



ارزیابی عملکرد تدریس در دانشگاه‌های ایران

سید علی استوار نامقی

دانشگاه صنعتی شاهرود

saostovarnamaghi@yahoo.com

چکیده

الگوهای نظری روش های متعددی را برای ارزیابی عملکرد مدرس ارائه می دهند. سیستم‌های ارزیابی محلی ارزیابی تدریس را عمدتاً به عهده ی دانشجویان گذاشته است. این روش کار مدرس را به طرق مختلف تعیین و تحدید می کند. با استفاده از روش نمونه گیری نظری و تحلیل و جمع آوری داده ها بر اساس نظریه ی مبتنی بر داده ها (نظریه ی زمینه ای)، این تحقیق بر آن است تاثیر روش ارزیابی بر عملکرد مدرس، برداشت مدرس از معیارهای ارزیابی و تاثیر روش ارزیابی بر زنده گی حرفه ای مدرس را مورد بررسی قرار دهد. یافته های تحقیق برای تمام طرف های ذینفع در امر ارزیابی مفید است چرا که شبکه ی مفهومی برگرفته از نظرات مدرسین کمک بسیار زیادی به بهبود روند ارزیابی عملکرد تدریس در دانشگاه های ایران می کند.

کلید واژه ها: ارزیابی تدریس، ارزیابی دانشجو از تدریس، معیارهای ارزیابی، برداشت مدرسین از معیارها

مقدمه ی تحقیق

دانیلسن (۲۰۰۰) برای ارزیابی از کار مدرس دو هدف تعیین میکند: کنترل کیفیت آموزشی و رشد حرفه ای مدرس. هدف اول از طریق ارزیابی پایانی برآورده می شود و هدف دوم از طریق ارزیابی تکوینی. هدف از ارزیابی پایانی ارایه ی مجوز تدریس ، استخدام، تبدیل وضعیت، ارتقای شغلی و یا اخراج مدرس می باشد. بر عکس ارزیابی تکوینی با ستفاده از ارایه ی بازخورد مناسب سعی در ایجاد انگیزه برای رشد حرفه ای، شکل دهی رفتار مناسب در مدرس، تغییر روش های موجود و ارایه ی روش های جدید دارد (پیترسن ۲۰۰۰).

هر چند هر دو روش سعی در اندازه گیری عملکرد مدرس دارند ولی روش تکوینی در صد ارایه ی راهکار هایی برای بهبود عملکرد آموزشی می باشد در حالی که ارزیابی پایانی سعی دارد بفهمد آیا بهبود آموزشی در حد تشویق هست یا نه. هر چند هر دو روش دارای ارزش بالقوه می باشد اما هیچ کدام به تنهایی نمی تواند باعث بهبود روند آموزشی شود. اما اگر این دو روش با هم تلفیق شوند، فرصتهای بهینه ای برای رشد حرفه ای مدرس ایجاد می کنند (برنت، مترز الیوا براون-سیمزو هس ۲۰۰۷). هر چند هر کدام از روشهای زیر منبع مفیدی از اطلاعات برای ارزیابی عملکرد معلم ارائه می دهد، هر کدام از آنها به تنهایی دارای محدودیت ها یی است:

