

The Comparison of Descriptive and Traditional Evaluation Methods Based on the Quartet Evaluation Standards from the Viewpoint of Elementary School Teachers in Tehran City

مقایسه روش‌های ارزشیابی توصیفی و سنتی بر اساس استانداردهای چهارگانه ارزشیابی از دیدگاه معلمان مدارس ابتدایی شهر تهران

Rasool Davoudi., Mahboobeh Shokrollahi

¹Department of Educational Sciences and Psychology, Zanjan Branch, Islamic Azad University, Zanjan, Iran

²Researcher and M.A in Curriculum Planning

رسول داودی، محبوبه شکراللهی

^۱گروه علوم تربیتی و روانشناسی، دانشگاه آزاد اسلامی، زنجان، ایران
^۲پژوهشگر و کارشناسی ارشد برنامه‌ریزی درسی

Abstract

The purpose of this study is to compare descriptive and traditional evaluation methods based on quartet evaluation standards (utility, propriety feasibility, and accuracy) from the viewpoint of elementary school teachers in Tehran city. In order to achieve this aim, and besides the review of literature, opinions of all teachers (40 persons) in executive schools of descriptive evaluation project were gathered through researcher-made questionnaire. The results of data analysis show that descriptive evaluation adjusts better to evaluation standards. In addition, among the mentioned standards, the efficiency standard represents more adjustments than feasibility standards with descriptive evaluation.

Keywords: Educational Evaluation, Descriptive Evaluation, Traditional Evaluation, Evaluation Standards

چکیده

هدف این پژوهش مقایسه روش‌های ارزشیابی توصیفی و سنتی بر اساس استانداردهای چهارگانه ارزشیابی (اخلاقی، کارآوری، قابلیت اجرا و دقت) از دیدگاه معلمان مدارس ابتدایی شهر تهران است. برای نیل به این هدف، علاوه بر مطالعات نظری، نظرات کلیه معلمان (۴۰ نفر) در مدارس مجری طرح ارزشیابی توصیفی از طریق پرسشنامه محقق ساخته جمع‌آوری شده است. داده‌های جمع‌آوری شده نشان می‌دهد که از دیدگاه معلمان، ارزشیابی توصیفی نسبت به ارزشیابی سنتی مطابقت بیشتری با استانداردهای ارزشیابی دارد. همچنین در میان استانداردهای مذکور، استاندارد کارآوری نسبت به استاندارد قابلیت اجرا مطابقت بیشتری را با ارزشیابی توصیفی نشان می‌دهد، ولی سایر استانداردها تفاوت معناداری با یکدیگر ندارند.

واژگان کلیدی: ارزشیابی آموزشی، ارزشیابی توصیفی، ارزشیابی سنتی، استانداردهای ارزشیابی

می‌شود. همان طور که اشاره شد ارزشیابی سنتی تا حد زیادی مبتنی بر مکتب روان‌شناسی رفتارگرایی و روان‌شناسی تجربی است، لیکن ارزشیابی توصیفی از دیدگاه روان‌شناسی شناختی نشأت گرفته و مکاتب تحولی و انسانی روان‌شناسی اغلب تغذیه کننده آن هستند (فرمehینی فراهانی، ۱۳۸۲).

در سال‌های اخیر نظام ارزشیابی سنتی انتقادات گوناگونی را به همراه داشته و همین امر موجب گرایش نظام آموزشی به رویکردهای نوین ارزشیابی شده است. از جمله این که ارزشیابی سنتی، سنجش و اندازه‌گیری را آخرین حلقه زنجیره فرایند آموزش تلقی نموده و آن را مهر خاتمه کار به شمار می‌آورد ولی ارزشیابی توصیفی (تکوینی) سنجش را حلقه میانی و واسط آموزش و یادگیری در نظر گرفته و بر اساس بازخورد حاصل از آن به اصلاح و تغییر ساختار آموزش، نیازسنجی، برنامه‌ریزی، اجرای فرایند آموزش و ارزیابی می‌پردازد و اساساً به همین منظور طراحی و ارائه شده است.

روش‌های سنتی و آزمون‌های رایج اندازه‌گیری به دلیل ویژگی ایستا و محافظه‌کارانه خود نمی‌توانند تأثیر قابل ملاحظه‌ای در بهبود تدریس و پیشرفت وضعیت تحصیلی شاگرد داشته باشند و دیگر آن که فرد را کمتر از آن چه هست و بسیار کمتر از آن چه می‌تواند باشد برآورد می‌کنند و قدرت پیش‌بینی ضعیفی دارند (لطف‌آبادی، ۱۳۸۲). از دیگر انتقادات وارد بر آزمون‌ها و روش‌های سنتی ارزیابی نادیده گرفتن تفاوت‌های فردی در یادگیری و سنجش، تأکید و تمرکز بر سنجش سطوح پایین هدف‌های آموزشی و شناختی، یکسان‌گری و همسان‌سازی فراگیران با استعدادهای متفاوت، یادگیری و سنجش حافظه به جای یادگیری در حد تسلط و سنجش عملکرد و غیره است (فرمehینی فراهانی، ۱۳۸۲).

ارزشیابی سنتی به دلیل ایجاد اضطراب، تنش و ناکامی در دانش‌آموزان، افسردگی و آسیب‌های روانی افراد جامعه را سبب شده است. این نوع ارزشیابی به علت انحراف در هدف‌های آموزش و پرورش و خلاصه نمودن آن در دانش اندوزی صرف و انتقال اطلاعات و مهارت‌ها و همچنین نادیده گرفتن رشد عاطفی، اجتماعی، جسمانی و حتی

در جهان امروز تحولات شگرف در علوم مختلف و ظهور دیدگاه‌های نوین پیرامون مسائل اجتماعی، سیاسی، اقتصادی و فرهنگی سبب شده است که راهبردها و روش‌های آموزشی و ارزشیابی دستخوش تغییر شوند. کیفیت مطلوب نظام‌های آموزشی نقش تعیین‌کننده‌ای در پیشرفت اقتصادی، اجتماعی، فرهنگی و توسعه پایدار کشورها دارند و اقدام برای بهبود آن مستلزم ایجاد دگرگونی در برنامه‌های درسی، فرایندهای یاددهی - یادگیری و شیوه‌های ارزشیابی آموخته‌هاست. در این میان ارزشیابی از پیشرفت تحصیلی دانش‌آموزان یکی از عناصر مهم و اساسی برنامه‌های درسی و نظام‌های آموزشی و پرورشی جهان به شمار می‌رود. آن چه ضرورت توجه به نظام ارزشیابی را در این نظام آشکار می‌سازد وجود مشکلاتی است که به ارزشیابی موجود در مدارس نسبت داده می‌شود (حسینی و کاظمی، ۱۳۸۲).

