

سوگیری و خطا در ارزش‌یابی‌های دانشجویان از تدریس

دکتر ناصر شیربگی*

چکیده

در زمینه ارزشیابی اثربخشی تدریس اعضای هیأت علمی، معمول‌ترین شیوه گردآوری داده‌ها به کارگیری مقیاس‌های ارزیابی توسط دانشجویان است. ارزشیابی دانشجویان از عملکرد مدرسان تنها یک نوع داده را برای داوری درباره اثربخشی تدریس فراهم می‌کند. از یک سو، ابزارهای سنتی ارزیابی توانایی تدریس، شاید در موقعیت‌هایی که فرایند حل مسأله یا رویکردهای نوین تدریس مدنظر باشند پاسخ‌گو نباشند، از این‌رو نظام‌های ارزشیابی کنونی برای انعکاس پیوستار گسترده‌تری از ویژگی‌های تدریس در تمام گونه‌های موجود تدریس نیازمند تغییر و بازنگری هستند. از سوی دیگر، پتانسیل‌های موجود برای نادرستی و ابهام در این منبع به‌تنهایی می‌تواند به حداقل هشت منبع سوگیری و خطای ارزشیابی منجر شود. این مقاله به بررسی منابع سوگیری و خطاهای مذکور می‌پردازد و الگویی برای حل این مشکل پیشنهاد می‌نماید.

واژگان کلیدی: آموزش عالی، اثربخشی تدریس، راهبردهای ارزشیابی، خطاهای ارزشیابی

مقدمه

اگرچه تدریس در فرهنگ دانشگاهی از اهمیت و ارزش خاصی برخوردار است اما ضروری است که رؤسای دانشگاه‌ها، مدیران گروه‌ها، اعضای هیأت علمی و دست‌اندرکاران

* استادیار گروه علوم تربیتی دانشگاه کردستان (مسئول مکاتبات: Nshirbagi@uok.ac.ir)

نظارت و ارزیابی درباره اثربخشی تدریس داوری‌هایی به عمل آورند. این داوری‌ها بر بررسی نتایج عملکرد سالانه، فرایند ارتقا و اشتغال تأثیر می‌گذارد. مهمتر از این، اگر هدف آن باشد که مدرسان به نظام ارزشیابی عملکرد آموزشی اطمینان داشته و عملکرد تدریس خود را براساس بازخورد فراهم‌شده توسط چنین نظام‌های ارزشیابی بهبود بخشند، باید خطاها در ارزشیابی به حداقل برسد. از این رو همچنان که صفری (۱۳۸۹) نیز تأکید می‌نماید استفاده از روش‌های مناسب ارزیابی عملکرد آموزشی اعضای هیأت علمی از مهمترین و چالش‌برانگیزترین موضوعاتی است که دانشگاه‌ها با آن مواجهند. موضوع ارزشیابی دانشجویان از تدریس استادان دارای ادبیات غنی و پیشینه تحقیقاتی درخشانی است. بیش از دو هزار مقاله و کتاب در هفتاد سال گذشته منتشر شده‌اند که در برگرفته نتایج پژوهش‌های پیچیده در این خصوص هستند (شیربگی، ۱۳۸۹). ارزشیابی دانشجویان از عملکرد مدرسان تنها یک نوع داده را برای قضاوت در مورد اثربخشی تدریس فراهم می‌کند. پتانسیل‌های موجود برای نادرستی و ابهام در این منبع به تنهایی می‌تواند به حداقل هشت منبع خطای ارزشیابی منجر شود. این مقاله قصد دارد به بررسی جامع منابع خطاهای مذکور بپردازد و الگویی پیشنهادی برای حل این مشکل ارائه نماید.

منابع خطا و سوگیری در ارزشیابی

۱- خطا در محتوی ابزار ارزشیابی

بیشتر ابزارهای تهیه‌شده برای ارزشیابی فرایند آموزش براساس تدریس کلاسی بنا نهاده شده است. اغلب ممکن است تدریس در محیط‌های آزمایشگاهی، عملی یا صحرایی ارزشیابی نشوند. علاوه بر این، امکان دارد که حیطه‌های رفتاری تدریس کلاسی و گویه‌های مربوط به آن به صورت نامناسب برای ارزشیابی تدریس در چنین موقعیت‌هایی به کار گرفته شود. بدون وجود یک پیوند منطقی بین رفتار تدریس کلاسی و رفتارهایی که در سایر موقعیت‌های غیرکلاسی به کار می‌رود، نمی‌توان فرض کرد که ابزارهای ارزشیابی تدریس قابلیت تعمیم دارند. از آنجا که حیطه‌های تدریس متفاوتند، به مقیاس‌های ارزشیابی جداگانه‌ای برای هر موقعیت تدریس، از جمله سخنرانی در کلاس، محیط آزمایشگاهی و کار عملی نیاز است.

