

تأثیر ارزشیابی توصیفی بر ویژگیهای شناختی، عاطفی و روانی - حرکتی دانش آموزان

دکتر اکبر رضایی *

دکتر علی اکبر سیف **

چکیده

این پژوهش با هدف بررسی اثربخشی طرح ارزشیابی توصیفی و با توجه به ویژگی‌های شناختی، عاطفی و روانی - حرکتی دانش آموزان پایه سوم ابتدایی شهر تهران انجام گرفت. طرح ارزشیابی توصیفی با استفاده از ارزشیابی تکوینی، ارزشیابی عملکردی، و دادن بازخورد توصیفی از سال تحصیلی ۸۳-۱۳۸۲ تا کنون در چهار منطقه شهر تهران (مناطق ۱، ۵، ۹ و ۱۶) به صورت آزمایشی اجرا شده است. بنابراین، جامعه آماری پژوهش حاضر محدود به این مناطق بوده است. در این پژوهش ۹۶ دانش آموز پایه سوم ابتدایی از مدارس مشمول طرح ارزشیابی توصیفی به عنوان گروه آزمایش (۴۸ دختر و ۴۸ پسر) و همچنین ۹۶ دانش آموز از مدارس هم سطح آنها که تحت پوشش طرح ارزشیابی توصیفی نبودند، به صورت تصادفی، به عنوان گروه گواه (۴۸ دختر و ۴۸ پسر) انتخاب شدند. ابزارهای اندازه گیری مورد استفاده در این پژوهش عبارت بودند از: (۱) آزمون پیشرفت تحصیلی ریاضی، (۲) آزمون پیشرفت تحصیلی علوم تجربی، (۳) آزمون پیشرفت تحصیلی زبان فارسی، (۴) پرسشنامه آگاهی فراشناختی برای سنجش ویژگیهای شناختی، (۵) پرسشنامه کیفیت زندگی در مدرسه برای سنجش ویژگیهای عاطفی و (۶) آزمونهای عملکردی علوم و ریاضی برای سنجش ویژگیهای روانی - حرکتی. برای تحلیل داده‌ها از

* عضو هیئت علمی دانشگاه پیام نور تبریز

** استاد دانشگاه علامه طباطبائی تهران

روشهای آماری تحلیل واریانس چند متغیری (MANOVA) و تحلیل واریانس یک سویه (ANOVA) استفاده شد. نتایج نشان داد که سطح پیشرفت دانش آموزان پایه سوم ابتدایی مدارس مشمول طرح ارزشیابی توصیفی در متغیرهای شناختی از قبیل دانش ریاضی، علوم تجربی، زبان فارسی و همچنین میزان آگاهی فراشناختی، به طور معناداری، بیشتر از دانش آموزان پایه سوم ابتدایی مدارس عادی هم سطح آنهاست. بین گروههای آزمایشی و گواه در متغیرهای حوزه عاطفی، از قبیل رضایت کلی از مدرسه، عاطفه منفی نسبت به مدرسه، نگرش در مورد معلمان، همبستگی اجتماعی، فرصت (باور نسبت به مفید بودن مدرسه)، موفقیت (احساس اطمینان به توانایی برای کسب توفیق در انجام کارهای مدرسه) و ماجرا (احساس خودانگیختگی و لذت بردن از یادگیری) تفاوت‌های معنی داری مشاهده نشد. در متغیرهای حوزه روانی-حرکتی بین دانش آموزان مدارس مشمول طرح ارزشیابی توصیفی و دانش آموزان مدارس عادی تفاوت‌های معنی داری به دست آمد. نتایج تحلیل‌های تکمیلی نشان داد که دانش آموزان مدارس مشمول طرح در آزمون درخت و برگ، نسبت به دانش آموزان مدارس عادی، عملکرد بهتری دارند. در آزمون‌های عملکردی دیگر، از قبیل آزمون اشیای دور ریختنی و آزمون دماسنج، دانش آموزان مشمول طرح ارزشیابی توصیفی، در مقایسه با دانش آموزان مدارس عادی، عملکرد بهتری داشتند؛ اما این تفاوت‌ها از نظر آماری معنی دار نبودند. در این مقاله ضمن بحث درباره یافته‌ها، پیشنهادهایی نیز برای پژوهشهای بعدی مطرح شده است.

واژه های کلیدی: ارزشیابی، سنجش، توصیفی، عملکردی، تکوینی، نگرش، پیشرفت تحصیلی، فراشناخت و دوره ابتدایی.

مقدمه

در سال های اخیر، بحثهای زیادی درباره اصلاحات آموزشی و بهبود کیفیت آموزشی مدارس مطرح شده است. سنجش و ارزشیابی^۱ دانش آموزان در سطح ملی، یک اهم و

۱ در این مقاله، سنجش و ارزشیابی اغلب مترادف به کار رفته‌اند.

2 Stiggins

3 Diez

4 Walberg

5 Linn

عنصر کلیدی برای اصلاح مدارس و بهبود آموزش و یادگیری در نظر گرفته می‌شود (سیف، ۱۳۸۲؛ استیگینز^۲، ۲۰۰۴؛ دی از^۳، ۱۹۹۷؛ والبرگ^۴ و همکاران، ۱۹۹۴ و لین^۵، ۱۹۹۴). انتظار می‌رود ارزشیابی درست، مقدمات اصلاحات آموزشی را فراهم کند.

در آموزش و پرورش سنتی، ارزشیابی، آخرین گام است و برای تصمیم‌گیری درباره ارتقای دانش‌آموزان به پایه‌های تحصیلی بالاتر به کار می‌رود. اکنون ارزشیابی، بخش جدایی‌ناپذیر و همگام فرایند آموزش-یادگیری است که به جای طبقه‌بندی دانش‌آموزان، بر هدایت یادگیری آنها متمرکز است (شریفی، ۱۳۸۳).

تغییرات نظام ارزشیابی کلاسی برای بهبود وضعیت یادگیری دانش‌آموزان، پدیده‌ای است که امروزه در اکثر کشورهای جهان مشاهده می‌شود. این پدیده را می‌توان حداقل از دو ریشه مرتبط به هم تبیین کرد: یکی رویکرد‌های جدید روان‌شناسی یادگیری، نظیر شناخت‌گرایی و سازنده‌گرایی یا سازندگی^۱ و دیگری نهضت اصلاحات آموزشی و تغییرات برنامه درسی و مقایسه‌های بین‌المللی در موضوع‌های مختلف درسی و نیاز به افزایش استانداردهای آموزشی (سیف، ۱۳۸۲؛ شپرد^۲، ۲۰۰۰؛ اگن^۳ و کاوچک^۴، ۲۰۰۱؛ رازدوسک-پوکو^۵، ۱۹۹۸؛ بلک^۶ و ویلیام^۷، ۱۹۹۸).

در سالهای اخیر در کشور ما نیز همگام با سایر کشورها، اقداماتی در این زمینه صورت گرفته است. برای مثال، شورای عالی آموزش و پرورش در ششصد و هفتاد و نهمین جلسه خود در مورخه ۱۳۸۱/۸/۳۰ به معاونت آموزش عمومی و امور تربیتی وزارت آموزش و پرورش ماموریت داد تا نسبت به تغییر مقیاس کمی (۲۰-۰) به مقیاس کیفی (توصیفی) در ارزشیابی دانش‌آموزان اقدام کنند و آن را از سال تحصیلی ۸۲-۸۱ به صورت آزمایشی در تعدادی از مدارس دوره ابتدایی، که از نظر نیروی انسانی و امکانات، ویژگی‌های لازم را دارند، اجرا نمایند (مجموعه مصوبات شورای عالی آموزش و پرورش، ۱۳۸۲). در راستای این تغییر، گروه مطالعات دفتر ارزشیابی تحصیلی و تربیتی، طرح ارزشیابی توصیفی را تهیه کرد و از نیمه دوم بهمن ماه ۱۳۸۱ عملاً اجرای پیش‌آزمایشی طرح مقیاس کیفی را در ۲۰ کلاس پایه اول ابتدایی در ۱۰ مدرسه و با کمک ۳۰ تن از معلمان آغاز کرد. این طرح، بعد از اجرای پیش‌آزمایشی و همچنین برگزاری پودمان‌های

1. Constructivism
4. Kauchak
7. Wiliam

2. Shepard
5. Razdevsek-pucko
8. formative evaluation

3. Eggen
6. Black
9. performance evaluation

آموزشی برای معلمان مجری طرح، عملاً از مهرماه سال ۱۳۸۲ در تعداد محدودی از مدارس ابتدایی کشور اجرا شده است.

در دویست و نود و ششمین جلسه کمیسیون معین شورای عالی آموزش و پرورش در مورخه ۱۳۸۲/۶/۱۲ هدف‌هایی برای طرح آزمایشی ارزشیابی توصیفی پیشرفت تحصیلی دانش‌آموزان دوره ابتدایی تصویب شد که یکی از آنها اصلاح روند آموزش - یادگیری در کلاس درس بود. افزایش ماندگاری ذهنی (دوام و پایداری یادگیری)، توجه به هدف‌های سطوح بالاتر حوزه و یا حیطه شناختی، تعمیق یادگیری از طریق افزایش درگیری با تکالیف یادگیری، افزایش علاقه به یادگیری، توجه به هدف‌های حیطه‌های غیر شناختی و ... از جمله خرده هدفها بوده‌اند. راهکارهای مورد توجه در طرح ارزشیابی توصیفی برای رسیدن به هدفهای فوق، عبارت‌اند از: ارزشیابی تکوینی^۱، ارزشیابی عملکردی^۲، بازخورد و کارنامه توصیفی (حسنی و کاظمی، ۱۳۸۲).