اخیراً تلاش‌های مربوط به پاسخ‌گویی نظام‌های آموزشی بر جمع‌آوری، تحلیل و استفاده از نتایج ارزیابی دانش‌آموزان جهت بهبود آموزش متمرکز شده است. بدین منظور مدارس روش‌هایی را جهت ارزیابی جستجو می‌کنند که بتوانند اطلاعات کافی را درباره عملکرد واقعی دانش‌آموزان فراهم نمایند. در پاسخ به این امر نظام ارزیابی جدیدی شکل گرفته است.

(ملی‌تلو، شوید و سربسی (Militelto, Schweid & Sireci, 2010)

چارچوب و مبانی نظری نظام ارزشیابی سنتی تحت تأثیر روان‌سنجی کلاسیک و مکتب روان‌شناسی رفتارگرایی یا تداعی‌گرایی شکل گرفته است. سال‌هاست که کارکرد این نظام ارزشیابی مورد بررسی و تأمل دانشمندان و مربیان واقع شده تا نقاط ضعف و آسیب‌زای آن شناسایی گردد. تداوم این نقدها و بروز ضعف‌های جدی این نظام به نوعی موجبات رویگردانی از آن را فراهم نموده است. این جریان موجب شده است تا جستجوی رویکردهای بدیل ارزشیابی تحصیلی از جمله رویکرد کیفی و توصیفی در دستور کار قرار گیرد (حسینی، ۱۳۸۴). روش‌های سنتی و جدید ارزشیابی از دیدگاه‌ها و فلسفه‌های متفاوتی برخوردارند. همین خاستگاه‌های متفاوت بنیادی و فلسفی است که موجب تفاوت‌های روشی، راهبردی، کلی و جزئی شده و

جلب نمود. در سال های دهه ۱۹۳۰ با توجه به تغییر نگرش نسبت به ماهیت آموزش، تحصیل و روان شناسی دانش آموز، ایده ارتقای خود به خود (Automatic Promotion) طرح شد که در آن کودکان به طور جمعی به همراه همسالان خود به کلاس بالاتر ارتقا می یافتند. در سال های دهه ۱۹۹۰ الگوی ارتقای خود به خود در آمریکا به چالش کشیده شد، این کار موجب طرح مباحث جدی در محافل علمی تربیتی این کشور گردید. تردیدهای جدی درباره الگوی ارتقای خود به خود از سال های ۱۹۷۰ موجب شد تا بسیاری از ایالات آمریکا عملاً از ارتقای خود به خود روی گردانده و به دنبال شکل سومی از شیوه ارتقا باشند، برخی هم تکرار پایه را جایگزین آن نمودند. این در حالی بود که هیچ پژوهشی از تکرار پایه به ویژه در سه سال نخست ابتدایی حمایت نمی کرد (استاینر ۱۹۹۰).

انگیزه اساسی برای رویگردانی از الگوی ارتقای خود به خود، مسأله استانداردهای آموزشی را به همراه داشت. جریانی که به شدت در فضای آموزشی آمریکا و برخی کشورهای دیگر جلوه کرده بود. این جریان تحت عنوان نهضت استانداردها (Standards Movement) پا به عرصه گذاشت و به دنبال راه حل عملی برای افت آموزشی کشورها بود. سابقه این طرح در برخی از کشورها حداقل به سال ۱۹۳۰ باز می گردد که با عنوان ارتقای خود به خود مطرح گردید. در ایران نیز این طرح در سال ۱۳۳۴ به اجرا آمد، اما پس از مدتی بدون آن که دلایل مثبت شده ای داشته باشد، دچار توقف شد (شهنازی، ۱۳۸۱).

به طور کلی در کشور ما فرایند سنجش و ارزشیابی تنها بر ارزیابی مجموعی یا نهایی استوار بوده است که در پایان دوره آموزش به منظور تشخیص میزان پیشرفت تحصیلی دانش آموزان و انجام نمره گذاری صورت می گیرد. در این روش تمامی تلاش معلم، دانش آموز و کل نظام آموزشی معطوف به آزمون های نهایی و نمره آنها است و نتیجه همه تلاش ها در یک عدد یک یا دو رقمی خلاصه می شود (بنی اسدی، ۱۳۸۴). بر همین اساس به منظور جلوگیری از آثار سوء و تبعات یاد شده نظام ارزشیابی موجود، دفتر ارزشیابی تحصیلی و تربیتی با شعار «ارزشیابی در خدمت تعلیم و تربیت» و با نگاه به آخرین رویکردهای

مهارت های عقلانی، چارچوب گرایی و تأکید بر نتایج مانع تقویت فرایندهای مهمی چون تفکر و خلاقیت در دانش آموزان این نظام شده و ما را به سمت جامعه ای بدون ابداع سوق داده است. کنترل بیرونی قدرت انتخاب افراد را کاهش داده و آنها را نیازمند انگیزه ها، پاداش ها یا تنبیه های بیرونی کرده است. بنابراین دانش آموز کمتر فرصت یا انگیزه پرداختن به مطالعات در زمینه علاقه خود و خارج از چارچوب را امتحانات پیدا می کند (حسنی و کاظمی، ۱۳۸۲).