جدول شماره ۱ مشهورترین حیطه‌های تدریس برگرفته‌شده از پژوهش‌های اصیل و مبانی نظری حوزه علوم تربیتی و روانشناسی را خلاصه کرده است. وتوروبا و رایت^۱ (۱۹۷۵) نتایج ۲۱ پژوهش منتشرشده در خصوص ارزیابی دانشجویان از تدریس را مرور کردند. از چهل معیاری که آنها فهرست کرده بودند، نه مورد توسط محققان دیگر هم ذکر شده است (سنتر^۲، ۱۹۷۶).

جدول شماره ۱- حیطه‌های تدریس برگرفته شده از مبانی نظری حوزه علوم تربیتی

حیطه‌ها	منابع و حیطه‌ها			
	دیکنسون ^۳	ووتروبا ^۴	سنتر ^۲	داس ^۵
سازماندهی دوره	-	موضوع و دوره.	سازمان-ساختار- شفافیت.	رتوس مطالب- ساخت‌دهی.
توانایی تدریس	روش‌های تدریس.	سازماندهی- بیان و تفسیر.	توانایی خطابه.	سبک و روش‌های تدریس.
نمره‌دهی و بازخورد	آزمون به عنوان تجارب یادگیری.	انصاف در نمره‌دهی.	نمره‌دهی و امتحان.	ارزشیابی عینی، بازخورد.
تعامل استاد و دانشجو	تعامل استاد- دانشجو.	انعطاف‌پذیر.	تعامل- سازگاری دانش آموز.	حساسیت، دسترسی، پاسخگویی.
بارکاری و سختی دوره	-	-	بارکاری، سختی دوره.	-
اشتیاق و انگیزش	اشتیاق.	اشتیاق- تشویق تفکر.	-	پویایی، اشتیاق.
تسلط بر موضوع	شایستگی.	دانش موضوع.	-	-

منبع: گروسینگ، والوک و ویلیامز (۱۹۹۴)

هفت حیطه از مواردی که اغلب مورد توافق بیشتر صاحب‌نظران علوم تربیتی و روان‌شناسی است در ستون اول جدول آمده است. ستون‌های بعدی حیطه‌هایی را که توسط

1- Wotruba & Wright

2- Centra

3- Dickinson

4- Wotruba

5- Das

6- Hilderbrand

پژوهش‌های تحلیل عامل مورد تأیید قرار گرفته‌اند نشان می‌دهد. حیطه‌هایی از تدریس که در بین موقعیت‌های مختلف تدریس مشابه هستند می‌تواند به صورت مشترک در مقیاس‌های ارزشیابی مورد استفاده قرار گیرد. گفتنی است که صفری (۱۳۸۹) در پژوهش خود که یکی از جدیدترین بررسی‌هایی است که در این خصوص در کشور انجام گرفته با بررسی نظرات مشارکت‌کنندگان در یک تحقیق توصیفی - تحلیلی با استفاده از چهار منبع ارزیابی (دانشجویان، همکاران، مدیران گروه و خود فرد) به بررسی میزان اهمیت مؤلفه‌های تشکیل‌دهنده فرایند تدریس و یادگیری پرداخت. وی در پایان هشت مؤلفه اصلی مورد توافق و با اهمیت را پیشنهاد کرده است که عبارتند از: تسلط بر محتوا، تدوین طرح درس، مهارت‌های عرضه محتوا، مدیریت درس، راهنمایی و مشاوره، ارزیابی یادگیری، مهارت‌های ارتباطی و مسایل اخلاقی در تدریس.

در جدول شماره ۲ توصیف مختصری از مقیاس‌هایی که برای تدریس در سه موقعیت مختلف می‌توان مورد استفاده قرار داد ارائه شده است. شوراهاى نظارت و ارزشیابی آموزشی یا هیأت‌های ممیزه دانشگاه‌ها ممکن است وزن‌های مختلفی را برای حیطه‌ها و مقیاس‌های گوناگون تعیین نمایند تا به یک ابزار ارزشیابی جامع برای فرایند آموزش دست یابند.

