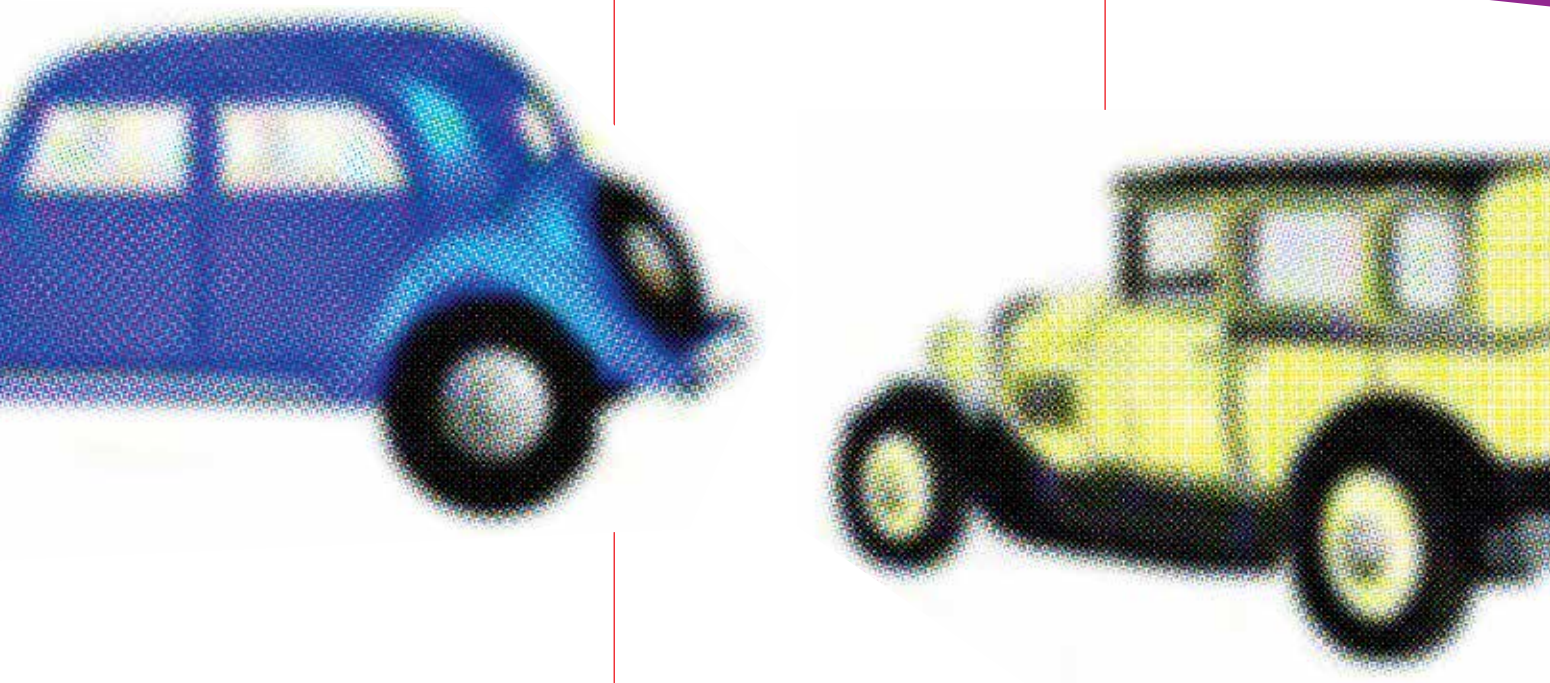
ترجمهٔ حمید تقی‌پور ارمکی

اصلاحات ارزشیابی: گذشته، حال و آینده

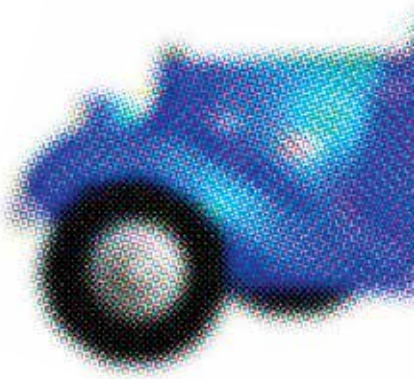


کلیدواژه‌ها:
ارزشیابی، پژوهش، کارکرد ارزشیابی، یادگیری، آزمون

مقدمه

«ارزشیابی» یکی از پرمعناترین واژه‌ها در فرهنگ آموزش است. این واژه برای افراد مختلف معانی متفاوتی دارد. برای مثال، برای دانش‌آموزان مترادف با اضطراب، فشار جسمی و روحی، حس رقابت، موفقیت، شکست، قضاوت، عدالت، انصاف، استاندارد و مسئولیت است و برای معلمان، مسئولیت‌پذیری در اجرای ارزشیابی، داشتن قضاوت عادلانه در مورد عملکرد دانش‌آموزان و تعیین نمره معنی می‌دهد.

اهمیت ارزشیابی در حدی است که دفاتر ثبت‌نام مراکز فنی‌وحرفه‌ای، برای گزینش افراد به اطلاعات مربوط به ارزشیابی نیاز دارند. علاوه براین، نتایج ارزشیابی برای مقامات بالا دستی نیز اهمیت فراونی دارد. زیرا براساس اطلاعات جمع‌آوری شده از ارزشیابی، می‌توانند میزان اثربخشی نظام آموزشی اجرا شده در کشور را با سایر کشورها مقایسه کنند و این تنها بخشی از ویژگی و کارایی ارزشیابی را نشان می‌دهد.



هر چند ارزشیابی همیشه با تنش و هیجان زیادی همراه اســت، اما می‌توان آن را به شیوه‌ای هدفمند ارائــه کرد تا مؤثر و کارآمد واقع شــود. برای مثال، ارزشــیابی را می‌توان به‌صــورت