

ارزشیابی در آموزش عالی: حرکت از جزمیت به کثرت در روش^۱

در حال حاضر در اکثر رشته‌های دانشگاهی، روش غالب ارزشیابی^۲ همان روش‌های سنتی امتحان و تکلیف کلاسی است، اگرچه استادان برخی رشته‌ها- هر چند محدود- به استفاده از دامنه وسیع‌تری از ارزشیابی اقدام نموده‌اند. هدف از این فصل کتاب، شکستن انحصار روشی در ارزشیابی و ترغیب و تشویق اساتید به گرایش به سمت استفاده از روش‌های بدیل در کنار روش‌های تثبیت شده ارزشیابی است. بنابراین این جستار را می‌توان به‌عنوان منبعی برای اساتید لحاظ نمود که هنگام ارزشیابی دانشجویان می‌توانند به آن رجوع نموده و از وجود دامنه گسترده‌ای از روش‌های ارزشیابی مطلع شوند و خود را به روش‌های متداول محدود نکنند. اگرچه این حقیقت هم وجود دارد که برخلاف تمایل زیادتر استادان به شکستن انحصار روش تدریس خویش و آزمایش روش‌های جدید تدریس، در خصوص به‌کارگیری روش‌های جدید ارزشیابی چنین تمایل و کششی وجود ندارد (روست،^۳ ۲۰۰۲). بنابراین به‌منظور افزایش اعتماد

۱. این جستار به سفارش و با همفکری پدیدآورنده اختصاصاً برای استفاده در این اثر به‌وسیله آقای دکتر ابراهیم طلایی، استاد دانشگاه تربیت مدرس به رشته تحریر در آمده است.

۲. در سراسر این فصل از واژه ارزشیابی به همان معنای سنجش آموخته‌ها استفاده می‌شود.

اساتید به روش‌های جدید ارزشیابی، نمونه‌هایی عینی از موارد موفق در این فصل ذکر شده است. ساختار این فصل بدین صورت است که ابتدا ضرورت روی آوردن به روش‌های متنوع ارزشیابی و فاصله گرفتن از جزمیت روشی تبیین می‌گردد و سپس روش‌های بدیل ارزشیابی در قالب شش چالش اصلی و نمونه‌های عینی بیان می‌شود.

۱-۱۰ تبیین ضرورت استفاده از روش‌های متنوع ارزشیابی

۱-۱-۱۰ نیل به اهداف مختلف ارزشیابی از دانشجویان

عمل ارزشیابی تحصیلی برخلاف آنچه به ظاهر تصور می‌شود، بسته به بافت و موقعیت استفاده شده، اهداف متنوعی دارد که شامل آگاهی از میزان یادگیری دانشجویان، اطلاع از میزان اثربخشی تدریس استاد، تضمین کیفیت، یادگیری مستمر و ... است. اگرچه در این کتاب تمرکز بر ارزشیابی به‌منظور سنجش آموخته‌های دانشجویان است، اما به‌دلیل هزینه‌بر و وقتگیر بودن فرایند ارزشیابی این امکان وجود دارد که دیگر اهداف مترتب بر ارزشیابی و امکان‌های بهره‌برداری نیز مورد نظر قرار گیرد. در طول برنامه کامل آموزشی برای یک درس یا یک رشته، نکته‌حائز اهمیت این است که استاد بتواند بین این اهداف مختلف، حالت متعادل و متوازی را ایجاد نماید. یکی از راه‌های نیل به این هدف به‌ناچار روی آوردن به کثرت روشی در خصوص نحوه ارزشیابی از دانشجویان است و نمی‌توان با اتکا به یک یا دو روش متداول به این اهداف چند لایه و متنوع دست یافت.

۱-۱-۲ همخوانی با تنوع و سطوح مختلف اهداف یادگیری در آموزش عالی

اهداف یادگیری در رشته‌ها و مقاطع مختلف آموزش عالی تنوع و پیچیدگی خاص خود را دارد و دسته‌بندی مرسوم اهداف به اهداف شناختی، مهارتی و عاطفی تنها بخشی از این تنوع را بیان می‌کند. چرا که در هر یک از این دسته‌بندی‌ها، سطوح مختلفی را برای اندیشه (شناخت)، عمل (مهارت) و احساس (عاطفه) در مراحل مختلف یک دوره تحصیلی می‌توان قائل شد. برای هر یک از سطوح نیز برنامه‌ریزان درسی از افعال ویژه‌ای استفاده می‌کنند (بلوم^۱ و همکاران، ۱۹۵۶؛ کراول^۱؛ ۲۰۰۲). برای مثال «تشخیص دادن» و «توصیف کردن» را

1. Bloom

برای سطوح پایین تر اهداف به کار می‌برند و افعال «تعمیم دادن» و «بازاندیشی و تحلیل» را برای سطوح پیشرفته تر استفاده می‌کنند. بنابراین لازم است استادان دامنه وسیع تری از روش‌ها و راهبردهای ارزشیابی را در برنامه خود لحاظ کنند تا با تنوع و پیچیدگی اهداف یادگیری همخوانی به وجود آید؛ روش‌های محدود متداول ارزشیابی تحصیلی از ایجاد چنین همخوانی ناتوان است. همچنان که رامسدن^۲ (۲۰۰۳) بیان داشته است: «اگرچه به‌کارگیری روش‌های متنوع ارزشیابی از نظر اجرایی ممکن است سخت به نظر آید، چنین راهبردی برای دانشجویان گستره وسیع تری را برای بروز دانش کسب شده فراهم می‌کند و این پتانسیل را نیز به وجود می‌آورد تا استاد تصویر صحیح‌تر - اگرچه پیچیده‌تری - از میزان پیشرفت تحصیلی و دستاوردهای دانشجویان به دست آورد». در واقع روش‌های ارزشیابی سنتی تنها بخشی از دانسته‌ها، مهارت‌ها و نگرش‌های دانشجویان را منعکس می‌کند.

۱۰-۱-۳ اصل فراگیر شدن ارزشیابی

هر یک از سبک‌ها و قالب‌های ارزشیابی به‌تنهایی می‌تواند به نفع گروهی از دانشجویان و ضرر گروهی دیگر باشد. بنابراین به‌کارگیری روش‌های متنوع یا فراهم کردن امکان انتخاب برای دانشجویان را می‌توان اتخاذ رویکردی فراگیر در ارزشیابی تحصیلی تلقی نمود. هر چند در آموزش نیز رویکرد فراگیری که بتواند قابلیت‌ها و خصوصیات همه دانشجویان را لحاظ کند مورد تأکید متخصصان آموزش عالی است. برخی روش‌های ارزشیابی مانع از نشان دادن یادگیری‌ها به‌وسیله گروهی از دانشجویان می‌شود. به‌عبارت دیگر، سبک‌های یادگیری مختلف، سبک‌های آشکارسازی یا بازنمایی یادگیری متفاوتی را اقتضا می‌کند. آیزنر^۳ (۱۹۹۴) از زاویه‌ای دیگر - و با تأکید بر ضرورت مورد بحث - آموزشگران را به اتخاذ رویکرد کثرت‌گرایی در ارزشیابی ترغیب می‌کند. او معتقد است اگر قرار باشد مواد و منابع آموزشی را بر اساس میزان پرداختن و به‌کارگرفتن شکل‌های متنوعی از بازنمایی معنا و انتقال آن، مورد ارزیابی قرار دهیم - که اصل شایسته‌ای است - لازم و ضروری است همان طیف وسیع اشکال بازنمایی یا بیان دانسته‌ها و آموخته‌ها را برای یادگیرنده نیز لحاظ کنیم (آیزنر، ۱۹۹۴: ۱۸۲). اما آنچه در واقعیت اتفاق می‌افتد اتکا به قالب شفاهی برای تدریس و انحصار به قالب زبانی (شفاهی یا

1. Krathwohl
2. Ramsden
3. Eisner

کتبی) برای ارزشیابی است و دیگر اشکال انتقال معنا یا بازنمایی آموخته‌ها (معانی) همانند دیداری، شنیداری و حرکتی مورد توجه قرار نمی‌گیرند. به زبان دیگر «تفاوت‌های فردی» نه تنها در آموزش و روش‌های تدریس خود را جلوه‌گر می‌نماید، بلکه در ارزشیابی نیز از اهمیت ویژه برخوردار است؛ اگرچه اغلب مورد غفلت قرار می‌گیرد و با اتخاذ رویکرد جزم‌گرای ارزشیابی تنها به روش‌های محدود و متداول بسنده می‌گردد. همزمان و همراستا با آیزنر، گاردنر^۱، (گاردنر، ۱۹۹۹) نیز با ارائه نظریه «هوش چندگانه» بر اصل تفاوت‌های فردی هم در مرحله آموزش و هم در مرحله سنجش آموزش تأکید مؤکد می‌ورزد.