- ۱- بررسی طرح درس: طراحی لازمه ی آمادگی معلم می باشد و همبستگی بالایی با یادگیری درد (استرانج، ۲۰۰۷). اما اشکال عمده این است که طرح درس ضمن اجرا دچار تعدیل و تغییر میشود. بنابراین طرح درس نمی تواند کیفیت تغییرات را توضیح دهد.
- ۲- مشاهده کلاس: مشاهده توصیف چیزی است که در واقع ضمن تدریس اتفاق می افتد (موجس ۲۰۰۶). اما افراد غیر متخصص در امر مشاهده و همچنین مشاهده کلاس در یک فاصله زمانی کوتاه معمولاً نتایج مغرضانه ای در بر دارد.
- ۳- خود ارزیابی: تحلیل تدریس و تامل و تفکر بعد از تدریس باعث یاد گیری معلم می شود (آلن بک و دیگران، ۲۰۰۲). هر چند این روش اطلاعات مفیدی در مورد تدریس ارائه می دهد، موفقیت آن نیازمند صرف زمان زیاد و حمایت سیستم مدیریتی است. بنابراین استفاده از این روش وابسته به اولویتهای مدیران دارد (پترسون و کاموکس، ۱۹۹۰).
- ۴- ارزیابی پورت فوی: پورت فوی به مسوول ارزیابی کمک می کند نکات قوت و ضعف استاد را پیدا کند. علاوه بر آن این روش باعث رشد حرفه ای مدرس می شود (ایتلو، و سورس، ۲۰۰۶). هر چند این روش مفید است باید مخطاطانه مورد استفاده قرار گیرد چرا که پایایی آن قابل قبول نیست. مشکل دیگر این است که تهیه و ارزیابی آن نیاز به زمان زیادی دارد و در عمل این شرایط بکار گیری آن را دچار مشکل می کند (تاکر ، استرانج و گاریز، ۲۰۰۲).
- ۵- داده های مربوط به پیشرفت تحصیلی دانشجویان: استفاده از نمره دانشجویان در آزمونهای استاندارد می تواند معیار مناسبی برای اندازه گیری کارایی تدریس باشد. اما مسئله این است که در بسیاری از سیستمهای آموزشی آزمون های استاندارد موجود نمی باشد. علاوه بر آن این آزمون ها جزئی از فعالیت کلاس و سر فصل را می سنجد (بری، ۲۰۰۷).

به طور خلاصه دیدگاه های نظری موجود به ما کمک می کند که در امر ارزیابی عملکرد تدریس از دو اشتباه اجتناب کنیم: (۱) اتکا به روش ارزیابی پایانی برای قضاوت شخصی و اخراج مدرسین (۲) اتکا به یک منبع برای ارزیابی تدریس. چنین شیوه های ناقص و محدود کننده هستند چرا که این روش ها منجر به رشد حرفه ای مدرس نمی شوند. ادبیات پیشینه مرتبط بیانگر این است که ارزیابی کیفی نه تنها مستلزم استفاده از دو روش تکوینی و پایانی می باشد، بلکه همچنین نیاز به جمع آوری و تحلیل اطلاعات از چندین منبع دارد. تغییر روشهای ارزیابی عملکرد با استفاده از مدل های نظری چشم انداز خوبی ایجاد می کند ولی اتکا به مدل های نظری عاری از اشکال نیست چرا که این روش اصلاحی نظرات افرادی را که مورد ارزیابی قرار می گیرند نادیده می گیرد. بنابراین ارزیابی کیفی باید هم بر اساس الگوهای نظری و هم بر اساس دیدگاه های مدرسین باشد.

هدف تحقیق

برای اصلاح سیستم ارزیابی فرد ارزیاب باید از دو منبع اساسی استفاده کند: دیدگاه های نظری و دیدگاه های افرادی که مورد ارزیابی قرار می گیرند. بدین منظور با استفاده از الگوهای کد گذاری نظریه ی مبتنی بر داده ها، این تحقیق سعی در ارائه یک شبکه مفهومی برگرفته از نظر مدرسین می باشد. به طور دقیق تر، با استفاده از مصاحبه کیفی محقق سه هدف را دنبال می کند: (۱) کاوش معیارهای ملی برای ارزیابی تدریس (۲) بررسی برداشت مدرسین از معیارهای ارزیابی و (۳) بررسی تاثیر این معیارها بر زندگی حرفه ای مدرسین.