تکوینی و تراکمی، رســمی و غیررســمی، داخلی و خارجی، معتبر و نامعتبر، شفاهی و کتبی و بر محوریت شاخص‌های مرجع به‌طــور متمرکز و غیرمتمرکــز برگزار کرد. ارزشیابی را می‌توان برای رتبه‌بندی، گزینش افراد در زمینه‌های گوناگون، و تعیین میزان شایستگی افراد در یــک مهارت خاص به‌کار برد.با توجه به تئوری و تفکر غالب در هر نظام آموزشی، ملاک‌ها و معیارهای ارزشــیابی متفاوت هستند و موضوع ارزشیابی مورد توجه بسیاری از محققان این حوزه در هر گوشه‌ای از دنیاست.

دخالت‌های ذی‌نفعان متعدد در فرایند ارزشیابی، مشــکلات بســیاری را در نظام آموزشــی به وجود می‌آورد. برای مثال، سیاســت‌گذارانی که به دنبال ایجاد یک نظام آموزشی تراز بین‌المللی هستند که نیازهای اجتماعی، اقتصادی و سیاســی را برآورده کند، به این نتیجه رسیده‌اند که ظرفیت کلاس‌ها و نیازهای دانش‌آموزان با مدل‌های متفاوت ارزشیابی محدود شــده اســت. اگرچه ذی‌نفعــان این حوزه بســیاری از هدف‌های کلی مربوط به آموزش را در راســتای ایجاد یک نظام یادگیری با سطوح کیفی بالا به اشتراک می‌گذارند، اما برای رسیدن به سطح کیفی بالا در یادگیری، باید ظرفیت و نیازهای واقعی کلاس درس در نظر گرفته شود.

در مقالهٔ حاضر این موضوعات مورد بحث و بررسی قرار گرفته و همچنین راه‌حل‌هایی در این خصوص ارائه شده‌اند.

