

A Survey on the Success Prediction of Cultural Exercise Students in the Teachers Assessment Competency Test (ASLAH) Using Classifier Variable EncodingsH. Meshkati^{۱*}, A. Delavar^۲, N. Farrokhi^۳, A. Minaee^۴, GH. Kiyani^۵

۱. Ph.D student of Psychometric, Allameh Tabataba'i University of Tehran And a lecturer at the Farhangian University of Iran, Khorasan Razavi; ۲. Corresponding Author: Professor in assessment & measurement group, Allameh Tabataba'i University, Tehran, Iran; ۳. Associate Professor in assessment & measurement group, Allameh Tabataba'i University, Tehran, Iran; ۴. Assistant Professor in assessment & measurement group, Allameh Tabataba'i University, Tehran, Iran; ۵. Associate Professor in English Language group, Tarbiat Modarres University, Tehran, Iran

بررسی پیش‌بینی موفقیت مهارت‌آموزان دانشگاه فرهنگیان در آزمون ارزیابی صلاحیت معلمی (اصلاح) با استفاده از کدگذاری متغیرهای طبقه‌ایحسن مشکاتی^۱، علی دلوار^۲، نورعلی فرخی^۳، اصغر مینایی^۴، غلامرضا کیانی^۵

۱. دانشجوی دکتری سنجش و اندازه‌گیری دانشگاه علامه طباطبائی تهران و مدرس دانشگاه فرهنگیان ایرن، خراسان رضوی؛ ۲. استاد گروه سنجش و اندازه‌گیری، دانشگاه علامه طباطبائی، تهران، ایران (نویسنده مسئول)؛ ۳. دانشیار گروه سنجش و اندازه‌گیری، دانشگاه علامه طباطبائی، تهران، ایران؛ ۴. استادیار گروه سنجش و اندازه‌گیری، دانشگاه علامه طباطبائی؛ ۵. دانشیار گروه زبان انگلیسی، دانشگاه تربیت مدرس، تهران، ایران

Abstract

Purpose: The purpose of this study was to predict the success of the students of the Farhangian University in the ASLAH test based on class variables using coding in multivariate regression.

Method: The present research is applied in terms of purpose and is a correlation type. The research community of all students in ۲۸th Madeh in Farhangian University (۹۴-۹۵ years), with a total of ۲۸۲۳ people. The sample size is ۳۴۱ people. Samples were selected using clustering method. The data of this research are the results of the ASLAH test of Farhangian University in Shahriyar ۱۳۹۵.

Finding: The results showed that gender predicts the success of students in the ASLAH test. The study of subscales showed that gender significantly predicts success in cultural and functional subscales, but does not have a significant role in predicting the scores of written scientific subscales of four choices and micro teaching. The specialty field predicts the success of the students in the ASLAH test. The specialty field significantly predicts the success of the four-choice written, performance, and micro teaching, but does not play a significant role in predicting the scores of cultural subscales.

The Provincial Student predicts the success of the students in the ASLAH test. The study of subscales showed that the province significantly predicts the success of the four-choice, functional and cultural subtest, but did not play a significant role in predicting the scores of the micro teaching.

Key words: ASLAH test, assesment of teacher's competency, Farhangian University, coding in multivariate regression

چکیده

هدف: هدف پژوهش حاضر بررسی پیش‌بینی‌پذیری موفقیت مهارت‌آموزان دانشگاه فرهنگیان در آزمون اصلاح متغیرهای طبقه‌ای با استفاده از کدگذاری در رگرسیون چند متغیری می‌باشد.

روش: پژوهش حاضر از نظر هدف کاربردی و از حیث روش، از نوع همبستگی است. جامعه پژوهشی کلیه مهارت‌آموزان ماده ۲۸ سال ۹۴-۹۵ دانشگاه فرهنگیان کشور است که تعداد آنها ۲۸۲۳ نفر می‌باشد. حجم نمونه ۳۴۱ نفر تعیین شد. نمونه با استفاده از روش خوشه‌ای انتخاب شد. داده‌های این پژوهش نتایج حاصل از آزمون اصلاح دانشگاه فرهنگیان در شهریور ۱۳۹۵ می‌باشد.

یافته‌ها: نتایج نشان داد که جنسیت، موفقیت مهارت‌آموزان در نمره کلی آزمون اصلاح را پیش‌بینی می‌کند. بررسی خرده‌مقیاس‌ها نشان داد که جنسیت موفقیت در خرده‌آزمون‌های فرهنگی و عملکردی را به‌طور معناداری پیش‌بینی می‌کند و اما در پیش‌بینی نمرات خرده‌مقیاس‌های علمی کتبی چهار گزینه‌ای و تدریس خرد نقش معناداری ندارد. همچنین رشته تخصصی، موفقیت در آزمون اصلاح را پیش‌بینی می‌کند. بررسی خرده‌مقیاس‌ها نشان داد که رشته تخصصی، موفقیت در خرده‌آزمون‌های علمی کتبی چهار گزینه‌ای، عملکردی و تدریس خرد را به‌طور معناداری پیش‌بینی می‌کند و اما در پیش‌بینی نمرات خرده‌مقیاس فرهنگی نقش معناداری ندارد. استان محل مهارت‌آموزی، موفقیت در آزمون اصلاح را پیش‌بینی می‌کند. بررسی خرده‌مقیاس‌ها نشان داد که استان محل مهارت‌آموزی، موفقیت در خرده‌آزمون‌های علمی کتبی چهار گزینه‌ای، عملکردی و فرهنگی را به‌طور معناداری پیش‌بینی می‌کند اما در پیش‌بینی نمرات خرده‌مقیاس تدریس خرد نقش معناداری ندارد.

کلید واژه‌ها: آزمون اصلاح، ارزیابی صلاحیت معلم، رگرسیون، کدگذاری متغیرهای طبقه‌ای، دانشگاه فرهنگیان

Accepted Date: ۲۰۱۸/۰۵/۰۳

Received Date: ۲۰۱۷/۱۱/۱۳

دریافت مقاله: ۱۳۹۶/۰۸/۲۲

پذیرش مقاله: ۱۳۹۷/۰۲/۱۳

Email: delavarali@yahoo.com

*نویسنده مسؤول:

مقاله برگرفته از رساله دکتری نویسنده اول می‌باشد.

مقدمه و بیان مسأله

براساس گزارش اینگاورسون و همکاران (Ingvarson and et al., ۲۰۱۳) و طبق نظر هماهنگ‌کنندگان تحقیقات ملی^۱ (NRCs) در طرح تربیت‌معلم و توسعه ریاضیات^۲ (TEDS-M)، قبلاً بسیاری از کشورها تصور می‌کردند فارغ‌التحصیل شدن از دوره تربیت‌معلم به‌خودی‌خود، ضامن کیفیت تدریس معلم است. در سال‌های اخیر، این اعتماد تضعیف شده است. از گزارشات هماهنگ‌کنندگان تحقیقات ملی به‌خوبی مشخص بود که سیاست‌ها و اقدامات مربوط به اعطای گواهینامه در زمان اجرای طرح تربیت‌معلم و توسعه ریاضیات، تفاوت‌های گسترده‌ای دارد. پس از بررسی گزارشات کشورها مشخص شد که شرکت‌کنندگان طرح تربیت‌معلم و توسعه ریاضیات، به سه دسته تقسیم می‌شوند که در جدول ۱ نشان داده شده است. درعین‌حال، همچنان در میان کشورهای هر گروه، تفاوت‌هایی وجود دارد.