روش تحقیق

بافت اجتماعی تحقیق

این تحقیق در دانشگاه صنعتی شاهرود انجام شد. در این دانشگاه مدرسین با یک فرم ارزیابی شامل ۱۵ سوال مورد ارزیابی قرار می گیرند. این فرم کلی است. بنابراین جنبه های تخصصی مربوط به تدریس رشته های مختلف نادیده گرفته می شود. اگر چه ادبیات پیشینه مربوط به ارزیابی تدریس روشهای متعددی را ارائه می دهد، ارزیابی در این دانشگاه متکی بر یک منبع است: ارزیابی دانشجویان از تدریس. چون دفتر ارزیابی دانشگاه متوجه اشکالات و محدودیتهای این روش شده است سعی بر آن است که در آینده از منابع و روش های دیگر از جمله ارزیابی همکار از همکار و ارزیابی دانشجویان دارای معدل بالاتر استفاده شود. شایان ذکر است که نتایج ارزیابی برای ارتقاء و ترفیع و تبدیل وضعیت استفاده می شود. با توجه به پنج سال تدریس در این دانشگاه محقق حساسیت لازم برای این تحقیق را داشت. در واقع گفتگوهای غیر رسمی با دوستان در نیم ساعت بین کلاسها تا حدی حساسیت نظری لازم برای نوشتن طرح تحقیق را فراهم آورد. با توجه به حساسیت موضوع این طرح با اقبال و پذیرش معاونت پژوهشی دانشگاه روبرو شد.

روش جمع آوری داده ها

با استفاده از تئوری مبتنی بر داده ها (استراس و کربین، ۱۹۹۸) محقق برای جمع آوری داده ها از روش نمونه گیری نظری استفاده کرد تا مفاهیم و معیارهای نظری مرتبط را از طریق محاسبه کیفی جمع آوری کند. با پیروی از اسیرمن (۱۹۹۱) محقق برای اهداف زیر از محاسبه استفاده کرد: (۱) آشنا کردن شرکت کنندگان با کم و کیف تحقیق (۲) برقراری رابطه دوستانه و ایجاد بافت لازم بر انجام تحقیق (۳) فهم دقیق دیدگاه های مدرسین در ارتباط با ارزیابی تدریس. این تحقیق با این سوال کلی شروع شد: "شما ارزیابی عملکرد تدریس در دانشگاه را چگونه می بینید؟" تحلیل داده های اولیه مبین نارضایتی مدرسین از نحوه ارزیابی بود. به منظور فراتر رفتن از سطح توصیف و تبیین، محقق سعی بر آن داشت که تاثیر این روش ارزیابی بر زندگی حرفه ای مدرسین را نیز مورد کاوش قرار دهد. بنابراین هدف شکل دهی یک شبکه ی مفهومی بود که از قدرت توصیف، تبیین و پیش بینی بالایی برخوردار باشد. به جایی نمونه گیری آماری که با یک نمونه از پیش تعیین شده شروع می شود، محقق از نمونه گیری نظری استفاده کرد. در این روش سوالات تحقیق از تحلیل داده های اولیه بدست می آید و نمونه گیری تا حد اشباع نظری ادامه می یابد. در این تحقیق بعد از مصاحبه هایی که از دوازده مدرس که حداقل شش سال سابقه تدریس داشتند، اشباع نظری حاصل شد.