ارزشـــیابی را می‌توان یک پژوهش در نظر گرفت. در هر پژوهش می‌توان اطلاعات به‌دســـت آمده را با به‌کارگیری ابزارهای معتبر و قابل اعتماد طبقه‌بندی کرد و ســـپس مورد تحلیل و بررســـی قرار داد و در نهایت همه نتایج به‌دســـت آمده را برای یک هدف مشـــخص به‌کار گرفت. روش‌های متنوعی در حوزهٔ پژوهش وجود دارند که شامل پژوهش‌های ارزیابانه (ارزیاب‌محور)، تحقیق‌محور و تفسیرمحور می‌باشد. در هر یک از این روش‌ها، سؤالات پژوهشی معینی مطرح هستند.

معمول‌تریـــن نـــوع پژوهش در حوزهٔ ارزشـــیابی، پژوهش «ارزیابانه» یا «ارزیاب‌محور» اســـت. در این روش ابتدا شـــواهد و اطلاعات طبقه‌بندی می‌شوند و ســـپس براســـاس تحلیل‌های انجام شده و نتایج به‌دست آمده، تصمیم‌گیری صورت می‌گیرد. اما در روش پژوهش «تحقیق‌محور» باید به دنبال علت و معلول یادگیری و روابط متقابل بین متغیرها باشیم. ایـــن روش رویکردی علمی در رابطه با ارزشـــیابی دانش‌آمـــوزان ارائه می‌دهد که شـــامل شناســـایی متغیرهای مرتبـــط، طبقه‌بندی داده‌هـــا و تحلیل آن‌هاست. از این طریق می‌توان ارتباط بین متغیرها و عامل‌های متخلف را با یکدیگر تعیین کرد. به‌عنوان یک نتیجـــه، از این روش می‌تـــوان برای روش‌های آماری استفاده کرد.

روش پژوهش «تفسیرمحور» (تفسیری) همواره به دنبال توصیف و تشـــریح یک پدیدهٔ اجتماعی است. ارزشـــیابی تحصیل دانش‌آموزان مدرسه، یک مثال ساده از کاربرد این روش است. براساس علوم رفتاری و اجتماعی، در روش پژوهش تفســـیری، ارزشیابی میزان یادگیری دانش‌آموزان می‌تواند هدف پژوهش در نظر گرفته شود.

در کل محققـــان تمایل کمتری بـــه نتایج واقعی ارزشـــیابی در مقایســـه با عامل‌هـــای اجتماعی که طبیعت ارزشیابی را شکل می‌دهند، دارند. بنابراین یک مطالعهٔ تفســـیری ـ مثلاً در رابطـــه با امتحان موســـیقی ـ ممکن اســـت همراه با تحلیل و بررسی طراحی آزمون، تعییـــن محدودهٔ ابزارهای مورد نیاز برای آزمون، شـــیوهٔ انجام ارزشیابی توسط آزمونگر، نقش و مســـئولیت بخش‌های متفـــاوت، پیامدهای خروجـــی، دلایل انتخـــاب یک نوع ابزار موســـیقی مشخص، نحوهٔ تعیین معیارهای ارزشیابی و... باشد.

روش پژوهـــش «بحرانی» موضوعـــات را از لحاظ عدالت و انصاف اجتماعی بررسی می‌کند.

در این مقاله، رویکردهای مختلف پژوهشـــی مورد بررســـی قرار گرفته‌اند تا «فرایند ارزشـــیابی» مورد ارزیابی قرار گیرد. عـــلاوه بر این، در مقالهٔ حاضر ما به دنبال شناســـایی و توصیف طبیعت سیاست‌های ارزشیابی هســـتیم تا نقاط ضعف و قوت فلسفه‌های مربوط به ارزشیابی را شناسایی کنیم.

در بخـــش بعـــدی مقالـــه، برخـــی از فرضیه‌ها و موضوع‌های مرتبط با هدف‌ها و طبیعت ارزشـــیابی بررسی می‌شوند و ســـپس نگاهی کلی به تاریخچهٔ ارزشـــیابی در آموزش، اصلاحات ایجاد شده و تأثیر ایده‌های متعدد در نظام آموزشی ارائه خواهد شد.