جدول (۱): پیش‌نیازهای حاکم بر ورود به عرصه تدریس در کشورهای مشارکت‌کننده در (TEDS-M)

کشور	مقررات تربیت‌معلم
بوتسوانا، شیلی، گرجستان، مالزی، نروژ، لهستان، فدراسیون روسیه، سنگاپور، اسپانیا (مدارس خصوصی)، سوئیس، تایلند	گروه ۱. کشورهایی که فارغ‌التحصیلی از دانشگاه به‌طور خودکار به ورود رسمی به عرصه تدریس می‌انجامد.
کانادا، عمان، فیلیپین، اسپانیا (مدارس دولتی)	گروه ۲. کشورهایی که در آن‌ها ورود به حرفه آموزش، منوط به قبولی در آزمون‌های دیگری است که نهادهای بیرونی اجرا می‌کنند (همچون آزمون‌های گواهی دانش حرفه‌ای).
چین و تایپه، آلمان، امریکا	گروه ۳. کشورهایی که در آن‌ها ورود به این حرفه و استخدام در آن، منوط به موفقیت در امتحانات دانش شغلی و ارزیابی عملکرد است.

در ایران، تا سال ۱۳۹۴، دانشجویان تربیت‌معلم و دانشگاه فرهنگیان هنگامی که دانش‌آموخته می‌شدند، در آموزش و پرورش مشغول به خدمت می‌شدند، تا این که معاونت نظارت، ارزیابی و تضمین کیفیت دانشگاه فرهنگیان، طرحی برای ارزشیابی صلاحیت‌های حرفه‌ای دانش‌آموختگان دانشگاه فرهنگیان که به‌طور مخفف «اصلح» نامیده شد، مصوب کرد. این آزمون، در شهریور ۹۵ برای دانش‌آموختگان ماده ۲۸ دانشگاه فرهنگیان اجرا شد. دانش‌آموختگان ماده ۲۸، داوطلبان شغل معلمی هستند که دارای مدرک حداقل کارشناسی از دانشگاه‌های معتبر هستند، در آزمون استخدام

۱. national research coordinators (NRCs)

۲. Teacher Education and Development Study in Mathematics (TEDS.M)

آموزش و پرورش پذیرفته شده‌اند و دو ترم در دانشگاه فرهنگیان، مهارت‌های اساسی معلمی را فرامی‌گیرند. شرط اشتغال این دانش‌آموختگان در آموزش و پرورش، موفقیت در آزمون اصلاح است. سیستم‌های ارزشیابی معلم در سراسر جهان، از معلم با شایستگی بالا به تدریس با کیفیت بالا و معلمان با تأثیرگذاری بالا تغییر یافته است. تلاش‌ها برای بهبود آموزش و عملکرد معلم در شغلش سیاست جدی است که به صورت تمرکز بر تعیین‌کننده‌های کلیدی پیشرفت دانش‌آموز در شرایط کنترل مدرسه، درخواست می‌شود (Hahnel and Jackson, ۲۰۱۲). با این حال، این سیاست‌ها به دلیل این که تدریس امری پیچیده، چندبُعدی، پویا و بسیار وابسته به خلق‌وخو و شرایط اجتماعی و فرهنگی می‌باشد، با چالش‌های مهمی مواجه هستند. بنابراین اندازه‌گیری عملکرد یا اثربخشی معلم بسیار مشکل‌تر از اندازه‌گیری صلاحیت‌های معلمی است (Darling-Youngs and Grissom, ۲۰۱۵; Shulman, ۱۹۹۸; Hammond, Amrein-Beardsley, Haertel and Rothstein, ۲۰۱۲). اگرچه ایده چندین شاخص، روایی صوری قابل توجهی دارد، پیاده‌سازی آن، سؤالات مهمی را درباره مفهوم‌سازی و ویژگی‌های فنی این شاخص‌ها و استفاده مناسب از آنها برای ارزیابی تکوینی و پایانی عملکرد معلم مطرح می‌کند (Brookhart, ۲۰۰۹; Goe and Croft, ۲۰۰۹).

بررسی عمیق عملکرد حرفه‌ای معلمان، به‌طور فزاینده‌ای به‌عنوان یک ابزار و سیاست برای بهبود آموزش و یادگیری با استفاده از تمایز بین معلمان از سطوح مختلف عملکرد استفاده می‌شود (Taylor and Taylor, ۲۰۱۲; Glazerman and et al., ۲۰۱۱). نظام‌های ارزیابی برای حمایت از دامنه‌ای از اهداف مختلف در حال توسعه هستند. برای مثال، شناسایی معلمان فعال در جهت کمک یا مداخله، ایجاد انگیزه برای معلمان با کارایی بالا، توسعه مدل‌هایی برای تدریس اثربخش برای اجرای گسترده‌تر، بهبود سیاست‌ها و شیوه‌های توسعه حرفه‌ای و مانند آن. یک چالش مفهومی جدی برای این تلاش‌ها، شامل تصمیم‌گیری درباره جنبه‌های عملکرد معلم برای ارزیابی است. روش‌های مشاهده کلاسی برای ارزیابی مهارت آموزشی و ارائه بازخورد به معلمان، به‌طور گسترده‌ای استفاده شده است (Pianta and Hamre, ۲۰۰۹). نظرسنجی از معلمان، والدین و مدیران برای ارزیابی خلق‌وخو، نگرش‌ها، اخلاق شهروندی و مشارکت در امور اجتماعی معلمان به‌کار می‌رود (Mayer, ۱۹۹۹). نظرسنجی از دانش‌آموزان، به‌طور فزاینده‌ای برای ارزیابی ادراک دانش‌آموزان از آموزش و مدیریت کلاس استفاده می‌شود که پایایی آن، به اندازه نظرسنجی معلمان است و قوی‌ترین پیش‌بینی‌کننده موفقیت دانش‌آموزان است (Kane and et al Martínez, ۲۰۱۲; Ferguson, ۲۰۱۱). آزمون‌های استاندارد (Hill and et al., ۲۰۰۸) و ارزیابی عملکرد باز^۱ (Dwyer, ۱۹۹۸) برای ارزیابی دانش موضوعی یا دانش درباره تعلیم و تربیت استفاده شده است. کارپوشه و ابزارهای سنجش مبتنی بر کاردستی، برای سنجش دانش تخصصی و تجربه استفاده می‌شود