اصطلاح ارزشیابی تکوینی را برای اولین بار اسکریون^۱ در مباحث مربوط به تنظیم برنامه‌های درسی به کار برده است. بلوم^۲، هستینگز^۳ و ماداوس^۴ (۱۳۵۵) نیز معتقد بودند که ارزشیابی تکوینی، نه تنها برای تنظیم برنامه درسی، بلکه برای تدریس و نیز یادگیری دانش‌آموزان مفید است. همچنان که توجه به دیدگاه‌های سازندگی تغییر می‌کند، سنجش کلاسی، به عنوان ابزاری برای بهبود فرایندهای یادگیری، به طور فزاینده‌ای مورد تاکید قرار می‌گیرد (استیگینز، ۲۰۰۵؛ بری^۵، ۲۰۰۵؛ ترز^۶، ۲۰۰۲؛ شپرد، ۲۰۰۰ و بلک و ویلیام، ۱۹۹۸). سیف (۱۳۸۳) معتقد است که هدف از کاربرد سنجش تکوینی در رابطه با پیشرفت تحصیلی دانش‌آموزان، آگاهی یافتن از میزان و نحوه یادگیری آنان برای تعیین نقاط قوت و ضعف یادگیری و نیز تشخیص مشکلات روش آموزشی معلم در رابطه با هدف‌های آموزشی است. شواهد پژوهشی نیز نشان می‌دهد که سنجش تکوینی با کیفیت بالا می‌تواند تاثیر قدرتمندی بر عملکرد تحصیلی و انگیزش دانش‌آموزان داشته باشد (برای مثال، وان اورا^۷، ۲۰۰۴؛ بلک و ویلیام، ۱۹۹۸؛ وادل^۸، ۲۰۰۴ و حیدری، ۱۳۷۵). نتایج پژوهش وان اورا (۲۰۰۴) نشان داد که بازخورد حاصل از سنجش تکوینی افزایش معنی‌داری را در

1. Scriven
5. Berry

2. Bloom
6. Taras

3. Hastings
7. Van Evera

4. Madaus
8. Waddell

کارآمدی شخصی دانش‌آموزان دوره راهنمایی به وجود می‌آورد. بلک و ویلیام (۱۹۹۸) نیز پس از مرور ۲۵۰ مقاله در زمینه سنجش کلاسی، اثرات مثبت سنجش تکوینی را بر عملکرد تحصیلی (در تمام سطوح سنی و توانایی) گزارش کرده‌اند. آن‌ها گزارش کرده‌اند که مطالعات سنجش تکوینی، به اندازه ۰/۴ تا ۰/۷ در آزمونهای استاندارد شده اثر دارد. این اندازه اثر، بیش‌تر از میزانی است که برای مداخلات آموزشی به دست آمده است. آن‌ها همچنین به این نتیجه رسیده‌اند که سنجش تکوینی، استانداردهای پیشرفت را ارتقاء می‌بخشد و این امر به گونه‌ای خاص برای دانش‌آموزانی که پیشرفت کمی دارند (کم‌اندوزها) موثرتر است. بنابراین از این طریق می‌توانیم شکاف بین دانش‌آموزان کم‌اندوز و بیش‌اندوز را بکاهیم و در عین حال، پیشرفت کلی را افزایش بدهیم.

بلک و ویلیام (۱۹۹۸) معتقدند که یکی از عوامل اساسی در ارتقاء آموزش از طریق سنجش کلاسی، بازخورد^۲ موثر به دانش‌آموزان است. بازخورد به عملکرد یادگیرنده، برای او اطلاعاتی فراهم می‌کند که این اطلاعات منجر به اصلاح شخصی و بهبود یادگیری می‌شود. شواهد پژوهشی، تاثیر بازخورد اسنادی بر کارآمدی شخصی (مک‌گی^۳، ۱۹۹۶) و همچنین تأثیر بازخورد شایستگی بر افزایش علاقه و عملکرد (حویزای، ۱۳۷۴) را نشان داده‌اند. در واقع، تاثیر بازخورد، چنان قوی است که اگن و کاوچک (۲۰۰۱) آن را یک اصل اساسی در یادگیری و انگیزش نامیده‌اند. در رویکرد ارزشیابی توصیفی، معلمان، پیشرفت دانش‌آموزان را با بیانات توصیفی، به صورت کلامی و گاهی اوقات به صورت نوشتاری، مورد سنجش قرار می‌دهند (حسینی، ۱۳۸۴). تعامل بین معلم و دانش‌آموزان در حین ارائه بازخورد، می‌تواند در رشد و توسعه یادگیری خودنظم‌دهی موثر باشد. اگر عملکرد دانش‌آموزان پایین‌تر از معیارها باشد، معلم آن‌ها را راهنمایی می‌کند تا بیش‌تر تلاش کنند. او در این حال می‌تواند راهبردهای یادگیری متفاوتی را پیشنهاد دهد. اگر عملکرد دانش‌آموزان با معیارها هماهنگ و یا از آن بالاتر باشد، احساس موفقیت خواهند داشت و کارآمدی شخصی شان افزایش خواهد یافت؛ که این امر نیز به نوبه خود بر نگرش آنان نسبت به آموزشگاه محل تحصیل، تاثیر مثبتی خواهد گذاشت.

1. low achiever

2. feedback

3. McGee

مرحله تکمیلی سنجش تکوینی، مستلزم این است که هم در نقش معلمان و هم در نقش دانش‌آموزان تغییراتی ایجاد شود (تری سا، ۲۰۰۴). از این رو در سنجش تکوینی، خودسنجی و سنجش همسالان از اهمیت ویژه‌ای برخوردار است. با این حال، این شیوه‌های سنجشی، در کلاس‌های درس متداول نیست و موضوعی است که غالباً در ادبیات سنجش کلاسی نیز مورد غفلت قرار گرفته است (بلک و ویلیام، ۱۹۹۸). هدف اساسی در این دو روش، ایجاد احساس مسئولیت بیش‌تر نسبت به یادگیری خود و همسالان است (حسینی و احمدی، ۱۳۸۴). مطالعات پژوهشی نیز اثرات مثبت سنجش شخصی و سنجش همسالان را بر عملکرد تحصیلی دانش‌آموزان در گروه‌های سنی و دروس مختلف نشان داده است (فونتانا و فرناندز، ۱۹۹۴؛ یونگ، ۲۰۰۰). بری (۲۰۰۵) و بلک و همکاران (۲۰۰۴) اعتقاد دارند، دانش‌آموزان با این دو روش سنجشی، یاد می‌گیرند که چگونه بر یادگیری خود و دیگران نظارت کنند و آن را کنترل و مدیریت نمایند. از این رو می‌توان انتظار داشت که دانش‌آموزان تحت پوشش طرح ارزشیابی توصیفی، با به کارگیری این روش‌ها، علاوه بر پیشرفت تحصیلی، از دانش‌فراشناختی بالاتری نیز برخوردار شوند.

سنجش عملکردی، شیوه دیگری از سنجش کلاسی است که متخصصان ارزشیابی پیشرفت تحصیلی بر آن تاکید می‌کنند. در سنجش عملکردی، فرایندها و فرآورده‌های یادگیری دانش‌آموزان، به طور مستقیم سنجیده می‌شود (سیف، ۱۳۸۴). در این نوع سنجش، دانش‌آموزان با استفاده از دانش قبلی و فعلی خود و بهره‌گیری از مهارت‌های مناسب برای حل مسائل واقعی، به صورت فعال، با مسئله، آزمون یا سؤال داده شده برخورد می‌کنند و پاسخ لازم را می‌یابند (کیامنش، ۱۳۷۷). بنابراین این نوع تکالیف، تنها برای سنجش نیستند؛ بلکه به صورت عملی و واقعی طراحی شده‌اند (در بافتی که برای دانش‌آموزان آشناست) و از این رو دانش‌آموزان را به چالش برمی‌انگیزد.

1. Teresa

2. Fontana

3. Fernandez

4. Young

فیوکر^۱ و همکارانش (۱۹۹۹) دریافتند، کلاس‌های آموزشی که با سنجش عملکردی هدایت می‌شوند، در مقایسه با گروه‌های که با سنجش عملکردی هدایت نمی‌شوند، مهارت‌های بیشتری در حل مسئله دارند. علاوه بر این، تأثیری که سنجش عملکردی بر علاقه و نگرش دانش‌آموزان نسبت به مدرسه و فعالیت‌های آموزشی دارد، قابل تأمل است.

یکی از فعالیت‌های تکمیلی که در رویکرد جدید سنجش عملکرد دانش‌آموزان و در جریان فرایند آموزش - یادگیری مورد استفاده قرار می‌گیرد، کارپوشه^۲ است. بنا به تعریف ولفلک (۲۰۰۴) کارپوشه، مجموعه‌ای از کارهای دانش‌آموز است که رشد، خود اندیشی^۳ و پیشرفت او را در یک زمینه خاص نشان می‌دهد. پژوهش‌ها حکایت از آن دارد که کارپوشه‌ها پیشرفت دانش‌آموزان را بهتر و واقعی‌تر نشان می‌دهند (گریس^۴، ۱۹۹۲؛ ادیجر^۵، ۲۰۰۰). هنگامی که دانش‌آموزان به طور فعال در رشد ابزارهای سنجشی خود درگیر می‌شوند، شروع به انتخاب هدف می‌کنند و پیشرفت خود را در رسیدن به هدف‌ها و ارسی می‌کنند و یاد می‌گیرند که توانایی‌های خود را ارزیابی نمایند (پالسون^۶ و پالسون، ۱۹۹۴؛ تیرنی^۷، ۱۹۹۲). استفاده از کارپوشه، به عنوان یک روش آموزش و سنجش، نه تنها به رشد مهارت‌های سنجش شخصی و پیشرفت تحصیلی کمک می‌کند، بلکه راهبردهای فراشناختی را نیز گسترش می‌دهد (سیف، ۱۳۸۴؛ کلناوسکی^۸، ۲۰۰۲؛ کتول^۹، ری^{۱۰} و شریل^{۱۱}، ۱۹۹۹). اهمیت روش کارپوشه برای رشد و پرورش مهارت‌های فراشناختی در این است که یادگیرندگان با اطلاع از آموخته‌ها، نحوه تفکر و فرایند یادگیری خود و نیز چگونگی به کارگیری دانش و مهارت در حل مسائل، می‌توانند یادگیری و تفکر خود را جهت دهند و هدایت کنند و از طریق خودنظم‌دهی و نظارت بر فرایندهای ذهنی خویش، پیشرفتشان را بهبود بخشند.

در اجرای طرح جدید ارزشیابی توصیفی در مدارس مشمول طرح، تلاش می‌شود تا با تأکید بر ارزشیابی تکوینی، ارزشیابی عملکردی و همچنین دادن بازخورد توصیفی، شرایط بهتری برای یادگیری فراهم شود. بنابراین، انتظار می‌رود که این نظام ارزشیابی در عملکرد تحصیلی، باورها و نگرش‌های دانش‌آموزان تأثیر مثبتی داشته باشد. هدف پژوهش حاضر بررسی چگونگی تحقق این انتظار بود. سؤال اصلی پژوهش این بود: "اجرای ارزشیابی

1. Fuchs
5. Ediger
9. Katula

2. portfolio
6. Paulson
10. Ray

3. self-reflection
7. Tierney
11 Sherrill

4. Grace
8. Klenowski.

توصیفی در ایران، تا چه اندازه توانسته است در دستیابی دانش‌آموزان دوره ابتدایی به هدف های آموزشی، در حیطه‌های شناختی (پیشرفت تحصیلی ریاضی، علوم تجربی و زبان فارسی و همچنین میزان آگاهی فراشناختی)، عاطفی (نگرش دانش‌آموزان نسبت به خودشان، دانش‌آموزان دیگر، معلمان، مدرسه و فعالیت‌های مدرسه) و روانی- حرکتی (پیشرفت عملکرد ریاضی و علوم تجربی) موثر باشد؟"

روش

آزمودنی‌ها یا شرکت‌کنندگان

در این پژوهش، تعداد کل افراد نمونه ۱۹۲ دانش‌آموز پایه سوم ابتدایی بود. ۹۶ دانش‌آموز (۴۸ دختر و ۴۸ پسر) در گروه آزمایش و ۹۶ دانش‌آموز (۴۸ دختر و ۴۸ پسر) در گروه گواه شرکت داشتند میانگین سنی گروه نمونه ۹ سال و ۶ ماه و با انحراف معیار ۰/۵۸ بود. برای انتخاب گروه نمونه، ابتدا کارشناسان مقطع ابتدایی اداره های آموزش و پرورش مناطق ۱، ۵، ۹ و ۱۶ مدارس عادی هم سطح با مدارس مشمول طرح را با در نظر گرفتن شرایط آنها از نظر درون داد (فضا، تجهیزات، نیروی انسانی و ویژگیهای فردی و خانوادگی دانش‌آموزان) شناسایی کردند و سپس از بین آنها، به طور تصادفی، یک مدرسه را برگزیدند. در مدارس تحت پوشش طرح ارزشیابی توصیفی و همچنین مدارس عادی، از میان دانش‌آموزان پایه سوم ابتدایی، ۲۴ نفر به صورت تصادفی انتخاب شدند. علاوه بر این، در هر مدرسه ۲ نفر دانش‌آموز نیز به صورت تصادفی و به عنوان افراد ذخیره در نظر گرفته شدند، تا در صورت غیبت دانش‌آموزان انتخاب شده اصلی، به آزمونهای مربوط به آنها پاسخ دهند.