کارکرد ارزشیابی

واژهٔ «ارزشـــیابی» به مجموعه فعالیت‌هایی گفته می‌شود که برای طبقه‌بندی اطلاعات و به‌کارگیری اطلاعـــات مرتبط با دانش، مهارت و نگرش یک فرد انجام می‌گیرند. روش‌های متعددی برای گروه‌بندی کارکرد ارزشیابی وجود دارد. به‌طور عمده دو کارکرد اصلی دارد:

۱. ارزیابی عملکرد افراد یا تعیین میزان اثربخشی نظام آموزشی؛

۲. بهبود یادگیری.

ارزشیابی در راســـتای سنجش توانایی و یا تعیین میزان ظرفیت دانش‌آموزان بـــرای ورود به حرفه و مدرســـهٔ خاص صورت می‌گیرد. همچنین، در حوزهٔ دیگری می‌توان از ارزشیابی برای تعیین شایستگی افراد در یک شـــغل یا حرفهٔ مشخص استفاده کرد؛ مانند آزمون‌های خلبانی. براساس ارزشیابی پیشرفت تحصیلی دانش‌آموزان می‌تـــوان یک مبنای کیفی تعییـــن کرد و به کمـــک آن، مقـــدار منابعی را که باید به‌طور مســـتقیم از طرف دولت به یک مدرسه اختصـــاص داده شـــود، تعیین کـــرد. خانواده‌ها نیز می‌تواننـــد از نتایج حاصل از ارزشـــیابی به منظور انتخاب مدرســـهٔ مناســـب برای بچه‌های خود بهره بگیرند.

در مقیـــاس کلان، «گرایش‌ها در علوم و ریاضیات بین‌المللـــی»[1] (TIMSS)، «انجمن بین‌المللی برای ارزیابی دستاوردهای آموزش»[2] (IEA)، «برنامه برای ارزشیابی دانش‌آموزان»[3] (PISA) و مطالعات تطبیقی مشابه، در واقع ارزشیابی عملکرد دانش‌آموزان کشور برای ارزیابی اثربخشـــی یک نظام آموزشی محسوب می‌شـــوند. البته نحوهٔ ارائهٔ تحلیل‌هـــای مربوط به ارزشیابی، به هدف‌های شنودگان آن بستگی دارد.

اخیراً در پژوهش‌های صورت گرفته، جدول‌های رتبه‌بندی تدوین می‌کنند؛ به همان شکلی که برای رتبه‌بندی تیم‌های ورزشی و بازیکنان یک رشتهٔ ورزشی نیز صورت می‌گیرد. برای مثال، در سیستم رتبه‌بندی، مقایسه‌ای از عملکرد دانش‌آموزان براساس نوع مدرسه‌ای که در آن تحصیل کرده‌اند، و یا نظام آموزشی که در آن حضور دارند، ارائه می‌شود و برای اینکه نتایج به آسانی قابل هضم باشند، پیچیدگی‌های ذاتی ارزشیابی دانش‌آموزان و مدرسه یا عملکرد سیستم در تحلیل و بررسی مورد ارزیابی قرار نمی‌گیرند و حذف می‌شوند.

به هرحال، اگر هدف ارزشیابی فراهم آوردن داده و اطلاعات از عملکرد دانش‌آموزان برای خانواده‌هایشان، معلمان و یا دیگر ذی‌نفعان باشد ـ که این اطلاعات ممکن است باعث توسعهٔ یادگیری دانش‌آموزان شود ـ تنها ساده‌سازی نحوهٔ ارائهٔ داده‌ها کافی نیست. برای تصمیم‌گیری درست، به شواهدی نیاز داریم که با استفاده از ابزارهای ارزشیابی کیفی و کمی دسته‌بندی شده باشند و به شکلی تحلیل و ارائه شوند تا یک مجموعه‌ای از راهبردها و تئوری‌های آموزشی را فراهم آورند که از طراحی و ایجاد زمینه‌هایی برای توسعه یادگیری دانش‌آموزان پشتیبانی کنند.