تحلیل داده ها

تمام مصاحبه ها به صورت صوتی ضبط شده و متعاقب آن به صورت لغت به لغت به فرم نوشتاری تبدیل شدند. نسخه نوشتاری هر مصاحبه به فرد مصاحبه شونده برگردانده شده تا از صحت آن اطمینان حاصل شود. قبل از انجام مصاحبه به افراد این اطمینان داده شد که نام افراد در گزارش نهایی ذکر نخواهد شد. چرا که در این روش مفاهیم استخراجی حائز اهمیت هستند نه افراد مصاحبه شونده (گلیسر، ۱۹۷۸). برای تحلیل داده ها فرم نوشتاری مصاحبه ها در معرض کد گذاری آزاد قرار گرفت. هدف از این مرحله از تحقیق، کشف مفاهیم نظری و تعیین ویژگی ها و ابعاد این مفاهیم می باشد (استراس و کربین، ۱۹۹۸). کد گذاری محوری سعی در شکل دهی یک ماتریس که هدف آن تعیین رابطه بین مفاهیم و طبقات و زیر طبقه ها ی مفهومی است می باشد. در مرحله نهایی کد گذاری محقق سعی بر تعیین طبقه مفهومی مرکزی دارد. این طبقه مفهومی نقطه اتصال تمام مفاهیم و طبقات مفهومی می باشد. در این تحقیق طبقه مفهومی مرکزی عبارت بود از "ریشه های نگرانی

مدرسین". برای تعیین روایی شبکه مفهومی مستخرج از داده ها، این شبکه به رویت شرکت کنندگان رسید و مورد تایید آنان قرار گرفت. علیرغم روش های دقیق کد گذاری در تئوری زمینه ای و تایید شبکه مفهومی توسط شرکت کنندگان، یافته های تحقیقات کیفی نمی توانند انعکاس دقیقی از واقعیت باشند چرا که واقعیت همیشه جزئی از تصویر کلی می باشد (کلیفورد، ۱۹۸۶).

نتایج تحقیق

ریشه نگرانی مدرسین

توصیف اولین گام در جهت فهم یک پدیده است. در این مورد پدیده ی مورد مطالعه نا رضایتی مدرسین نسبت به ارزیابی عملکرد تدریس می باشد. اما توصیف محدود می شود به معلول، یعنی بخش مرئی پدیده. گام دوم تبیین پدیده از طریق کشف علت یا علت های نامرئی پدیده می باشد. بنابراین در این تحقیق سوال اساسی این است که علل نگرانی در سیستم ارزیابی عملکرد مدرسین در دانشگاه صنعتی شاهرود چیست؟ جمع آوری و تحلیل داده ها چهار ریشه اصلی برای نگرانی مدرسین را نمایان کرد. نظرات مغرضانه دانشجویان، ابزار ارزیابی ناقص، اجرای ناقص ابزار ارزیابی و کاربرد نا مناسب نتایج. هدف از قسمتهای بعدی این تحقیق عبارت است از تایید و مستند سازی این یافته ها با توجه به بریده هایی از نظرات شرکت کنندگان.

نظرات مغرضانه ی دانشجویان

نظرات دانشجویان نمی تواند منبع خوبی برای تصمیم گیری باشد چرا که این نظرات انعکاسی از عامل های دیگری نیز می باشد. اندیشمندان استناد به نظر دانشجویان را با توجه به عدم وجود روایی در آنها رد می کنند (اسپودل، ۲۰۰۲). آنها نظر دانشجویان را چیزی بیش از " کمیت های بی معنی " که به جای ارزیابی عملکرد تدریس منجر به " رقابتهای شخصیتی " می شود نمی دانند (کولیک، ۲۰۰۱). در این دانشگاه استفاده از نظر دانشجویان به عنوان تنها منبع تصمیم گیری نگرانیهای زیادی را ایجاد کرده است. در این راستا یکی از شرکت کننده گان می گوید:

در بیشتر موارد دانشجویانی که وارد رشته ریاضی می شوند دانش پیش نیاز لازم برای تحصیل در این رشته را ندارند. با توجه به نیاز و فشار اجتماعی دانشجویان مستعد تر رشته های مهندسی را در اولویت قرار می دهند. امتحان ورودی دانشجویان را به دو گروه قوی و ضعیف تقسیم می کند. در بیشتر موارد کسانی که در رشته های مهندسی قبول نمی شوند وارد رشته های علوم پایه از جمله ریاضی می شوند. چون این دانشجویان دانش پیش نیاز لازم را ندارند ، تدریس مدرس را نمی فهمند و در پایان ترم عملکرد مدرس را منفی ارزیابی می کنند.