ابزارهای پژوهش

در این پژوهش برای جمع آوری اطلاعات مورد نیاز از ابزارهای زیر استفاده شد:

- (۱) آزمون پیشرفت تحصیلی ریاضی، (۲) آزمون پیشرفت تحصیلی علوم تجربی، (۳) آزمون پیشرفت تحصیلی زبان فارسی، (۴) آزمونهای عملکردی علوم تجربی و ریاضی، (۵) پرسشنامه آگاهی فراشناختی و (۶) پرسشنامه کیفیت زندگی در مدرسه. در ادامه، اطلاعات لازم درباره هر یک از این آزمونها و پرسشنامه‌ها ارائه شده است.

آزمون پیشرفت تحصیلی ریاضی

در این پژوهش برای سنجش پیشرفت تحصیلی ریاضی دانش‌آموزان پایه سوم ابتدایی، از آزمون محقق ساخته استفاده شد. برای تدوین سؤالهای این آزمون، ابتدا جدول مشخصات^۱ سنجش، به منظور تعیین هدفها و محتوای کتاب درسی پایه سوم ابتدایی تهیه شد. سپس با استفاده از اطلاعات جدول مشخصات سنجش و منابع موجود، ۱۰۵ سؤال برای آزمون مقدماتی طراحی شد. برای اندازه‌گیری دامنه گسترده‌تری از توانایی دانش‌آموزان، با توجه به محدودیت زمانی، از انواع مختلف سؤالات استفاده شد. این سؤالات با استفاده از نظر کارشناسی معلمان پایه سوم ابتدایی، استادان متخصص روانشناسی تربیتی و روان سنجی، مجدداً بررسی و ویرایش شدند. پس از بررسیهای اولیه، دانش‌آموزان پایه سوم ابتدایی در دبستانهای کوثر و شهید بهشتی در منطقه ۶ به سؤالات تدوین شده پاسخ دادند. در این مرحله، سؤالهای مبهم شناسایی شد. همچنین پاسخ‌های داده شده، پس از جمع‌آوری و تصحیح، سؤال به سؤال مورد تجزیه و تحلیل قرار گرفت و با توجه به سطح دشواری و ضریب تمیز، سؤالهای مناسب انتخاب شدند. سپس به کمک معلمان پایه سوم ابتدایی، از بین سؤالهای مناسب، ۲۹ سؤال چند گزینه‌ای، ۴ سؤال کوتاه پاسخ و ۲ سؤال محدود پاسخ، و در مجموع، ۳۵ سؤال با توجه به جدول مشخصات، انتخاب و دفترچه آزمون تدوین شد.

روایی محتوایی: با توجه به تهیه جدول مشخصات و تدوین سؤالهای آزمون پیشرفت تحصیلی ریاضی بر اساس هدف‌های آموزشی و همچنین محتوای برنامه این درس، می‌توان گفت که این آزمون از روایی محتوایی^۲ لازم برخوردار است.

پایایی: در این پژوهش، پایایی^۳ سؤالات آزمون پیشرفت ریاضی با استفاده از روش آلفای کرانباخ، ۰/۸۸ به دست آمد.

1. table of specifications

2. content-related validity

3. reliability

آزمون پیشرفت تحصیلی علوم تجربی

در این پژوهش برای سنجش پیشرفت تحصیلی علوم تجربی دانش‌آموزان، از آزمون محقق ساخته استفاده شد. سؤالهای آزمون علوم نیز هم زمان و همانند سؤالهای آزمون ریاضی، با استفاده از جدول مشخصات سنجش، تدوین شد و بعد از اجرای آزمایشی و تحلیل نتایج به کار گرفته شد. آزمون پیشرفت تحصیلی علوم تجربی از ۲۹ سؤال چند گزینه‌ای، ۶ سؤال کوتاه پاسخ، و در مجموع، ۳۵ سؤال تشکیل شده است.

روایی محتوایی: محتوای آزمون پیشرفت تحصیلی علوم تجربی، همانند آزمون ریاضی، بر اساس خانه‌های جدول مشخصات سنجش (هدفهای آموزشی و محتوای برنامه این درس) تدوین گشت. بر این اساس می‌توان گفت که این آزمون نیز از روایی محتوایی لازم برخوردار است.

پایایی: در این پژوهش ضریب پایایی آزمون پیشرفت علوم تجربی با استفاده از روش آلفای کرانباخ، ۰/۸۱ به دست آمد.

آزمون عملکردی علوم تجربی و ریاضی

آزمونهای سنجش عملکردی، با بهره‌گیری از پیشینه نظری، نمونه‌های موجود و تجارب معلمان پایه سوم ابتدایی تهیه شد. ابتدا ۱۰ آزمون به صورت مقدماتی در یک نمونه کوچک ($N = 20$) مورد آزمایش قرار گرفت. پس از اجرا و تجزیه و تحلیل اطلاعات، تغییراتی در برخی از این آزمونها اعمال شد. برخی از سؤالهای مبهم حذف شد و در پاره‌ای از موارد، سؤالهای جدید جایگزین گشت. با توجه به معیارهایی از قبیل قابلیت اجرایی، سطح دشواری و مطابقت با جدول مشخصات (با توجه به بعد محتوا و اهداف آموزشی)، آزمونهای مناسب برای اجرای نهایی مشخص شدند. برای ایجاد تعادل بین آزمونهای ریاضی و علوم، یک آزمون در حوزه ریاضی (آزمون درخت و برگ)، یک آزمون در حوزه علوم (آزمون دماسنج) و یک آزمون ترکیبی با محتوای علوم و ریاضی (آزمون اشیای دور ریختنی) انتخاب شد.

پایایی مصححان: سؤالهای آزمون عملکردی اشیای دور ریختنی و درخت و برگ، علاوه بر تصحیح اولیه، توسط یکی از معلمان پایه سوم ابتدایی نیز به طور مستقل و بر

اساس طرح نمره‌گذاری تصحیح شد. ضریب همبستگی بین نمرات مصححان برای آزمون اشیای دور ریختنی و آزمون درخت و برگ به ترتیب ۰/۹۸ و ۰/۹۳ بود. برای افزایش پایایی، موارد اختلافی پس از بحث مصححان برطرف شد. در آزمون دماسنج نیز عملکرد ۴۸ دانش‌آموز به طور مستقل توسط دو مشاهده‌گر نمره‌گذاری شد. ضریب توافق مشاهده‌گران ۰/۹۸ به دست آمد.

آزمون پیشرفت تحصیلی زبان فارسی

در این پژوهش برای اندازه‌گیری پیشرفت تحصیلی زبان فارسی دانش‌آموزان پایه سوم ابتدایی، از آزمون استاندارد شده شریفی (۱۳۸۱) استفاده شد. این آزمون شامل خرده‌آزمون‌های درک شنیداری، خواندن و فهمیدن، کاربرد کلمه در جمله، درک معنای کلمه‌ها و تشخیص کلمه‌های متضاد است. در زیر ویژگی‌های هر یک از این خرده‌آزمون‌ها ارائه شده است.

درک شنیداری: این بخش شامل یک متن فارسی همراه با یک آزمون ۷ سؤالی است که مجری پس از خواندن متن، آزمون را در اختیار دانش‌آموزان قرار می‌دهد تا دانش‌آموزان به سؤالات چند گزینه‌ای آن پاسخ دهند.

خواندن و فهمیدن: این بخش شامل یک متن فارسی است که دانش‌آموز باید آن را بخواند و پاسخ درست پرسش‌های آن را از بین پاسخهای پیشنهادی انتخاب کند.

کاربرد کلمه در جمله: این آزمون شامل تعدادی جمله است که در هر جمله، یک کلمه حذف شده است. آزمودنی باید با انتخاب کلمه مناسب از بین گزینه‌های داده شده، معنای جمله را کامل کند.

درک معنای کلمه‌ها: این آزمون شامل تعدادی سؤال به صورت کلمه است که آزمودنی باید کلمه هم معنی را از بین کلمه‌های پیشنهادی انتخاب کند.

تشخیص کلمه‌های متضاد: در این آزمون، تعدادی کلمه نوشته شده و در زیر هر کلمه، چهار کلمه پیشنهادی به عنوان مخالف آن آمده است. آزمودنی باید کلمه‌ای را که معنای مخالف دارد، انتخاب کند.

روایی آزمون: محتوای آزمون پیشرفت تحصیلی زبان فارسی بر اساس اهداف آموزشی، با در نظر گرفتن محتوای برنامه این درس و با نظر معلمان و متخصصان تدوین شده است. بر این اساس می‌توان گفت که این آزمون از روایی محتوایی لازم برخوردار

است (شریفی، ۱۳۸۱). در پژوهش حاضر نیز برای برآورد روایی محتوایی، سؤالات آزمون در اختیار معلمان پایه سوم ابتدایی قرار گرفت تا آنها را با اهداف آموزشی این درس مطابقت دهند. با استفاده از نظر کارشناسی معلمان، تعدادی از سؤالات خرده آزمونهای کاربرد کلمه در جمله، درک معنای کلمه‌ها و تشخیص کلمه‌های متضاد، ویرایش و اصلاح شدند. سپس فرم جدید آزمون روی یک نمونه ۵۷ نفری از دانش‌آموزان دختر و پسر پایه سوم ابتدایی مدارس کوثر و شهید بهشتی در منطقه ۶ اجرا شد. در این مرحله سؤالات مبهم و همچنین سؤالاتی که فاقد مشخصات لازم روان‌سنجی بودند حذف و با نظر معلمان پایه سوم ابتدایی سؤالات جدیدی جایگزین آنها شد.

پایایی آزمون: شریفی (۱۳۸۱) پایایی آزمون پیشرفت تحصیلی زبان فارسی پایه سوم را با استفاده از روش دو نیمه کردن، آلفای کراباخ و گاتمن، به ترتیب ۰/۸۹، ۰/۸۸ و ۰/۸۸ گزارش کرده است. در پژوهش حاضر نیز ضریب آلفای ۰/۸۷ برای کل آزمون به دست آمد.