ارزشیابی را می‌توان از زوایای مختلف سیاسی، اجتماعی، آموزشی و غیره نگاه کرد که هر کدام کارکردهای متفاوت شامل نمره‌دهی، گزینش افراد، تعیین مقدار شایستگی، تعیین درجهٔ استادی دارند. حتی می‌توان از آن در راستای تدوین یک راهنمای تحصیلی برای دانش‌آموزان بهره برد. علاوه‌بر مواردی که ذکر شد، از ارزشیابی می‌توان برای پیش‌بینی توسعهٔ اجتماعی، آموزشی، صنعتی و مانند آن استفاده کرد.

نمره‌دهی

نمره‌دهی رتبه‌بندی دانش‌آموزان از لحاظ پیشرفت و موفقیت‌های تحصیلی و فیلتر کردن دانش‌آموزان موفق از دیگر دانش‌آموزان را شامل می‌شود. «ارزشیابی تشخیصی» کمک می‌کند، ضعف‌ها و توانایی‌های دانش‌آموزان در یادگیری معلوم شود. در این صورت می‌توان یک برنامهٔ جامع جبرانی مناسب برای دانش‌آموزان به منظور پوشش دادن ضعف‌ها و بهبود توانایی آن‌ها طراحی کرد.

با به‌کارگیری «ارزشیابی تجربی» می‌توان تعیین کرد که آیا دانش‌آموزان یک شایستگی معینی از سطوح از پیش تعیین شده برای هدف‌های موردنظر ـ برای مثال گرفتن مدرک ـ را کسب کرده‌اند یا نه؟

قابل ذکر است که کتابچهٔ راهنما، سلسله اطلاعاتی را برای دانش‌آموزان به منظور اتخاذ تصمیمات مهم در دوران تحصیلی فراهم می‌آورد که در رسیدن به پاسخ پرسش‌هایی مانند کدام درس را برداریم؟ یا: کدام مسیر شغلی را انتخاب کنیم؟ آن‌ها را راهنمایی می‌کند.

«ارزشیابی پیشگویانه» به گونه‌ای طراحی می‌شود که در آن ارتباط بین استعداد و توانایی دانش‌آموزان با دانش و مهارت خاص در یک شغل مشخص را نشان می‌دهد، و نیز کمک می‌کند تا آیندهٔ دانش‌آموزان را در آن مهارت خاص پیش‌بینی کنیم.

آزمونی که از دانش‌آموزان به‌عمل می‌آید، ممکن است تنها برای تعیین صلاحیت آن‌ها و مدرک دادن باشد. تعیین محدودهٔ کارکرد ارزشیابی به مدل‌های ارزشیابی ارائه شده بستگی دارد. اگر هدف از طراحی ارزشیابی، گزینش دانش‌آموزان باشد، ارزشیابی را می‌توان به‌گونه‌ای طراحی کرد که دانش‌آموزان در شرایط تجربی به‌صورت عادلانه و منصفانه ارزیابی شوند.

به‌طور کلی، اگر یک برنامهٔ درسی و فرایند ارزشیابی واحد وجود داشته باشد، می‌توان همهٔ دانش‌آموزان را براساس یک سلسله وظایف مشابه در یک زمان مشخص و در شرایط تکراری ارزیابی کرد. همچنین می‌توان سیستم نمره‌دهی یا رتبه‌بندی معینی را به‌طور مداوم در ارزیابی آن‌ها به کار برد.

معمولاً ارزشیابی در پایان هر واحد درسی اجرا می‌شود که ممکن است به‌صورت تکوینی ـ تراکمی باشد.

اما «ارزشیابی گزینشی» (گزینش‌محور) به یک سیستم نمره‌دهی نیاز دارد، تا عملکرد دانش‌آموزان را براساس یک سلسله معیارهای مرجع در شرایط تجربی بسنجد. با این حال، اگر هدف از ارزشیابی شناسایی باشد، مدل آن می‌تواند با ارزشیابی گزینشی تفاوت بسیاری داشته باشد، و نیازی نیست که ارزشیابی در شرایط استاندارد شده به‌طور رسمی همراه با ایجاد استرس فراوان برگزار شود. شاخص‌های یادگیری دانش‌آموزان در یک دورهٔ مطالعاتی معمول می‌تواند پایهٔ ارزشیابی را بدون نیاز به آزمون‌ها و ابزارهای دیگر فراهم سازد. با تحلیل اطلاعات جمع‌آوری شده از پژوهش‌ها می‌توان به

مجموعه‌ای از راهبردهای مؤثر در توسعهٔ یادگیری دانش‌آموزان دست یافت که در طراحی یک برنامهٔ جامع به منظور بهبود یادگیری دانش‌آموزان اثربخش خواهد بود.