گاهی نظر دانشجویان تحت تاثیر وضعیت ناهمگن کلاس قرار می گیرد. چون قبل از شروع ترم امتحان تعیین سطح وجود ندارد، دانشجویان ضعیف و قوی هر دو درس را انتخاب می کنند. ناهمگن بودن سطح کلاس هم بر یادگیری و هم بر تدریس اثر منفی می گذارد. توضیحات زیر بهتر این وضعیت را توصیف می کند:

در درس زبان عمومی برخی از دانشجویان از سطح تسلط خوبی برخوردارند. بر عکس تعدادی هستند که سطح زبانی ضعیفی دارند. صرف نظر از نحوه ی تدریس مدرس، در چنین کلاسی همواره یک گروه ناراضی هستند: اگر نیاز های گروه قوی در نظر گرفته شود ، گروه ضعیف درس را نمی فهمند و بر عکس اگر سطح کلاس با سطح دانشجویان ضعیف هماهنگ شود، گروه قوی خسته شده و شاکمی می شوند. هر چند ناهمگن بودن کلاس در کنترل استاد نیست، این امر ارزیابی او را تحت اشعاع قرار می دهد.

مشکل دیگر این است که اکثر دانشجویان موفقیت را بر اساس نمره پایان ترم تعیین می کنند. یعنی به جای اینکه نگران یاد گیری باشند بیشتر نگران نمره پایانی هستند. بنابراین به جای ارزیابی تدریس، مدرس را بر اساس سطح نمراتش در آزمون پایانی ارزیابی می کنند. یکی از شرکت کنندگان می گوید:

مشکل دانشجویان این است که تعداد زیادی از آنها از همان ابتدا برای پاس کردن برنامه ریزی می کنند نه برای یادگیری. ضمن تدریس یک سوال به طور متداوم پرسیده می شود، " آیا این نکته در امتحان پایانی می آید؟" اگر جواب مثبت باشد همگی یادداشت برداری می کنند. اگر نکته ای توضیح داده شود که در کتاب نباشد، دانشجویان آن را جدی نمی گیرند چون می دانند در امتحان نمی آید. در واقع اگر سطح نمرات دانشجویان در آزمون پایانی بالا باشد، باید به طور طبیعی انتظار نمره ارزیابی بالایی را داشته باشی. عکس این قضیه نیز طبیعتاً صادق است.

تعداد دانشجویان در یک کلاس هم میتواند نظر آنها در مورد تدریس را تحت الشعاع قرار دهد. در کلاس های شلوغ زمان لازم برای پرسش و پاسخ وجود ندارد. یعنی دانشجویان فرصت شرکت در فعالیتهای کلاسی را ندارند. وقتی کلاس بیش از حد شلوغ است مدرس بیشتر نگران نظم و مدیریت کلاس است تا کیفیت تدریس. یکی از شرکت کنندگان توضیح می دهد:

ارزیابی دانشجویان از عملکرد تدریس من بستگی به تعداد دانشجویان در کلاس دارد. این طبیعی است چرا که در کلاس های شلوغ جایی برای تعامل با دانشجویان نیست. هیچ زمانی برای دریافت بازخورد و تطبیق روش تدریس با نیازهای آنان نیز وجود ندارد. در کلاسی که دارای ۶۰ دانشجو می باشد، تنها چیزی که می توانم چک کنم حضورشان است. در چنین کلاسی وقتی برای چک کردن عملکرد دانشجویان وجود ندارد. بنا بر این من از سطح دانش و مهارت دانشجویانم اطلاعی ندارم. این قضیه تا امتحان پایان ترم ادامه دارد. نتایج امتحان پایان ترم تنها بازخوردی است که در چنین کلاسهای شلوغی دستگیر دانشجو می شود.