۱۰-۱-۴ احساس و انگیزه دانشجویان

پژوهش‌های اندکی در خصوص احساس و ادراک دانشجویان از روش‌های مختلف ارزشیابی انجام شده است، اما طی تحقیق مروری که استروون^۲ و همکارانش (۲۰۰۲) انجام دادند، به این جمع‌بندی رسیدند که به‌طور کلی دانشجویان در خصوص به‌کارگیری روش‌های «بدیل» ارزشیابی نگرش و احساس مثبتی دارند. برخی از ویژگی‌های این روش‌های بدیل که مورد رغبت دانشجویان قرار گرفت عبارتند از: تلفیق ارزشیابی با فرایند تدریس، تعیین تکالیفی که «واقعی» و «معنادار»^۳ باشند و اجازه دخالت به دانشجویان به‌عنوان مشارکت‌کنندگان آگاه و فعال در فرایند ارزشیابی. در این تحقیق نشان داده شد که دانشجویان تنوع روشی در ارزشیابی را به این دلیل معتبرتر می‌دانستند که بدین طریق کیفیات، شایستگی‌ها و توانمندی‌های مختلف آنان در ارتباط معناداری با زندگی واقعی مورد اندازه‌گیری قرار می‌گرفت و این بسیار فراتر از صرفاً گذراندن درس یا قبولی در امتحان است. یافته‌های این تحقیق مؤید این نظر است که انگیزش درونی دانشجویان - که به‌مراتب اثرگذار و ماندگارتر از انگیزش بیرونی است - به‌واسطه به‌کارگیری تکالیف و آزمون‌هایی تحریک و تقویت می‌گردد که آن تکالیف ذاتاً ارزشمند باشند؛ همانند تکالیف «واقعی» و معنادار که متأسفانه روش‌های سنتی ارزشیابی به‌ندرت از معیار «واقعی بودن» برخوردارند. برخی از نمونه‌های واقعی ارزشیابی عبارتند از: نوشتن برنامه برای بازاریابی یک شرکت، تدوین پیشنهاد تحقیق برای اخذ حمایت مالی از مراکز پژوهشی، تدوین خبر برای خبرگزاری‌ها و طراحی خبرنامه.

1. Gardner
2. Struyven
3. Authentic

۱۰-۱-۵ قابلیت اجرا

افزایش روزافزون تعداد دانشجویان حجم کاری استادان را افزایش داده است. برخی متخصصان مانند گیبز^۱ (۲۰۰۶) بر این باورند که شاید بتوان با افزایش پذیرش دانشجو و تدریس در گروه‌های بزرگ، به پدیده «مزیت مقیاس»^۲ در آموزش دست یافت. اما در ارزشیابی افزایش تعداد دانشجو نه تنها مدیران را به آن مزیت مورد انتظار نمی‌رساند، بلکه به واسطه لزوم ارزیابی کار تک تک دانشجویان، که زمان زیادی را به خود اختصاص می‌دهد، از مزیت مورد انتظار دور می‌شویم. مزیت مقیاس یا «صرفه‌جویی‌های مقیاس» مفهومی است که در علم اقتصاد - به خصوص اقتصاد خرد - به کار می‌رود و به این مضمون اشاره می‌کند که یکی از راه‌های کم نمودن هزینه‌های تولید، افزایش میزان (تعداد یا مقیاس) محصول است. به زبان دیگر، افزایش تعداد خروجی، عاملی برای کاهش هزینه‌های یک بنگاه تجاری است. در آموزش عالی نیز به واسطه هزینه‌بر بودن آموزش، یکی از راه‌های کاهش این هزینه را افزایش محصول یا همان دانشجو می‌انگارند، اگرچه بر این رویکرد نسبت به آموزش نقدهای جدی و عمیقی وارد شده (ویلیامسن و کوفیلد^۳، ۱۹۹۷)، اما اتخاذ چنین نگاه تجاری نه تنها در ایران بلکه در سطح بین‌الملل نیز وجود دارد (بارنت^۴، ۱۹۹۰). در حوزه ارزشیابی همان‌طور که در بالا بیان شد، افزایش تعداد دانشجو برای نظام آموزش عالی محدودیت ایجاد می‌کند و حجم کاری استادان را به شدت افزایش می‌دهد. یکی از راه‌های مواجهه با این چالش که قابلیت اجرایی هم داشته باشد، تمسک به کثرت روش‌های ارزشیابی است؛ چراکه روش‌های متداول همانند تکلیف کلاسی انرژی و وقت بسیاری از هیأت علمی طلب می‌کند (رامسدن^۵، ۲۰۰۳).

حال که در بندهای فوق در خصوص ضرورت به‌کارگیری روش‌های متنوع ارزشیابی صحبت شد، لازم است در انتهای این بخش دو نکته مهم برای آموزشگران تبیین و تشریح شود. نکته اول در خصوص مقاومت برخی استادان در برابر استفاده از روش‌های بدیل و جدید ارزشیابی است. یکی از استدلال‌های مورد اشاره این گروه شباهت و ایراداتی است که به رویه

1. Gibbs
2. Economy of scale
3. Coffield & Williamson
4. Barnett
5. Ramsden

نمره‌دهی و احتمال عدم رعایت یکپارچگی - یا همان کاهش پایایی - وارد می‌شود (مکلان^۱)، در پاسخ به این ایراد باید گفت چنین نگرانی‌هایی در خصوص روش‌های «باز پاسخ» در ارزشیابی‌های متداول نیز صادق است. بنابراین باید به دنبال راه‌هایی برای افزایش پایایی در ارزشیابی‌های بدیل باشیم که به این منظور افرادی مانند فالچیکو^۲ (۲۰۰۵) راهبردهایی را پیشنهاد می‌کنند. لازم است استادان با آزمودن این راهکارها، قابلیت اجرایی آنان را مورد ارزیابی قرار دهند و در عین حال به دنبال خلق راهبردهای نوین باشند. برخی از این راهبردها عبارتند از: ارزشیابی اهداف یادگیری از طریق فعالیت‌های مختلف دانشجویی، آموزش دقیق و عمیق ارزیاب‌ها (همانند دستیاران استاد)، به‌کارگیری چندین ممتحن و استفاده از ملاک‌های معین و دقیق.

نکته دوم هم این‌که در هرگونه نوآوری در آموزش از جمله به‌کارگیری روش‌های بدیل ارزشیابی، خطر مقاومت در برابر تغییر از طرف دانشجویان وجود دارد چراکه آنان با این روش‌ها آشنا و راحت نیستند و هنوز نسبت به معنادار بودن و ارتباطش با یادگیری‌شان قانع نشده‌اند (هونسل و همکاران^۳، ۲۰۰۶). علاوه بر این از آنجا که روش‌های بدیل ارزشیابی تغییراتی در بارم‌بندی، رتبه‌بندی دانشجویان در کلاس و دانشگاه، مهلت گزارش نمرات، جریمه‌های تأخیر و به‌طور کلی آیین‌نامه ارزشیابی دانشگاه به‌وجود می‌آورد، این خطر وجود دارد که در ابتدای اجرا هم دانشجویان و هم کارکنان دانشگاه را دچار سردرگمی کند. بنابراین پیشنهاد می‌شود که رویکرد «اجرای مرحله‌ای» اتخاذ گردد و همه افراد ذی‌دخل و ذی‌ربط - خصوصاً دانشجویان - به‌طور کامل و روشن توجیه شوند. به زبان دیگر لازم است تدابیری برای اجرا اتخاذ گردد که مشروعیت نظام جدید ارزشیابی از طریق راهبردهای اثربخش در درون نظام دانشگاهی به‌وجود آید.

۱۰-۲ راهبردهایی برای ایجاد تکثر در روش‌های ارزشیابی دانشجویان

اگرچه می‌توان سیاهه‌ای از روش‌های ارزشیابی ارائه نمود، به‌منظور پیشگیری از شکل‌گیری این تلقی که چنین سیاهه‌ای می‌تواند جامع باشد و همچنین خطر این‌که استادان خود را به آن محدود کنند و مانع بروز خلاقیت فردی آنان شود، در این قسمت مباحثی در قالب چند راهبرد

1. Maclellan
2. Falchikov
3. Hounsell et al

اصلی برای قدم گذاشتن در حوزه توسعه روش‌های ارزشیابی مطرح می‌گردد. هدف از ارائه این راهبردها این است که استادان را نسبت به تفکر در خصوص رویکردهای بالقوه‌ای که می‌توان در هر رشته دانشگاهی در خصوص تنوع ارزشیابی دانشجویان اتخاذ کرد، ترغیب و تشویق نماید. این مباحث در قالب شش راهبرد اصلی تکثر روشی در ارزشیابی در دوره آموزش عالی مطرح می‌شود و انتظار می‌رود بدین طریق قوه تخیل و ابتکار استادان نسبت به سناریوهای متنوعی که می‌تواند برای ارزشیابی دانشجویانش ترسیم کنند، تقویت گردد.