پرسشنامه آگاهی فراشناختی

پاناورا^۱ و فیلیپو^۲ (۲۰۰۵) این پرسشنامه را برای سنجش آگاهی فراشناختی دانش‌آموزان پایه‌های سوم، چهارم و پنجم ابتدایی تدوین کرده‌اند. پرسشنامه آگاهی فراشناختی از ۳۰ سؤال تشکیل شده است که رفتار دانش‌آموزان را در فعالیت‌های حل مسئله منعکس می‌کند. برای مثال: "وقتی که به مشکلی در حل مسئله برخورد می‌کنم، آن مسئله را دوباره می‌خوانم." دانش‌آموزان با انتخاب یکی از پنج گزینه هرگز (۱)، خیلی کم (۲)، بعضی وقتها (۳)، بیش تر وقتها (۴) و همیشه (۵) به سؤالات پاسخ می‌دهند. پاناورا و فیلیپو (۲۰۰۵) پایایی پرسشنامه را ۰/۷۸ گزارش کرده‌اند.

پرسشنامه آگاهی فراشناختی پس از ترجمه و مقابله توسط استادان زبان انگلیسی، در یک مطالعه مقدماتی ($n = 29$) آزمایش شد. اجرای مقدماتی نشان داد که دانش‌آموزان بدون هیچ مشکلی در خواندن، توانستند به خوبی از عهده آن برآیند. در این پژوهش همچنین پرسشنامه دانش فراشناختی سوانسون^۳ و همکاران (۱۹۹۰)، ترجمه آوانسیان، (۱۳۷۷) به منظور بررسی روایی ملاکی، روی دانش‌آموزان اجرا شد. ضریب همبستگی بین

1. Panaoura

2. Philippou

3. Swanson

نمرات دو پرسشنامه برابر با ۰/۵۳۳ بود. این ضریب نشان می‌دهد که پرسشنامه آگاهی فراشناختی از روایی ملاکی لازم برخوردار است.

پایایی پرسشنامه: در این پژوهش، ضریب پایایی کل مجموعه ۳۰ سؤالی با استفاده از ضریب آلفای کرانباخ برابر ۰/۸۸ به دست آمد. حذف سؤالات ۲، ۱۳، ۲۷ و ۳۰ که بار فرهنگی داشتند و همبستگی آنها با نمره کل ضعیف بود، باعث شد که ضریب پایایی آزمون به ۰/۸۹ افزایش یابد. از این رو با حذف چهار پرسش فوق، تحلیل‌های آماری بر اساس ارزشهای ۲۶ پرسش باقی مانده انجام گرفت.

پرسشنامه کیفیت زندگی در مدرسه

پرسشنامه کیفیت زندگی در مدرسه^۱ توسط اینلی^۲ و بورک^۳ در سال ۱۹۹۲ و برای سنجش ویژگیهای عاطفی دانش‌آموزان دبستانی تهیه شده است (سیف، ۱۳۸۴). این پرسشنامه از ۳۹ ماده (سؤال) تشکیل شده و دربرگیرنده هفت مقیاس است که دو تا از آنها کلی (رضایت کلی و عاطفه منفی) و پنج تای دیگر اختصاصی‌اند (همبستگی اجتماعی، فرصت، معلم، ماجرا و پیشرفت یا موفقیت). یک نمونه سؤال برای مقیاس رضایت کلی عبارت است از: "من در مدرسه احساس شادمانی می‌کنم؟" سؤالیهای این پرسشنامه دارای چهار گزینه یا چهار طبقه پاسخ کاملاً موافق، موافق، مخالف و کاملاً مخالف است که به ترتیب ۴، ۳، ۲ و ۱ نمره به آنها تعلق می‌گیرد. در پژوهشهای گذشته، ضریب آلفای کرانباخ (پایایی آزمون با تاکید بر همسانی درونی) برای خرده مقیاسهای این پرسشنامه در دامنه ای از ۰/۶۷ تا ۰/۸۹ گزارش شده است (ماک^۴ و فلین^۵، ۲۰۰۲). در این پژوهش، ضریب پایایی آلفای کرانباخ برای کل پرسشنامه (۳۹ سؤال) برابر ۰/۹۰ به دست آمد. و ضریب پایایی آلفا برای خرده مقیاسهای پرسشنامه، از ۰/۶۸ تا ۰/۸۵ بود.

روش اجرا

این پژوهش در سه مرحله اجرا شد. در مرحله اول، مدارس هم سطح با مدارس مشمول طرح ارزشیابی توصیفی در مناطق ۱، ۵، ۹ و ۱۶ به کمک کارشناسان مقطع ابتدایی مناطق شناسائی شدند و مجوز لازم برای مراجعه به آنها دریافت شد. در مرحله دوم، آزمونها و پرسشنامه‌ها به صورت مقدماتی در مدارس کوثر و شهید بهشتی، واقع در

1. the Quality of School Life Questionnaire (QSL)

2. Ainley

3. Bourke

4. Mok

5. Flynn

منطقه ۶، اجرا شد و پس از رفع اشکالات، برای اجرای اصلی آماده گشت. در مرحله سوم نیز پس از هماهنگی با مدیران و معلمان مدارس گروههای آزمایش و گواه، در نیمه دوم اردیبهشت ماه سال ۱۳۸۵، آزمونها و پرسشنامهها بر اساس دستورالعملهای مربوطه و در شرایط یکسان، روی دانش آموزان پایه سوم ابتدایی اجرا شد.

نتایج

در جدول ۱، میانگین و انحراف معیار متغیرهای شناختی، عاطفی و روانی - حرکتی دانش آموزان مشمول طرح ارزشیابی توصیفی (گروه آزمایش) و دانش آموزان مدارس عادی (گروه گواه) ارائه شده است.

جدول ۱. شاخص های توصیفی متغیرهای شناختی، عاطفی و روانی - حرکتی در دو گروه آزمایش و گواه

گواه		آزمایش		شاخص ها	
Sd	M	Sd	M	آزمون و مقیاس ها	
۸/۰۰	۲۲/۸۸	۴/۹۱	۲۸/۵۵	پیشرفت تحصیلی ریاضی	شناختی
۶/۲۲	۲۵/۵۵	۴/۳۸	۲۷/۴۸	پیشرفت تحصیلی علوم تجربی	
۸/۲۱	۳۰/۳۱	۵/۴۶	۳۵/۰۸	پیشرفت تحصیلی زبان فارسی	
۱۵/۱۰	۱۰۴/۲۹	۱۳/۲۶	۱۱۰/۹۰	آگاهی فراشناختی	
۳/۸۵	۱۹/۹۷	۴/۴۲	۲۰/۵۵	رضایت کلی از مدرسه	عاطفی
۳/۷۴	۹/۸۳	۳/۷۰	۸/۷۶	عاطفه منفی نسبت به مدرسه	
۲/۸۹	۱۷/۷۱	۳/۰۵	۱۷/۵۸	نگرش در مورد معلمان	
۴/۶۸	۲۵/۴۷	۴/۴۸	۲۶/۱۵	همبستگی اجتماعی	
۳/۴۸	۲۱/۴۲	۲/۹۱	۲۲/۱۸	فرصت (باور نسبت به مفید بودن مدرسه)	
۲/۵۶	۱۳/۷۴	۱/۹۵	۱۴/۴۹	موفقیت یا پیشرفت	
۲/۸۹	۱۶/۷۵	۲/۸۸	۱۷/۳۱	ماجرا (احساس خودانگیزگی در یادگیری)	
۱/۷۱	۲/۱۵	۱/۵۴	۲/۷۹	آزمون درخت و برگ	روانی-حرکتی
۲/۷۳	۶/۰۷	۲/۰۹	۶/۶۸	آزمون اشیای دور ریختنی	
۰/۹۴	۱/۴۸	۱/۰۳	۱/۶۱	آزمون دماسنج	

بر اساس شاخصهای فوق می‌توان گروههای آزمایش و گواه را به صورت ظاهری مقایسه کرد؛ ولی برای نشان دادن این که آیا این تفاوت‌ها از نظر آماری معنی‌دار است یا نه، از روشهای مناسب آمار استنباطی، نظیر تحلیل واریانس چند متغیری (MANOVA) و تحلیل واریانس یک سویه (ANOVA) استفاده شد.

برای تعیین تفاوت بین گروه آزمایش (دانش‌آموزان مشمول طرح ارزشیابی توصیفی) و گروه گواه (دانش‌آموزان مدارس عادی هم سطح) در متغیرهای حوزه شناختی، تحلیل واریانس چند متغیری (MANOVA) انجام شد و نمرات آزمونهای پیشرفت تحصیلی ریاضی، علوم تجربی و زبان فارسی، به عنوان متغیرهای وابسته، وارد تحلیل شدند. نتایج تحلیل واریانس چند متغیری نشان داد که بین گروههای آزمایش و گواه در متغیرهای حوزه شناختی، تفاوت معنی‌داری وجود دارد ($P < 0/001$ و $F_{(3,188)} = 15/445$ و $\lambda = 0/802$) لامبدای ویلکز).

تحلیل واریانس یک سویه (ANOVA) برای هر متغیر وابسته، به عنوان آزمون تکمیلی تحلیل واریانس چند متغیری، انجام شد. جدول ۲ نتایج آزمونهای تک متغیری تحلیل واریانس را در متغیرهای شناختی نشان می‌دهد.

جدول ۲. نتایج حاصل از آزمونهای F تک متغیری بر اساس گروههای مورد مطالعه در متغیرهای شناختی

سطح معنی‌داری	F	MS	d.f	SS	شاخص‌های آماری	
					منابع تغییرات	
0/000	35/116	1547/005	1	1547/005	پیشرفت تحصیلی ریاضی	گروه
0/000	22/495	1092/521	1	1092/521	پیشرفت تحصیلی زبان فارسی	
0/014	6/160	178/255	1	178/255	پیشرفت تحصیلی علوم تجربی	
		44/054	190	8370/240	پیشرفت تحصیلی ریاضی	گروه
		48/568	190	9227/958	پیشرفت تحصیلی زبان فارسی	
		28/935	190	5497/698	پیشرفت تحصیلی علوم تجربی	
			191	9917/245	پیشرفت تحصیلی ریاضی	گروه
			191	10320/479	پیشرفت تحصیلی زبان فارسی	
			191	5675/953	پیشرفت تحصیلی علوم تجربی	

1 Wilks' Lambda

2. Bonferonroni

برای کنترل خطای نوع اول هر تحلیل، از روش بن فرونی^۱ در سطح ۰/۰۱۷ استفاده شد ($0/017 \div 3 = 0/005$). همان طور که در جدول ۲ مشاهده می‌شود، تفاوت‌های معنی‌داری بین گروههای آزمایش و گواه در پیشرفت تحصیلی ریاضی ($P < 0/001$) و $(F_{(1,190)} = 35/116)$ ، علوم تجربی ($P = 0/014$ و $F_{(1,190)} = 6/160$) و زبان فارسی ($P < 0/001$ و $F_{(1,190)} = 22/495$) وجود دارد. مقایسه میانگین نمرات دانش‌آموزان گروههای آزمایش و گواه (جدول ۱) نشان می‌دهد که سطح پیشرفت تحصیلی دانش‌آموزان مشمول طرح ارزشیابی توصیفی در دروس ریاضی، علوم تجربی، زبان فارسی، بالاتر از دانش‌آموزان پایه سوم ابتدایی مدارس عادی هم سطح آنهاست.