از طرفی ارزشیابی توصیفی نیز با ضعف هایی چون؛ عدم آگاهی معلمان از شیوه اجرای آن، وجود ابهامات در ارائه کارنامه و چگونگی ثبت و تحویل آن به اولیاء، وقت گیر بودن به دلیل تنوع ابزارها، عدم تناسب با حجم کتاب درسی، کاهش انگیزه و بی توجهی به یادگیری، امکان قضاوت سطحی معلمان و غیره همراه است. این موارد نشان می دهد که هر دو نوع روش ارزشیابی از ضعف ها و کاستی های گوناگونی رنج می برند و نیازمند اتخاذ رهیافتی نو در ارزشیابی است. این رهیافت را می توان در نقش و اهمیت استانداردهای چهارگانه ارزشیابی جستجو کرد تا این که معلوم شود کدام یک از روش های ارزشیابی مطابقت بیشتری با استانداردهای مذکور دارند. در این پژوهش سعی بر آن است تا ضمن مقایسه دو نوع روش ارزشیابی بر اساس استانداردهای چهارگانه، میزان تطابق آنها را با این استانداردها مشخص کرده تا بتواند راهنمای خوبی برای کسب اطمینان از روش مطلوب ارزشیابی تحصیلی در مدارس باشد.

مروری بر مبانی نظری و پیشینه پژوهش نشان می دهد که از زمانی که مدارس پایه بندی، و به جای مدارس یک کلاسه، کلاس هایی با گروه های همگن تأسیس شدند، موضوع نحوه ارتقای آنها به کلاس های بالاتر مطرح گردید. این امر حدوداً در میانه قرن نوزدهم اتفاق افتاد. در آن هنگام میزان ها و استانداردهای مشروط و غیر قابل انعطافی وجود داشت که ارتقای دانش آموزان بر همان اساس صورت می گرفت. به تدریج موضوع تکرار پایه با هدف ارائه فرصت اضافی به دانش آموزان ضعیف برای افزایش دانش خود و یا کسب ارتقا مطرح گردید و توجه پژوهشگران را به خود

سنتی اتخاذ شده ولی دارای ویژگی‌هایی است که با ویژگی‌های رویکرد سنتی تفاوت‌های قابل ملاحظه‌ای دارد. در جدول (۱) برخی از ویژگی‌های آنها با هم مورد مقایسه قرار گرفته است:

استانداردهای ارزشیابی (Evaluation Standards) شامل مجموعه‌ای از اصول و ملاک‌های مورد توافق متخصصان و دست‌اندرکاران ارزشیابی آموزشی است که در صورت به کارگیری آنها، ارتقای کیفی سنجش پیشرفت تحصیلی را در جهت تحقق اهداف آن سبب می‌شود. استانداردهای سنجش پیشرفت تحصیلی را می‌توان به چهار دسته و هر دسته را به چند مؤلفه طبقه‌بندی کرد (گالیک سان ۲۰۰۱، Gullickson).

جدید در نظام ارزشیابی دنیا در سال ۱۳۸۱ طرح ارزشیابی توصیفی را تدوین و به شورای عالی آموزش و پرورش پیشنهاد داد. طرح مذکور در مورخ ۸۱/۸/۳۰ تصویب و مقرر گردید تا در سه پایه اول، دوم و سوم ابتدایی هر دانش‌آموز پس از یک سال حضور در یک پایه تحصیلی به پایه تحصیلی بالاتر ارتقا یابد. پس از آن به دلیل تأکید بر این نوع ارزشیابی، در تاریخ ۱۳۸۳/۶/۳ استمرار اجرای آزمایشی ارزشیابی توصیفی در دوره ابتدایی به تصویب شورای عالی آموزش و پرورش کشور رسید. در مورخ ۱۳۸۷/۴/۱۸ با تصویب شورای عالی آموزش و پرورش مقرر گردید تا طرح ارزشیابی توصیفی (کیفی) دوره ابتدایی در سال تحصیلی ۸۸-۱۳۸۷ در آن دسته از مدارس که شرایط و امکانات لازم را

جدول ۱- مقایسه ویژگی‌های ارزشیابی سنتی و جدید

مؤلفه‌ها	ویژگی‌های ارزشیابی سنتی	ویژگی‌های ارزشیابی جدید (توصیفی)
غایت و نهایت	تمرکز بر مسئولیت و پاسخ‌گویی	تمرکز بر بهبود و پیشرفت
ماهیت اهداف	تمایل به اهداف مجزا و منفرد سطح پایین	مجموعه اهداف متحد و یکپارچه از برون‌دادهای سطح بالا
ماهیت تکالیف	گرایش به تکالیف ساختارمند و محدود	روش‌های متعدد و متنوع به ویژه باز پاسخ
شرایط اجرا	از نظر زمان به صرفه است	در مواردی زمان بر است
ماهیت پاسخ	پاسخ‌های ثابت و از پیش تعیین‌شده، انتخابی (چندگزینه‌ای)	معمولاً توسط فراگیر ساخته می‌شود به علاوه پاسخ‌های شفاهی، فرایندی، کتبی، برون‌داده و مشاهده عملکرد
نمره‌گذاری و گزارش‌دهی	روش‌های عینی، ماشین‌های نمره‌دهی و نمرات قابل فهم برای همه	ممکن است پیچیده باشد و چند بعدی
اعتبار (پایایی)	تقریباً همیشه در سطح بالاست	آموزش و کارورزی می‌تواند به حد قابل قبول برسد
هزینه	در حد متوسط و نرمال است	ممکن است خیلی بالا باشد
فشار و اضطراب بر آزمودنی	ممکن است فشار شکننده باشد به ویژه در آزمون‌های سرنوشت‌ساز و نهایی	ترس و اضطراب کم است

(با اقتباس از پین، ۲۰۰۳ به نقل از فراهانی، ۱۳۸۱)

۱- استانداردهای اخلاقی و قانونی (Utility Standards): استانداردهایی هستند که با در نظر گرفتن موازین قانونی و اصول اخلاقی، پیشرفت تحصیلی دانش‌آموزان را ارزشیابی می‌کنند به گونه‌ای که جانب عدالت، حرمت، سلامت و بهداشت روانی دانش‌آموز و همه کسانی که از سنجش تأثیر می‌پذیرند، مراعات شود.