۱۰-۲-۱ واقعی و معنادار نمودن ارزشیابی

افزایش میزان واقعی و معنادار بودن ارزشیابی، یکی از اهداف کثرت بخشی به روش‌های ارزشیابی در دانشگاه است. واقعی بودن ارزشیابی بدین معنی است که قادر باشد تصویری از خود نشان دهد که در آن دانش و مهارت‌های کسب شده به وسیله دانشجویان در دنیای واقعی کاربرد داشته باشد (مکلان^۱، ۲۰۰۱: ۳۰۸). همچنین ارزشیابی معنادار آن است که بافتمند بودن آموزش، ریشه در فرهنگ و بوم داشتن تعلیم و تربیت (یعنی آموختن وابسته به موقعیت است و در خلأ اتفاق نمی‌افتد) را به رسمیت بشناسد و این حقیقت را از خود متجلی نماید. بنابراین بازتولید دانش به وسیله دانشجو در آزمون غیر بافتمند^۲ (غیر وابسته به موقعیت) دانشگاهی، هیچ تضمینی را برای این که دانشجو بتواند آن دانش را در دنیای واقعی نیز به کاربرد، ایجاد نمی‌کند. البته در برخی رشته‌های دانشگاهی همانند فنی و حرفه‌ای، ارزشیابی در دنیای واقعی و محل کار انجام می‌شده است، اما تأکید مورد بحث در این قسمت بر امکان و توسعه ارزشیابی معنادار و واقعی در خارج از محیط کار یا در محیط دانشگاه است.

چند نمونه از این نوع ارزشیابی در این قسمت شرح داده می‌شود. پیکفورد^۳ (۲۰۰۶) نمونه‌ای را بدین شرح آورده است که ارزشیابی معنادار و واقعی برای ۳۰۰ دانشجوی سال اول دانشگاه در دانشکده فناوری اطلاعات به انجام رسیده است.

1. Maclellan
2. Decontextualised
3. Pickford

دانشجویان به گروه‌های پنج نفری مشاوران تجاری دسته‌بندی شدند و سناریوی تجاری از یک شرکت دامداری که با بحران مواجه شده و وضعیت بغرنج و پیچیده‌ای پیدا کرده بود، برای‌شان تشریح شد. آنان به‌عنوان مشاور باید به شناسایی و واکاوی مسائل شرکت فرضی پرداخته و یافته‌های خود را به‌صورت گروهی در فرصت سه هفته‌ای ارائه می‌کردند. گروه‌ها برای برنده شدن در عقد قرارداد فرضی با شرکت مزبور با هم به رقابت پرداختند. قرارداد فرضی با گروهی منعقد شد که توانسته بودند در رویکرد به‌کار گرفته شده بیش‌ترین اصالت و خلاقیت را از خود نشان دهند. این امر (موضوع) در پایان نیمسال طی برنامه ویژه‌ای که گروه‌ها باید کارهای خود را به‌ترتیب در مدت سه دقیقه در مقابل چند ممتحن ارائه کنند، مشخص شد. زمان اندک ارزشیابی (سه دقیقه برای هر گروه) نیز هم بهره‌وری را از نظر حجم کاری استادان و هم توان ارائه و اثرگذاری مشاوران تجاری آینده بر مشتریان فرضی‌شان را لحاظ کرده بود. روند ارزیابی هم به‌صورت چند ممتحنی انجام شد و علاوه بر استادان این درس، ممتحنان میهمانی از شرکت‌های تجاری محلی و کارآفرینان در دسترس، برای افزایش میزان اعتبار ارزشیابی دعوت شدند. ممتحنان هم تنها باید با لحاظ سه معیار اصلی مندرج در برگه‌ها، تأثیرپذیری خود از ارائه‌های دانشجویان را علامت بزنند. دانشجویان هم بلافاصله پس از ارائه، قادر به رؤیت برگه بازخوردشان از طرف ممتحنان بودند. پیکفورد اظهار می‌دارد که اگرچه این فعالیت در راستای هدف آموزشی درس مورد بحث یعنی رشد مهارت‌های تجاری و مشاوره‌ای آنان تعریف شده است، اما به وضوح نشان می‌دهد که چنانچه ارزشیابی تحصیلی به‌صورت دقیق و اثربخش طراحی شود، فواید جانبی بسیاری دارد که خود در متن اهداف تعلیم و تربیتند. برای مثال، کار در گروه به‌صورت فشرده و پیاپی، فراگیری از یکدیگر، جامعه‌پذیر شدن دانشجویان که خود عامل مهمی برای نگه‌داشت دانشجو و جلوگیری از ترک تحصیل و بی‌رغبتی آن‌هاست (تینتو^۱، ۱۹۹۳). در این فعالیت ارزشیابی یادگیری و ارزشیابی یک کل درهم تنیده‌اند (پیکفورد^۲، ۲۰۰۶). براساس نظر دانشجویان نتیجه این ارزشیابی همراه با رضایت و لذت بیش‌تری بوده است و به احتمال بسیار باعث ایجاد انگیزه نیز شده است، چرا که کار مورد انتظار را باید به‌صورت مرحله‌ای طبق معیارهای از پیش تعیین شده به جلو

1. Tinto
2. Pickford

ببرند و همین درگیر شدن آنان با فعالیت مورد انتظار، ماندگاری آنان در کار و نهراسیدن از فعالیت بزرگ شده بود.

رویکردهای دیگری - که شاید دامنه محدودتری هم داشته باشند- نیز برای خلق ارزشیابی واقعی و معنادار وجود دارد، همانند این که دانشجویان برای عناوین طرح‌های تحقیقی مراکز و شرکت‌ها، پیشنهاد تحقیق بنویسند و یا در مناقصه‌های واقعی شرکت کنند. البته این نمونه‌ها فعالیت‌های پیچیده‌ای هستند که دانشجویان باید بر اساس رشته خویش، موضوع کار را به خوبی بشناسند، پیشینه آن را بدانند، با روش‌های مختلف آشنا شوند، عوامل ریسک یا خطر را تحلیل کنند. بنابراین این نوع ارزشیابی‌ها نه تنها واقعی و معنادارند، بلکه می‌توانند به عنوان شاخص‌های معتبری از یادگیری عمیق شناختی هم عمل کنند. برای کمک به کاهش حجم کاری استادان در چنین مواقعی محدودیت قائل شدن برای تعداد واژگان مجاز به وسیله دانشجویان می‌تواند کمک نماید.

نمونه دیگری از این نوع ارزشیابی، ایجاد وبگاه برای اهداف خاص است که به عنوان نمونه نایت و یورک^۱ (۲۰۰۳) این روش را ارزیابی کردند. آن‌ها پیشنهاد می‌کنند از آنجایی که ایجاد وبگاه کار پیچیده‌ای است، توصیه می‌شود به عنوان کار گروهی انجام شود، اگرچه برخی نرم‌افزارهای ایجاد وبگاه این کار را اندکی تسهیل کرده است. وبگاه‌ها این امکان را به وجود می‌آورند که هم دانشجویان توان خویش در نگارش برای مخاطبان متنوع را نشان دهند و هم میزان عمق دانشی خویش از موضوع درسی را بیان کنند. شواهد تحقیقی هم وجود دارد که نشان می‌دهد دانشجویان، امروزه این نوع فعالیت را بیش‌تر از تکالیف سنتی می‌پسندند (ربر^۲، ۲۰۰۵).

به‌طور کلی، تحقیقات نشان می‌دهند که انتخاب مسائل «واقعی» و تمرکز بر موارد خاص، امکان انعکاس بافت و موقعیت واقعی زندگی انسانی را بیش‌تر از امتحان، آزمون و تکالیف درسی سنتی فراهم می‌آورد. این نوع ارزشیابی، علاوه بر توان بالقوه خویش برای انگیزه‌مند نمودن دانشجویان از طریق انجام فعالیت معنادار، امکان ارزشیابی دامنه وسیع‌تری از مهارت‌ها و دریافت‌های دانشجویان را به وجود می‌آورد.