۱. open.ended performanceassessments

(Youngs and Whittaker, ۲۰۱۵; Martínez, Borko and Stecher, ۲۰۱۲; Shulman, ۱۹۹۸). سرانجام مدل‌های ارزش افزوده^۱ (VAMs) برای ساختن شاخص‌های کلیدی میزانی از رشد یادگیری دانش‌آموزان که به معلمان نسبت داده می‌شود، استفاده می‌شود (Baker and et al., ۲۰۱۰). با استفاده از شاخص‌های چندگانه انتظار می‌رود یک تصویر کامل‌تری از عملکرد معلم ارائه شود (Goe and Croft, ۲۰۰۹) و مهم‌تر این‌که طبقه‌بندی‌های با ثبات‌تر برای عملکرد معلمان ارائه می‌دهد (De Pascale, ۲۰۱۲; Steele, Hamilton, and Stecher, ۲۰۱۰) و باعث بهبود بازخورد به معلمان و افزایش رشد حرفه‌ای آنان می‌شود (Duncan, ۲۰۱۲; Baker and et al., ۲۰۱۰). اگرچه این کاربردها به لحاظ مفهومی قانع‌کننده است، اما در عمل، توافق بسیار کمی در مورد این‌که دقیقاً سیستم باید این شاخص‌ها را در تنظیم سیاست واقعی برای نظارت و بهبود عملکرد معلمان استفاده کند وجود دارد (Mehrens, ۱۹۸۹).

مارتینز، اسچویگ و گلداسمیت (Martínez, Schweig and Goldschmidt, ۲۰۱۶)، به این نتیجه رسیدند که دقت پیش‌بینی عملکرد معلم در مدل‌های مختلف و نقاط برش مختلف متفاوت است و اینکه مدل‌هایی با دقت مشابه ممکن است به طبقه‌بندی‌های متفاوتی از معلمان بینجامد. دقت و ثبات، هنگامی بیشترین مقدار است که نمرات ترکیبی برای حداکثرکردن پایایی ساخته شوند و هنگامی که به دنبال پیش‌بینی بهینه نمرات دانش‌آموزان باشیم، دقت و ثبات، کمترین مقدار خواهد بود. نتایج پژوهش نیوتون (Newton, ۲۰۱۰) نشان داد که بین نمرات معلمان تازه‌کار در آزمون ارزشیابی عملکرد معلمان کالیفرنیا^۲ (PACT)، سنجش عملکرد قبل از آموزش در دامنه گسترده‌ای از مهارت‌های تدریس و اندازه‌ای از اثربخشی تدریس آن‌ها که با پیشرفت تحصیلی دانش‌آموزان آن‌ها در زبان انگلیسی به دست آمد، رابطه وجود دارد. مرتبط سازی نمرات ارزشیابی عملکرد معلمان کالیفرنیا (PACT) و چهار برآورد جداگانه از ارزش افزوده، نشان داد که همبستگی نمره کلی PACT و ارزش افزوده معلمان برابر ۰/۵ می‌باشد. نمرات خرده‌آزمون‌ها با ارزش افزوده بین ۰/۵۸ تا ۰/۶۶ همبستگی داشت. نتایج پژوهش قنبری‌نیا (Ghanbaryniya, ۲۰۱۱) درباره بررسی تطبیقی وضعیت برنامه ریزی درسی و صلاحیت‌های حرفه‌ای معلمان آموزش ابتدایی در کشورهای ایران و ژاپن، نشان داد که در بُعد صلاحیت‌های حرفه‌ای معلمان، متمرکز بودن نحوه‌گزینش و استخدام معلمان، انجام مصاحبه‌های عمومی و تخصصی و همچنین آزمون‌های کتبی و شفاهی در گزینش و استخدام معلمان از ویژگی‌های مشابه این دو کشور می‌باشد درحالی‌که آزاد بودن معلمان در انتخاب شغل معلمی، انتخاب معلمان از بین با صلاحیت‌ترین اشخاص جامعه از لحاظ علمی و اخلاقی و همچنین واگذار کردن برخی از اختیارات در جذب و استخدام معلمان به ایالات محلی در نظام آموزش و

۱. value.added models (VAMs)

۲. Performance Assessment for California Teachers (PACT)

پرورش ژاپن از تفاوت‌های این نظام آموزشی با نظام آموزشی ایران است. ضرورت کسب گواهینامه معلمی جهت تصدی شغل معلمی، موقعیت برجسته اجتماعی و اقتصادی معلمان در کشور ژاپن از تفاوت‌های دیگر این نظام آموزشی با نظام آموزش ایران است. کاکوکتوران (Kucukturana, ۲۰۱۱) آزمونی با ۱۷۱ سؤال و ۷ خرده‌مقیاس برای سنجش صلاحیت معلمان ابتدایی ساخت. خرده‌مقیاس‌ها عبارتند از: صلاحیت عمومی برای معلمی، توانایی هنر، ویژگی‌های شخصیتی، مهارت‌های زبانی-بیانی، توانایی ریاضی، اطلاعات عمومی و صلاحیت معلم در ایجاد یک محیط یادگیری با نشاط. نتایج پژوهش سروندی (Sarvandi, ۲۰۱۱) نشان داد که تأثیر مؤلفه‌های دانش، ارزش‌ها، مهارت‌ها و ویژگی‌های شخصیتی بر صلاحیت حرفه‌ای مدیران از دیدگاه معلمان و مدیران بالاتر از حد انتظار بود. در مورد تفاوت عوامل مؤثر بر صلاحیت حرفه‌ای مدیران نتایج نشان دهنده این است که بین این عوامل تفاوت معناداری وجود دارد. همچنین در مورد تفاوت دیدگاه مدیران و معلمان می‌توان چنین گفت که بین نظرات این دو گروه تفاوت معناداری وجود دارد و در تمامی مؤلفه‌ها، نظرات مدیران، بیشتر و بالاتر از نظرات معلمان است. نتایج پژوهش مهاجری، شریف و مهاجری (Mohajeri, Sharif and Mohajeri, ۲۰۱۱) نشان داد که رابطه معناداری بین قابلیت‌های حرفه‌ای معلمان در شش مؤلفه با عملکرد تحصیلی دانش‌آموزان وجود دارد. شش مؤلفه ۴۸ درصد از تغییرات عملکرد تحصیلی دانش‌آموزان را می‌توانند پیش‌بینی کنند. مقایسه بین قابلیت حرفه‌ای معلمان نشان داده است، معلمان با مدرک فوق‌لیسانس، سابقه خدمت زیر ۱۰ سال از قابلیت حرفه‌ای بیشتری برخوردارند. دانش‌آموزان معلمان فوق‌لیسانس از عملکرد تحصیلی بهتری برخوردارند. نتایج پژوهش ایزدی، صالحی‌عمران و منصوری‌بککی (Ezadi, salehi omran, and mansoori bekaki, ۲۰۱۱) نشان داد که معلمان از صلاحیت‌های حرفه‌ای مبتنی بر رویکرد سازنده‌گرایی برخوردارند، هرچند، نتایج در بخش کیفی (مصاحبه و مشاهده) نشان داد که اساس عملکرد معلمان بر رویکرد سازنده‌گرایی قرار ندارد. نتایج پژوهش عقیلی (Aghili, ۲۰۱۳) نشان داد که میزان تحقق صلاحیت‌های معلم در برنامه درسی تربیت معلم در کلیه مؤلفه‌ها در سطح نسبتاً مطلوب قرار دارد. نتایج پژوهش خرم (Khorram, ۲۰۱۵) نشان داد که صلاحیت‌های آموزشی، رفتاری، فکری، مدیریتی و صلاحیت‌های طراحی محیط فیزیکی معلمان فارغ‌التحصیل از مراکز تربیت معلم بیشتر از معلمان ورودی از طریق آزمون بود، ولی این دو گروه در صلاحیت‌های اجتماعی، حقوقی، اخلاق حرفه‌ای، توسعه حرفه‌ای، فناوری و صلاحیت طراحی محیط عاطفی تفاوتی با یکدیگر نداشتند. نیجولت و همکاران (Nijveldt and et al., ۲۰۰۵) در پژوهشی راهکاری برای سنجش شایستگی‌های بین فردی معلمان ایجاد و روایی آن بررسی کردند. این راهکار شامل استفاده از پرسشنامه تعاملات معلم و ابزار مشاهده و خود ارزیابی بود. روایی این راهکار با روش‌های مختلف بررسی شد و رضایت بخش بود و استفاده از ابزارهای مختلف به‌طور آشکاری اندازه‌گیری را بهبود بخشیده بود.