جدول شماره ۲- محتوای مقیاس ارزشیابی برای حیطه‌های تدریس در سه محیط

	محیط تدریس		
	کلاس.	لابراتوار.	عملیات میدانی.
۱ توانایی تدریس - سخنرانی	X	-	-
۲ توانایی تدریس - آزمایشگاه	-	X	-
۳ توانایی تدریس - عملیات میدانی	-	-	X
۴ سازماندهی دوره	X	X	X
۵ انتخاب و استفاده از رسانه	X	X	-
۶ ارزشیابی عملکرد دانشجویان	X	X	X
۷ تعامل استاد و دانشجو	X	X	X
۸ اشتیاق/ انگیزش	X	X	X
۹ تسلط بر موضوع	X	X	X

منبع: گروسینگ، والوک و ویلیامز (۱۹۹۴)

خطا در محتوی ابزار ارزشیابی ممکن است در ارزشیابی موقعیت‌ها یا شرایطی رخ دهد که خارج از کنترل مدرس است. پرسش از دانش‌آموزان در موارد ذیل از آن جمله است: ارزیابی مسائل مربوط به برنامه درسی (مانند تناسب پیشنیاز دروس)؛ ارزیابی محیط یادگیری (مانند روشنایی کلاس، دما و تهویه)؛ یا گزارش خود-ارزیابی علایق یادگیری و نتایج عملکرد آنان. هرچند ممکن است موارد بیان شده حیطه‌های مفیدی را برای ارزشیابی برنامه‌های درسی باشد اما آنها بر مبنای متغیرهای مربوط به دانشکده‌ها و دانشجویان نهاده شده‌اند و نه بر عملکرد مدرسان. از این رو، ناظران آموزشی باید از تفاوت مقیاس‌های رفتاری تدریس مانند: «مدرس از فهرست مطالب درس پیروی می‌کند» با سایر رفتارهای مربوط به دانشجو مانند: «من درخواست کمک می‌کنم وقتی که مطالبی را نمی‌فهمم» یا توصیف‌کننده‌های فضای فیزیکی کلاس مانند: «محیط کلاس یک فضای مناسب را فراهم ساخت» آگاه باشند.

۲- خطا در تفسیر معانی و نتایج ارزشیابی‌ها

استفاده از هنجارها در تفسیر ارزشیابی دانشجویان از تدریس برای مقایسه‌های معنادار و منصفانه، اساسی است. با این حال، ماهیت هنجارهای گروهی ممکن است در خطای تفسیرها مشارکت داشته باشد. برای مثال، چون پژوهش‌ها نشان داده است که مدرسان در کلاس‌های دروس علوم پایه ارزشیابی‌های کمتری نسبت به افرادی که در علوم انسانی و اجتماعی تدریس می‌کنند دریافت می‌کنند بنابراین معنی هنجارهای دانشگاه مورد پرسش است (گروسینگ، والوک و ویلیامز^۱، ۱۹۹۴). علاوه بر متغیرهای مربوط به محتوی برنامه درسی متغیرهای متعدد دیگری ممکن است در خطاهای تفسیر هنجاری مشارکت داشته باشند که سطح پختگی دانشجویان، مقطع تحصیلی، نوع دروس-اختیاری/پایه و متغیرهای آموزشی از آن جمله‌اند (براسکمپ، براندنبرگ و اوری^۲، ۱۹۸۴). هنجارهای مقیاس ارزشیابی باید در چندین نرم گروهی تهیه شود تا مقایسه‌ها معنادارتر شوند. برای مثال: نرم‌های دروس پایه/اختیاری، دروس نظری/عملی؛ مقطع تحصیلی و نرم‌هایی برای هر گروه، دانشکده و دانشگاه. پیامدهای هر نوع خطایی ممکن است خیلی گسترده‌تر از تفسیرهای فردی ارزشیابی‌ها باشد.

1- Grussing, Valuck, & Williams

2- Braskamp, Brandenburg & Ory

تجارب منفی از نظام ارزشیابی ممکن است سبب کاهش اطمینان از فرایند ارزشیابی دانشکده شود و اعضای هیأت علمی و مدیران را دل‌سرد کند.

۳- نمایش یا معجری‌گری

تأثیر نحوه ارائه درس توسط یک مدرس بر یادگیری دانشجویان و ادراک آنان از رضایت از مدرس در تحقیقات مختلف اثبات شده است (وار و ویلیامز^۱، ۱۹۷۵). ارائه‌هایی که به دلیل ویژگی‌های اغواگری و سرگرم‌کنندگی به خاطر می‌مانند ممکن است خطاهایی را در ارزشیابی برخی از جنبه‌های اثربخشی تدریس مانند «تحت پوشش قرار دادن محتوای برنامه درسی» ایجاد نمایند. اما جنبه‌هایی از شخصیت و سبک مدرس همانند مقدار مراقبت و پشتیبانی تلقینی از یادگیری دانشجویان و همچنین رفتارهای انگیزشی که قابل مشاهده است به‌طور واقعی در ایجاد محیط یادگیری مهیج و شاداب مشارکت می‌کند و بنابراین ممکن است به‌عنوان جزئی از عوامل مرتبط با تدریس اثربخش به‌نظر رسد. (هیلدبراند، ویلسون و دینت^۲، ۱۹۷۱). مقیاس‌های تهیه‌شده باید از ارزشیابی دانشجویان از شخصیت، جذبه (کاریزما) یا صفات مشابه با آن خودداری کند و تنها آن دسته از ویژگی‌های مدرسان که ثابت شده با تدریس اثربخش مرتبط هستند. (مانند: تعامل معلم و دانشجویان آموز یا دغدغه برای یادگیری آنها) باید مورد تأکید قرار گیرند.