دارند اجرا شود و پس از رفع نارسایی‌های موجود از سال تحصیلی ۸۹ - ۱۳۸۸ در کلیه مدارس ابتدایی کشور مورد اجرا قرار گیرد.

واقعیت آن است که رویکرد نوین ارزشیابی با نگاهی متفاوت از رویکرد سنتی به دنبال ارزیابی فراگیران است. با این که این رویکرد به منظور رفع نارسایی‌های رویکرد

۲- استانداردهای کارآوری و سودبخشی (Propriety Standards)

بر اساس این استانداردها، در فرایند یاددهی - یادگیری باید به گونه‌ای از پیشرفت تحصیلی دانش‌آموزان ارزشیابی به عمل آید که نتایج آن مفید و سود بخش بوده و بتوان از آن استفاده بهینه به عمل آورد.

۳- استانداردهای قابلیت اجرا (Feasibility Standards)

استانداردهایی هستند که پیشرفت تحصیلی دانش‌آموزان را به صورتی واقع بینانه با توجه به محدودیت‌ها و امکانات موجود مورد سنجش قرار می‌دهند.

۴- استانداردهای دقت و صحت (Accuracy Standards)

استانداردهایی هستند که نتایج ارزشیابی از پیشرفت تحصیلی و عملکرد واقعی دانش‌آموز را با درجه قابل قبولی از دقت و صحت و به دور از خطای سوگیری توصیف می‌کنند. در جدول زیر مؤلفه‌های مربوط به هر یک از استانداردهای چهارگانه، به تفکیک آورده شده است:

۵- مطلوب بودن دو روش ارزشیابی سنتی و توصیفی می‌تواند بر اساس استانداردهای چهارگانه مورد مقایسه و بررسی قرار گیرد تا معلوم شود کدام یک از آنها از نظر استانداردهای مذکور کارآیی لازم را در مدارس دارا هستند. متأسفانه تا به حال پژوهش‌هایی که به بررسی تطبیقی استانداردهای مذکور برای دو روش ارزشیابی انجام شده باشد، یافت نشده است، بلکه اکثر تحقیقات متمرکز بر بررسی تأثیر ارزشیابی تکوینی در کنار ارزشیابی پایایی است.

۶- مرکوفر (Merkhofer, 1995) و استیفنز (stephens, 1991)

به نقل از شاه زمانی (۱۳۸۰) نشان دادند که چنانچه آزمون‌های تکوینی و یا سایر آزمون‌ها، بتوانند بازخورد لازم را در اختیار دانش‌آموزان قرار دهند عملکرد آنان افزایش می‌یابد و در صورتی که دانش‌آموزان را برای آزمون‌های مرحله‌ای برانگیزیم، پیش از آزمون‌های پایانی آمادگی‌های لازم را کسب خواهند نمود. نتایج تحقیقات بالک (Balk, 1990) نشان داد که در کلاس گروه گواه دانش‌آموزان تنها توانستند نیمی از نمره مربوط به آزمون‌های پایانی برای هر یک از تکالیف یادگیری را به دست آورند و در مقابل روش استفاده از

آزمون‌های تکوینی و فراهم آوردن بازخورد مناسب و به دنبال آن اصلاح مشکلات یادگیری که در مورد دانش‌آموزان گروه تجربی به کار رفت سبب شد که این دانش‌آموزان تا ۹۰ درصد نمرات همان آزمون‌ها را به دست آورند. اوسگود (Osgood, 1996) و پتر (Peter, 1992) نیز در تأیید تأثیر آزمون‌های مرحله‌ای بر پیشرفت تحصیلی به نتایج مشابهی دست یافتند (شاه زمانی، ۱۳۸۰).

اندرسون و همکاران (Anderson, 2006) استفاده ارزشیابی توصیفی بدون ساخت را برای رشد و پیشرفت دانش‌آموزان کافی نمی‌دانند. از نظر آنها ارزشیابی توصیفی ساختمان‌ساز می‌تواند موجبات رشد و پیشرفت آنها را فراهم آورد. مک فایل و هالبرت (MacPhail, & Halbert, 2010) با طرح رویکرد «ارزیابی برای یادگیری» (Assessment for learning) روش ارزیابی برای نمره گرفتن را روش مناسبی برای نیل به پیشرفت دانش‌آموزان نمی‌دانند.

از نظر آنان ارزیابی تکوینی اگر به گونه درستی اجرا شود، می‌تواند روش مناسبی برای نیل به رویکرد ارزیابی برای یادگیری باشد.

بروخت، موس و لانگ (Brookhart, Moss&Long, 2010) در بررسی نقش ارزیابی تکوینی در کلاس‌های جبرانی نشان دادند که انعکاس نتایج پیشرفت به دانش‌آموزان و مشاهدات نظارتی در طول اجرای آموزش می‌تواند نقش قابل توجهی را در پیشرفت دانش‌آموزان - هم در درس علمی و هم نظری - داشته باشد.

حیدری (۱۳۷۵) در پژوهشی که به بررسی تأثیر اجرای ارزشیابی تکوینی بر افزایش پیشرفت تحصیلی انجام داده، نشان داد که اگر ارزشیابی تکوینی همراه با بازخورد باشد با میزان یادگیری همبستگی مستقیم دارد در صورتی که بدون بازخورد باشد یادگیری در سطح پایین‌تر خواهد بود.