برای بررسی تفاوت بین گروههای آزمایش و گواه در میزان آگاهی فراشناختی، از تحلیل واریانس یک سویه استفاده شد. جدول ۳ نتایج تحلیل واریانس یک سویه نمرات آگاهی فراشناختی دانش‌آموزان گروههای آزمایش و گواه را نشان می‌دهد.

جدول ۳. نتایج تحلیل واریانس یک سویه در متغیر آگاهی فراشناختی با توجه به گروههای مورد مقایسه

منابع تغییرات / شاخصهای آماری	SS	d.f	MS	F	سطح معنی‌داری
بین‌گروهی	۲۰۹۳/۵۲۱	۱	۲۰۹۳/۵۲۱	۱۰/۳۷۱	۰/۰۰۲
درون‌گروهی	۳۸۳۵۴/۷۹۲	۱۹۰	۲۰۱/۸۶۷		
کل	۴۰۴۴۸/۳۱۲	۱۹۱			

همان طوری که در جدول ۳ مشاهده می‌شود، در میزان آگاهی فراشناختی تفاوت معنی‌داری بین دانش‌آموزان گروههای آزمایش ($P = 0/002$ و $F_{(1,190)} = 10/371$)، تفاوت معنی‌داری بین دانش‌آموزان گروههای آزمایش (دانش‌آموزان مشمول طرح ارزشیابی توصیفی) و گواه (دانش‌آموزان مدارس عادی) وجود دارد. مقایسه میانگین نمرات دانش‌آموزان گروههای آزمایش و گواه (جدول ۱) نشان می‌دهد که میزان آگاهی فراشناختی دانش‌آموزان تحت پوشش طرح ارزشیابی توصیفی، بالاتر از دانش‌آموزان مدارس عادی است.

برای تعیین تفاوت بین گروه آزمایش (دانش‌آموزان مشمول طرح ارزشیابی توصیفی) و گروه گواه، در متغیرهای حوزه عاطفی، تحلیل واریانس چند متغیری انجام شد و نمرات مقیاسهای رضایت کلی از مدرسه، عاطفه منفی نسبت به مدرسه، نگرش در مورد معلمان، همبستگی اجتماعی، فرصت (باور نسبت به مفید بودن مدرسه)، موفقیت (احساس اطمینان به توانایی برای کسب توفیق در انجام کارهای مدرسه) و ماجرا (احساس خودانگیزگی و لذت بردن از یادگیری) به عنوان متغیرهای وابسته وارد تحلیل شدند. نتایج تحلیل واریانس چند متغیری نشان داد که بین گروههای آزمایش و گواه، در متغیرهای حوزه عاطفی، تفاوت معنی‌داری وجود ندارد ($P = 0/061$ و $F_{(7,184)} = 1/971$ و $\lambda = 0/93$).

نتایج تحلیل واریانس چند متغیری نشان داد که بین گروههای آزمایش و گواه، در متغیرهای حوزه روانی- حرکتی، تفاوت معنی‌داری وجود دارد ($P = 0/028$ و $\lambda = 0/96$ و $F_{(3,188)} = 0/953$). تحلیل واریانس یک سویه بر هر متغیر وابسته، به عنوان آزمون تکمیلی تحلیل واریانس چند متغیری انجام شد. جدول ۴، نتایج آزمونهای تک متغیری تحلیل واریانس را در متغیرهای حوزه روانی- حرکتی نشان می‌دهد.

جدول ۴. نتایج حاصل از آزمونهای F تک متغیری بر اساس گروههای مورد مطالعه در متغیرهای حوزه روانی- حرکتی

سطح معنی‌داری	F	MS	d.f	SS	شاخصهای آماری	
					منابع تغییرات	
0/007	7/551	20/021	1	20/021	آزمون درخت و برگ	تکرار
0/087	2/958	17/521	1	17/521	آزمون اشیای دور ریختنی	
0/343	0/905	0/880	1	0/880	آزمون دماسنج	
		2/652	190	503/792	آزمون درخت و برگ	گروه
		5/924	190	1125/479	آزمون اشیای دور ریختنی	
		0/972	190	184/698	آزمون دماسنج	
			191	523/813	آزمون درخت و برگ	کلاس
			191	1143/000	آزمون اشیای دور ریختنی	
			191	185/578	آزمون دماسنج	

نتایج آزمونهای تک متغیری F (جدول ۴) در آزمون درخت و برگ ($P = ۰/۰۰۷$) و $F_{(۱,۱۹۰)} = ۷/۵۵۱$ با آلفای میزان شده بن فرونی در سطح $۰/۰۱۷$ ، تفاوت‌های معنی‌داری را بین گروههای آزمایش و گواه نشان داد؛ در حالی که تفاوت بین گروههای آزمایش و گواه در آزمون اشیای دور ریختنی ($P = ۰/۰۸۷$ و $F_{(۱,۱۹۰)} = ۲/۹۵۸$) و آزمون دماسنج ($P = ۰/۳۴۳$ و $F_{(۱,۱۹۹)} = ۰/۹۰۵$) معنی‌دار نبود. مقایسه میانگین نمرات آزمون درخت و برگ دانش‌آموزان گروههای آزمایش و گواه (جدول ۱) نشان می‌دهد که دانش‌آموزان گروه آزمایش، یعنی دانش‌آموزان مشمول طرح ارزشیابی توصیفی در این آزمون، عملکرد بهتری دارند.

بحث و نتیجه‌گیری

در این بخش، یافته‌های حاصل از تجزیه و تحلیل داده‌های این پژوهش در قالب ویژگیهای شناختی، عاطفی و روانی - حرکتی مورد بحث قرار گرفته است.

ویژگیهای شناختی

نتایج حاصل نشان داد که سطح پیشرفت تحصیلی دانش‌آموزان پایه سوم ابتدایی مدارس مشمول طرح ارزشیابی توصیفی در متغیرهای حوزه شناختی از قبیل دانش ریاضی، علوم تجربی، زبان فارسی و همچنین میزان آگاهی فراشناختی، به طور معناداری، بیش‌تر از دانش‌آموزان پایه سوم ابتدایی مدارس عادی است. متوسط درصد نمره دانش‌آموزان مدارس تحت پوشش طرح ارزشیابی توصیفی در آزمون پیشرفت تحصیلی ریاضی $۱۴/۹۲$ درصد، در آزمون پیشرفت تحصیلی علوم تجربی $۵/۳۶$ درصد، در آزمون پیشرفت تحصیلی زبان فارسی $۱۰/۹۴$ درصد و در میزان آگاهی فراشناختی $۵/۰۹$ درصد از متوسط نمره دانش‌آموزان مدارس عادی بیش‌تر بود. نتایج پژوهش حاضر حاکی از آن است که راهبردهای سنجشی مورد استفاده در طرح ارزشیابی توصیفی، شرایط لازم برای رشد دانش و معلومات و تواناییها و مهارت‌های شناختی و فراشناختی را فراهم کرده است.

برخی از مولفه‌های طرح ارزشیابی توصیفی از پشتوانه تجربی محکمی برخوردار است. برای مثال، پژوهشهای متعددی، اثربخشی ارزشیابی تکوینی و ارائه بازخورد در پیشرفت تحصیلی دانش‌آموزان را مورد تایید قرار داده‌اند (کرد، ۱۳۸۱؛ سپاسی، ۱۳۷۸؛

دانشیان، ۱۳۷۵؛ حیدری، ۱۳۷۵؛ شادمهر، ۱۳۵۹؛ ون اورا، ۲۰۰۴ و بلک و ویلیام، ۱۹۹۸). علاوه بر اینها، تأثیر بازخورد اطلاعاتی بر افزایش انگیزه درونی و عملکرد (احسانی، ۱۳۷۴)، تأثیر بازخورد شایستگی بر افزایش عملکرد و علاقه (حویزروی، ۱۳۷۴)، تأثیر بازخورد اسنادی و هدف‌گزینی بر کارآمدی شخصی برای نوشتن و پیشرفت نوشتاری از قبیل سیالی، سازمانبندی و رعایت قواعد دستوری (مک‌گی، ۱۹۹۶) و تأثیر بازخورد نوشتاری در سنجش تکوینی بر جهت‌گیری هدفی تسلطی (وادل، ۲۰۰۴) نشان داده شده است. به رغم اهمیت ارزشیابی تکوینی و ضرورت به کارگیری آن در کلاس درس، این شیوه ارزشیابی در کلاسهای عادی به خوبی مورد استفاده قرار نمی‌گیرد (یادگارزاده، ۱۳۸۴؛ پرویزیان و کاظمی، ۱۳۸۴؛ کوثری، ۱۳۷۹ و جای‌ا‌کا، ۲۰۰۶).

در پژوهشهای دیگر، اثربخشی راهبردهای سنجشی و آموزشی متأثر از رویکرد سازندگی، بر متغیرهای حوزه شناختی، مورد پژوهش قرار گرفته است. برای مثال، فیوکز و همکاران (۱۹۹۹) نشان دادند که دانش‌آموزان کلاسهایی که با سنجش عملکردی هدایت می‌شوند، از مهارت‌های حل مسئله غنی‌تری برخوردارند. فونتانا و فرناندز (۱۹۹۴) دریافتند که دانش‌آموزان مقطع ابتدایی که سنجش شخصی را آموخته‌اند، نسبت به کسانی که چنین آموزشی ندیده بودند، عملکرد بهتری در ریاضیات داشتند. یونگ (۲۰۰۰) نیز که آزمون خود را با دانش‌آموزان پایه چهارم در درس انشا اجرا کرد، در زمینه سنجش شخصی به نتایج مشابهی دست یافت. یافته‌های پژوهش وو^۲ و سی^۳ (۲۰۰۵) نشان داد که دانش‌آموزان در گروه آموزشی مبتنی بر سازندگی، به طور معنی‌داری، بهتر از گروه آموزشی سنتی به بازده‌های یادگیری (بر حسب میزان و یکپارچگی ساختار شناختی شان، درگیری فراشناختی و استفاده از راهبردهای پردازش اطلاعات) رسیده بودند. پژوهش حاضر نیز در تایید پژوهشهای فوق نشان داد که راهبردهای سنجشی می‌تواند در بهبود فرایند آموزش - یادگیری و تعمیق یادگیری نقش مهمی داشته باشد.