۴- خطا در پایایی ابزار

وقتی که گویه‌های متعددی در پرسشنامه‌های معمولی که به‌طور عددی نمره‌گذاری شده‌اند، ظاهر می‌شوند، گویه‌های مشابه، یا آنهایی که معانی مشابه دارند به‌عنوان مقیاس‌های مشخص اندازه‌گیری حیطه‌های عملکرد تدریس جمع بسته می‌شوند و یا میانگین آنها محاسبه می‌شود و بنابراین باید همبستگی‌های بالایی از ارزشیابی دانشجویان را نشان دهد. برآوردهای مربوط به سازگاری درونی پرسشنامه‌ها باید به استفاده‌کنندگان ابزارها گزارش شود تا پایایی چنین مقیاس‌های چندگویه‌ای را اثبات کند. به‌علاوه، استفاده از ابزارهای درجه‌بندی باید ثبات

1- Ware & Williams

2- Hildebrand, Wilson & Dienst

را در به‌کارگیری‌های متوالی نشان دهد. داده‌های به‌دست‌آمده از آزمون پایایی به روش بازآزمایی اجازه می‌دهد که تضمین‌های بیشتری از کیفیت ابزار ارائه شود. به‌ویژه هنگامیکه داده‌ها نشان داده‌اند که صفات یا رفتارهای اندازه‌گیری‌شده در طول زمان ثابت می‌مانند. مقیاس‌های ارزشیابی‌های دانشجویی باید نشان دهند که همبستگی بالایی با سایر ابزارهای استاندارد و مشهور که تهیه‌کنندگان آن ادعا می‌کنند همان حیطه‌های عملکرد تدریس را اندازه‌گیری می‌کنند داشته باشند. این نوع تخمین پایایی به شکل فرم‌های موازی، با مفهوم ایجاد روایی هم‌زمان از معانی نمرات حاصل از مقیاس‌های مشابه مرتبط است. افزایش اطمینان از معنی ارزشیابی می‌تواند با توسعه مقیاس‌ها (خوشه‌هایی از گویه‌های ارزیابی مرتبط) براساس حیطه‌های عملکرد تدریس که روایی آنها به اثبات رسیده تقویت شود.

۵- اهداف چندگانه ارزشیابی

طبیعی است که اغلب مدرسان به دنبال خود-بهبودی و کاربرد روش‌های ارزشیابی برای دستیابی به بازخورد درباره اثربخشی نوآوری‌ها یا سایر جرح و تعدیل‌ها در تدریس باشند. این رویکرد تکوینی^۱ ممکن است مورد تأکید مدیران گروه‌ها یا استادان برجسته باشد. چنین خودارزیابی‌هایی حساس بوده و یک فرایند شخصی است. استقلال در ارائه ابتکارات در تدریس مستلزم آن است که مدرسان آزاد باشند که نتایج چنین ارزیابی‌هایی را در میان‌ترم یا پایان ترم انتخاب و تفسیر نمایند. همچنین مدرسان در ارزشیابی‌هایی که به‌صورت رسمی لازم است انجام شود شرکت می‌کنند ماهیت تراکمی^۲ دارد و برای بررسی‌های دوره‌ای عملکرد سازمان یا ارتقای افقی و عمودی اعضای هیأت علمی و سایر تصمیمات استخدامی به‌کار می‌رود. اگر داده‌های تکوینی و تراکمی به‌وسیله مدیران در هم ادغام شده و هر دو بر مبنای داده‌های تراکمی در تصمیم‌های استخدامی و کارگزینی در نظر گرفته شوند؛ به دلیل استفاده نامناسب و پریشان از داده‌های شخصی تکوینی خطاهایی در ارزشیابی رخ می‌دهد و شاید

۱- ارزشیابی تکوینی به ارزشیابی‌های یک فرایند یا محصول برای فراهم ساختن بازخورد به منظور انجام اصلاحات میان فرایندی اشاره می‌کند.

۲- ارزشیابی تراکمی یا پایانی به این منظور انجام می‌شود که کیفیت یا اثر یک فرایند تکمیل شده یا محصول نهایی را بیازماید.

ارزشیابی عملکرد کلی را تنزل دهد (مانینگ^۱، ۱۹۸۸). نظام‌های ارزشیابی آموزشی باید برای هر دو هدف فراهم شوند؛ یک مجموعه ثابت از گویه‌ها و مقیاس‌ها برای ارزشیابی تراکمی و یک مجموعه از گویه‌ها به انتخاب مدرسان برای اهداف خودکاو شخصی در ارزشیابی تکوینی.