خوش خلق (۱۳۸۳) در بررسی مقایسه‌ای خود از امتحانات دو نوبته با سه نوبته نتیجه می‌گیرد که هر چه سایه امتحان، به ویژه امتحان تراکمی و نمره از محیط کلاس و آموزش کم شود، تأثیرات مثبتی به همراه خواهد داشت. بنابراین اگر امتحان در مواردی که این چنین آسیب‌رسان باشد، بهتر است که در چگونگی اجرای آن یا حتی در اصل اجرای آن شک و تردید نمود و سعی کرد با بازنگری آن،

جدول ۲- استانداردهای چهارگانه ارزشیابی و مؤلفه‌های مربوط به آنها

استانداردهای اخلاقی	استانداردهای کارآوری	استانداردهای	استانداردهای
و قانونی	و سودبخشی	قابلیت اجرا	دقت و صحت
<ul style="list-style-type: none"> • خدمت به دانش آموز • سیاست‌ها و رویه‌های مناسب • دسترسی به اطلاعات سنجش • رعایت حقوق دانش آموزان در • تمامی مراحل سنجش پیشرفت • تحصیلی 	<ul style="list-style-type: none"> • جهت گیری سازنده ارزشیابی • مشخص بودن موارد کاربردی نتایج و کاربران آن • دامنه اطلاعات جامع • صلاحیت‌های لازم ارزشیاب • ملاک‌های درجه‌بندی و ارزش‌گذاری صریح و روشن • گزارش سود بخش و مؤثر • پی‌گیری نتایج 	<ul style="list-style-type: none"> • جهت‌گیری عملی • هماهنگی با سیاست‌ها • منابع پشتیبانی لازم 	<ul style="list-style-type: none"> • روایی • انتظارات تعریف‌شده برای دانش آموز • تحلیل تأثیر بافت سنجش • اطلاعات و نتایج پایا و قابل دفاع • مدیریت اطلاعات و کنترل کیفیت • تجزیه و تحلیل اطلاعات • نتیجه‌گیری توجیه‌پذیر • فرار ارزشیابی

(با اقتباس از گالیک سان، ۲۰۰۳)

خوی‌نژاد (۱۳۸۰؛ به نقل از زمانی، ۱۳۸۱) در تحقیقی تحت عنوان «استفاده از استانداردها به عنوان سنجش حداقل شایستگی در نظام آموزش و پرورش ایران» نشان داد که در نظام ارزشیابی ما، استانداردهای موفقیت به نمره‌های خاصی اطلاق می‌شود که به وسیله ابزارها کسب می‌شوند و هیچ‌گونه اعتمادی بر روایی و اعتبار آنها نیست. اساساً شواهدی در دست نیست که بتواند تأکید کند، نمره خام تا چه حد مبین تسلط فرد در آن مهارت است. بنابراین، استفاده از شیوه‌های جدیدتر اندازه‌گیری و ارزشیابی در تعلیم و تربیت در جایگزین کردن نظام سنتی امتحان‌های پایان سال تحصیلی، و به ویژه امتحان‌های پایان سطح‌های تحصیلی و امتحان‌های نهایی، که هنوز تأکید آن‌ها بر محفوظات دانش‌آموزان است، ضروری است.

روش پژوهش

تحقیق حاضر به لحاظ هدف از نوع تحقیقات کاربردی است. زیرا که هدف آن، کمک و کاربرد نتایج برای اتخاذ تصمیم‌های مربوط به نظام آموزشی است. از سوی دیگر این پژوهش از لحاظ گردآوری داده‌ها از نوع تحقیقات توصیفی - پیمایشی است. چون پژوهشگر قصد دارد با جمع‌آوری اطلاعات پیرامون مطابقت روش‌های ارزشیابی سنتی و

محیط یاددهی و یادگیری را به محیطی سالم و با نشاط تبدیل کرد.

شاه زمانی (۱۳۸۰) در بررسی «ارزشیابی مستمر نظام سالی واحدی دوره متوسطه استان اصفهان» نشان داده است که از نظر دبیران، ارزشیابی‌های تکوینی به میزان زیادی در آمادگی دانش‌آموزان برای ارزشیابی پایانی مؤثر است. تراکم کلاس، مفاهیم و محتوای کتب درسی، کمبود امکانات و تجهیزات آموزشی و عدم فضای مناسب برای اجرای ارزشیابی‌های مستمر تأثیر بسزایی دارد.

زمانی (۱۳۸۱) نیز با بررسی تطبیقی «روش‌های جدید ارزشیابی تحصیلی دانش‌آموزان دوره ابتدایی کشورمان با کشورهای کانادا و استرالیا» نشان داده که ارزشیابی در کشورمان با این کشورها کاملاً متفاوت است. معلمان در این کشورها بر اساس توصیف‌ها و ویژگی‌های شخصیتی و... پیشرفت تحصیلی دانش‌آموزان را می‌سنجند. کارنامه تحصیلی توصیفی است و معلم راجع به نقاط قوت و ضعف دانش‌آموزان اظهار نظر می‌کند. در کارنامه، فرصتی برای خود ارزیابی دانش‌آموزان در نظر گرفته شده و شناخت بسیار دقیق دانش‌آموزان و هدایت آنها به رشته‌های متناسب از دستاوردهای نظام ارزشیابی آنهاست.

سعی شد تا ضرایب آلفای هر یک از مؤلفه‌ها به طور مستقل و در نهایت ضریب آلفای کل پرسشنامه (۰/۹۲) به دست آید. ضرایب آلفای به دست آمده از این تحلیل حاکی از پایایی مورد قبول پرسشنامه بود. پس از اطلاع از میزان روایی و پایایی، پرسشنامه نهایی شده توزیع و جمع آوری گردید.

برای تجزیه و تحلیل داده‌های جمع‌آوری شده از شاخص‌های آمار توصیفی مواردی چون؛ فراوانی، میانگین و انحراف استاندارد و از شاخص‌های آمار استنباطی ضمن در نظر داشتن فرض‌های مربوط به آزمون‌های آماری پارامتریک شاخص‌هایی چون: آزمون t یک گروهی جهت آزمون فرض، تحلیل واریانس یک راهه و آزمون تعقیبی شفه برای تعیین تفاوت یا عدم تفاوت میان بیش از دو میانگین استفاده شده است.

نتایج و یافته‌ها

فرضیه اول پژوهش: ارزشیابی توصیفی و سنتی از نظر انطباق با استانداردهای چهارگانه متفاوتند.