1. Knight and Yorke

2. Reber

۱۰-۲-۲ رعایت اصل اقتصاد در ارزشیابی (یا افزایش نسبت ستانده به نهاده)

اهداف برنامه‌های درسی در رشته‌های مختلف دانشگاهی اکثراً خواستار ارزشیابی از طیف و لایه‌های متنوعی از آموخته‌های دانشجویانند. این طیف از آموخته‌ها شامل مهارت‌ها و دانسته‌های حوزه خاص آن رشته و همچنین دانش و مهارت‌هایی هستند که صبغه عمومی‌تری دارند. روش‌های بدیل ارزشیابی که در این فصل بر آن‌ها تأکید می‌شود، امکان چنین امری را بیش‌تر فراهم می‌کند. برای نمونه، (دان و همکاران^۱، ۲۰۰۴) بر اهمیت ارزشیابی که هم تسلط دانشجو بر مبانی دانشی و نظری موضوع درس و هم چیرگی در مهارت و اجرایی نمودن مهارت کسب شده را ارزیابی می‌کنند، تأکید می‌ورزند.

برای تقریب ذهن نمونه‌ای از این روش‌های ارزشیابی بدیل بیان می‌گردد:

«در درس تاریخ هنر، استاد از دانشجویان می‌خواهد به‌عنوان ارزشیابی تکوینی، نمایشگاه هنری را به‌صورت مجازی روی وب ترتیب دهند. در این فعالیت دانشجویان باید نظر و عمل را با هم تلفیق کرده، از طریق رسانه نمایشگاه، تفاسیر خود از آنچه در این درس فراگرفته‌اند را نمایش دهند. دانشجویان با علاقه و انگیزه نمایشگاه‌های مجازی خویش را آماده می‌کنند و این فرصت را نیز غنیمت می‌شمرند که می‌توانند روی هنرمند خاصی، دوره تاریخی ویژه‌ای، یا سبک هنری مورد علاقه‌شان برای مدتی متمرکز شوند و نمایشگاه‌شان را به میل خویش برپا کنند. آنچه فراگیری دانشجویان را در این فعالیت راهبری می‌کرد، فرایند یادگیری و تعاملات آنان بود نه صرف محصول کار و ارائه آن به‌عنوان تکلیف درسی (کریستوفر بیل، گزارش شده در مک‌دول و همکاران^۲، ۲۰۰۵: ۱۲)».

یکی از مهارت‌هایی که به‌طور ضمنی در تکلیف درسی ارزشیابی می‌شود مهارت‌های فناورانه دانشجویان تا بتوانند محصولی تولید کنند که قدرت تعامل بصری معنادار و اثربخش با مخاطب داشته باشد. بنابراین راه‌های متنوعی برای ارائه بازخورد به دانشجویان هم از طرف همکلاسی‌های‌شان و هم از طرف استادان در نظر گرفته شد. این مورد نمونه خوبی از ارزشیابی واقعی و معنادار است که عناصر دیگری از مطلوبیت همانند ارزشیابی دامنه وسیعی از فراگرفته‌ها و همچنین دادن امکان انتخاب به دانشجو را در خود دارد. این در حالی است که همزمان هدف اصلی یعنی ارزشیابی مهارت‌های درس مورد نظر - تاریخ هنر - را هم در سطوح

1. Dunn & et al
2. Christopher Bailey, Reported in McDowell & et al

عمیق آن - و نه در سطوح اولیه شناختی - لحاظ کرده است. علاوه بر این ایجاد فرصت برای دادن بازخورد به دانشجویان در خصوص کارهایشان و آن هم در حین کار - و نه در پایان - اصل تلفیق صحیح ارزشیابی برای یادگیری و ارزشیابی از یادگیری را نیز رعایت کرده است.

۱۰-۲-۳ افزایش روایی آزمون‌ها یا پاسخگو نمودن آن‌ها به اهداف یادگیری در

سطوح بالاتر شناختی

یکی از موارد مورد مناقشه در خصوص روش‌های ارزشیابی سنتی - و در واقع پاشنه آشیل این نوع از سنجش - که تمرکز بیش‌تری بر بازتولید دانش دارند، پایین بودن درجه اعتبار یا روایی آن‌هاست. به‌خصوص آزمون‌های میان و پایان ترم دانشگاهی در معرض چنین اتهامی هستند، اما در عین حال در خصوص روش‌های جایگزین ارزشیابی در دانشگاه، همانند نوشتن مقاله، نقد و تکالیف نوشتاری دیگر، نگرانی‌هایی از باب گسترش دزدی ادبی و کپی‌برداری نیز مطرح است. از این‌رو آنچه هم‌اکنون در ارزشیابی دانشگاهی، اصل و بدون تغییر مانده است، همان آزمون‌های نیمسال دانشگاهی می‌باشد. بنابراین لازم است به‌دنبال راه‌هایی برای افزایش روایی این آزمون‌ها باشیم تا از ارزشیابی مفاهیم و لایه‌های پیچیده‌تر یادگیری به‌وسيله این آزمون‌ها - و نه فقط بسنده نمودن به یادگیری بازتولید دانش آموخته شده در طول نیمسال - اطمینان حاصل کنیم.

یک نمونه برای مواجهه با چالش روایی آزمون‌های دانشگاهی را جانسون و بارتمن^۱ (۲۰۰۶) در رشته تربیت معلم و با عنوان «آزمون تعاملی»^۲ بدین صورت آورده‌اند. دانشجویان در رایانه روبه‌روی‌شان، سه فیلم کوتاه که هر یک سناریویی از تدریس در کلاس درس واقعی بود، تماشا کردند. آن‌ها امکان دسترسی به متن گفتگوهای درون فیلم و اطلاعات بیش‌تر در خصوص کلاس‌های درس نمایش داده شده نیز داشتند. وظیفه دانشجویان این بود که موقعیت‌های کلاسی نمایش داده شده را توصیف و تحلیل کرده و سپس توصیه‌های خویش را برای معلم آن کلاس‌ها بنویسند. پس از این‌که دانشجویان اولین مرحله آزمون را طی کردند و نوشته‌های خویش را تحویل دادند، توصیه‌ها و راه‌حل‌های متخصصان همراه با تجربه‌هایشان به دانشجویان داده شد تا بتوانند ظرف یک هفته آن‌ها را با آنچه خود نوشته‌اند، مقایسه تحلیلی

1. Jonsson & Baartman
2. Interactive examination

و انتقادی کنند و در خصوص تفاوت‌ها دلایل متقن بیاورند و همچنین نیازهای یادگیری خویش را برای آینده با این مقایسه استخراج کنند. استفاده از سه سناریو کلاسی به‌جای یکی امکان بهبود میزان پایایی آزمون را نیز فراهم می‌آورد.

چنین رویکرد واقع‌گرایانه و معناداری که بافتی غنی برای ارزشیابی ایجاد نماید را می‌توان در رشته‌های دیگری که نیازمند حل مسئله به هر شکلی است، به‌کار برد. چنین روشی همچنین امکان ارزشیابی از لایه‌های عمیق‌تر شناختی را فراهم می‌آورد، چراکه بدین طریق دانشجویان باید عناصر کلیدی را از بطن سناریوی واقعی زندگی - یعنی کلاس درس - استخراج و آن‌ها را ارزیابی و تحلیل کنند. اما در روش سنتی ارزشیابی، سؤالات امتحانی، خود مواردی را برای بحث مشخص و تجویز می‌نماید (بود و فلاچیکو^۱، ۲۰۰۶). از طرف دیگر نمونه‌ای که در بالا ذکر شد را می‌توان به‌عنوان نمونه‌ای از آزمون تکوینی^۲ نیز برای خود ارزیابی دانشجویان تلقی کرد، چرا که بلافاصله پس از اتمام کار دانشجویان می‌توانند تحلیل‌های خود را با تحلیل‌های متخصصان مقایسه و محک بزنند. به‌طور خلاصه یافته‌های جانسون و بارتمن (۲۰۰۶) از اجرای این نوع ارزشیابی تعاملی این است که از منظر «همراستایی با هدف ارزشیابی، مقایسه‌پذیری، واقع‌گرایی و معناداری، لحاظ نمودن لایه‌های عمیق یادگیری و هم‌راستایی با اصل خود ارزیابی دانشجویان»، از کیفیت بالایی برخوردار بوده است.

آزمون‌های کتاب آزاد^۳ هم‌اکنون در برخی رشته‌ها مثل ادبیات انگلیسی کاربرد وسیعی پیدا کرده است. چنین آزمون‌هایی تأکید بر حفظ کردن را کاهش می‌دهد و سؤالات پیچیده‌تری را می‌توان طرح کرد (فریمن و لويس^۴، ۱۹۹۸) تا مهارت‌های لایه‌های عمیق‌تر یادگیری همانند حل مسئله، تفسیر، ارج نهادن به اثر ادبی، کاربرد و ارزیابی انتقادی را نیز مورد سنجش قرار داد. اما از طرف دیگر امکان دسترسی به کتاب در جلسه امتحان ممکن است دانشجویان را از ایجاد آمادگی برای امتحان وا دارد (میلر و همکاران^۵، ۱۹۹۸) یا این‌که در جلسه امتحان دسترسی به کتاب تمرکز را از برخی دانشجویان سلب کند که نتوانند عملکرد با کیفیتی را از خود نشان دهند (نایت و یورک^۶، ۲۰۰۳).