نتایج این پژوهش با یافته‌های حقیقی (۱۳۸۴) هم‌خوانی دارد. او سطوح مختلف یادگیری یعنی دانش، ادراک - کاربرد و قوه خلاقه (نوآوری، سیالی، اصالت) را در دانش‌آموزان تحت پوشش طرح ارزشیابی توصیفی و دانش‌آموزان عادی مناطق ۱ و ۹ تهران مورد مقایسه قرار داده بود. نتایج حاکی از این بود که عملکرد گروه آزمایش در کل

1. Gioka

2. Wu

3. Tsai

آزمون و همچنین در سطوح مختلف یادگیری، بهتر از گروه کنترل است. با این حال، پژوهشگران دیگر در جوامع آماری دیگر به نتایج متفاوتی رسیده اند. برای مثال، نتایج خوش خلق (۱۳۸۵) از اجرای آزمایشی طرح ارزشیابی توصیفی در ۱۰ استان کشور، حاکی از آن است که در املا و خواندن، بین دانش آموزان مشمول طرح ارزشیابی توصیفی و دانش آموزان غیر مشمول طرح، در پایه های اول، دوم و سوم ابتدایی تفاوت معنی داری وجود ندارد. در درس علوم، تفاوت بین این دو گروه در هر سه پایه اول، دوم و سوم معنی دار مشاهده شده است. در درس انشای پایه های اول و دوم و در جمله بندی پایه های دوم و سوم و در درس ریاضی پایه اول، تفاوت معنی داری بین دو گروه گزارش شده است. کلهر (۱۳۸۴) نیز در پایان دومین سال اجرای آزمایشی طرح ارزشیابی توصیفی، به بررسی میزان تحقق اهداف آن در استان قزوین پرداخته و گزارش کرده است که بین پیشرفت تحصیلی دانش آموزان تحت پوشش طرح ارزشیابی توصیفی در دروس املا و ریاضی با دانش آموزان همتای غیر تحت پوشش، تفاوت وجود ندارد. یافته های جانبی پژوهش حاضر نیز نشان داد که بین پیشرفت تحصیلی دانش آموزان مدارس تحت پوشش طرح ارزشیابی توصیفی در مناطق ۱، ۵، ۹ و ۱۶ تفاوت معنی داری وجود دارد. عملکرد دانش آموزان پایه سوم ابتدایی مشمول طرح ارزشیابی توصیفی منطقه آموزشی ۱ در آزمون پیشرفت تحصیلی زبان فارسی، به طور معنی داری، بهتر از دانش آموزان مناطق ۵، ۹ و ۱۶ است. در آزمون پیشرفت تحصیلی علوم تجربی نیز عملکرد دانش آموزان مشمول طرح ارزشیابی توصیفی منطقه آموزشی ۵، به طور معنی داری، بهتر از دانش آموزان مناطق ۹ و ۱۶ بود.

مبانی نظری و راهکارهای اجرا شده طرح ارزشیابی توصیفی در مناطق مختلف یکسان بوده است؛ اما طرح در برخی مناطق، از نظر پیشرفت تحصیلی، موفقیتی در پی نداشته است. عدم موفقیت طرح در مناطق مختلف را می توان به عوامل مختلفی نسبت داد: از قبیل عدم آشنایی معلمان و مدیران با طرح، حجم زیاد کتابها، تعداد زیاد دانش آموزان در کلاس درس، عدم استقبال طرح از سوی برخی مجریان به ویژه معلمان مرد، عدم تعهد برخی از مدارس مجری و نبود سازماندهی و حمایت کافی (مرتضایی نژاد، ۱۳۸۴؛ کلهر، ۱۳۸۴ و محقق معین، ۱۳۸۲). به نظر می رسد حذف ارزشیابی پایانی در طرح ارزشیابی توصیفی، یکی دیگر از عوامل موثر در عدم موفقیت آن در مناطق مختلف باشد. در مرور ادبیات پژوهشی، هیچ پژوهشی نبود که نشان دهد ارزشیابی پایانی تاثیر منفی بر پیشرفت تحصیلی

دارد. اغلب پژوهش‌های انجام گرفته در زمینه ارزشیابی، بر تحلیل آزمون متمرکز هستند و نقاط ضعف آزمونها از قبیل عدم توجه به سطوح بالاتر یادگیری را گوشزد می‌کنند (مهاجر، ۱۳۸۴؛ تاجیک اسماعیلی، ۱۳۸۲؛ شاه حسینی، ۱۳۷۶؛ جعفری، ۱۳۷۶؛ اکبری، ۱۳۷۵ و روشنفکر دزفولی و هومن، ۱۳۷۳). بنابراین، حذف ارزشیابی هیچ مبنای پژوهشی نداشته و حتی برخی از صاحب نظران، از جمله بلوم، که خود از مدافعان ارزشیابی تکوینی است، حذف ارزشیابی پایانی را توصیه نمی‌کنند. در این زمینه بلوم، هستینگز و ماداوس (۱۳۵۵) در کتاب خود تحت عنوان «راهنمای ارزشیابی تکوینی و ارزشیابی مجموعی» چنین می‌گویند: "ما به هیچ وجه معتقد نیستیم که می‌توان از ارزشیابی مجموعی صرف نظر کرد و آرزو هم نمی‌کنیم که این نوع ارزشیابی متوقف گردد" (ص ۲۶۳). عدم اجرای ارزشیابی پایانی در طرح ارزشیابی توصیفی باعث می‌شود که نظام آموزشی دوره ابتدایی در درازمدت به یک نظام ارتقای خود به خودی تبدیل شود و نتیجه آن، بی سواد بار آمدن دانش‌آموزان خواهد بود. بررسی آیین نامه و دستورالعمل‌های امتحانات هم نشان می‌دهد که در سال تحصیلی ۳۶-۱۳۳۵ طرح مشابهی برای پایه‌های اول، دوم و سوم دبستان اجرا و پس از سه سال متوقف شد (شهبازی، ۱۳۸۱). در مورد این که چرا و به چه دلیل این طرح بعد از سه سال اجرا لغو شد، سند مکتوبی وجود ندارد؛ ولی بنا به گفته برخی سالخوردگان، این طرح به دلیل افت شدید در مهارت‌های اساسی دانش‌آموزان مورد بازبینی قرار گرفت (حسینی، ۱۳۸۲).

ویژگیهای عاطفی

نتایج تحلیل واریانس چند متغیری نشان داد که بین دانش‌آموزان پایه سوم ابتدایی مشمول طرح ارزشیابی توصیفی و مدارس عادی در متغیرهای حوزه عاطفی تفاوت معنی‌داری وجود ندارد. با این حال، مقایسه متوسط درصد نمرات نشان داد که در مقیاسهای عاطفه منفی نسبت به مدرسه و موفقیت، تفاوت بین دو گروه، اندکی بیش تر از سایر مقیاسهاست؛ ولی نه به اندازه ای که از نظر آماری معنی‌دار باشد. متوسط درصد نمره دانش‌آموزان مشمول طرح ارزشیابی توصیفی و مدارس عادی برای مقیاس رضایت کلی از مدرسه به ترتیب ۸۵/۶۲ و ۸۳/۲۱ درصد، برای مقیاس عاطفه منفی نسبت به مدرسه به ترتیب ۴۳/۸ و ۴۹/۱۵ درصد، برای مقیاس نگرش درباره معلمان به ترتیب ۸۷/۹ و ۸۸/۵۵ درصد، برای مقیاس همبستگی اجتماعی به ترتیب ۸۱/۷۲ و ۷۹/۶۰ درصد، برای مقیاس

فرصت (باور نسبت به مفید بودن مدرسه) به ترتیب ۹۲/۴۲ و ۸۹/۲۵ درصد، برای مقیاس موفقیت یا پیشرفت به ترتیب ۹۰/۵۶ و ۸۵/۸۸ درصد و برای مقیاس ماجرا (احساس خودانگیختگی در یادگیری) به ترتیب ۸۶/۵۵ و ۸۳/۷۵ درصد بود. متوسط نمره دانش‌آموزان مدارس تحت پوشش طرح ارزشیابی توصیفی در مقیاس های رضایت کلی از مدرسه ۲/۴۱ درصد، همبستگی اجتماعی ۲/۱۲ درصد، فرصت ۳/۱۷ درصد، موفقیت یا پیشرفت ۴/۶۸ درصد و ماجرا ۲/۸ درصد، از متوسط نمره دانش‌آموزان مدارس عادی بیشتر است؛ در عوض در مقیاس های عاطفه منفی نسبت به مدرسه ۵/۳۵ و معلمان ۰/۶۵ کمتر است.

نتایج، حاکی از آن است که دانش‌آموزان مدارس عادی به اندازه دانش‌آموزان مدارس مشمول طرح ارزشیابی توصیفی از مدرسه راضی هستند و نگرش آنها نسبت به خودشان، دانش‌آموزان دیگر و معلمان و فعالیت‌های مدرسه مثبت است. در مطالعه بین‌المللی پیشرفت سواد خواندن (پرلز) در سال ۲۰۰۱ نتایج مشابهی به دست آمد. نتایج پرلز نشان داد که نگرش دانش‌آموزان ایرانی نسبت به درس خواندن، از تمام کشورها بالاتر است و ایران از نظر نگرش مثبت دانش‌آموزان ابتدایی به درس خواندن، رتبه اول را دارد. در حالی که در همان سال، عملکرد دانش‌آموزان در پیشرفت سواد خواندن پایین بود (کریمی، ۱۳۸۳). در پژوهش حاضر نیز هر چند در میزان پیشرفت تحصیلی، تفاوت معنی‌داری بین دانش‌آموزان مشمول طرح ارزشیابی توصیفی و دانش‌آموزان مدارس عادی مشاهده شد، ولی در ویژگیهای عاطفی، از جمله نگرش دانش‌آموزان درباره خودشان، دانش‌آموزان دیگر، معلمان، مدرسه و فعالیت‌های مدرسه تفاوت‌ها معنی‌دار نبود. این بدین معناست که بین ویژگیهای عاطفی و پیشرفت تحصیلی دانش‌آموزان، رابطه معنی‌داری وجود ندارد.

بر اساس پژوهشها تجربه دانش‌آموزان از محیط کلاس (ماک و فلین، ۲۰۰۲)، عوامل شخصیتی به ویژه عزت نفس^۱ آموزشیگاهی (کرتزیس^۲ و همکاران، ۲۰۰۲) و رفتار کنترل دانش‌آموزان در کلاس (اشمیت^۳، ۱۹۸۷) از مهمترین عوامل تبیین کننده کیفیت زندگی در مدرسه است. به نظر می‌رسد در ایران، به غیر از تجارب کلاسی، عوامل دیگری در شکل‌گیری نگرشهای دانش‌آموزان نقش دارند؛ از جمله می‌توان به تاثیر نگرش والدین و جامعه اشاره کرد که از سنین کودکی یک نگرش مثبت قوی نسبت به تحصیل و کسب علم ایجاد

1. self-esteem

2. Karatzias

3. Schmidt

می‌کنند. به هر حال، این حوزه جای مطالعه بیشتری دارد و پژوهشهای بعدی لازم است عوامل تبیین‌کننده ویژگیهای عاطفی در دانش‌آموزان ایرانی را مورد بررسی قرار دهند.

ویژگیهای روانی - حرکتی

نتایج حاصل از تجزیه و تحلیل داده‌ها نشان داد که در آزمونهای عملکردی، دانش‌آموزان مدارس مشمول طرح ارزشیابی توصیفی، بهتر از دانش‌آموزان مدارس عادی هستند. آزمونهای تعقیبی حاکی از آن بود که تفاوت بین گروههای آزمایش و گواه، فقط در آزمون درخت و برگ معنی‌دار است. در آزمون اشیای دور ریختنی و همچنین آزمون دماسنج تفاوت‌ها معنی‌دار نبود.