این فرضیه شامل چهار زیر مجموعه بر حسب استانداردهای چهارگانه ارزشیابی است. نتایج حاصل از تحلیل آنها در جدول زیر قابل مشاهده است. در این جدول، میانگین فرضی به معنای عدم تفاوت انطباق بین دو روش ارزشیابی با استانداردها، پایین‌تر از میانگین فرضی به معنای برتری مطابقت ارزشیابی سنتی نسبت به توصیفی و بالاتر از میانگین فرضی به معنای برتری مطابقت ارزشیابی توصیفی نسبت به سنتی است.

بر اساس داده‌های جدول بالا میانگین کلیه استانداردهای ارزشیابی با توجه به سطوح لیکرت در طیف «زیاد» قرار گرفته است. میانگین به دست آمده حاکی از آن است که معلمان میزان انطباق ارزشیابی توصیفی را نسبت به ارزشیابی سنتی در همه استانداردها در سطح زیاد و بالاتر از میانگین فرضی اعلام نموده‌اند. از سوی دیگر آزمون فرض انجام شده پیرامون استانداردهای مذکور نشان می‌دهد که در تمام استانداردها تفاوت میانگین‌های به دست آمده با میانگین‌های فرضی معنادار است. این بدین معنی است که از نظر معلمان هنگام مقایسه دو نوع ارزشیابی با استانداردهای چهارگانه، مطابقت ارزشیابی

توصیفی با استانداردهای چهارگانه، وضعیت موجود را آن گونه که هست گزارش کند.

جامعه آماری پژوهش حاضر شامل کلیه معلمان است که در سال تحصیلی ۸۶ - ۱۳۸۵ در مناطق ۱، ۵، ۹، ۱۲ و ۱۶ به عنوان نمونه منتخب سازمان آموزش و پرورش شهر تهران در هفت مدرسه ابتدائی مشغول اجرای آزمایشی طرح ارزشیابی توصیفی هستند که تعداد آنها ۴۰ نفر است. بدین ترتیب در این پژوهش به دلیل محدود بودن تعداد اعضای جامعه همه آنها مورد مطالعه قرار گرفته و به عبارتی حجم جامعه و نمونه با هم برابر است. ویژگی‌های جامعه مذکور به شرح زیر است:

جدول ۳- جامعه آماری پژوهش

منطقه	نام آموزشگاه	دخترانه	پسرانه
۱	عدالت	۵	
۱	مهدوی	۶	
۵	غیرانتفاعی رهیار		۷
۹	آزادگان		۷
۹	عاصمی	۴	
۱۲	برهان		۵
۱۶	غنچه‌های انقلاب	۶	

در این پژوهش ابزار مورد نظر برای گردآوری داده‌ها پرسشنامه است. علاوه بر درج اطلاعات جمعیت شناختی در ابتدای پرسشنامه، سعی شد تا بر اساس هر یک از استانداردهای چهارگانه گویه‌های مربوط به آن با توجه به مطالعات نظری تدوین شود. برای اطلاع از میزان روایی (Validity) پرسشنامه، نظرات اصلاحی چند نفر از استادان مربوط، متخصصان امر و چند نفر از افراد جامعه آماری دخیل گردید. برای محاسبه پایایی نیز پرسشنامه تدوین شده به طور مقدماتی بر روی ۳۰ نفر از اعضای جامعه آماری اجرا، سپس داده‌های جمع آوری شده از طریق نرم‌افزار SPSS تحلیل و میزان ضریب همسانی درونی سوالات به وسیله آلفای کرانباخ محاسبه شد. در این باره

جدول ۴- میانگین مطابقت ارزشیابی توصیفی و سنتی با استانداردهای چهارگانه و آزمون فرض انجام شده درباره آنها

نوع استاندارد	تعداد	انحراف معیار	میانگین	t	df	سطح معناداری
اخلاقی و قانونی	۴۰	۰/۵۸	۳/۹۳	۳	۳۹	۰/۰۰۰
کارآوری و سودبخشی	۴۰	۰/۵۸	۴/۱۲	۳	۳۹	۰/۰۰۰
قابلیت اجرا	۴۰	۰/۷۰	۳/۶۶	۳	۳۹	۰/۰۰۰
دقت و صحت	۴۰	۰/۶۴	۳/۸۴	۳	۳۹	۰/۰۰۰
کل استانداردها	۴۰	۰/۵۳	۳/۹۳	۳	۳۹	۰/۰۰۰

جدول ۵ - تحلیل واریانس یک راه جهت تعیین تفاوت میانگین استانداردهای چهارگانه

نوع استاندارد	تعداد	میانگین	انحراف معیار	F	سطح معناداری
اخلاقی و قانونی	۴۰	۳/۹۳	۰/۵۸	۳/۵۹۶	۰/۰۱۵
کارآوری و سودبخشی	۴۰	۴/۱۲	۰/۵۸		
قابلیت اجرا	۴۰	۳/۶۶	۰/۷۰		
دقت و صحت	۴۰	۳/۸۴	۰/۶۴		

جدول ۶ - نتایج آزمون تعقیبی شفه برای تعیین تفاوت میانگین استانداردهای چهارگانه

مؤلفه‌ها	تفاوت میانگین‌ها	سطح معناداری
کارآوری و سودبخشی	-۰/۱۸۱۰	۰/۶۴۸
قابلیت اجرا	۰/۲۷۰۹	۰/۲۹۹
دقت و صحت	۰/۰۹۷۱	۰/۹۲۴
قابلیت اجرا	۰/۴۵۱۹	۰/۰۱۸
دقت و صحت	۰/۲۷۸۱	۰/۲۷۶
دقت و صحت	-۰/۱۷۳۸	۰/۶۷۷

مقایسه میانگین‌های فوق در جدول (۶) بیانگر آن است که در میان استانداردهای مذکور، تنها میانگین استاندارد کارآوری و قابلیت اجرا با همدیگر تفاوت معناداری دارند. بدین معنی که مطابقت استاندارد کارآوری با ارزشیابی توصیفی نسبت به استاندارد قابلیت اجرا بیشتر است. با این حال در بین میانگین استانداردهای دیگر تفاوت معناداری مشاهده نمی‌شود. بنابراین فرضیه دوم پژوهش برای دو استاندارد کارآوری و قابلیت اجرا رد شده ولی برای سایر استانداردها مورد تأیید است.