1. Boud & Falchikov
 2. Formative
 3. Open-book examinations
 4. Freeman & Lewis
 5. Miller & et al
 6. Knight & Yorke

یک راه بدیل ارزشیابی، برگزاری امتحان به صورت مطالعه موردی از پیش اطلاع داده شده^۱ است. برای مثال، مطالعه موردی دو هفته قبل از امتحان با جزئیات کامل در اختیار دانشجویان رشته مدیریت قرار می‌گیرد. آنان باید روی آن مطالعه متمرکز شده و تکلیف کوچکی در حد ۲۰۰ کلمه در خصوص برداشت کلی خود از مطالعه مورد نظر، پیش از موعد امتحان تحویل بدهند. روز امتحان، هم کار تحویل داده شده و هم نسخه‌ای دیگر از آن مطالعه موردی با اندکی تغییر به دانشجویان برگردانده می‌شود. سؤال اصلی امتحان این است که دانشجویان برخی جنبه‌های خاصی از مطالعه موردی را تحلیل دقیق نمایند و برای سازمان فرضی راه حل ارائه کنند. سؤال دوم، اطلاعات جدیدی را به شرایط سازمانی مورد مطالعه اضافه می‌کند که در آن مدیر مربوط باید در اسرع وقت اتخاذ تصمیم کند. دانشجویان باید پاسخ‌های مدلل و راه‌حل‌های جایگزین را برای رفع چالش به وجود آمده پیشنهاد دهند. در چنین ارزشیابی حتی امکان و فضایی برای بروز خطر بازتولید سطحی یادداشت‌های کلاسی به وسیله دانشجویان در جلسه امتحانی به وجود نمی‌آید. بلکه دانشجویان را برای کاربرد عملی و معنادار نظریه‌های مطالعه شده در طول نیمسال ترغیب و تشویق می‌نماید. چنین فراگیری بر اساس یافته‌های ولنچیک (۱۹۹۵) منجر به یادگیری عمیق‌تری از نظریه‌ها می‌گردد. برای علاقه‌مندان به استفاده اثربخش از مطالعات موردی در ارزشیابی، اثر لاینز (۲۰۰۵) توصیه می‌شود. او علاوه بر خلاصه نمودن فواید این روش در مقایسه با روش‌های متداول امتحان محور در آموزش عالی، نسبت به محدودیت‌های آن نیز مواردی را متذکر شده است. برای مثال افزایش زمان مورد نیاز برای آمادگی استاد، کاهش حجم موضوعات که در طول نیمسال فراگرفته می‌شود و خطر بی‌نتیجه بودن روش در صورتی که دانشجویان موفق به کشف نکات کلیدی در آن مطالعه موردی نشوند را می‌توان نام برد. بنابراین ارائه چند مطالعه موردی، خطر اخیر را کاهش می‌دهد و اگر باز هم دانشجویی با ارائه چند مطالعه موردی، هنوز نکات کلیدی آن را دریابد و مسیر غلط را پیش گیرد، لازم است میزان فهم و درک دانشجویان از مسئله و یا مهارت‌های حل مسئله آنان بازبینی شود.

نمونه‌ای از ارزشیابی‌هایی که در آنها تمرکز ارزیاب هم بر مهارت‌های عمومی آکادمیک دانشجویان، هم بر موضوع تخصصی رشته و هم بر مسائل روش‌شناختی است، استفاده از مقاله‌ای

چاپ شده در مجلات تخصصی و حذف عنوان، نویسنده و خلاصه مقاله است. دانشجویان باید برای مقاله مورد نظر عنوان و چکیده مناسب تدوین کنند و در انتها نیز نقد مددلی را برای مقاله بنویسند. برای مثال، در رشته‌های علوم طبیعی، در قسمت اول امتحان دانشجویان لازم است پیام اصلی مقاله، روش به‌کار گرفته شده و نتیجه‌گیری را از مقاله استخراج کنند و در قسمت دوم ارتباط و انسجام مطالعه مورد نظر، روش علمی به‌کار رفته، اعتبار یافته‌ها و میزان تطابق نتایج ارائه شده از تحقیق با شواهد موجود در آن را تحلیل نمایند. تنها هشدارى که در این نوع ارزشیابی وجود دارد، مسئله حقوق معنوی مقاله استفاده شده در آزمون است.

۱۰-۲-۴ ادغام ارزشیابی در تدریس

نگارنده تا اینجا تلاش نموده است نقش مهم ارزشیابی در بهبود یادگیری دانشجویان و آثار منفی آن در صورت به‌کارگیری روش‌های نادرست ارزشیابی را مورد تأکید ویژه قرار دهد. تنوع‌دهی به روش‌های ارزشیابی خود یکی از راهبردهای بهبود تدریس است، به عبارت دیگر «تدریس» و «ارزشیابی» در هم تلفیق شدنی‌اند، به این معنی که استاد در حین تدریس خویش دست به ابتکارات آموزشی می‌زند که از این ابتکارات، داده‌هایی برای ارزشیابی نیز به دست می‌آید. در این نوع از ارزشیابی، آن را به صورت فرایندی یا در فضایی فرایندی - و نه بروندادی - تعقیب کرده و ارزشیابی را در تدریس منحل یا ادغام می‌کنیم.

برای نمونه اگر در آزمون‌های چند گزینه‌ای - که به علت محدودیت‌شان برای کمک به یادگیری عمیق مورد انتقادند - تغییری به وجود آید، می‌توان آن‌ها را به عنوان آزمون‌های تکوینی که یادگیری دانشجویان و آموزش استاد را بهبود می‌بخشد، مورد استفاده اثربخش قرار داد، چرا که می‌توان از آزمون‌های چند گزینه‌ای برای ارزشیابی گروه کثیری از دانشجویان استفاده کرد و از طریق تصحیح به کمک رایانه در وقت صرفه‌جویی بسیاری نمود. گاردنر-مدوین^۱ (۲۰۰۶) برای این موضوع پیشنهاد اتخاذ رویکردی موسوم به «نمره‌دهی مبتنی بر میزان اطمینان دانشجو از جواب» را ارائه نموده‌اند که در آن تلاش بر این است که دانشجو به این باور برسد که حدس‌های صحیح در آزمون‌های چند گزینه‌ای هم ارزش دانش نیستند و یا این‌که پاسخ‌های نادرستی که با اطمینان به نادرستی آنان داده می‌شود،

1. Gardner-Medwin

نه تنها بی‌ارزش نیستند، بلکه مستحق توجه بیش‌ترند (گاردنر-مدوین، ۲۰۰۳). روش اجرایی آن هم بدین صورت است که دانشجویان علاوه بر ارائه پاسخ‌های‌شان به سؤالات، میزان اطمینان خود از پاسخ‌شان را هم ثبت می‌کنند. بیش‌ترین امتیاز برای زمانی است که جواب صحیح را با اطمینان کامل انتخاب می‌کنند (۳)، در حالی که اگر با اطمینان بالا پاسخ سؤالی را علامت بزنند و آن پاسخ نادرست باشد، نمره منفی زیادی را خواهد داشت (۶-). همچنین اگر دانشجویی نسبت به پاسخ سؤالی اظهار اطمینان کامل نکند، برای پاسخ صحیح، امتیاز کم‌تری نصیب خود می‌کند؛ اگرچه برای پاسخ نادرست جریمه نمی‌شود. چنین نظام نمره‌دهی متفاوتی، دانشجو را ترغیب می‌کند نسبت به پاسخی که می‌دهد با دقت تفکر کند؛ اگرچه چنین رویکردی در ارزشیابی متداول نیست، چراکه عمدتاً روش‌های مورد استفاده متأسفانه به نحوی دانشجویان را به سمت پوشاندن فهم عمیق و بروز یادگیری سطحی سوق می‌دهد. بدین روش همچنین دانشجویان تلاش می‌کنند سؤالات سخت را نیز پاسخ گویند؛ چراکه اگر هم با اطمینان پایین آن‌ها را جواب دهند، به‌واسطه پایین بودن اطمینان از جواب امتیازی را از دست نمی‌دهند. بازخورد دانشجویان در این خصوص نیز گویای نظر مثبت آنان نسبت به عادلانه و مفید بودن آن است. نرم‌افزار این نظام ارزشیابی نیز به‌صورت رایگان قابل احصاست.^۱

