاتی جدید ارائه شد. در نهایت این مطالعه مؤلفه‌های بیشتری را برای سنجش آگاهی معلمان از ارزشیابی تکوینی در نظر گرفت. درحالی‌که ابزار و مدل نیاز به اصلاح بیشتری دارند، آنها پایه‌ای قوی برای تحقیق و تمرین آینده در سواد ارزیابی تکوینی ارائه می‌دهند.

موازین اخلاقی

در این مطالعه اصول اخلاق در پژوهش شامل اخذ رضایت آگاهانه از شرکت کنندگان و حفظ اطلاعات محرمانه آنها رعایت گردیده است.

تشکر و قدردانی

پژوهشگران مراتب قدردانی و تشکر خود را از کلیه شرکت کنندگان این پژوهش که با استقبال و بردباری، در روند استخراج نتایج همکاری نمودند، اعلام می‌دارند.

تعارض منافع

نویسندگان این مطالعه هیچ گونه تعارض منافی در انجام و نگارش آن ندارند.

واژه نامه

۱. صریح ساختن یادگیری
۲. ارتقای استقلال یادگیری
۳. جهت‌گیری عملکرد
۴. یادگیری را در مسیر ادامه دهید
۵. رضایت آگاهانه
۶. آزمون کرویت بارلت
۷. ماتریس همانی
۸. اندازه‌گیری کفایت نمونه‌گیری کایزر - میر - الکین
۹. واریانس مشترک
۱۰. تحلیل مؤلفه اصلی
۱۱. نمودار اسکری
۱۲. ارزش ویژه
۱۳. ملاک کایزر
۱۴. واریانس تبیین
۱۵. تحلیل موازی
۱۶. میانگین ولیسر
۱۷. چرخش اکوامکس
۱۸. آلفای کرونباخ
۱۹. رویکرد اصلاح مدل تعقیبی
1. Making learning explicit
2. Promoting learning autonomy
3. Performance orientation
4. Keep Learning on Track
5. Informed consent
6. Bartlett's Test of Sphericity
7. identity matrix
8. The Kaiser-Meier-Olkin (KMO) Measure of Sampling Adequacy
9. common variance
10. Principal Component Analysis (PCA)
11. Scree Plot
12. eigenvalues
13. Kaiser's Criterion
14. Percentage of Variance Explained
15. Parallel Analysis
16. Velicer's Minimum Average Partial (MAP) Test
17. equamax, varimax و quartimax
18. Cronbach's alpha
19. Post Hoc Model Modification (PMM)

فهرست منابع

- Ahmadi, A. and Jahangard, E. (2020). Ranking of the neighborhoods of Tehran metropolis in terms of development and quality of life: A fuzzy
- Arrafii, M.A., & Sumarni, B. (2018). Teachers' Understanding of Formative Assessment. *Lingua Cultura*, 12, 45-52.
- Beauducel, A. (2015). Searching for parsimony in loading patterns with non-zero secondary loadings. *arXiv: Applications*.
- Bennet, R. E. (2011). Formative Assessment: A critical review. *Assessment in Education: Principles, Policy and Practice*, 18(1), 5-15.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, 5(1), 7-74.
- Cattell, R.B. (1988). *The Meaning and Strategic Use of Factor Analysis*.