بحث و نتیجه‌گیری

نتایج به دست آمده در این پژوهش نشان می‌دهد که معلمان مجری طرح ارزشیابی توصیفی، این طرح را نسبت

توصیفی نسبت به ارزشیابی سنتی در همه استانداردها بیشتر است. همچنین این امر برای میانگین کل استانداردها نیز صادق است. بنابراین می‌توان گفت که فرضیه اول پژوهش تأیید می‌گردد. بدین معنی که مطابقت ارزشیابی توصیفی نسبت به ارزشیابی سنتی در تمام استانداردها بیشتر است. به عبارت دیگر، از نظر افراد مطالعه شده ارزشیابی سنتی به اندازه ارزشیابی توصیفی با استانداردهای ارزشیابی مطابقت ندارد.

فرضیه دوم پژوهش: میانگین استانداردهای چهارگانه ارزشیابی تفاوت معناداری با یکدیگر ندارند.

نتایج جدول بالا نشان می‌دهد که در بین میانگین استانداردهای چهارگانه ارزشیابی تفاوت معناداری وجود دارد. نتایج به دست آمده از آزمون تعقیبی شفه درباره

یکدیگرند و در واقع همه تأکید و حساسیت علمی برای آن است که اولاً استفاده از آن‌ها به شرایط ویژه و اقتضایی محدود شود و ثانیاً این روش‌ها در کنار روش‌ها و رویکردهای تعاملی و پویا و انسان مدار به کار رود و تصمیمات و اقدامات حیاتی و سرنوشت ساز صرفاً مبتنی بر روش‌های سنتی اتخاذ نگردد.

بنا بر مبانی نظری و نتایج پژوهش‌های انجام شده، ارزشیابی توصیفی زمانی کارآمد است که با بازخورد همراه باشد. لذا در این زمینه باید تدابیر عملی اندیشیده شود و از بازخورد حاصل از مرور نتایج سنجش پیشرفت تحصیلی دانش‌آموزان، اقدامات اساسی در جهت ارتقای روش‌های تدریس و اصلاح برنامه‌های آموزشی صورت گیرد. بنابراین سعی شود تا دیدگاه والدین، دانش‌آموزان و کارکنان آموزشی نیز به عنوان ذی‌نفعان دیگر طرح ارزشیابی آموزشی به طور مستمر مورد بررسی گیرد. زیرا که هر طرحی به نوبه خود دارای محاسن و معایب خاصی است که می‌تواند از طریق بازخوردهای اصلاحی صحیح نتیجه بخش باشد. علاوه بر این در این پژوهش به لحاظ ماهیت آن امکان کنترل برخی از متغیرها میسر نبود، پیشنهاد می‌شود که تحقیقات آزمایشی پیرامون دانش‌آموزانی که به صورت توصیفی و سنتی ارزشیابی می‌شوند به عمل آید تا با توجه به نتایج به دست آمده به میزان قابلیت و کارآمدی هر دو روش دست یافت. هم‌چنین برخی از مشکلات و منشأ آسیب‌پذیری نظام ارزشیابی ممکن است ناشی از موقعیت خود نظام ارزشیابی و یا ناشی از تعارض‌ها و تضادهای موجود میان ارزشیابی و دیگر ساختارها باشد. لذا با انجام پژوهش‌هایی در رابطه با آسیب‌های ساختار سازمانی و اداری سعی شود تا دانش‌آموزان به اهداف نهایی و اصیل آموزش و پرورش نزدیک شوند.

به ارزشیابی سنتی از نظر مطابقت با استانداردهای چهارگانه بهتر ارزیابی نموده‌اند. به نظر می‌رسد که از دیدگاه معلمان اگر قصد اصلاح نظام ارزشیابی بر اساس رعایت موازین مربوط به استانداردهای اخلاقی، کارآوری، قابلیت اجرا و دقت در میان باشد، کاربست و اجرای این نوع ارزشیابی در مقاطع ابتدایی مناسب‌تر از ارزشیابی سنتی است. بر این اساس جهت توسعه طرح ارزشیابی توصیفی، نظام آموزش و پرورش می‌تواند آن را به تدریج در کل مدارس ابتدایی و همان‌طور در پایه‌ها و مقاطع بالاتر نیز گسترش دهد. نتایج و یافته‌های این پژوهش با نتایج پژوهش‌های پیشین نیز قابل مقایسه است. به طوری که صرف نظر از مقایسه دو نوع ارزشیابی با استانداردهای چهارگانه، تحقیقات افرادی چون؛ حیدری (۱۳۷۵)، شاه زمانی (۱۳۸۰) و زمانی (۱۳۸۱) کاربرد ارزشیابی توصیفی را به ویژه مواقعی که دارای بازخورد باشد، در کنار ارزشیابی سنتی مثبت قلمداد نموده‌اند. هم‌چنین خوش خلق (۱۳۸۳) و خوی نژاد (۱۳۸۰؛ به نقل از زمانی، ۱۳۸۱) به منظور اصلاح و بهبود نظام ارزیابی حاکم بر مدارس در جستجوی گزینه بدیل و روش‌های جدید ارزشیابی هستند. بنابراین می‌توان گفت از آن جا که نتایج پژوهش حاضر نیز مطابقت ارزشیابی توصیفی را بالاتر از ارزشیابی سنتی در استانداردهای چهارگانه نشان می‌دهد، لذا می‌تواند گزینه بهتری برای رفع نواقص ارزشیابی سنتی در نظام آموزشی کشورمان باشد. مقایسه نتایج این پژوهش با پژوهش‌های انجام یافته در خارج از کشور نیز نشان می‌دهد که یافته‌های پژوهش حاضر به نوعی با یافته‌های آنها نیز مطابقت دارد، به طوری که افرادی چون؛ بالک (۱۹۹۰)، استیفنز (۱۹۹۱)، پیتر (۱۹۹۲)، مرکوفر (۱۹۹۵)، اوسگود (۱۹۹۶) به نقل از شاه زمانی (۱۳۸۰)، اندرسون و همکاران (۲۰۰۶) و بروخارت، موس و لانگ (۲۰۱۰) نتایج ارزشیابی تکوینی و توصیفی را نسبت به ارزشیابی سنتی بهتر قلمداد نموده‌اند. هم‌چنین فایل و هالبرت (۲۰۱۰) جهت نیل به رویکرد «ارزشیابی برای یادگیری»، ارزیابی تکوینی را روش مناسبی برای نیل به این هدف تلقی می‌کنند. با این حال این امر بدین معنا نیست که روش‌های سنتی باید از نظام آموزش و سنجش به طور کلی حذف شوند، زیرا این دو نوع ارزشیابی مکمل