۶- روش‌های ناهماهنگ به کارگیری ابزارها

عملیات اجرایی متعددی در نادرستی و اشتباهات ارزشیابی دانشجویان از تدریس و عدم اطمینان در گردآوری داده‌ها دخالت دارند از جمله: نبود استانداردسازی در مدیریت زمان (مانند انجام ارزشیابی‌ها زودهنگام در طول ترم، خیلی نزدیک به دوره آزمون نهایی، بدون زمان کافی برای تکمیل ابزار) و انجام ارزشیابی در زمان‌های متفاوت در ترم برای هر مدرس. همچنین از آنجاکه خطاها در ارزشیابی ممکن است به عدم بیان اهداف فرایند ارزشیابی نسبت داده شود (تکوینی در مقابل تراکمی) بنابراین باید برای ارزیابان و ارزش‌یابی‌شوندگان توضیح داده شود که چه استفاده ویژه‌ای از داده‌ها می‌شود و چه کسانی از داده‌ها استفاده می‌کنند (ستترا، ۱۹۷۹؛ براسکمپ، براندنبرگ و اوری، ۱۹۸۴). اگر ابزارهای ارزشیابی به وسیله خود مدرسان به دانشجویان معرفی شود و پخش و تکمیل آنها در حضور مدرسان یا با نظارت آنها انجام شود پرسش‌هایی در مورد تأثیرات نامناسب ممکن است مطرح شود. به کارگیری ابزارهای ارزشیابی توسط خود دانشجویان به عنوان راه‌حل چنین تأثیرات نامناسبی از طرف برخی صاحب‌نظران سفارش شده است (براسکمپ، براندنبرگ و اوری، ۱۹۸۴). در نهایت چون ابزارها در یک حالت نظام‌دار و امن جمع‌آوری نشده‌اند، خطاهای اندازه‌گیری رخ خواهد داد. دانشجویان گزارش داده‌اند که در یک فعالیت ارزشیابی مشارکت کرده‌اند، بعد از آن کلاس را ترک کرده و قبل از آنکه ابزار به یک دانشجو که به عنوان نماینده تعیین شده بازگردانند. هنگامی که تکمیل فرم ارزشیابی در یک محیط کلاسی استاندارد شده به انجام نرسد، درستی آنها مورد سوءظن قرار می‌گیرد. به دلیل اینکه نماینده همه دانشجویان نیست چراکه تنها بخشی از دانشجویان آنها را تکمیل می‌کنند و برمی‌گردانند. یک تهدید نهایی در ادغام ارزشیابی‌ها رخ می‌دهد هنگامی که چند مدرس یک کلاس با یک ابزار یکسان ارزشیابی می‌شوند.

برای چنین کلاس‌های تدریس‌شده توسط چند استاد، مدت ناکافی در معرض دید قرار گرفتن مدرس برای دانشجویان و دوره زمانی ناسازگار بین زمان ارائه درس توسط مدرس و ارزشیابی توسط دانشجویان در خطای ارزشیابی از تدریس مشارکت می‌کند. خطاهای نسبت داده‌شده به روش‌های اجرایی ارزشیابی می‌تواند از طریق استفاده از دستورالعمل‌های استاندارد و به‌کارگیری کارشناسان توجیه‌شده که هدف و کاربرد فرایند ارزشیابی را هنگام به‌کارگیری مقیاس‌ها در زمان‌های معمول و منظم پایان دوره آموزش در تمام کلاس‌ها بیان کنند، به حداقل برسد.

۷- خطاهای معمول تأثیرات ارزشیابی

در صورتیکه دانشجویان دیدگاه مثبتی نسبت به یک مدرس داشته باشند، ممکن است منجر به ارزیابی مثبت‌تر بدون توجه به سطح واقعی اثربخشی تدریس او شود. پژوهش‌ها نشان داده‌اند که صفات افراد بر قضاوت‌های آنها مؤثر بوده است. پشتیبانی از این ایده از نظریه شخصیت ضمنی آچ^۱ (۱۹۶۶) و برونر و تاگیوری^۲ (۱۹۵۴) ناشی می‌شود. همچنین کلی^۳ (۱۹۵۰) در پژوهش‌های خود دریافت که دستکاری ویژگی‌های دو قطبی مانند گرم-سرد تأثیر فراوانی در قضاوت‌های دانشجویان از استادان دارند (به نقل از شیرگی، ۱۳۸۹). دست‌کم پنج نوع از تأثیرات ارزشیابی ممکن است سبب خطای ارزشیابی شود. اصطلاح اثرات‌هاله‌ای (ورنون^۴، ۱۹۶۴) به موقعیت‌هایی اطلاق می‌شود که در آن ارزشیابی‌های ویژه براساس مشاهده عملکرد مربوطه نیست بلکه مبنای آن مشاهده تجارب یا نگرش‌های پیشین است (گروسینگ، والوک و ویلیامز، ۱۹۹۴). این تأثیر به‌طور معمول زمانی رخ می‌دهد که تعبیر عملکردهای موفقیت‌آمیز در یک حیطة به مبالغه در حیطة دیگری منجر شود. در مقابل تأثیر هاله‌ای معکوس ممکن است وقتی رخ می‌دهد که دانشجویان متمایل به ارائه ارزیابی پایین از عملکرد مدرس در یک حیطة به دلیل آن که ممکن است آنها نگرش‌های نامناسبی درباره حیطة دیگری از عملکرد وی داشته باشند. برای مثال سخت‌گیری گروه آموزشی مربوط. تأثیر آسان‌گیری نیز در خطای ارزشیابی سهیم است که در مقابل آن تأثیر