مروری تاریخی

ارزشیابی رسمی اولین‌بار قبل از میلاد مسیح در چین ثبت شد. این آزمون‌ها به‌طور منظم و با قاعده برای گزینش مقامات به منظور استخدام در بخش خدمات اجتماعی در امپراتوری چین به‌عمل می‌آمدند.

برای دست یافتن به این منصب‌ها رقابت شدیدی بین افراد جریان داشت و فرایند آزمون به صورت فکری و فیزیکی دقیق انجام می‌گرفت. خانواده‌های ثروتمند برای آمادگی دانش‌آموزان پسرشان در این امتحان، از معلم‌های خبره استفاده می‌کردند.

در طول قرن‌ها، امتحان و آزمون به‌طور فزاینده‌ای فرموله شد. در انتهای سلسلهٔ «چینگ»[4] در چین در قرن نوزدهم، آزمون‌ها بر محور توانایی افراد در نوشتن یک «انشای هشت‌بخشی»[5] در رابطه با تجربیاتشان طرح‌ریزی می‌شدند. امتیاز بیشتر به افرادی تعلق می‌گرفت که به مهارت‌های بلاغی (بیان مطلب با سخن شیوا و رسا)، ادبی و ساختار شاعرانه و مهارت خطاطی تسلط بیشتری داشتند.

تأکید روی ادبیات کلاسیک و نوشتن بلاغی، از ویژگی‌های ارزشیابی در مدارس معتبر، مانند مدرسهٔ «هارو»[6] در انگلیس در ابتدای قرن نوزدهم بوده است. از این آزمون‌ها برای ثبت‌نام دانش‌آموزان در مدرسه و برای ترفیع به درجات و سطوح بالاتر استفاده می‌شود. انگیزهٔ برگزاری آزمون در این مدارس تلاش برای ایجاد سیستم رقابتی و شایسته سالاری به جای ارزشیابی براساس امتیاز و حمایت افراد بالادستی بوده است. قرن بیستم با افزایش آزمون‌ها و امتحانات همراه بوده است. ایده این‌که یادگیری انسان می‌تواند به‌طور هدفمند سنجیده شود، برخاسته از مطالعهٔ علمی محققان روان‌شناسی است.

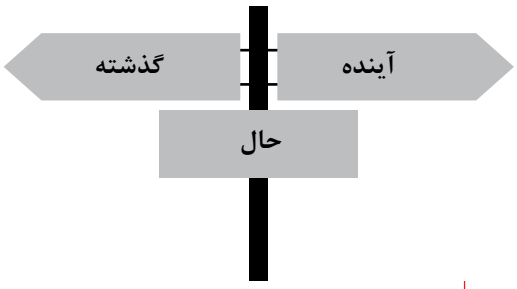
ارزشیابی به‌عنوان ابزار تشخیصی، ضعف و قوت دانش‌آموزان را از لحاظ هوش، شخصیت، استعداد و رفتار اجتماعی با ایجاد شناخت عاطفی و روانی شناسایی می‌کند. ابتدا از سال ۱۹۳۰، هیئت آموزش در آمریکا تطبیق آزمون‌های استاندارد شده را آغاز کرد و در مقیاس کلان و بین‌المللی، «یونسکو» و دیگر مراکز آموزشی به مطالعات مقایسه‌ای دستاوردها و یافته‌های دیگر ملت‌ها پرداختند. در ادامه، ایده‌های مربوط به شاخص‌ها و استانداردهایی که در قرن نوزدهم در آموزش ایجاد شده بودند، در انتهای قرن بیستم به یک محرک قدرتمند برای اصلاح برنامه و فرایند ارزشیابی مبدل شدند. براساس نتایج مقایسه‌های بین‌المللی انجام شده و با استقرار یک چارچوب بین‌المللی برای اهداف یادگیری و جدول رتبه‌بندی از عملکرد دانش‌آموزان نتایج قابل قبولی به‌دست آمد. اگرچه این نتایج واکنش‌های شدید را از سوی مدعیان این حوزه به دنبال داشته است.