روش ارزشیابی بدیلی که می‌توان از آن برای تقویت اثربخشی یادگیری دانشجویان بهره جست، استفاده از مجله به‌معنای یادداشت‌های تحلیلی^۲ است (گیبز،^۳ ۱۹۹۵). یادداشت‌های تحلیلی را می‌توان برای مقاصد مختلفی همانند پیگیری قدم به قدم امور در طرح‌های بزرگ و پیچیده و بازنگری نقادانه مراحل آن به‌کار برد (نایت و یورک، ۲۰۰۳). این روش به‌خصوص برای دوره‌های تخصصی کاربرد فراوان و اثربخش دارد، چراکه دانشجویان فرصتی برای بازاندیشی جوانب مختلف پیشرفت کار خویش پیدا می‌کنند. البته ممکن است آنان فراتر از جنبه توصیفی امور نروند (لاینز،^۴ ۲۰۰۵) و لازم است استادان آنان را برای کشف و تعمیق نظریه‌ها و الگوهای عملی ترغیب نمایند. در واقع از دانشجویان انتظار می‌رود در این یادداشت‌های تحلیلی، پیشینه کار را در ارتباط با رویکردهای معاصر نقد و بررسی نمایند تا

1. www.ucl.ac.uk/lapt/

2. A Journal, or a log, or diary

3. Gibbs

4. Lines

برای پیشرفت حرفه‌ای خویش در دوره‌ای که در حال گذراندن آن هستید، یافته‌های متقنی ارائه دهند (دان و همکاران، ۲۰۰۴). بخش تحلیلی چنین یادداشت‌های روزانه‌ای نیز باید اساس و مبنای ارزشیابی استاد از کار دانشجو قرار گیرد.

یادداشت‌های تحلیلی علاوه بر دوره‌های تخصصی، در ارزشیابی دروس دیگر که نیازمند ثبت و ضبط پیشرفت پژوهش آن درس است، نیز کاربرد دارد. در این موارد، از یادداشت‌های تحلیلی به‌عنوان مبنای بخشی از گفتگوها و مباحث تکوینی که در طول ترم دانشجو و استاد در جلسات مشاوره با یکدیگر دارند، نیز استفاده می‌گردد و یا حتی به‌عنوان بخشی از تکلیف پایان ترم نیز می‌تواند تحویل استاد شود.

به‌طور کلی، یادداشت‌های تحلیلی به‌عنوان ابزار یادگیری و ارزشیابی ارزش والایی دارد، اما استادان لازم است هدف از استفاده، نحوه ارزشیابی از طریق یادداشت‌ها و معیارهای ارزشیابی را به‌طور شفاف برای دانشجویان بیان کنند و حتی برای سازمان‌دهی یا یادداشت، ساختار پیشنهادی را معرفی کنند (مون^۱، ۱۹۹۹). گیبز (۱۹۹۵) هم برای اثربخش بودن یادداشت‌ها بر این نکات تأکید می‌کند: دستورالعملی برای طول (یا حجم) یادداشت‌های تحلیلی، سبک نگارش، فرصت بازنگری و تصحیح یادداشت‌ها، آنچه قرار است ارزشیابی شود، محرمانه بودن برخی اطلاعات و ارجاعات به افراد از جمله خود استاد و دیگر دانشجویان.

البته یادداشت‌های تحلیلی زمانی که به‌عنوان ابزاری برای ارزشیابی دانشجویان استفاده شود، با چالش احتمال کاهش پایایی این ابزار مواجه است (مون، ۱۹۹۹). بنابراین عاقلانه است که از دانشجویان خواسته شود تنها بخش‌های انتخابی یادداشت‌های خود را به‌عنوان ارزشیابی نیمسال تحویل دهند، اگرچه تجربه نشان داده است برای افزایش تعداد دانشجویانی که اقدام به نگارش منظم یادداشت می‌کنند، لازم است با ارزشیابی نیمسال گره زده شود. برای اطلاعات کامل‌تر در خصوص این روش به مون (۱۹۹۹) مراجعه کنید.

روش دیگری از ارزشیابی که بر یادگیری دانشجویان اثر مثبتی دارد «مناظره‌های علمی است که مورد سنجش قرار می‌گیرد»^۲. در این روش دانشجویان به‌صورت انفرادی یا گروهی نظریات، رویکردها و نگرش‌های بدیلی را نسبت به موضوع مورد بحث بیان و برای آن

1. Moon

2. Assessed debates

استدلال می‌کنند. این روش ارزشیابی، کاربرد وسیعی در رشته‌های مختلف دارد و دانشجویان را به تفکر انتقادی از طریق نگریستن به موضوع از زوایای مختلف و سنجش شواهد علمی ترغیب می‌کند. برای ارزیابی عملکرد دانشجویان می‌توان هم مهارت‌های ارائه آنان و هم کیفیت استدلال و شواهدی که ارائه می‌کنند را مورد سنجش قرار داد. به‌منظور افزایش مشارکت حضار و ایجاد انگیزه برای مناظره کنندگان می‌توان اعلام نمود که فرایند مناظره، روش و روند استدلال‌های طرفین به رأی نهایی عموم - همه افراد حاضر در جلسه مناظره - گذارده خواهد شد.

در نهایت، مشارکت کلاسی دانشجویان نیز می‌تواند مورد ارزیابی استاد قرار گیرد تا بدین وسیله پشتوانه‌ای برای یادگیری آنان شود؛ چراکه درگیر شدن فعال دانشجو با مواد و منابع آموزشی، یادگیری را افزایش می‌دهد، به‌خصوص برای دانشجویانی که در کلاس غیرفعالند و توان/علاقه مفهوم‌سازی فردی از موضوعات درسی را ندارند (بیگز^۱، ۲۰۰۳). تحقیقات نشان داده است نمره‌دهی به مشارکت کلاسی دانشجویان و دادن بازخورد به آنان در طول ترم در خصوص وضعیت‌شان، مشارکت فعال آنان را افزایش می‌دهد (دانسر و کامونیا^۲، ۲۰۰۵). مطالعه مذکور همچنین توصیه به خود ارزیابی و دیگر-ارزیابی (یعنی ارزیابی همکلاسی‌ها) از میزان مشارکت کلاسی هر یک از دانشجویان می‌نماید که در آن علاوه بر حضور در کلاس، معیارهای آمادگی برای کلاس، سهیم شدن در مباحث کلاسی، مهارت‌های گروهی و ارتباطی نیز لحاظ می‌گردد.

۱۰-۲-۵ گره زدن ارزشیابی تکوینی و پایانی

مفهوم ارزشیابی برای یادگیری (در مقابل ارزشیابی از یادگیری) بر نقش ارزشیابی تکوینی در بهبود یادگیری دانشجویان تأکید می‌ورزد، اما نگرانی که معمولاً در خصوص ارزشیابی‌های تکوینی (همانند ارزشیابی میان ترم یا ارزشیابی‌های ماهیانه) وجود دارد، این است که دانشجویان آن‌ها را به اندازه ارزشیابی پایانی جدی تصور نمی‌کنند، بنابراین اندیشیدن در خصوص چگونگی گره زدن این دو نوع ارزشیابی با هم، ارزشمند است. یکی از این روش‌ها، می‌تواند تنوع‌بخشی به روش‌های ارزشیابی موجود باشد.

1. Dancer and Kamvounias
2. Biggs

برای نمونه «ارزشیابی جورچینی»^۱، نوآوری در روش‌های ارزشیابی است که برای تقویت درک انتقادی دانشجویان در رشته‌های مختلف علوم انسانی، هنر، علوم اجتماعی، حرفه‌های خدماتی و حتی علوم طبیعی - خصوصاً در مسائل پرمنازعه مثل مباحث اخلاق علمی - کاربرد دارد (اسکوگینز و ونیتر^۲، ۱۹۹۹). در این روش، دانشجویان باید در طول ترم تکالیف نوشتنی را انجام دهند که بسته به رشته دانشگاهی متفاوت خواهد بود، برای مثال نقد کتاب، شعر، داستان، تابلو، واقعه تاریخی، یا واکنش علمی در خصوص سخنرانی یا مقاله‌ای چاپ شده را می‌توان نام برد. از بعد تکوینی، چنین ارزشیابی بدین صورت توجیه می‌گردد که دانشجویان لازم است نوشته‌های‌شان را برای هم‌کلاسی‌های خود ارسال و پاسخ آنان را درباره کارشان دریافت و در صورت توجیه، اعمال کنند. همچنین آنان از پیشرفت امور در درس مورد نظر تکالیف و بازخوردهای‌شان، یادداشت‌های تحلیلی منظمی را تدوین می‌کنند. در نیمه ترم، دانشجویان تمام نوشته‌های خویش (تکالیف و یادداشت‌های تحلیلی) را مرور و بازبینی می‌کنند تا امکان استخراج «مضمونی مشترک»^۳ از آن‌ها را بررسی نمایند. این مضمون یا مضامین با دیگر هم‌کلاسی‌ها مطرح و نظرخواهی می‌گردد. تکلیف نهایی ترم نیز منتخبی از کارهای تحویل داده شده پیش است که امکان دارد دوباره مورد بازبینی دانشجو قرار گیرد تا چارچوب تفسیری مدونی را به خود گیرد و از طریق این چارچوب، دانشجو ارتباط مضمون مشترک کشف شده را با هر یک از تکالیف نوشته شده در طول ترم تشریح نماید (اسکوگینز و ونیتر، ۱۹۹۹: ۴۸۷).