مقایسه نمرات آزمون درخت و برگ دانش‌آموزان گروههای آزمایش و گواه نشان داد که دانش‌آموزان مشمول طرح ارزشیابی توصیفی در این آزمون عملکرد بهتری دارند. در آزمون درخت و برگ، متوسط نمره دانش‌آموزان مدارس تحت پوشش طرح ارزشیابی توصیفی ۹/۱۵ درصد از متوسط نمره دانش‌آموزان مدارس عادی بیش تر بود. این آزمون، توانایی اندازه‌گیری طول به میلی‌متر، شناسایی روند و یا الگوها در داده‌ها و نیز منابع احتمالی خطا در اندازه‌گیری را مورد سنجش قرار می‌داد. این آزمون برای دانش‌آموزان هر دو گروه، یک تکلیف تقریباً ناآشنا بود؛ زیرا سؤالات آزمون را در کلاس تمرین نکرده بودند؛ به همین خاطر متوسط نمرات دانش‌آموزان هر دو گروه پایین بود. متوسط نمره دانش‌آموزان مشمول طرح ارزشیابی توصیفی ۳۹/۸۶ درصد و دانش‌آموزان مدارس عادی ۳۰/۷۱ بود.

هر چند در آزمون اشیای دور ریختنی و همچنین آزمون دماسنج تفاوت بین دو گروه معنی‌دار نبود، ولی متوسط عملکرد دانش‌آموزان مشمول طرح ارزشیابی توصیفی بیش تر از دانش‌آموزان عادی بود. متوسط نمره دانش‌آموزان مدارس تحت پوشش طرح ارزشیابی توصیفی در آزمون اشیای دور ریختنی ۷/۶۲ درصد و در آزمون دماسنج ۴/۳۳ درصد بیش تر از دانش‌آموزان مدارس عادی بود. آزمون اشیای دور ریختنی، توانایی سازمان‌دادن به اطلاعات و ترسیم نمودار و آزمون دماسنج، توانایی دانش‌آموزان برای اندازه‌گیری دمای آب در شرایط مختلف را مورد سنجش قرار می‌داد. با توجه به این که این آزمونها برای دانش‌آموزان هر دو گروه آشنا بود، درصد بالایی از هر دو گروه پاسخ درست داده بودند.

متوسط نمره دانش آموزان مشمول طرح ارزشیابی توصیفی و مدارس عادی برای آزمون اشیای دور ریختنی، به ترتیب ۸۳/۵ و ۷۵/۸۸ درصد و برای آزمون دماسنج، به ترتیب ۵۳/۶۷ و ۴۹/۳۳ درصد بود.

مطالعات گذشته نشان داده است که عملکرد دانش آموزان در آزمونهای سنجش عملکرد از سؤالهای آزمون تاثیر می پذیرد (شیولسون^۱، بکستر^۲ و پاین^۳، ۱۹۹۲) و عملکرد، از یک آزمون به آزمون دیگر، با توجه به تجارب قبلی دانش آموزان، نوع فعالیت های خواسته شده و علاقه آنان به انجام فعالیت تغییر می کند (کیامنش، ۱۳۷۷). با این حال، نتایج دور از انتظار نبود. بدون شک استفاده از انواع متعدد سنجش عملکردی در برنامه های روزمره کلاسی مدارس تحت پوشش طرح ارزشیابی توصیفی، در این نتایج بی تاثیر نبوده است. در هیچ یک از پژوهشهای قبلی ویژگیهای روانی - حرکتی دانش آموزان مشمول طرح ارزشیابی توصیفی و عادی مورد سنجش و مقایسه قرار نگرفته است؛ بنابراین امکان مقایسه نتایج وجود نداشت. مرور ادبیات پژوهشی نشان داد که در ایران، سنجش عملکردی در مقیاس وسیع، برای اولین بار در سومین مطالعه بین المللی ریاضی و علوم به اجرا در آمده است. نتایج این پژوهش حاکی از آن است که متوسط نمره دانش آموزان ایرانی پایه بالای جمعیت ۱ (سال چهارم ابتدایی) در شش آزمون سنجش عملکرد در بخش علوم، ۷ درصد از متوسط نمره دانش آموزان ۹ کشور شرکت کننده کم تر است. در آزمونهای عملکردی علوم، دانش آموزان ایران در بین ۹ کشور، رتبه هشتم را کسب کرده اند. متوسط نمره دانش آموزان سال چهارم ابتدایی در ۶ آزمون سنجش عملکرد در بخش ریاضی، ۴ درصد از متوسط نمره دانش آموزان ۹ کشور شرکت کننده در جمعیت یک بیش تر بود. در آزمونهای عملکردی ریاضی، دانش آموزان ایران در بین ۹ کشور، رتبه سه و نیم را کسب کرده بودند (کیامنش، ۱۳۷۷).

بدیهی است که دانش آموزان برای کسب توانایی در به کارگیری دانش و مهارت های خود در دنیای واقعی، به سنجش های عملکردی نیاز دارند. از طرفی دیگر پژوهشها نشان می دهد که سنجش های عملکردی و روند جدید در مدارس ایران کم تر مورد توجه قرار گرفته است (بهنام و نریمانی، ۱۳۷۳ و مهاجر، ۱۳۸۴). از این رو آشنایی آموزگاران و دبیران با

1. Shavelson

2. Baxter

3. Pine

انواع سنجش عملکردی، چگونگی تهیه و اجرای آنها ضروری به نظر می‌رسد. این کار می‌تواند از طریق مراکز تربیت معلم و دوره‌های آموزشی ضمن خدمت انجام شود.

محدودیت‌ها: محدودیت‌های پژوهش حاضر عبارت اند از:

- ۱- در این پژوهش برای انتخاب مدارس هم سطح با مدارس مشمول طرح ارزشیابی توصیفی، تلاش‌های نسبتاً زیادی به عمل آمد. با این حال، رعایت احتیاط درباره هم سطح بودن مدارس مورد مقایسه و تفسیر نتایج ضرورت دارد.
- ۲- جامعه آماری پژوهش حاضر محدود به مدارس انتخاب شده در مناطق ۱، ۵، ۹ و ۱۶ تهران و همچنین مدارس هم سطح با آنهاست. بنابراین، یافته‌های این پژوهش را تنها به جوامع مشابه می‌توان تعمیم داد.
- ۳- بر اساس مجوز شورای عالی آموزش و پرورش، طرح ارزشیابی توصیفی، در مدارس که امکانات و نیروی انسانی لازم را داشتند اجرا شده است؛ بنابراین تعمیم نتایج به سایر مدارس، که از این ویژگی‌ها را ندارند، باید با احتیاط صورت گیرد.

منابع

- آوانسیان، اما. (۱۳۷۷). نقش آموزش راهبردهای فراشناختی بر درک مطلب و سرعت یادگیری دانش‌آموزان دختر مدارس روزانه مقطع راهنمایی شهر تهران. پایان‌نامه کارشناسی ارشد، دانشگاه علامه طباطبایی.
- احسانی، پرنیان. (۱۳۷۴). اثر انواع پسخوراند بر انگیزه درونی و عملکرد دانش‌آموزان کلاس پنجم دبستان. پایان‌نامه کارشناسی ارشد، دانشگاه تهران.
- اکبری، ابوالقاسم. (۱۳۷۵). ارزشیابی و تحلیل محتوای سؤالات امتحان نهایی پنجم ابتدایی اداره کل شهرستان‌های استان تهران خرداد ماه ۱۳۷۴. گزارش پژوهشی، سازمان آموزش و پرورش شهرستان‌های تهران.
- بلوم، اس. بی؛ هستینگز، ت. ج؛ ماداوس، اف. ج. (۱۳۵۵). راهنمای ارزشیابی تکوینی و ارزشیابی مجموعی از آموخته‌های دانش‌آموزان (ترجمه ابراهیم کظیمی). تهران: انتشارات دانشگاه تربیت معلم. (تاریخ انتشار به زبان اصلی، ۱۹۷۱).
- بهنام، محمود و نریمانی، فاطمه. (۱۳۷۳). بررسی دلایل بی‌توجهی به فعالیت‌های آزمایشگاهی و نحوه ایجاد علاقه در دبیران و دانش‌آموزان در جهت استفاده از امکانات آزمایشگاهی. گزارش پژوهشی، سازمان آموزش و پرورش استان آذربایجان شرقی.

- پرویزیان، محمد علی و کاظمی، مریم. (۱۳۸۴). بررسی وضع ارزشیابی های مستمر در دوره ی ابتدایی. فصلنامه تعلیم و تربیت، ۸۲، صص ۵۹-۸۶.
- تاجیک اسماعیلی، عزیزاله. (۱۳۸۲). عملکرد معلمان پایه پنجم ابتدایی با توجه به میزان آگاهی آنان از اهداف، روش تدریس، مهارت تدریس، کاربرد مواد آموزشی، روش ارزشیابی و کیفیت سؤالیهای امتحانی در درس تعلیمات اجتماعی. رساله دکتری، دانشگاه تربیت معلم.
- جعفری، علی برار. (۱۳۷۶). بررسی شیوه های ارزشیابی موجود دانش آموزان سال سوم راهنمایی قائم شهر و مقایسه نظری آن با الگوی مطلوب ارزشیابی از دیدگاه متخصصان تعلیم و تربیت. پایان نامه کارشناسی ارشد، دانشگاه تربیت معلم.
- حسینی، محمد و احمدی، حسین. (۱۳۸۴). ارزشیابی توصیفی: الگویی نو در ارزشیابی تحصیلی. تهران: انتشارات مدرسه.
- حسینی، محمد و کاظمی، یحیی. (۱۳۸۲). طرح ارزشیابی توصیفی (اهداف، اصول و راهکارها). تهران: آثار معاصر.
- حسینی، محمد. (۱۳۸۲). الگوی ارتقای خود به خود تحصیلی: از آرمان تا واقعیت. مجموعه مقالات اولین همایش ارزش یابی تحصیلی (صص ۱۳۳-۱۱۳). تهران: تزکیه.
- حسینی، محمد. (۱۳۸۴). راهنمای اجرای ارزشیابی توصیفی: پایه های اول، دوم و سوم دبستان. تهران: آثار معاصر.
- حقیقی، فهیمه السادات. (۱۳۸۴). نقش ارزشیابی مستمر در تعمیق یادگیری دانش آموزان پایه دوم مقطع ابتدایی شهر تهران در سال تحصیلی ۸۴-۱۳۸۳. پایان نامه کارشناسی ارشد، دانشگاه پیام نور تهران.
- حویزوی، عبدالامام. (۱۳۷۴). تاثیر سطوح پاداش و پسخوراند بر عملکرد و علاقه به درس ریاضی در دانش آموزان پسر سال سوم راهنمایی شهرستان سوسنگرد. پایان نامه کارشناسی ارشد، دانشگاه شهید چمران اهواز.
- حیدری، شعبان. (۱۳۷۵). تاثیر ارزشیابی تکوینی بر افزایش پیشرفت تحصیلی دانش آموزان سال اول دبیرستانهای نظام جدید منطقه فریدون کنار در سال تحصیلی ۷۵-۷۴. پایان نامه کارشناسی ارشد. دانشگاه علامه طباطبایی.
- خوش خلق، ایرج. (۱۳۸۵). گزارش ارزشیابی اجرای آزمایشی ارزشیابی توصیفی در مدارس ابتدایی برخی از مناطق آموزشی کشور در سال تحصیلی ۸۵-۸۴ (مرحله سوم). گزارش پژوهشی، پژوهشکده تعلیم و تربیت.
- دانشیان، محمد حسن. (۱۳۷۵). بررسی شیوه های رایج ارزشیابی و تاثیر آن بر پیشرفت تحصیلی دانش آموزان ابتدایی شهرستان یزد. گزارش پژوهشی، سازمان آموزش و پرورش استان یزد.
- روشنفکر دزفولی محمد جواد ؛ هومن، عباس. (۱۳۷۳). ارزشیابی محتوای سؤالات امتحانات هماهنگ مقاطع مختلف تحصیلی در خردادماه ۱۳۷۳ از لحاظ مطابقت با هدف های حیطه شناختی. گزارش پژوهشی. گزارش پژوهشی، آموزش و پرورش منطقه ۵ شهر تهران.