دیگر مشکلی که در مورد ارزیابی توسط دانشجویان وجود دارد این است که آنها با کمترین زحمت بیشترین بازدهی را می خواهند. به این معنی که آنها یک جزوه کوتاه که میان بر امتحانات پایان ترم باشد را به یک سر فصل مفصل و سنگین ترجیح می دهند. خصوصاً آنها روشی را ترجیح می دهند که موفقیت در آزمون پایانی را تضمین کند. با هدف دریافت مدرک دانشگاهی، با کوتاه کردن مسیرها دنبال میان بر می گردند. بعضی اساتید در برابر این وسوسه مقاومت می کنند. مثال زیر را در نظر بگیرید:

ارتقاء مدرسین به نظر دانش جویان بستگی دارد. آنها نیز از آن اطلاع دارند. آنها روند تدریس را با تحمیل خواسته هایشان شکل می دهند. آنها مرا منفی ارزیابی می کنند نه به این خاطر که خوب تدریس نمی کنم بلکه به این دلیل که تسلیم خواسته های آنها نمی شوم. درست مانند یک پدر مسئول، یک استاد باید به دانشجویانش چیزی بدهد که احتیاج دارند نه چیزی که دوست دارند. کودک دوست ندارد دارو مصرف کند. اما پدر او را وادار به خوردن دارو می کند. به عنوان استاد، برخی مواقع باید دانشجویان را مجبور کنیم تا در حد توانایی خود مطالعه کنند هر چند خودشان پذیرای این امر نباشند.

برخی اساتید به واسطه عدم آشنایی دانشجویان با ارزیابی، ارزیابی دانشجویان از عملکرد مدرس را مورد انتقاد قرار می دهند. از آنجا که آنها هیچ وقت در طول دبیرستان، معلمان خود را مورد ارزیابی قرار نداده اند، در ورود به دانشگاه نیز نمیدانند چگونه این کار را انجام دهند. هرچه دانشجویان با معیارها و ارزشهای دانشگاهی آشناتر می شوند، تصمیماتشان نیز قابل اعتمادتر میشوند. به این نقطه نظر توجه کنید:

در طول دبیرستان، آنها هیچ نقشی در تصمیم گیری ها ندارند. مسئولین آنها را وادار می کنند کاری را که به آنها گفته شده انجام دهند. آنها هرگز به ارزیابی معلمان فکر نمی کنند. ناگهان وارد دانشگاه می شوند و از آنها خواسته می شود تا اساتید را ارزیابی کنند. آنها هرگز چیزی را ارزیابی نکرده اند. اساتیدی که به دانشجویان ترمهای بالا تدریس می کنند نمرات ارزیابی بالاتری دارند چون این دانشجویان با رسوم و تصمیم گیری های دانشگاهی آشنایی پیدا کرده اند.

ابزار نامناسب ارزیابی اساتید

پیش از استفاده از هر گونه ابزار اندازه گیری و ارزیابی، باید از روایی ابزار اطمینان حاصل کنیم. شرکت کنندگان شکایت دارند که ابزار ارزیابی تدریس در این دانشگاه دارای روایی مناسب نمی باشد. به همین ترتیب، در سطح سوالات، شرکت کنندگان شکایت دارند که بعضی سوالات عملکرد بالا و پایین در تدریس را تمیز نمی دهند. پر تکرار ترین شکایت این است که سوالات اکثراً کلی هستند. از آنجا که این ابزار ارزیابی برای ارزیابی عملکرد تدریس در جنبه های عمومی طراحی شده، مواردی را که نشانگر عملکرد کیفی در یک رشته خاص مثل تدریس زبان و تربیت بدنی باشد پوشش نمی دهد. تفاوت های اساسی زیادی بین روش و تکنولوژی تدریس این دو رشته وجود دارد. هیچ کدام از این تفاوتها در سوالات ابزار ارزیابی کنونی در نظر گرفته نشده اند. اگر این تفاوتها در نظر گرفته نشوند این ارزیابی نمی تواند بازخورد مؤثر مناسبی ایجاد کند. یکی از شرکت کنندگان می گوید:

اول ابزار ارزیابی باید بهبود یابد. مانیزا به یک فرم خاص داریم که در راستای اهداف آموزشی باشد. فرم کلی نمی تواند بسیاری از جنبه های تخصصی یک درس خاص را اندازه بگیرد. یک ابزار کلی نمی تواند برای اندازه گیری عملکرد استاد در تربیت بدنی و مهندسی برق مورد استفاده قرار گیرد. مشکل اصلی چنین ارزیابی این است که نمی تواند بازخورد مناسبی برای بهبود عملکرد به ما بدهد. برای بهبود عملکرد تدریس، بازخورد باید خاص باشد.

برخی از شرکت کنندگان بر این باورند که دانشجویان دانش و مهارت های مورد نیاز برای ارزیابی ابعاد تکنیکی تدریس مثل کفایت سرفصل، روش تدریس، و دانش و مهارت های حرفه ای مدرس را ندارند. این ابعاد باید در هر دانشکده توسط کمیته حرفه ای با یک فرم خاص مورد ارزیابی قرار گیرد. یک شرکت کننده لزوم یک فرم خاص را به صورت زیر شرح می دهد:

سیستم ارزیابی کنونی نمیتواند دانش حرفه ای تدریس از جمله سرفصل و میزان پوشش آن را ارزیابی کند. این ابعاد باید توسط یک جمع حرفه ای ارزیابی شود. قبل از شروع ترم اساتید باید سرفصل و طرح درس هایشان را به کمیته ارائه دهند و آنها باید تصمیم بگیرند که آیا سرفصل یا طرح درسها در راستای اهداف درس هستند یا خیر. بعلاوه کمیته باید امتحان های میان ترم و پایان ترم اساتید را مورد بررسی قرار داده تا درجه تطبیقشان را با اهداف بسنجند.

مشکل دیگر این است که در فرم ارزیابی پرسش ها عملیاتی نشده اند. بنابراین شرکت کنندگان سوالات را نامفهوم، دو پهلو و سلیقه ای توصیف می کنند. یک شرکت کننده سوالات دو پهلو را به صورت زیر شرح می دهد:

من فکر می کنم باید سوالات بیشتری وجود داشته باشد. برای جلوگیری طولانی شدن پرسشنامه، آنها معیارهای زیادی را در یک سؤال گنجانده اند. این سؤال که آیا می تواند کلاس را مدیریت و کنترل کند را در نظر بگیرید. این سؤال دو معیار متضاد را شامل می شود. کنترل یک مفهوم منفی دارد در حالیکه مدیریت مفهومی مثبت. هرچه است آنها یک چیز نیستند گرچه در یک سؤال مطرح می شوند. من اصلاً با مفهوم کنترل موافق نیستم برای کنترل کلاس یک دیکتاتور لازم است به همین منظور مدرس می تواند مانند یک فرمانده عمل کند و ممکن است دانشجویان را به شکل سرباز ببیند. بنابراین او می تواند کلاس را کنترل کند، اما آیا این از لحاظ آموزشی قابل پذیرش است؟

علاوه بر این موارد، بعضی از سوالات در معرض تعبیر و تفسیر سلیقه ای دانشجویان قرار دارند. دانشجویان مختلف این موارد را به روشهای مختلف تفسیر می کنند چون آنها دارای یک مفهوم عینی نیستند. یکی از شرکت کنندگان سؤال مربوط به رعایت اصول و معیارهای حرفه ای را در نظر بگیرید. یکی از شرکت کنندگان سلیقه ای بودن این پرسش را این گونه توصیف می کند:

این سؤال چند پهلو است. یعنی چیزهای مختلف زیادی را اندازه گیری می کند ولی مشخص نمی کند که این چیزها چه هستند. برای هر دانشجو ممکن است نمایانگر یک یا چند معنی خاص باشد بنابراین دانشجویان آن را به راه های نسبتاً زیادی

تفسیر می کنند. برای یک نفر ممکن است نشانگر رفتار کلامی، برای شخص دیگر نشانگر نوع لباس و حتی برای یک نفر دیگر نشانگر مرتب بودن یا رعایت نظم باشد.