1- Asch

2- Bruner & Tagiuri

3- Kelley

4- Vernon

سخت‌گیری وجود دارد. این نوع تأثیرات به ارزیابی‌های یکنواخت پایین یا بالای دانشجویان اشاره می‌کند که بدون در نظر گرفتن تفاوت‌ها عملکرد مشخصی مورد ارزیابی قرار می‌گیرد. دانشجویان ممکن است احساس کنند که وادار شده‌اند که مدرسانشان را به صورت بالا ارزشیابی کنند یا چنین دریافته‌اند که دادن ارزشیابی‌های پایین بر شخصیت آنها تأثیر منفی خواهد گذاشت چراکه مرتبط با عملکرد یادگیری یا میزان توانایی درک آنان است. دادن ارزشیابی پایین به صورت ثابت وقتی که بر اساس معیارهای مقایسه‌ای معتبری نباشد ممکن است ریشه در باورهای دانشجویان درباره سخت‌گیری آموزشی داشته باشد یا از اثرات انگیزشی ناشی شود. مدرسان همچنین شواهدی از تأثیر تمایل به حد وسط مشاهده کرده‌اند که نتیجه آن ارزشیابی‌های نه‌چندان بالا و نه پایین است (اسمیت^۱، ۱۹۸۳). چنین رویکرد ارزشیابی ممکن است با موضوع ترس دانشجویان حمایت شود که ارزشیابی‌های پایین یا بالا ممکن است نیازمند توجهات اضافی باشد یا اینکه ارزشیابی‌های پایین بر روابط حرفه‌ای شخصی مدرس با دانشجو تأثیر بگذارد. ارائه دستورالعمل‌های استاندارد مناسب به دانشجویانی که ارزیابی می‌کنند می‌تواند خطاهای تأثیرات را در ارزشیابی به کمترین حد ممکن برساند.

۸- خطا در کاربرد داده‌ها

برخی مدیران مراکز آموزش عالی ممکن است احساس ناراحتی کنند وقتی نتایج ارزشیابی‌ها را به اعضای هیأت علمی بازخورد می‌دهند (ایوانسویچ^۲، ۱۹۹۲). شاید چنین سرپرستانی از استفاده و تفسیر داده‌های مربوط به عملکرد آموزشی بیمناک باشند چون از تربیت یا تجربه کافی مدیریتی برخوردار نیستند و یا آنها مقیاس‌های ارزشیابی (یا در حقیقت تمام فرایند ارزشیابی عملکرد اعضای هیأت علمی) را عینی در نظر نمی‌گیرند. بنابراین از ارزش فرایند ارزیابی و داده‌ها می‌کاهند. خطاها ممکن است خود را در یک بازخورد ناقص یا اشتباه یا در نبود هر بازخورد مفید در کل هویدا کنند. بخشی از فرایند بهسازی مدیران کم‌تجربه باید شامل آموزش آمادگی نحوه ارائه بازخورد عملکرد به کارکنان و استفاده از داده‌های عملکردی روا و پایا باشد.