با توجه به این که سنجش صلاحیت معلمان در کشورهای پیشرفته دنیا مرسوم است و در ایران هم برای اولین بار در شهریور ۱۳۹۵ برای مهارت‌آموزان ماده ۲۸ اجرا شد. نتایج این آزمون هم برای نظام آموزشی و هم برای مهارت‌آموزان دانشگاه فرهنگیان بسیار مهم و حیاتی است. برای مهارت‌آموزان از این جهت مهم است که سرنوشت شغلی آنها را مشخص می‌کند، یعنی آیا بعد از گذراندن آزمون استخدامی و طی کردن دوره مهارت‌آموزی، سطحی از مهارت را دارا می‌باشند که بتوانند به تدریس موفق در آموزش و پرورش بپردازند یا باید دوباره به جستجوی خود برای شغلی جدید ادامه دهند. اهمیت این آزمون برای نظام آموزشی کشور از این جهت است که به‌عنوان یک صافی برای عبور معلمان با صلاحیت عمل می‌کند. اهمیت نظام آموزشی در جهت تربیت نسل آینده بر هیچ کس پوشیده نیست و مهم‌ترین عامل در این نظام معلمی است که می‌خواهد دانش‌آموز را تربیت کند لذا آزمون‌هایی که می‌خواهد صلاحیت ورودی‌های آموزش و پرورش را تعیین کند بسیار مهم و سرنوشت‌ساز است. بررسی عوامل موفقیت یا شکست مهارت‌آموزان در این آزمون می‌تواند راهگشای تصمیم‌گیران در جهات مختلف باشد. مثلاً از جهت انتخاب و گزینش مهارت‌آموزان می‌تواند کمک‌کننده باشد؛ اگر مهارت‌آموزان زن در کل کشور موفقیت، صلاحیت و مهارت بیشتری داشته‌باشند می‌توان تا حد ممکن از زنان برای تدریس استفاده کرد. اگر مهارت‌آموزان کل کشور در یک رشته موفقیت بیشتر یا کمتری دارند، می‌توان با بررسی علل توفیق یا عدم توفیق، در جهت موفقیت بیشتر و پیشگیری از عدم توفیق در سال‌های آینده برنامه‌ریزی کرد. یا اگر مهارت‌آموزان یک استان موفقیت بیشتری دارند و این در سال‌های بعدی هم تکرار می‌شود، می‌توان رشته‌های کشوری را بیشتر به این استان‌ها سوق داد و از پتانسیل این استان‌ها بیشتر استفاده کرد. لذا سؤالاتی که در این پژوهش، مطرح است، عبارتند از: آیا جنسیت، موفقیت در آزمون اصلح را به‌طور معناداری پیش‌بینی می‌کند؟ آیا رشته تخصصی، موفقیت در آزمون اصلح را به‌طور معناداری پیش‌بینی می‌کند؟ آیا استان محل مهارت‌آموزی، موفقیت در آزمون اصلح را به‌طور معناداری پیش‌بینی می‌کند؟

روش شناسی پژوهش

پژوهش حاضر از نظر هدف کاربردی و از حیث روش، از نوع همبستگی است. جامعه پژوهشی کلیه مهارت‌آموزان ماده ۲۸ سال ۹۴-۹۵ دانشگاه فرهنگیان کشور می‌باشند که تعداد آنها ۲۸۲۳ نفر می‌باشند. حجم نمونه با استفاده از جدول مورگان ۳۴۱ نفر تعیین شد. برای انتخاب نمونه از روش خوشه‌ای استفاده شد. به این صورت که از بین استان‌های مختلف کشور از هر منطقه یک استان انتخاب کردیم از شرق، استان‌های خراسان و سیستان و بلوچستان از مرکز استان‌های تهران و اصفهان و از غرب کشور استان خوزستان انتخاب شدند.

ابزار پژوهش: داده‌های این پژوهش نتایج حاصل از آزمون اصلح دانشگاه فرهنگیان در شهریور ۱۳۹۵ می‌باشد. این آزمون در پایان دوره مهارت‌آموزی دانشجویان ماده ۲۸ دانشگاه فرهنگیان جهت صدور گواهی صلاحیت تدریس اجرا می‌شود و برای اولین بار در شهریور ۱۳۹۵ اجرا شد. شامل چهار بخش بود. الف) آزمون تدریس خرد برای ارزیابی مهارت عملی معلمان استفاده شد. ب) آزمون علمی کتبی براساس نمره‌ای بود که آزمودنی‌ها در آزمون کتبی هماهنگ کشوری کسب کردند. ج) آزمون عملکردی کتبی براساس نمره‌ای بود که آزمودنی‌ها در آزمون کتبی هماهنگ کشوری کسب کردند. د) نمره پوشه کار فرهنگی براساس نظر داوران درباره گزارش‌هایی بود که دانشجویان درباره فعالیت‌های فرهنگی در طول دوره مهارت آموزی ارائه دادند.