دو موضوع اصلی در این واکنش‌ها نمایان بود:

موضوع اول به این نکته اشاره داشت که ارزشیابی باید به‌صورت طبیعی و معتبر صورت گیرد. بر این اساس، آزمون‌ها تمایل دارند که جنبه‌های جامع یادگیری را نادیده بگیرد که در این حالت با موقعیت زندگی واقعی نامرتبط است.

موضوع دوم ارزشیابی برای یادگیری است و تئوری‌هایی مانند تئوری‌های **گیبس** و **سیمسون** را مورد بحث قرار می‌دهند که ارزشیابی برای یادگیری فراگیرنده را قادر می‌سازد خودارزیابی، خودارجایی، خوداصلاحی و خودتحلیلی را انجام دهد.

به‌طور کلی، به فرایند تمرکز روی ارزشیابی در جهت یادگیری، به‌جای جمع‌آوری داده و اطلاعات، «ارزشیابی در خدمت یادگیری» می‌گویند. در نظام آموزشی به این شکل، معلمان تشویق می‌شوند، شیوهٔ ارزشیابی برای یادگیری را برای دانش‌آموزان تشریح کنند و به دنبال ایجاد روحیهٔ خودارزیابی در دانش‌آموزان از طریق دادن بازخورد ارزشیابی به آن‌ها باشند. با انجام این کار، دانش‌آموزان فرامی‌گیرند که چگونه یادگیری‌شان را ارزیابی کنند، چگونه توانایی‌های خود را توسعه دهند، چطور فعالیت‌های خود را مورد قضاوت قرار دهند، و در ادامه چه کاری انجام دهند تا یادگیری‌شان توسعه یابد. توانایی دانش‌آموزان در ارزیابی کارهایشان، با به‌دست گرفتن کنترل یادگیری‌شان مرتبط است. همهٔ این موارد پایه‌گذار یک اصل برای یادگیری مادام‌العمر می‌باشد. به‌طور کلی، مدافعان روش ارزشیابی برای یادگیری، ادعا می‌کنند اگر این روش به درستی اجرا شود، می‌تواند افراد شایسته و توانمند تربیت کند.



سیاست‌گذاری ارزشیابی و سیاست‌های پیاده‌سازی

سیاست‌گذاری در آموزش چندان آسان نیست. به دلیل اینکه سیاست‌ها غالباً متأثر از اهداف متناقض و مختلفی هستند که به واسطهٔ ذی‌نفعان در سطوح محلی، منطقه‌ای و بین‌المللی ارزش‌گذاری می‌شوند، سیاست‌گذاران مجبور هستند، در راستای به دست آوردن رضایت گروه‌های دارای علاقه‌های متفاوت، با آن‌ها بحث و گفت‌وگو داشته باشند. سیاست‌ها غالباً یک سلسله از ارزش‌ها و جهات آموزش را منعکس می‌کنند. اگر بر سیاستی تأکید می‌شود، هیچ ضمانتی وجود ندارد که آن سیاست به‌صورت صادقانه در هر کلاسی در نظام آموزشی اجرا شود. از آنجا که سیاست بر زمینه‌های مختلفی تأثیر می‌گذارد، روند تهیهٔ برنامه غالباً بسیار پیچیده است. برای مثال، توسعه‌دهندگان برنامهٔ درسی سیاست را طبق فرض‌های متداول از یک مجموعه از موضوعات مرتبط تفسیر می‌کنند. در حالی‌که مسئولان ارزشیابی، ابزارهای مربوط به آزمون را براساس نقطه نظرها، منابع و ظرفیت‌های موجود طراحی می‌کنند. همچنین منتشر کنندگان، مواد آموزشی را با نگاه به قابلیت عرضه در بازار توسعه می‌دهند.