1- Smith

2- Ivanceovich

بحث و نتیجه‌گیری

به نظر می‌رسد نظام موجود ارزشیابی تدریس در کشور هنوز بیشتر بر مبنای نظرات دانشجویان انجام می‌شود. این فرایند دارای مشکلات متعددی است که به مهم‌ترین آنها اشاره شد. روی هم‌رفته نتایج تحقیقات به عمل آمده در زمینه اثرات متغیرهای نامربوط بر روایی ارزشیابی دانشجویان از مدرسان نشان داده است که در تفسیر این داده‌ها لازم است شرط احتیاط رعایت شود. چالش کاهش خطا برای همه مدرسان، مدیران گروه‌ها، مدیران عالی و مسئولان ارزشیابی مشابه است. به دلیل علاقه ذاتی مدرسان برای خودبهبودی لازم است که ابزارهای بازخوردهای تکوینی روا، پایا و محرمانه باشند. اعضای هیأت علمی از تقویت عینیت در فرایند کنترل و نظارت همگنان سود می‌برند. دست‌اندرکاران در گروه‌ها و دانشکده‌ها نیازمند نوعی ابزارهای ارزشیابی عملکرد هستند که استانداردهای قانونی حرفه‌ای برای روایی و پایایی را تأمین کند. علاوه بر این مرکز خدمات دانشگاه که درگیر فرایند ارزشیابی آموزشی هستند مجبورند که رواترین و پایاترین ابزارهای موجود را فراهم کنند.

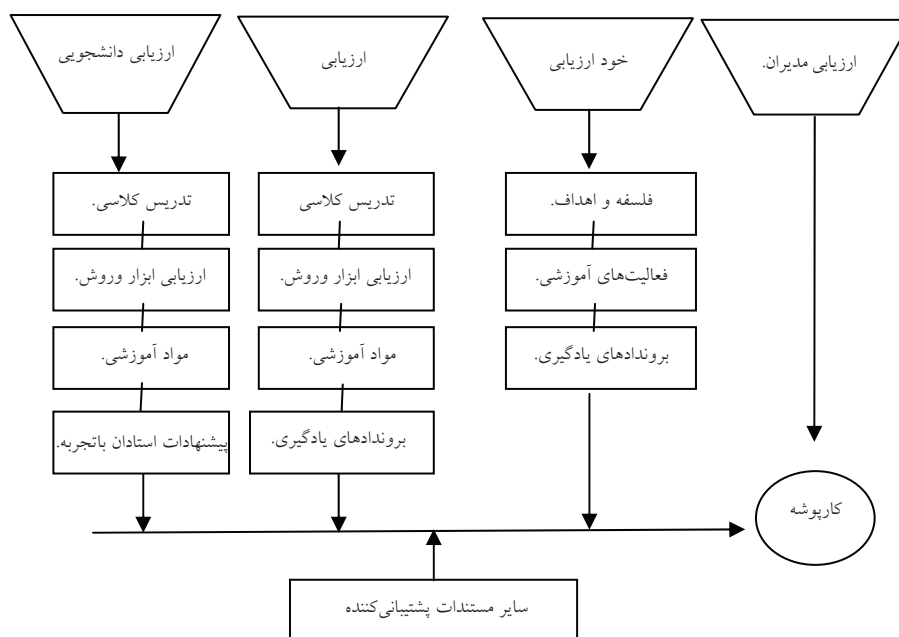
یکی از راه‌حل‌های کم کردن سوگیری و خطای ارزشیابی استفاده از راهبردهای متنوع تهیه مستندات اثربخشی تدریس است. نمونه‌ای از یک نظام جامع ارزشیابی با استفاده از منابع مختلف در نمودار شماره یک به تصویر کشیده شده است. اگرچه در پیشینه تحقیقاتی و مبانی نظری مربوط به حوزه ارزشیابی از تدریس حداقل چهارده راهبرد مجزا برای تهیه مستندات و شواهد اثربخشی تدریس ارائه شده است (شیربگی، ۱۳۸۹)، اما فقط پنج مورد از آنها در نظام جامع پیشنهادی ذیل به کار گرفته شده است.

از مزیت‌های نظام جامع ارزشیابی نخست آن است که مبنای خوبی را برای ارزشیابی تکوینی فراهم می‌کند و دوم آن که اگر این الگو در چند سال اولیه مسیر شغلی اعضای هیأت علمی به کار گرفته شود امکان این که به صورت معنی‌داری ارزشیابی‌های تراکمی مناسبی را فراهم آورد بسیار زیاد خواهد بود.

باید به خاطر داشت که اگر دو یا چند مجموعه داده گردآوری شده مربوط به ارزشیابی تدریس یک مدرس کاملاً متفاوتی به دست دهد پس به احتمال زیاد مواردی در دست‌کم یکی از منابع ایراد دارد یا تدریس مدرس در کلاس ارزشیابی شده نماینده تدریس واقعی او نبوده

است، بنابراین باید بررسی‌های بیشتری انجام شود (برنت و فدلر^۱، ۲۰۰۴). یک شیوه برای افزایش پایایی ارزشیابی‌هایی که براساس چند منبع انجام می‌شود آن است که اطمینان حاصل شود که برخی همپوشانی‌ها در اطلاعات که منابع فراهم کرده‌اند وجود دارد. به‌عنوان مثال اگر برگه مشاهده کلاس توسط همگنان شامل گویه‌هایی درباره آمادگی برای سخنرانی، وضوح بیان و احترام به دانش‌آموزان باشد پس برگه‌های ارزشیابی که به‌وسیله دانش‌آموزان تکمیل می‌شود باید همان ویژگی‌ها را مورد پرسش قرار دهد.