یافته‌های پژوهش

عضویت در یک طبقه معین از متغیر را می‌توان با استفاده از متغیر تصنعی (یا ظاهری) نشان داد. متغیر تصنعی برداری است که در آن به اعضای طبقه معین یک عدد دلخواه و به بقیه یعنی آزمودنی‌های که عضو این طبقه معین نیستند عدد دلخواه دیگری نسبت داده می‌شود. برای مثال در مورد متغیر جنسیت می‌توان به مردان ۱ و به زنان صفر نسبت داد. استفاده از متغیرهای تصنعی در تحلیل داده‌های تحقیقاتی که در آنها متغیرهای مستقل طبقه‌ای اند می‌تواند بسیار سودمند باشد. با این وجود متغیرهای تصنعی دارای مقاصد سودمند دیگری هم هستند. مثلاً می‌توان آنها را در کنار متغیرهای مستقل پیوسته به کار برد. اصول و روش‌های تحلیل رگرسیون چند متغیره به همان صورت که در مورد متغیرهای پیوسته به کار می‌روند، در مورد متغیرهای طبقه‌ای هم می‌توانند به کار روند. وقتی سروکار محقق با یک متغیر مستقل طبقه‌ای است، آن را می‌تواند به صورت مقتضی با متغیرهای تصنعی بیان کند، سپس با متغیرهای تصنعی به عنوان متغیرهای مستقل رفتار کند و تحلیل رگرسیون را به کار ببرد. یکی از مزایای تحلیل رگرسیون با متغیرهای کدگذاری این است که وقتی فراوانی‌ها در خانه‌های طرح فاکتوریال نابرابرند یا نامتناسب‌اند، می‌توان از تحلیل رگرسیون با کدگذاری استفاده کرد. بنابراین وقتی n ها نابرابرند تحلیل رگرسیون با کدگذاری درست به همان صورتی که با n های برابر انجام می‌شود، پیش می‌رود (Saraei, ۲۰۱۷).

جهت کدگذاری تصنعی متغیری که سه طبقه دارد، باید دو بردار ایجاد کنیم در بردار اول یک گروه کد ۱ و دیگران کد صفر می‌گیرند. در بردار دوم گروه دوم کد ۱ و دیگران کد صفر می‌گیرند. یک بردار دیگر هم می‌توان ایجاد کرد. در این بردار، به آزمودنی‌هایی که عضو گروه سوم هستند ۱ و به دیگران صفر بدهیم. با این وجود ایجاد چنین برداری زاید است زیرا دو بردار نخست تمامی اطلاعات مربوط به عضویت گروهی را شامل می‌شود. به عبارت دیگر آگاهی از وضعیت یک آزمودنی در دو بردار نخست، یعنی آگاهی از عضویت یا عدم عضویت او در دو گروه اول و دوم برای شناخت

عضویت گروهی او کافی است. اگر آزمودنی عضو یکی از گروه‌های اول و دوم نباشد باید عضو گروه سوم باشد. همیشه تعداد بردارهای لازم و کافی برای کدگذاری عضویت گروهی معادل تعداد گروه‌ها یا طبقات منهای یک است.

سؤال اول پژوهش: آیا جنسیت، موفقیت در آزمون اصلح را به‌طور معناداری پیش‌بینی می‌کند؟ برای بیان متغیر جنسیت یک بردار لازم و کافی است چون دو طبقه دارد. بنابر این یک بردار بنام X_1 ایجاد شد. رگرسیون را اجرا کردیم به‌طوری که X_1 به‌عنوان متغیر مستقل و نمره کلی آزمون اصلح به‌عنوان متغیر وابسته وارد معادله شد. R برابر $0/225$ و R^2 برابر $0/051$ می‌باشد که همان‌طور که جدول زیر نشان می‌دهد به لحاظ آماری معنادار است.

جدول (۲): نتایج تحلیل واریانس رگرسیون سؤال اول

منبع واریانس	مجموع مجزورات	درجه آزادی	میانگین مجزورات	F	سطح معناداری
رگرسیون	۵۱۸۵۹/۵۸۶	۱	۵۱۸۵۹/۵۸۶		
باقی‌مانده	۹۶۸۹۹۵/۳۶۶	۳۳۹	۲۸۵۸/۳۹۳	۱۸/۱۴۳	$P < 0,01$
جمع	۱۰۲۰۸۵۴/۹۵۲	۳۴۰			

در جدول فوق مشاهده می‌نمایید، F محاسبه شده برابر $18/143$ و سطح معناداری از $0/01$ کمتر است لذا این ضریب رگرسیون به لحاظ آماری معنادار است. لذا نتیجه می‌گیریم که جنسیت، موفقیت مهارت‌آموزان در آزمون اصلح را پیش‌بینی می‌کند.

همین محاسبات را برای نمرات خرده‌مقیاس‌های مختلف انجام دادیم به این صورت متغیر کد گذاری X_1 به‌عنوان متغیر مستقل و نمرات خرده‌مقیاس‌ها هر بار یکی به‌عنوان متغیر وابسته تعریف شدند. برای خرده آزمون علمی کتبی چهار گزینه‌ای R برابر $0/112$ بود که به لحاظ آماری معنادار نبود. برای خرده‌آزمون فرهنگی R برابر $0/181$ بود و به لحاظ آماری معنادار می‌باشد. برای خرده‌آزمون تدریس خرد R برابر $0/035$ بود و به لحاظ آماری معنادار نمی‌باشد. برای خرده‌آزمون عملکردی R برابر $0/212$ بود و به لحاظ آماری معنادار می‌باشد.

سؤال دوم پژوهش: آیا رشته تخصصی، موفقیت در آزمون اصلح را به‌طور معناداری پیش‌بینی می‌کند؟

برای بررسی نقش رشته تخصصی، رشته‌ها را به سه گروه، علوم ریاضی، علوم تجربی و علوم انسانی طبقه‌بندی کردیم. برای بیان متغیر رشته، ۲ بردار لازم و کافی است چون ۳ طبقه دارد. بنابراین ۲ بردار بنام X_2 و X_3 ایجاد شد.

رگرسیون را اجرا می‌کنیم به طوری که X_2 و X_3 به عنوان متغیرهای مستقل و نمره کلی آزمون اصلح به عنوان متغیر وابسته بود. R برابر ۰/۳۹۹ و R^2 برابر ۰/۱۶۰ به دست آمد، همان طور که جدول زیر نشان می‌دهد به لحاظ آماری معنادار است.

جدول (۳): نتایج تحلیل واریانس رگرسیون سؤال دوم

منبع واریانس	مجموع مجذورات	درجه آزادی	میانگین مجذورات	F	سطح معناداری
رگرسیون	۳۵۷۶۳/۵۸۵	۲	۱۷۸۸۱/۷۹۲	۵/۹۵۴	$P < ۰,۰۱$
باقی مانده	۹۸۵۰۹۱/۳۶۷	۳۳۸	۳۰۰۳/۳۲۷		
جمع	۱۰۲۰۸۵۴/۹۵۲	۳۴۰			

در جدول فوق مشاهده می‌نمایید، F محاسبه شده برابر ۵/۹۵۴ و سطح معناداری از ۰/۰۱ کمتر است لذا این ضریب رگرسیون به لحاظ آماری معنادار است. و نتیجه می‌گیریم که رشته تخصصی موفقیت مهارت‌آموزان در آزمون اصلح را پیش‌بینی می‌کند.