در این میان، مدارس سیاست‌های برنامه را با اخلاق، صفات و شخصیت انسان تطبیق می‌دهند و معلمان به‌عنوان مهم‌ترین رکن اجرایی در فرایند آموزش، به‌وسیلهٔ طراحی و تدریس درس‌ها به شیوه‌ای که عقاید آموزشی‌شان و درک‌شان از نیازها ایجاب می‌کند، برنامه‌ها را اجرا می‌کنند. دانش‌آموزان نیز به شیوه‌ای بسیار فردی یاد می‌گیرند. معنی این شرایط آن است که مقصود سیاست ممکن است به واسطهٔ تغییر زمان دچار تغییر شود. به عبارت دیگر، سیاست‌هایی که آن‌ها در کلاس پیاده کرده‌اند، در طول زمان دستخوش تغییر می‌شود.

* پی‌نوشت‌ها

1. Trends in International Mathematics and Science Study
2. Programme for International Student Assessment
3. International Association for the Evaluation of Educational Achievement
4. Qing
5. “eight-legged” essay
6. Harrow

* منابع

1. Berry, R. (2006). Activating learners using the learner autonomy approach: An action research on the relevance of teaching to classroom practice. Curriculum Perspectives, 26(3), 34–43.
2. Berry, R. (2008a). Assessment for learning. Hong Kong: Hong Kong University Press.
3. Berry, R. (2008b). From theory to practice: Curriculum for autonomous learning. In M. F. Hui &
4. D. Grossman (Eds.), Improving teacher education through action research. New York: Routledge.
5. Berry, R. (2011). Assessment trends in Hong Kong: Seeking to establish formative assessment in an examination culture. Assessment in Education: Principles, Policy & Practice, 18(2).
6. Biggs, J., & Watkins, D. A. (2001). The paradox of the Chinese learner and beyond. Hong Kong: Comparative Education Research Centre.
7. Black, P. J. (1998). Testing: Friend or foe? Theory and practice of assessment and testing. London: The Falmer Press.
8. Broadfoot, P. (2009). Foreword. In C. Wyatt-Smith & J. J. Cumming (Eds.), Educational assessment in the 21st century: Connecting theory and practice (pp. v–xii). New York: Springer.
9. Bush, G. (2006). Learning about learning: From theories to trends. Teacher Librarian, 34(2),14–18.Clarke, S. (2001). Unlocking formative assessment. Abingdon: Hodder & Stoughton.
10. Gibbs, G., Simpson, C., & Macdonald, R. (2003). Improving student learning through changing assessment – a conceptual and practical framework.
11. Paper presented at the European Association for Research into Learning and Instruction Conference, Padova, Italy, Aug 2003.
12. Gipps, C., McCallum, G., & Hargreaves, E. (2000). What makes a good primary school teacher? Expert classroom strategies. London: Routledge Falmer.
13. Hsü, I. C. Y. (2000). The rise of modern China. New York: Oxford University Press.
14. McTighe, J., & Wiggins, G. (2004). The understanding by design handbook. Alexandra, VA: Association for Supervision and Curriculum Development.
15. Morris, P., & Adamson, B. (2010). Curriculum, schooling and society in Hong Kong. Hong Kong: Hong Kong University Press.
16. Roach, J. (1971). Public examinations in England 1850–1900. Cambridge: Cambridge University Press.
17. Sadler, D. R. (1998). Formative assessment: Revisiting the territory. Assessment in Education, 5(1), 77–84.
18. Stiggins, R. J. (2004). Classroom assessment for student learning. Portland, OR: Assessment Training Institute.
19. Sutherland, G. (1996). Assessment: Some historical perspectives. In H. Goldstein & T. Lewis (Eds.), Assessment: Problems, developments and statistical issues (pp. 9–20). Chichester, England: Wiley.