منابع

فرمهبینی فراهانی، محسن. (۱۳۸۲). *ارزشیابی استاندارد مدار تربیت شهروندی*. مجموعه مقالات همایش ارزشیابی تحصیلی و تربیتی، تهران: وزارت آموزش و پرورش، معاونت آموزش عمومی و تربیتی، انتشارات تزکیه.

لطف‌آبادی، حسین. (۱۳۸۲). *سنجش و اندازه‌گیری در علوم تربیتی و روانشناسی*. مشهد: انتشارات حکیم فردوسی.

یارمحمدیان، محمدحسین. (۱۳۸۹). *اصول برنامه‌ریزی درسی (ماهیت برنامه‌ریزی درسی - مبانی فلسفی، روانشناختی و جامعه‌شناختی برنامه، ارزشیابی برنامه، تحلیل محتوای برنامه یا کتاب درسی)*. تهران: انتشارات یادواره کتاب.

وزارت آموزش و پرورش. (۱۳۸۷). *استمرار اجرای طرح ارزشیابی توصیفی (کیفی)*. مصوبه هفتصد و شصت و نهمین جلسه شورای عالی آموزش و پرورش.

بنی‌اسدی، حسین. (۱۳۸۴). *آشنایی با طرح ارزشیابی توصیفی*. تهران: نشر آزمون نوین.

حسنی، محمد و کاظمی، یحیی. (۱۳۸۲). *طرح ارزشیابی توصیفی (اهداف، اصول و راهکارها)*. تهیه شده در دفتر ارزشیابی تحصیلی و تربیتی معاونت آموزش عمومی و امور تربیتی وزارت آموزش و پرورش، تهران: نشر آثار معاصر.

حسنی، محمد. (۱۳۸۴). *راهنمای اجرای ارزشیابی توصیفی پایه‌های اول، دوم و سوم ابتدایی*. تهران: نشر آثار معاصر.

خوش‌خلق، ایرج. (۱۳۸۳). *بازنگری در روش‌های سنجش و نقش آن در اصلاحات آموزشی*. تهران: انتشارات تزکیه.

حیدری، شعبان. (۱۳۷۵). *تأثیر اجرای ارزشیابی تکوینی بر افزایش پیشرفت تحصیلی سال اول دبیرستان‌های نظام جدید منطقه فریدون‌کنار*. پایان‌نامه کارشناسی ارشد، دانشگاه علامه طباطبایی.

زمانی، بی‌بی‌عشرت. (۱۳۸۱). *بررسی روش‌های جدید ارزشیابی تحصیلی دانش‌آموزان دوره ابتدایی*. چکیده مقاله‌های همایش ملی مهندسی اصلاحات در آموزش و پرورش، پژوهشکده تعلیم و تربیت.

شاه زمانی، محمد. (۱۳۸۰). *ارزشیابی مستمر نظام سالی واحدی دوره متوسطه استان اصفهان*. مجموعه مقالات اولین همایش ارزشیابی تحصیلی استان اصفهان.

شهبازی، رحمت‌الله. (۱۳۸۱). *سیر تحول آیین‌نامه امتحانات از سال ۱۳۰۲ تا ۱۳۸۱ در نظام آموزش و پرورش ایران دوره عمومی*. دفتر ارزشیابی تحصیلی وزارت آموزش و پرورش.

فراهانی، مهدی. (۱۳۸۱). *نگاهی به روان‌سنجی سنتی و رویکردهای جدید سنجش و ارزشیابی آموخته‌ها*. مجموعه مقالات همایش ارزشیابی تحصیلی و تربیتی، تهران: وزارت آموزش و پرورش، معاونت آموزش عمومی و تربیتی. انتشارات تزکیه.

Anderson, Cynthia M. English, Carie L & Hedrick, Theresa M. (2006). Use of the Structured Descriptive Assessment With Typically Developing Children. Behavior Modification, Beverly Hills: Vol. 30, Iss. 3; pg. 352.

Brookhart, Susan M. Moss, Connie M & Long. Beverly A. (2010). Teacher inquiry into formative assessment practices in remedial reading classrooms. Assessment in Education, Abingdon: Vol. 17, Iss. 1; pg. 41.

Gullickson, Arlen. (2003). The Student Evaluation Standards. USA: Corwin press, American Journal of Evaluation, Vol. 25, Iss. 1; pg. 90

MacPhail, Ann & Halbert, John. (2010). We had to do intelligent thinking during recent PE': students' and teachers' experiences of assessment for learning in post-primary physical education. Assessment in Education, Abingdon: Vol. 17, Iss. 1; pg. 23.

Militello, Matthew. Schweid, Jason & Sireci, Stephen G. (2010). Formative assessment systems: evaluating the fit between school districts' needs and assessment systems' characteristics.

Educational Assessment, Evaluation and Accountability. Dordrecht: Vol. 22, Iss. 1; pg. 29.

Steiner. K. (1990). Retention and Social Promotion. Eric: ED 207899

Archive of SID