همین محاسبات را برای نمرات خرده‌مقیاس‌های مختلف انجام دادیم به این صورت متغیرهای کد گذاری X_2 و X_3 به عنوان متغیر مستقل و نمرات خرده‌مقیاس‌ها هر بار یکی به عنوان متغیر وابسته تعریف شدند. برای خرده‌آزمون علمی کتبی چهار گزینه‌ای R برابر ۰/۴۰۰ بود که به لحاظ آماری معنادار می‌باشد. برای خرده‌آزمون فرهنگی R برابر ۰/۱۲۷ بود که به لحاظ آماری معنادار نمی‌باشد. برای خرده‌آزمون تدریس خرد R برابر ۰/۲۵۸ بود که به لحاظ آماری معنادار می‌باشد. برای خرده‌آزمون عملکردی R برابر ۰/۵۴۳ بود که به لحاظ آماری معنادار می‌باشد. سؤال سوم پژوهش: آیا استان محل مهارت‌آموزی، موفقیت در آزمون اصلح را به طور معناداری پیش‌بینی می‌کند؟

برای بیان متغیر استان محل مهارت‌آموزی ۴ بردار لازم و کافی است چون ۵ طبقه دارد. بنابراین ۴ بردار بنام X_4 تا X_7 ایجاد شد.

رگرسیون را اجرا کردیم به طوری که X_4 تا X_7 به عنوان متغیرهای مستقل و نمره کلی آزمون اصلح به عنوان متغیر وابسته وارد معادله شدند. R برابر ۰/۳۳۸ و R^2 برابر ۰/۱۱۴ به دست آمد و همان طور که جدول زیر نشان می‌دهد به لحاظ آماری معنادار است.

جدول (۴): نتایج تحلیل واریانس رگرسیون سؤال سوم

منبع	مجموع	درجه	میانگین	F	سطح
واریانس	مجذورات	آزادی	مجذورات		معنی‌داری
رگرسیون	۱۱۶۷۱۹/۸۹۰	۴	۲۹۱۷۹/۷۹۳	۱۰/۸۴۴	$P < ۰,۰۱$
باقی‌مانده	۹۰۴۱۳۵/۰۶۱	۳۳۶	۲۶۹۰/۷۸۷		
جمع	۱۰۲۰۸۵۴/۹۵۲	۳۴۰			

در جدول فوق مشاهده می‌نمایید، F محاسبه شده برابر ۱۰/۸۴۴ و سطح معناداری از ۰/۰۱ کمتر است لذا این ضریب رگرسیون به لحاظ آماری معنادار است. و نتیجه می‌گیریم که استان محل مهارت‌آموزی موفقیت مهارت‌آموزان در آزمون اصلح را پیش‌بینی می‌کند.

در اینجا نیز همین محاسبات را برای نمرات خرده مقیاس‌های مختلف انجام دادیم به این صورت که متغیرهای کد گذاری $X_۴$ تا $X_۷$ به‌عنوان متغیر مستقل و نمرات خرده‌مقیاس‌ها هر بار یکی به‌عنوان متغیر وابسته تعریف شدند. برای خرده‌آزمون علمی کتبی چهار گزینه‌ای R برابر ۰/۳۰۹ بود که به لحاظ آماری معنادار می‌باشد. برای خرده‌آزمون فرهنگی R برابر ۰/۱۵۸ بود که به لحاظ آماری معنادار می‌باشد. برای خرده‌آزمون تدریس خرد R برابر ۰/۱۸۵ بود که به لحاظ آماری معنادار نمی‌باشد. برای خرده‌آزمون عملکردی R برابر ۰/۲۴۱ بود که به لحاظ آماری معنادار می‌باشد. وقتی استان محل مهارت‌آموزی می‌تواند نمرات آزمون اصلح را به‌طور معناداری پیش‌بینی کند به این معنی است که بخش معناداری از واریانس نمرات آزمون اصلح به متغیر استان محل مهارت‌آموزی نسبت داده می‌شود و مربوط به تفاوت در امکانات و سبک‌های نمره‌دهی در استان‌های مختلف می‌باشد و می‌تواند برای مدیران دانشگاه قابل تأمل و بررسی باشد.

بحث و نتیجه‌گیری

سنجش صلاحیت‌های حرفه‌ای معلمان قبل از ورود به حرفه معلمی یک موضوع کاملاً جدید در ایران می‌باشد و برای اولین بار در شهریور ۱۳۹۵ برای مهارت‌آموزن یکساله دانشگاه فرهنگیان اجرا شد. با توجه به جدید بودن و اهمیت نتایج آن برای سرنوشت مهارت‌آموزان و نظام آموزشی ایران، پژوهش در این زمینه ضروری به نظر می‌رسد. بررسی پیش‌بینی‌کننده‌های نمرات مهارت‌آموزان در

این آزمون می‌تواند راهنمای مجریان برای آینده باشد مثلاً اگر استان محل مهارت‌آموزی در پیش‌بینی نمرات نقش معناداری داشته باشد به این معنی است که سهم معناداری از واریانس نمرات آزمون به متغیر استان نسبت داده می‌شود.

ابتدا نقش جنسیت در پیش‌بینی موفقیت در آزمون اصلح بررسی شد. با توجه به این که R برابر ۰/۱۸۶ و به لحاظ آماری معنادار است. می‌توانیم بگوییم که جنسیت، موفقیت مهارت‌آموزان در آزمون اصلح را پیش‌بینی می‌کند. بررسی خرده‌مقیاس‌ها نشان داد که جنسیت موفقیت در خرده‌آزمون‌های فرهنگی و عملکردی را به‌طور معناداری پیش‌بینی می‌کند و اما در پیش‌بینی نمرات خرده‌مقیاس‌های علمی کتبی چهار گزینه‌ای و تدریس خرد نقش معناداری ندارد. بررسی تفاوت زنان و مردان در نمره کل آزمون نشان می‌دهد که میانگین نمره زنان از مردان بیشتر می‌باشد. در تمامی خرده‌مقیاس‌ها نیز میانگین نمرات زنان از مردان بیشتر است. اما در دو خرده‌مقیاس علمی چهار گزینه‌ای و تدریس خرد این تفاوت‌ها به لحاظ آماری معنادار نیست و در خرده‌مقیاس‌های فرهنگی و عملکردی تفاوت معنادار است. از آن‌جا که نمرات زنان در کل آزمون و دو خرده‌آزمون از مردان بهتر است، می‌توان گفت احتمالاً صلاحیت، مهارت و انگیزه زنان در معلمی و فعالیت‌های تدریس از مردان بیشتر است. البته برای تصمیم‌گیری به بررسی بیشتر نیاز است لذا به پژوهشگران توصیه می‌شود که پیش‌بینی‌پذیری نتایج آزمون براساس متغیر جنسیت را با داده‌های آزمون در سال‌های آینده مورد بررسی قرار دهند.