ارزشیابی جورچینی به منظور پر کردن شکاف بین ارزشیابی تکوینی جزیره‌ای و نامرتب (همانند آزمون میان ترم، یادداشت‌های تحلیلی و ...) و ارزشیابی پایانی (همانند تکلیف پایان ترم) است. تفاوت بین تکالیف درسی در طول ترم و اولین نوشته‌ها برای ارزشیابی جورچینی در این است که اولی بیش‌تر دانشجو را به مدیریت مطالب درسی و منسجم ساختن آن‌ها ترغیب می‌نماید در حالی که دومی ماهیتی اکتشافی^۴ و نقادانه دارد که اتفاقاً هم‌راستا با مفاهیم و ماهیت دانش و تحلیل موضوعات آن (مانند تحلیل تجارب، آثار هنری ...) از جوانب مختلف است (پارکر^۵، ۲۰۰۳).

۱. (Patchwork texts assessment) نوعی از ارزشیابی که همانند یک جورچین، تکه‌های مختلفی دارد و هر کدام جایگاه خاص خویش را دارند و چنانچه به‌درستی در کنار هم قرار گیرند، تصویری کامل، واضح و همه‌جانبه از مفهوم مدنظر ارائه می‌کند.

2. Scoggins and Winter

3. Unifying theme

4. Exploratory

5. Parker

واژه جورچین از این منظر استفاده شده است که چنین ارزشیابی‌ای باید تکه‌های مختلف جدا از هم را در یک کل واحد کنار هم بیاورد. فواید بسیاری را برای این روش ارزشیابی و تأثیر آن بر یادگیری برشمرده‌اند که ما در اینجا تنها به ترغیب دانشجویان به نوشتن برای ارزشیابی‌های تکوینی و جدی گرفتن آن‌ها بسنده می‌کنیم و علاقه‌مندان را به خواندن کتاب اسکوگینز و وینتر (۱۹۹۹) ارجاع می‌دهیم.

نمونه‌ای دیگر از ارزشیابی که استاد می‌تواند آن را به‌طور مستقل و جامع برای ارزشیابی دانشجو در طول یک دوره تحصیلی لحاظ کند و هم با تدابیری آن را با ارزشیابی پایانی گره بزند، برگزاری آزمون‌های کوچک و منظم در فواصل زمانی کوتاه است. این آزمون‌ها که کوئیز^۱ نیز نامیده می‌شود، جزء ارزشیابی تکوینی محسوب می‌گردد، از این رو ابزارهای بسیار مطلوبی هم برای ارزشیابی و هم برای کمک به آموزش استاد و یادگیری دانشجو تلقی می‌شوند. آثار مثبت «سنجش‌های مداوم»^۲ و منظم بر یادگیری دانشجو و تدریس استاد در پژوهش‌های گسترده‌ای تبیین شده است (مانند مارسل^۳، ۲۰۰۸؛ شیروانی^۴، ۲۰۰۷ و ۲۰۰۹) و بیش‌تر آن‌ها بر این استدلال استوارند که کوئیزهای مداوم تأثیر میانجی^۵ مثبتی بر عملکرد تحصیلی دانشجو از طریق اثرگذاری بر میزان مشارکت و درگیر شدن^۶ او در فرایند یادگیری دارد. میزان مشارکت دانشجو در فرایند یاددهی و یادگیری نیز دومین عامل پراثر پس از توانایی‌های ذهنی و شناختی او در یادگیری قلمداد شده است (دنهام و لیبرمن^۷، ۱۹۸۰). علاوه بر درگیر ساختن فراگیر در روند یادگیری، تحقیقات تجربی نشان داده‌اند کوئیزهای مداوم و منظم تأثیرات زیر را دارند:

- خود تنظیمی و خود نظارتی^۸ یادگیری از طریق دادن بازخورد سریع به عملکرد دانشجو: دانشجویان پس از دریافت بازخورد نسبت به اتخاذ تدابیر و سازمان‌دهی برنامه برای جبران نقص‌ها اقدام می‌نمایند (لیمینگ^۹، ۲۰۰۲)، به‌طور نمونه ساعات و

-
1. Quiz
 2. Frequent assessment
 3. Marcell
 4. Shirvani
 5. Mediating effect
 6. Student engagement/involvement
 7. Denham & Liberman
 8. Self-regulation and self-supervision
 9. Leeming

عادات مطالعه خویش را بازنگری و اصلاح نموده‌اند در حالی که ارزشیابی‌های غیر مداوم منجر به مطالعه فشرده در روزهای نزدیک به امتحان^۱ شده است (جیست و سهرن^۲، ۱۹۹۷).

- کاهش میزان اضطراب امتحان، چرا که اضطراب امتحان با عملکرد تحصیلی دانشجوی رابطه مستقیم دارد (کسادی و جانسن^۳، ۲۰۰۲).

- افزایش نگره‌داشت آموخته‌ها و در نتیجه کمک به دانشجو برای رسیدن به درجه تسلط در یادگیری^۴ (ولف^۵، ۲۰۰۷).

۱۰-۲-۶ ارزشیابی در کلاس‌های پرجمعیت

در مقابله با معضل کلاس‌های شلوغ دانشگاهی لازم است واکنش حساب شده‌ای داشته باشیم تا به واسطه تلاش برای اجتناب از روش‌های ارزشیابی پرهزینه در دامان وجهه مصالحه قرار دادن کیفیت ارزشیابی نیفتیم. چالش این است که بتوان نقطه توازن بین روش‌های ارزشیابی بهره‌ور، کارا و مقتصد و جلوگیری از وارد شدن لطمه بر یادگیری دانشجویان پیدا نمود. بسنده کردن به کاهش تعداد ارزشیابی در این‌گونه کلاس‌های دانشگاهی لزوماً به افزایش مشارکت آنان در یادگیری نخواهد انجامید، از این‌رو به دنبال روش‌های بدیل ارزشیابی باید گشت.

ارزشیابی‌های خودکار^۶ همانند تست‌های برخط^۷، زمان مورد نیاز استادان برای ارزشیابی را تا حد قابل ملاحظه‌ای کاهش می‌دهد و به دانشجویان نیز بازخورد بلافاصله‌ای ارائه می‌نماید. همان‌طور که قبلاً در این بخش اشاره شد راه‌هایی برای مقابله با محدودیت این نوع ارزشیابی - یعنی فروکاست‌گرایی ارزشیابی به یادگیری‌های سطحی و غیر عمیق - و تلفیق آن با روش‌های دیگر وجود دارد که شامل استفاده از نظام نمره‌دهی مبتنی بر میزان اطمینان دانشجو از هر پاسخ، یا نوشتن جملاتی کوتاه در تشریح چرایی انتخاب گزینه علامت زده شده می‌باشد.

-
1. Cramming
 2. Giest & Soehren
 3. Cassady & Johnson
 4. Mastery learning
 5. Wolf
 6. Mechanised assessment
 7. Online test