- Khodabakhshzadeh, H., Kafi, Z., & Hosseinnia, M. (2018). Investigating EFL Teachers' Conceptions and Literacy of Formative Assessment: Constructing and Validating an Inventory. *International Journal of Instruction*, 11(1), 139-152.
- Kılıç, S. (2016). Cronbach's alpha reliability coefficient -. *Journal of Mood Disorders*, 6, 47.
- Lee, H., Chung, H.Q., Zhang, Y., Abedi, J., & Warschauer, M. (2020). The Effectiveness and Features of Formative Assessment in US K-12 Education: A Systematic Review. *Applied Measurement in Education*, 33, 124 - 140.
- Linn, R. L. & Miller, M. D. (2005). *Measurement and assessment in teaching* (9th Eds.) Upper Saddle River, New Jersey: Pearson Merrill Prentice Hall.
- Marzano, R., & Kendall, J. (2007). *The New Taxonomy of Educational Objectives*.
- McMillan, J. H. (2017). *Classroom assessment: Principles and practice that enhance student learning and motivation*. Pearson.
- Mertler, C. A. (2004). Secondary teachers' assessment literacy: Does classroom experience make a difference. *American Secondary Education*, 33(1), 49 – 64.
- Mertler, C. A. (2009). Teachers' assessment knowledge and their perceptions of the impact of classroom assessment professional development. *Improving Schools*, 12(2), 101 – 113.
- Miller - Carpenter, Serena. (2018). Ten Steps in Scale Development and Reporting: A Guide for Researchers. *Communication Methods and Measures*. 12. 25-44. 10.1080/19312458.2017.1396583.
- Mohd Matore, M.E., Zainal, M.A., Mohd Noh, M.F., Khairani, A.Z., & Abd Razak, N. (2021). The Development and Psychometric Assessment of Malaysian Youth Adversity Quotient Instrument (MY-AQi) by Combining Rasch Model and Confirmatory Factor Analysis. *IEEE Access*, 9, 13314-13329.
- Nie, Y. (2020). A Survey of the Effectiveness Model of College English Teachers' Classroom Discourse Based on Computer Assisted Technology from the Perspective of Speech Act Theory. *Journal of Physics: Conference Series*, 1648.
- Clark, L.A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7, 309-319.
- de Mesquita, J.W., & Kosteljik, E.J. (2021). *Exploratory Factor Analysis (EFA). Marketing Analytics*.
- Desimone, L. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181-199.
- Dixon, K. (2020). *Formative Assessment in a Teacher Education Course*.
- Entika Fani, P., & Prabowo, A. B. (2019). Constraints In Implementing Formative Assessments: Teachers' Voice Toward The Principle Of Assessment For Learning.
- Ghasemi, V. (2013). *Structural Equation Modeling in Social Research Using Amos Graphics*. Sociologists.
- Haugen, H. J., & de Lange, T. (2024). Multiple choice as formative assessment in dental education. *European journal of dental education : official journal of the Association for Dental Education in Europe*, 10.1111/eje.13002. Advance online publication. <https://doi.org/10.1111/eje.13002>
- James, M., & Pedder, D. (2006). Beyond method: Assessment and learning practices and values. *The Curriculum Journal*, 17(2), 109-138.
- Jaqes, L.E., Depoian, A.C., Murrell, E., Xie, D., Bailey, C.P., & Guturu, P. (2022). Novel L1 PCA informed K-means color quantization. *Defense + Commercial Sensing*.
- Jöreskog, K. G., Olsson, U. H., & Wallentin, F. Y. (2016). *Multivariate analysis with LISREL*. Basel, Switzerland: Springer.
- Kaya, G., Atasoy, V., Candan-Helvacı, S., & Pektaş, M. (2021). The Role of Science Teachers' Awareness in their Classroom Practice of Formative Assessment. *Education & Science/Eğitim ve Bilim*, 46(205).
- Kazinyirako, J. P., Ntawiha, P., Buhigiro, J. L., & Ndayambaje, I. (2024). Exploring Formative Assessment Practices in Public Secondary Schools in Kicukiro District. *African Journal of Empirical Research*, 5(2), 388-396.

- Multivariate behavioral research, 25(1), 89–95.
https://doi.org/10.1207/s15327906mbr2501_11
- Winter, J.D., & Dodou, D. (2012). Factor recovery by principal axis factoring and maximum likelihood factor analysis as a function of factor pattern and sample size. *Journal of Applied Statistics*, 39, 695 - 710.
- Yan, Z., & Brown, G. T. L. (2021). Assessment for learning in the Hong Kong assessment reform: A case of policy borrowing. *Studies in Educational*
- Yan, Z., & Pastore, S. (2022). Are teachers literate in formative assessment? The development and validation of the Teacher Formative Assessment Literacy Scale. *Studies in Educational Evaluation*, 74, 101183.
- ylie, E. C., & Lyon, C. J. (2015). The fidelity of formative assessment implementation: issues of breadth and quality. *Assessment in Education: Principles, Policy & Practice*, 22(1), 140- 160.
- Nkansah, B.K. (2018). On the Kaiser-Meier-Olkin's Measure of Sampling Adequacy. *Mathematical theory and modeling*, 8, 52-76.
- Okan, N., & Ekşi, H. (2020). The Adaptation of the Moral Integrity Scale into Turkish and Analysis of the Psychometric Features. *Spiritual Psychology and Counseling*, 5, 327-340.
- Santoso, P.H., Istiyono, E., & Haryanto, H. (2022). Principal Component Analysis and Exploratory Factor Analysis of the Mechanical Waves Conceptual Survey. *JP3I (Jurnal Pengukuran Psikologi dan Pendidikan Indonesia)*.
- Schmitt, T.A., & Sass, D.A. (2011). Rotation Criteria and Hypothesis Testing for Exploratory Factor Analysis: Implications for Factor Pattern Loadings and Interfactor Correlations. *Educational and Psychological Measurement*, 71, 113 - 95.
- Tobias, S., & Carlson, J. E. (1969). Brief Report: Bartlett's Test Of Sphericity And Chance Findings In Factor Analysis. *Multivariate behavioral research*, 4(3), 375–377.
https://doi.org/10.1207/s15327906mbr0403_8
- Widaman K. F. (1990). Bias in Pattern Loadings Represented by Common Factor Analysis and Component Analysis.