نتایج بررسی پیش‌بینی‌پذیری موفقیت در آزمون اصلح، براساس رشته تخصصی مهارت‌آموزان، نشان داد که R برابر ۰/۳۹۸ و به لحاظ آماری معنادار است لذا می‌توانی نتیجه بگیریم که رشته تخصصی موفقیت مهارت‌آموزان در آزمون اصلح را به‌طور معناداری، پیش‌بینی می‌کند. بررسی خرده‌مقیاس‌ها نشان داد که رشته تخصصی موفقیت در خرده‌آزمون‌های علمی کتبی چهار گزینه‌ای، عملکردی و تدریس خرد را به‌طور معناداری پیش‌بینی می‌کند و اما در پیش‌بینی نمرات خرده‌مقیاس فرهنگی نقش معناداری ندارد. در خرده‌آزمون علمی چند گزینه‌ای مهارت‌آموزان گروه علوم تجربی بیشترین میانگین و بعد گروه علوم ریاضی و در نهایت گروه علوم انسانی پایین‌ترین میانگین نمرات را داشتند. در خرده‌آزمون فرهنگی، مهارت‌آموزان گروه علوم ریاضی بالاترین میانگین و گروه علوم تجربی و انسانی تقریباً هم سطح بودند. در خرده‌آزمون تدریس خرد، مهارت‌آموزان گروه علوم تجربی بالاترین میانگین و بعد از آن علوم انسانی و در نهایت مهارت‌آموزان گروه ریاضی کمترین میانگین را داشتند و در خرده‌آزمون عملکردی، مهارت‌آموزان گروه ریاضی بیشترین میانگین و بعد گروه علوم تجربی و در نهایت گروه علوم انسانی پایین‌ترین میانگین نمرات را داشتند. گروه علوم تجری در مجموع بهترین وضعیت و گروه ریاضی در رتبه دوم و گروه علوم انسانی در آخرین جایگاه قرار دارند به تصمیم‌گیران و پژوهشگران توصیه می‌شود با داده‌های آزمون

در سال‌های آینده نمرات مهارت‌آموزان رشته‌های مختلف را با هم مقایسه نمایند و درباره علل بالاتر یا پایین‌تر بودن نمرات گروه‌ها بررسی‌های دقیق انجام دهند و از نتایج برای برنامه‌ریزی دوره مهارت‌آموزی در سال‌های آینده استفاده نمایند.

در نهایت، معناداری پیش‌بینی موفقیت در آزمون اصلاح براساس استان محل مهارت‌آموزی، بررسی شد. نتایج نشان داد که R برابر $0/306$ و به لحاظ آماری معنادار است لذا نتیجه می‌گیریم که استان محل مهارت‌آموزی موفقیت مهارت‌آموزان در آزمون اصلاح را پیش‌بینی می‌کند. بررسی خرده‌مقیاس‌ها نشان داد که استان محل مهارت‌آموزی، موفقیت در خرده‌آزمون‌های علمی کتبی چهار گزینه‌ای، عملکردی و فرهنگی را به‌طور معناداری پیش‌بینی می‌کند اما در پیش‌بینی نمرات خرده‌مقیاس تدریس خرد نقش معناداری ندارد.

بررسی میانگین‌های استان‌ها نشان می‌دهد که در نمره کل آزمون و خرده‌آزمون کتبی چند گزینه‌ای، استان اصفهان بالاترین میانگین و بعد از آن به ترتیب خراسان رضوی، تهران، خوزستان و سیستان قرار دارد. در خرده‌مقیاس فرهنگی و عملکردی، استان خراسان رضوی بالاترین میانگین نمرات و بعد از آن به ترتیب تهران، اصفهان، خوزستان و سیستان قرار دارد. در خرده‌مقیاس تدریس خرد، استان اصفهان بالاترین میانگین نمرات و بعد از آن به ترتیب سیستان، خراسان رضوی و خوزستان هم سطح و تهران پایین‌ترین میانگین را دارد. این وضعیت برای مجریان آزمون پیام مهمی دارد. ترتیب استان‌ها در همه خرده‌آزمون‌ها به جز تدریس خرد، تقریباً متناسب با میزان برخورداری استان از امکانات می‌باشد و این نشان‌دهنده اجرای درست آزمون می‌باشد. استان سیستان که در سایر خرده‌آزمون‌ها پایین‌ترین نمره را داشته است در خرده‌آزمون تدریس خرد که در استان‌ها برگزار شده است و نتایج آن به سازمان مرکزی در تهران، ارسال شده، نتیجه خوبی گرفته‌اند. به مجریان آزمون پیشنهاد می‌شود برای جلوگیری از خطاهای احتمالی دستور العمل‌ها و روش‌های نمره‌گذاری خرده‌آزمون تدریس خرد که در استان‌ها برگزار می‌شود، عینی و دقیق باشد و بیشتر بر آزمون‌های هماهنگ کشوری تأکید شود.

References:

- Aghili, A. (۲۰۱۳). Teacher Competencies in Teacher Training Curriculum: Identify the degree of realization and the strengths and weaknesses From the perspective of professors, students and graduates. The master's thesis of curriculum planning at Ferdowsi University of Mashhad has not been published by Amin Khandaghi's Meaning and Consultation, Saeedi Rezvani, Mahmoud, [In Persian].
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (۱۹۹۹). Standards for educational and psychological testing. Washington, DC: American Educational Research Association.
- Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L. and Shepard, L. A. (۲۰۱۰). Problems with the Use of Student Test Scores to Evaluate Teachers. EPI Briefing Paper ۲۷۸. Economic Policy Institute.
- Brookhart, S. M. (۲۰۰۹). The many meanings of multiple measures. *Education Leadership*, ۶۷(۳): ۶-۱۲.
- Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E. and Rothstein, J. (۲۰۱۲). Evaluating teacher evaluation. *Phi Delta Kappan*, ۹۳(۶): ۸-۱۵.
- De Pascale, C. (۲۰۱۲). Managing multiple measures. *Principal*, ۹۱(۵): ۶-۱۰.
- Duncan, A. (۲۰۱۲, August ۲۲). Duncan tells teachers: Change is hard. *Homeroom*. Retrieved from <http://www.ed.gov/blog/۲۰۱۲/۰۸/duncan-tells-teacherschange-is-hard/>
- Dwyer, C. A. (۱۹۹۸). Psychometrics of Praxis III: Classroom performance assessments. *Journal of Personnel Evaluation in Education*, ۱۲: ۱۶۳-۱۸۷.
- Ezadi, S., Salehi Omran, E. and Mansoori Bekaki, S. (۲۰۱۱). Examination of professional qualifications of high school male teachers of the branch of human sciences based on constructivist approach (case study: mazandaran prvince). *Journal of Educational and Learning Studies*, ۴ (۱): ۱-۲۷, [In Persian].
- Ferguson, R. (۲۰۱۱). Student perceptions of quality feedback in teacher education. *Assessment & Evaluation in Higher Education*, ۳۶: ۵۱-۶۲.
- Ghanbaryniya, M. (۲۰۱۱). A Comparative Study of the Current Situation of Curriculum Planning and Professional Competencies of Elementary Education Teachers in Iran and Japan in order to provide an Appropriate Pattern for Use in Iran. Master thesis, Allameh Tabatabai University of Tehran, Supervisor: Soghra Ebrahimi Ghavam; Consultant Professor Ahmad Aghazadeh, not published, [In Persian].
- Glazerman, S., Goldhaber, D., Loeb, S., Raudenbush, S., Staiger, D. O. and Whitehurst, G. J. (۲۰۱۱). *Passing muster: Evaluating evaluation systems*. Washington, DC: Brown Center on Education Policy at Brookings.
- Goe, L. and Croft, A. (۲۰۰۹). *Methods of evaluating teacher effectiveness*. Washington, DC: National Comprehensive Center for Teacher Quality.
- Hahnel, C. and Jackson, O. (۲۰۱۲). Learning denied: The case for equitable access to effective teaching in California's largest school district. Oakland, CA: The Education Trust-West.