راه دیگر استفاده از تکالیف گروهی است تا بتوان حجم کار تصحیح و نمره‌دهی استاد را کاهش داد. این روش خصوصاً برای دانشجویان سال اول توصیه می‌گردد، چراکه ارزشیابی‌های تکوینی برای فرایند یادگیری به اندازه دیگر ارزشیابی‌ها، قابل اعتماد اثبات شده، مهم و تأثیرگذار است. روش دیگر کاستن حجم کار ارزشیابی استاد، ارزشیابی دانشجو در زمان کلاسی است، همانند ارائه دانشجویی، بحث‌های کلاسی، پوستر و سامانه‌های پاسخ شخصی که این راهکارها در واقع در راستای راهبرد پنجم - یعنی ادغام ارزشیابی در تدریس - هستند. راهبرد دیگر برای چنین کلاس‌های پرجمعیتی استفاده از تکالیف کوچک‌تر و کوتاه‌تر است. اگرچه واژگان کم‌تر را معادل یادگیری کم‌تر و در نتیجه نمره پایین‌تر تصور می‌کنند. این در حالی است که به‌نظر ما آنچه باید مبنای مقایسه قرار گیرد، میزان جد و جهد فکری است که هر ارزشیابی (آزمونی یا تکلیف درسی) از دانشجو طلب می‌کند. منشأ ایده برای ساخت ارزشیابی که هم موجز باشد و هم جد و جهد فکر مطلوبی را طلب کند، نیز انواع نوشته‌هایی است که به‌طور طبیعی در رشته‌های مختلف به زیور طبع آراسته می‌شوند. مثال‌هایی از این قبیل عبارتند از: چکیده‌هایی اجرایی از گزارش‌های بلند، گزارش‌های خبری، پیشنهاد پژوهشی یا قراردادی، بروشور، پوستر و ... در خصوص انتخاب تکلیف کوتاه لازم است دقت کافی به‌کار رود تا دانشجویان را ترغیب به بازتولید واژه به‌واژه و سطحی‌یادداشت‌های کلاسی و مطالب کتاب‌ها نکنند، بلکه از آن‌ها و مهارت‌های سطوح عمیق شناختی استفاده نمایند تا اقدام به کاربرد نظریه‌ها، ارزیابی از پدیده‌ها، حل مسئله و خلاقیت نمایند. البته این چالش - یعنی تعداد واژگان و عدم به‌کارگیری سطوح عمیق یادگیری - در خصوص امتحانات نیز - علاوه بر تکالیف که شرح آن داده شد - صادق است چراکه آزمون‌هایی که از سؤالات کوتاه پاسخ استفاده کنند، نمره‌دهی آنان سریع‌تر خواهد بود. همچنین آزمونی که یک یا دو ساعت به‌طول انجامد ممکن است به همان اندازه از نظر ارزشیابی یادگیری و سطوح عمیق آن اثربخش باشد که آزمون سه ساعته است، خصوصاً این‌که اگر دانشجویان امکان انتخاب سؤال را نداشته باشند.

توصیه عام‌تر هم این‌که ممکن است بازبینی راهبردهای ارزشیابی حتی در حد رشته - و نه فقط درس خاص - لازم و ضروری باشد چراکه ممکن است اهداف یادگیری مشابه در دروس مختلف با هم همپوشانی زیادی داشته باشند و امکان کاهش ارزشیابی به‌طریقی به‌وجود

بباید. گیبز (۱۹۹۲) هم بر اهمیت آماده نمودن و توجیه خوب دانشجویان در خصوص ارزشیابی و نظام طراحی شده آن، انتظارات استاد و وظایف دانشجو در طول ترم تأکید ویژه دارد تا این که در اجرا با کم‌ترین مشکل برخورد کنیم. روش‌های دیگر برای ارزشیابی دانشجویان در کلاس‌های پرجمعیت و کاهش حجم کاری استاد در این حوزه را در آثار ذیل مطالعه نمایید: (روست، ۲۰۰۱؛ بیگز، ۲۰۰۳؛ گیبز، ۱۹۹۲)

منابع

- Barnett, R. (1990) The idea of higher education. UK: Open University Press.
- Biggs, J.B.(2003) Teaching for Quality Learning at University. 2nd ed., Buckingham: Open University Press.
- Bloom, B.S. (Ed.), Engelhart, M.D., Furst, E.J., Hill, W.H., & Krathwohl, D.R.(1956) Taxonomy of educational objectives: the classification of educational goals. New York: David McKay.
- Bloxham, S. & Boyd, P. (2007) Developing effective assessment in higher education: A practical guide. UK: Open University Press.
- Boud, D. and Falchikov, N. (2006) Aligning assessment with long term learning. Assessment and Evaluation in Higher Education, Vol 31 (4): 399-413.
- Cassady, J.C., & Johnson, R.E. (2002) Cognitive test anxiety and academic performance. Contemporary Educational Psychology, Vol. 27 (2), 270-295
- Coffield, F. & Williamson, B. (1997) Repositioning higher education. UK: SRHE and Open University Press.
- Dancer, D. and Kamvounias, P. (2005) Student involvement in assessment: a project designed to assess class participation fairly and reliably. Assessment and Evaluation in Higher Education, Vol. 30 (4).
- Denham, C. & Liberman, A. (1980) Time to learn. Washington, DC: National Institute of Education.
- Dunn, L., Morgan, C., O'Reilly, M. and Parry, S. (2004) The student assessment handbook: new directions in traditional and online assessment. London: RoutledgeFalmer.
- Eisner, E. W. (1994) The educational imagination: on the design and evaluation of school programs, 3rd ed. USA: Macmillan College Publishing Company.

- Falchikov, N. (2005) Improving assessment through student involvement. London: RoutledgeFalmer
- Gardner, H. (1983) Frames of mind: the theory of multiple intelligences. New York: Basic Books
- Gardner-Medwin. A.R. & Gahan, M. (2003) Formative and summative confidence-based assessment.
<http://www.eassessmentlive2009.org.uk/pastConferences/2003/proceedings/gardner-medwin.pdf> (accessed on 05/02/2012).
- Geist, J.R. & Soehren, S.E. (1997) The effect of frequent quizzes on short- and long-term academic performance. *Journal of Dental Education*, Vol. 61 (4), 339-345.
- Gibbs G. (Ed.) (1995) Improving student learning through assessment and evaluation. UK: Oxford Brooks University.
- Gibbs, G. (2006) How assessment frames student learning. *Innovative assessment in Higher Education*.
- Hounsell, D., McCune, V., Hounsell, J., and Litjens, J. (2006) The quality of guidance and feedback. Paper presented to the Northumbria EARLI SIG Assessment Conference, Darlington, 29 August- 1 September.
- Jonsson, A. and Baartman, J. (2006) Estimating the quality of new modes of assessment: the case of the 'interactive examination' for teacher competency. Paper presented to the Northumbria EARLI SIG Assessment Conference, Darlington, 29 August- 1 September.
- Knight, P.T. and Yorke, M. (2003) Assessment, learning and employability. Maidenhead: Open University Press.
- Krathwohl, D.R. (2002) A revision of Bloom's taxonomy: an overview. *Theory into Practice*. Vol. 41 (4), 212-218.
- Leeming, F.C. (2002) The exam-a-day procedure improves performance in psychology. *Teaching of Psychology*, Vol. 29 (3), 212.
- Lines, D. (2005) Developing a variety of assessment methods including peer and self-assessment: an overview. In *Quality Assurance Agency, Enhancing Practice: Reflections on Assessment, Volume 1*, Gloucester: QAA.
- Marcell, M. (2008) Effectiveness of regular online quizzing in increasing class participation and preparation, *International Journal for the Scholarship of Teaching and Learning*, Vol. 2(1).
- Maclellan, E. (2001) Assessment for Learning: the differing perceptions of tutors and students. *Assessment and Evaluation in Higher Education*. Vol. 26 (4).
- Moon, J. (1999) *Learning Journals*. London: Kogan Page.

- Parker, J. (2003) The patchwork text in teaching Greek tragedy. *Innovations in Education and Teaching International*, Vol. 40 (2), 180-193.
- Pickford, R. (2006) Assessing enterprise, enterprising assessment. Paper presented to the Northumbria EARLI SIG Assessment Conference, Darlington, 29 August- 1 September.
- Ramsden, P. (2003) *Learning to teach in higher education*. 2nd ed. London: RoutledgeFalmer.
- Reber, R. (2005) Assessing motivational factors in educational technology: the case of building a website as course assignment. *British Journal of Educational Technology*, Vol (36 (1): 93-5.
- Rust, C. (2002) The impact of assessment on student learning: how can the research literature practically help to inform the development of departmental assessment strategies and learner-centred assessment practices? *Assessment and Evaluation in Higher Education*, Vol. 3 (2): 145-58.
- Rust, C. (2001) *A briefing on assessment of large groups*. York: Learning and Teaching Support network.
- Scoggins, J. & Winter, R. (1999) The patchwork text: a coursework format for education as critical understanding. *Teaching in Higher Education*, Vol. 4 (4).
- Shafiq, F. & Siddiquah, A. (2011) Effects of classroom quizzes on graduate students' achievement. *International Journal of Academic Research*, Vol. 3 (5).
- Shirvani, H. (2009) Examining an assessment strategy on high school mathematics achievement: Daily quizzes vs. weekly tests. *American Secondary Education*, Vol. 38 (1)
- Shirvani, H. (2007) Effects of teacher communication on parents' attitudes and their children's behaviors at schools. *Education*, Vol. 128(1), 34-47.
- Struyven, K., Docy, F. & Janssens, S. (2002) Students' perceptions about assessment in higher education: a review. Paper presented to the Northumbria EARLI SIG Assessment Conference, Darlington, 29 August- 1 September.
- Tinto, V. (1993) *Leaving college: rethinking the causes and cures of student attrition*. 2nd ed., Chicago: University of Chicago Press.
- Velenchik, A.D. (1995) The case method as a strategy for teaching policy analysis. *Journal of Economic Education*. Vol 26.
- Wolf, P. (2007) Academic improvement through regular assessment. *Peabody Journal of Education*, Vol. 82(4), 690-702.