دورنمای تدریس همواره در حال تغییر است. ابزارهای سنتی ارزیابی توانایی تدریس شاید در موقعیت‌هایی که فرایند حل مسأله یا رویکردهای نوین تدریس مدنظر باشد پاسخ‌گو نباشند از این‌رو نظام‌های ارزشیابی کنونی برای انعکاس پیوستار گسترده‌تری از صفات تدریس در تمام گونه‌های موجود تدریس نیازمند تغییر و بازنگری خواهند بود.



نمودار ۱- نمونه‌ای از یک مدل جامع ارزشیابی عملکرد تدریس.

منابع

- ۱- شیربگی، ناصر (۱۳۸۹). راهبردهای اندازه‌گیری اثربخشی تدریس در نظام‌های آموزش عالی، نامه آموزش عالی. دوره جدید، سال سوم، (۱۰) ۳۳-۹.
- ۲- شیربگی، ناصر و همتی، آزاد (۱۳۸۹). مدل‌یابی تأثیر جذبه شخصیتی استادان بر ارزیابی دانشجویان از تدریس اثربخش در دانشگاه کردستان. چهارمین همایش سالانه ارزیابی درونی کیفیت در نظام دانشگاهی. دانشگاه تهران.
- ۳- صفری، ثنا. (۱۳۸۹). نقش منابع اطلاعات گوناگون در ارزیابی آموزشی اعضای هیئت علمی، فصلنامه پژوهش در برنامه‌ریزی آموزش عالی، ۱۶ (۱)، ۸۵-۶۹.
- 4- Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology*, 41(2), 258-290.
- 5- Braskamp, L. A., Brandenburg, D.C. & Ory, J.C., (1984). *Evaluating Teaching Effectiveness*, Sage Publications, Newbury Park CA 29-76.
- 6- Brent, R. & Felder, R. M. (2004). A Protocol for Peer Review of Teaching, *Proceedings of the 2004 American Society for Engineering Education Annual Conference & Exposition*, North Carolina State University
- 7- Bruner, J. S., & Tagiuri, R. (1954). the perception of people. In G. Lindzey (Ed.), *Handbook of Social Psychology* (Vol. 2). London Addison Wesley.
- 8- Centra, J. A. & Creech, F.R., (1976). "The Relationship between Student, Teachers, and Course Characteristics and Student Ratings of Teacher Effectiveness," Project Report 76-1, Educational Testing Service, Princeton NJ
- 9- Centra, J. A., (1979). *Determining Faculty Effectiveness*, Jossey-Bass. San Francisco CA.7-11.
- 10- Das, H., Frost, P.J. & Barnowe, J.T., (1979). "Behaviorally anchored scales for assessing behavioral science teaching," *Canadian Journal of Behavioral Science*, 11, 79-88
- 11- Dickinson, T. L. & Zelliger, P.M., (1980) "A Comparison of the behaviorally anchored rating and mixed standard scale formats." *Journal of Applied Psychology*, 65, 147-154.
- 12- Grussing, P.G., Valuck, R. & Williams, R.G., (1994). "Development and validation of behaviorally-anchored rating scales for student evaluation of pharmacy instruction", *American Journal of Pharmacia Education*, 58, 25-37N.B.33.
- 13- Hildebrand, M., Wilson, R.C. & Dienst, E.R., (1971). *Evaluating University Teaching*, Center for Research and Development in Higher Education, Berkely CA 18-20.
- 14- Ivancevich, J.M., (1992). *Human Resource Management*, 5th ed., Irwin, Homewood IL, 327-327.
- 15- Kelley, H. H. (1950). The warm-cold variable in first impressions of persons. *Journal of Personality and Social Psychology*, 18(3), 4 31-439. .

- 16- Manning, R.C., (1988). The Teacher Evaluation Handbook, Prentice-Hall, Englewood Cliffs NJ 4-5. (10) Ibid., p. 6-9.
- 17- Smith, P.C., (1983) "Behaviors, results and organizational effectiveness: The problem of criteria," in Handbook of Industrial Psychology, (edit. Dunnette, M.) Wiley & Sons, New York, NY. 757.
- 18- Vernon, P. E. (1964). Personality Assessment: a critical survey. London,: Methuen.
- 19- Ware, J.E. & Williams, R.G., (1975). "The Dr. Fox effect: A study of lecture effectiveness and ratings of instruction," Journal of Medical Education., 50, 149-156.
- 20- Wotruba, T.R., & Wright, P.L., (1975). How to develop a teacher-rating instrument: A research approach, Journal of Higher Education, 46, 653-63.