- Hill, H. C., Blunk, M., Charalambous, C., Lewis, J., Phelps, G. C., Sleep, L. and Ball, D. L. (۲۰۰۸). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction*, ۲۶: ۴۳۰-۵۱۱.
- Ingvarson, L. J. Schwille, M. T. T., Glenn R., Ray P. And Sharon L. S. (۲۰۱۳), An Analysis of Teacher Education Context, Structure, and Quality Assurance Arrangements in TEDS-M Countries, Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M), Australian Council for Educational Research Michigan State University, Amsterdam, the Netherlands, www.iea.nl
- Kane, T. J., Staiger, D. O., McCaffrey, D., Cantrell, S., Archer, J., Buhayar, S. and Parker, D. (۲۰۱۲). *Gathering feedback for teaching: Combining high quality observations with student surveys and achievement gains*. Seattle, WA: Bill & Melinda Gates Foundation.
- Khorram, N. (۲۰۱۵). Comparison of vocational qualifications of elementary school teachers, graduated from teacher training centers and recruitment through a test. Master's thesis, Allameh Tabataba'i University of Tehran, Supervisor: fariborz dortaj; Advisor: Farkhondeh Mofidi, [In Persian].
- Kucukturana, G. (۲۰۱۱). Qualification scale for early childhood teachers, *Procedia - Social and Behavioral Sciences* ۲۸ (۲۰۱۱) ۵۷۹ – ۵۹۰, Available online at www.sciencedirect.com.
- Martínez, J. F. (۲۰۱۲). Consequences of omitting the classroom in multilevel models of schooling: An illustration using opportunity to learn and reading achievement. *School Effectiveness and School Improvement*, ۲۳: ۳۰۵-۳۲۶.
- Martínez, J. F., Borko, H. and Stecher, B. (۲۰۱۲). Measuring instructional practices in middle school science using classroom artifacts. *Journal for Research in Science Teaching*, ۴۹: ۳۸-۶۷.
- Martínez, José Felipe., Schweig, Jonathan., Goldschmidt, P. (۲۰۱۶). Approaches for Combining Multiple Measures of Teacher Performance: Reliability, Validity, and Implications for Evaluation Policy Educational Evaluation and Policy Analysis December, ۳۸(۴):۷۳۸-۷۵۶
- Mayer, D. (۱۹۹۹). Measuring instructional practice: Can policy makers trust survey data? *Educational Evaluation and Policy Analysis*, ۲۱(۱): ۲۹-۴۵.
- Mehr Mohammadi, M. (۲۰۱۵). Evaluation of University Quality through the Output Quality Assessment (Comprehensive Test). Proceedings of the ۲nd National Conference and ۹th Conference on the Evaluation of Academic Systems Quality. Tehran, Farhangian University, Pages ۲-۴. [In Persian].
- Mehrens, W. (۱۹۸۹). Combining evaluation data from multiple sources. In J. Millman & L. DarlingHammond (Eds.), *The new handbook of teacher evaluation: Assessment of elementary and secondary school teachers* (pp. ۳۲۲-۳۳۶). Newbury Park, CA: SAGE.
- Mohajeri, M., Sharif, M. and Mohajeri, A. (۲۰۱۱). The Relationship Between Professional Capabilities of Primary Teachers and Students' Academic Performance in Isfahan. *New Educational Approaches*, University of Isfahan, Faculty of Education and Psychology. ۷ (۱): ۲۷-۴۸, [In Persian].

- National Center for Teacher Quality. (۲۰۱۵). State of the states: Evaluating teaching, leading, and learning. Washington, DC: Author.
- National Research Council. (۲۰۱۰). Getting value out of value-added (H. Braun, N. Chudowsky, & J. Koenig, Eds.). Washington, DC: National Academies Press.
- Newton, Stephen (۲۰۱۰) Preservice Performance Assessment and Teacher Early Career Effectiveness: Preliminary Findings on the Performance Assessment for California Teachers, Stanford University
- Nijveldt, Mirjam. Beijaard, Douwe. Brekelmans, Mieke. Verloop, Nico. Wubbels, Theo. (۲۰۰۵) Assessing the interpersonal competence of beginning teachers: The quality of the judgement process, *International Journal of Educational Research*, ۴۳: ۸۹-۱۰۲.
- Peterson, K. (۱۹۸۷). Teacher evaluation with multiple and variable lines of evidence. *American Education Research Journal*, ۲۴: ۳۱۱-۳۱۷.
- Pianta, R. C. and Hamre, B. K. (۲۰۰۹). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, ۳۸: ۱۰۹-۱۱۹.
- Reynolds, M. (۱۹۹۹). Standards and professional practice: The TTA and initial teaching training. *British Journal of Educational Studies*, ۴۷: ۲۴۷-۲۶۰.
- Saraei, H. (۲۰۱۷). Multiple Regression in Behavioral Research. Translation, original narrator Elazar J. Pedhazur, Fred Nichols Kerlinger, Ninth Edition, Samt Publication, [In Persian].
- Sarvandi, A. (۲۰۱۱). Assessing the Professional Competency of Managers from the Point of View of Teachers and Managers of Primary Schools in Pakdasht County, Thesis for Master Degree, orientation: Educational Management at Islamic Azad University, Central Tehran Branch, Supervisor: Dr Roxana Nasiri, Consultant Professor: Dr. Lotf Allah Abbasi, has not been published, [In Persian].
- Shulman, L. S. (۱۹۸۷). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, ۵۷(۱): ۱-۲۲.
- Shulman, L. S. (۱۹۹۸). Teacher portfolios: A theoretical activity. In N. Lyons (Ed.), *With portfolio in hand* (pp. ۲۳-۳۷). New York, NY: Teachers College Press.
- Steele, J., Hamilton, L. S. and Stecher, B. M. (۲۰۱۰). Incorporating student performance measures into teacher evaluation systems. Santa Monica, CA: The RAND Corporation.
- Taylor, E. S. and Tyler, J. (۲۰۱۲). The effect of evaluation on teacher performance. *American Economic Review*, ۱۰۲: ۳۶۲۸-۳۶۵۱.
- Youngs, P. and Grissom, J. (۲۰۱۵). Multiple measures in teacher evaluation: Lessons learned and guidelines for practice. In P. Youngs & J. Grissom (Eds.), *Improving teacher evaluation systems* (pp. ۱۶۹-۱۸۴). New York, NY: Teachers College Press.
- Youngs, P. and Whittaker, A. (۲۰۱۵). The role of EdTPA in assessing content specific instructional practices. In P. Youngs & J. Grissom (Eds.), *Improving teacher evaluation systems* (pp. ۸۹-۱۰۱). New York, NY: Teachers College Press.