- سپاسی، حسین. (۱۳۷۸). مقایسه تأثیر ارزشیابی تکوینی و ارزشیابی پایانی بر پیشرفت تحصیلی دانشجویان دانشگاه شهید چمران. مجله ی علوم تربیتی و روان شناسی، دوره سوم، شماره ۱، ۵۰-۳۹.
- سیف، علی اکبر. (۱۳۸۲). مفهوم سنجش و تأثیر آن بر روش آموزش معلم و کیفیت یادگیری دانش آموزان. مجموعه مقالات اولین همایش ارزش یابی تحصیلی (صص ۲۱۲-۱۹۹). تهران: تزکیه.
- سیف، علی اکبر. (۱۳۸۳). اندازه گیری، سنجش، و ارزشیابی آموزشی (ویرایش سوم). تهران: نشر دوران.
- سیف، علی اکبر. (۱۳۸۴). سنجش فرایند و فرآورده یادگیری: روشهای قدیم و جدید. تهران: نشر دوران.
- شادمهر، علی. (۱۳۵۹). تأثیر ارزشیابی مرحله ای، اطلاع از نتایج، رقابت، تقویت در یادگیری. پایان نامه کارشناسی ارشد، مدرسه عالی تلویزیون.
- شاه حسینی، محمد. (۱۳۷۶). بررسی سؤالات امتحانی دبیران و هنرآموزان و انطباق آن با شیوه های سنجش و اندازه گیری. گزارش پژوهشی، سازمان آموزش و پرورش استان سمنان.
- شریفی، حسن پاشا. (۱۳۸۱). تدوین و هنجاریابی آزمونهای پیشرفت تحصیلی زبان فارسی برای دوره های ابتدایی و راهنمایی تحصیلی. فصلنامه تعلیم و تربیت، ۶۹ و ۷۰، ۸۷-۶۷.
- شریفی، حسن پاشا. (۱۳۸۳). سنجش عملکرد در فرایند یاددهی-یادگیری. مجموعه مقالات همایش ملی-مهندسی اصلاحات در آموزش و پرورش (صص ۳۴۱-۳۲۳). تهران: پژوهشکده تعلیم و تربیت.
- شهبازی، سید رحمت اله. (۱۳۸۱). سیر تحول آیین نامه امتحانات از سال ۱۳۰۲ تا ۱۳۸۱ شمسی در نظام آموزش و پرورش ایران دوره آموزش عمومی. گزارش پژوهشی، دفتر ارزشیابی تحصیلی و تربیتی. مجموعه مصوبات شورای عالی آموزش و پرورش. (۱۳۸۲). تهران: انتشارات مدرسه.
- محقق معین، محمد حسن. (۱۳۸۲). ارزشیابی اجرای پیش آزمایشی طرح ارزشیابی توصیفی. گزارش پژوهشی، دفتر ارزشیابی تحصیلی و تربیتی.
- مرتضایی نژاد، عصمت. (۱۳۸۴). بررسی نگرش معلمان و والدین در خصوص کاربرست ارزشیابی توصیفی در پایه های اول و دوم ابتدایی شهر تهران در سال ۸۴-۸۳. پایان نامه کارشناسی ارشد، دانشگاه آزاد اسلامی واحد شهر ری.
- مهاجر، یحیی. (۱۳۸۴). ارائه یک الگوی مناسب ارزشیابی پیشرفت تحصیلی و مقایسه آن با وضع موجود در دبستانهای تهران. رساله دکتری، دانشگاه تربیت معلم تهران.
- کرد، بهمن. (۱۳۸۱). بررسی تأثیر بازخورد در ارزشیابی تکوینی بر پیشرفت تحصیلی درس علوم تجربی دانش آموزان پسر پایه پنجم ابتدایی شهرستان بوکان در سال تحصیلی ۸۲-۱۳۸۱. پایان نامه کارشناسی ارشد، دانشگاه تربیت معلم.
- کریمی، عبدالعظیم. (۱۳۸۳). بررسی نتایج مطالعه بین المللی پیشرفت سواد خواندن پرلز (PIRLS) ۲۰۰۱. تهران: پژوهشکده تعلیم و تربیت.

کلهر، منوچهر. (۱۳۸۴). بررسی میزان تحقق اهداف ارزشیابی توصیفی در استان قزوین. گزارش پژوهشی، سازمان آموزش و پرورش استان قزوین.

کوثری، رضا. (۱۳۷۹). بررسی میزان آشنایی و بکارگیری شیوه های ارزشیابی از آموخته های دانش آموزان توسط دبیران دوره متوسطه در شهرستان قروه (کردستان) در سال تحصیلی ۷۹-۷۸. گزارش پژوهشی، سازمان آموزش و پرورش استان مرکزی.

کیامنش، علیرضا و خیریه، مریم. (۱۳۸۰). سنجش و اندازه گیری در ریاضی و علوم همراه با سؤالهای ریاضی و علوم TIMSS در دوره ابتدایی. تهران: پژوهشکده تعلیم و تربیت.

کیامنش، علیرضا. (۱۳۷۷). سنجش عملکرد در سومین مطالعه بین المللی ریاضی و علوم TIMSS چهارم ابتدایی و سوم راهنمایی. تهران: پژوهشکده تعلیم و تربیت.

یادگارزاده، غلامرضا. (۱۳۸۴). بررسی میزان آگاهی معلمان از روشهای ارزشیابی تکوینی و میزان کاربرد آن در کلاس درس، نمونه موردی استان همدان. فصلنامه تعلیم و تربیت، ۸۳، ۸۵-۶۲.

Berry, R. (2005). Entwining feedback, self, and peer assessment. *Academic Exchange Quarterly*, 9, 225-230.

Black, P. J. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles Policy and Practice*, 5, 7-73

Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, 86, 8 -21.

Diez, M. E. (1997). Assessment as a lever in education reform. *Phi Kappa Phi Journal - National Forum*, 77, 27-30.

Ediger, M. (2000). *Qualitative assessment versus measurement of student achievement*. (ERIC Document Reproduction Service No. ED 44v 154).

Eggen, P, Kauchak, D. (2001). *Educational Psychology*. (5th Ed). Upper Saddle River, NJ: Prentice Hall.

Fontana, D., & Fernandez, M. (1994). Improvements in mathematics performance as a consequence of self assessment in portuguese primary school pupils. *British Journal of Educational Psychology*, 64, 407-417.

Fuchs, L., Fuchs, D., Karnes, K., Hanlett, C., & Katzaroff, M. (1999). Mathematics performance assessment in the classroom: effects on teacher planning and student problem solving. *American Educational Research Journal*, 36, 609-646.

Gioka, O. (2006). Assessment for learning in physics investigations: assessment criteria, questions and feedback in marking. *Physics Education*, 41, 341-346.

Grace, C. (1992). *The portfolio and its use: Developmentally appropriate assessment of young children*. Urbana, IL: Clearinghouse on Elementary and Early Childhood Education. (ERIC Document Reproduction Service No. ED 357 393).

Karatzias, A., Power, K. G., Flemming, J., Lennan, F., & Swanson, V. (2002). The role of demographics, personality variables and school stress on predicting school satisfaction/dissatisfaction. *Educational Psychology*, 22, 33-50.

Katula, A., Ray, S., & Sherrill, C. (1999). *Improving student motivation, parent communication, and assessment while implementing a portfolio program*. (ERIC Document Reproduction Service No. ED 436 309).

Klenowski, V. (2002). *Developing portfolios for learning and assessment*. London: Routledge.

Linn, R. L. (1994). *Assessment-based reform: challenges to educational measurement*. (ERIC Document Reproduction Service No. ED 393 875).

McGee, S.J. (1996). *The effect of goal setting and attributional feedback on self-efficacy for writing and writing achievement*. Doctoral Dissertation, Florida international University.

Mok, M. & Flynn, M. (2002) determinants of students Quality of School Life: A path model. *Learning Environments Research*, 5, 275-300.

Panaoura, A. & Philippou, G. (2005). *The measurement of young pupils' metacognitive ability in mathematics: the case of self-representation and self evaluation*. Sant Feliu de Guíxols : CERME 4.

Paulson, F. L., & Paulson, P. P. (1994). *Student led portfolio conferences*. (ERIC Document Reproduction Service No. ED 377 241).

Razdevsek-Pucko, C. (1998). *Changing the assessment paradigm in Slovenian schools*. Paper presented at the ECER conference, September 1998, Ljubljana.

Schmidt, L. J. (1987). *Relationships between pupil control ideology, pupil control behavior, and the quality of school life*. Doctoral Dissertation, Loyola University of Chicago.

Shavelson, R. J., Baxter, G. P., & Pine, J. (1992). Performance assessments: political rhetoric and measurement reality. *Educational Researcher*, 21, 22-27.

Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher*. 29, 1-14.

Stiggins, R. (2004). New Assessment Beliefs for a New School Mission. *Phi Delta Kappan*, 86, 22-32.

Stiggins, R. (2005). From formative assessment to assessment FOR learning: a path to success in standards-based schools. *Phi Delta Kappan*, 87, 324-332.

Taras, M. (2002). Using assessment for learning and learning for assessment. *Assessment and Evaluation in Higher Education*, 27, 501-510.

Teresa, W. L. (2004). Formative assessment in general studies classrooms. *Hong Kong Teachers Centre Journal*, 3, 141-154.

Tierney, R. (1992). Setting a new agenda for assessment. *Learning*, 21, 62-64.

Van Evera, W.C. (2004). *Achievement and motivation in the middle school science classroom: The effects of formative assessment feedback*. Doctoral Dissertation, George Maason University.

Waddell, C.A. (2004). *The effect of negotiated written feedback within formative assessment on fourth-grade students motivation and goal orientations*. Doctoral Dissertation, University of Missouri – Saint Louis.

Walberg, H. J., Haertel, G. D., & Gerlach-Downie, S. (1994). *Assessment reform: challenges and opportunities*. (ERIC Document Reproduction Service No. ED 387 243).

Woolfolk, A. E. (2004). *Educational Psychology*. (9th Ed.). Pearson, International Edition.

Wu, Y., & Tsai, C. (2005). Development of elementary school students' cognitive structures and information processing strategies under long-term constructivist-oriented science instruction. *Science Education*, 89, 822-846.

Young, E. A. (2000). *Enhancing student writing by teaching self-assessment strategies that incorporate the criteria of good writing*. Doctoral Dissertation, Rutgers the state U. of N.J. –New Brunswick and U.M.D.N.J.

Archive of SID