برخی از سؤالات به این دلیل که وابسته هستند رد می شوند. برای مثال انتقال مطالب وابسته به تسلط بر مفهوم است. به همین ترتیب انتقال به روش تدریس و تکنولوژی وابسته است. سؤال ۲ دو پهلو است. این سؤال روش تدریس و استفاده از تکنولوژی آموزشی را با هم مورد سؤال قرار می دهد. این روش تدریس است که تکنولوژی را مشخص می کند، بنابراین تکنولوژی به روش تدریس بستگی دارد. وابستگی سؤالات نواحی همپوشی ایجاد می کند. یکی از شرکت کننده ها می گوید:

من اعتقاد دارم ۸۰ درصد همپوشانی بین سؤالات ۱ و ۲ وجود دارد و روش تدریس و استفاده از تکنولوژی مناسب به شدت انتقال مطالب را تسهیل می بخشد. از سوی دیگر انتقال مطالب بدون استفاده از روش تدریس مناسب بسیار مشکل می گردد. من اعتقاد دارم همه چیز به روش تدریس بستگی دارد. یک روش به وضوح نقش تکنولوژی، سرفصل، معلم و دانشجو را مشخص می کند.

اما نگرانی اصلی شرکت کنندگان این است که ابزار ارزیابی مشخصه های متمایز کننده ی عملکرد تدریس را فقط در ۲ سؤال خلاصه می کند. روش تدریس و تهیه ی مطالب درسی و رعایت سرفصل توسط سؤالات ۲ و ۳ مورد ارزیابی قرار می گیرند. برای ارزیابی سرفصل، ابزارهای استاندارد و معتبر شامل بیش از ۱۰ سؤال وجود دارد. ساده کردن فعالیت های کلاسی در دو سؤال منجر به خطای اندازه گیری در ارزیابی شده است. یک شرکت کننده این تبعات را بصورت ذیل شرح می دهد:

دو مدرس را در نظر بگیرید: یکی اصول حرفه ای را رعایت می کند اما سرفصل او کاملاً قدیمی، نامناسب و بی تأثیر است در حالیکه دیگری اصول حرفه ای را رعایت نکرده - چیزی که در معرض قضاوت سلیقه ای دانشجویان است چون مشخص و تعریف نشده- ولی سرفصل او به روز، مناسب و مؤثر و مرتبط با نیاز دانشجویان است. حال در نظر بگیرید که دانشجویان اولی را در رابطه با سرفصل بسیار ضعیف و دیگری را در مورد اصول حرفه ای اش منفی ارزیابی کنند. اگر هر چیز دیگر در این دو استاد یکسان باشد، این دو استاد نمره عملکرد یکسانی خواهند داشت چون هر دو جنبه در یک سؤال مورد سنجش قرار گرفته است. این تصمیم از انصاف به دور است. گرچه اتیکت حرفه ای مهم است موارد آنها نیز باید مشخص گردند. بعلاوه این مورد باید ارزش کمتری نسبت به سرفصل درسی داشته باشد.

اجرای نادرست ابزار

یک ابزار ارزیابی جامع که از روایی و پایایی خوبی برخوردار است اگر تحت شرایط یکسان و استاندارد برگزار نگردد می تواند نتایج نادرست بدهد. اجرای بد می تواند واریانس خطا ایجاد نماید. شرکت کنندگان از زمان اجرا، محرمانه نبودن پاسخ دانشجویان و همچنین محرمانه نبودن سوابق ارزیابی راضی نیستند. شرکت کننده ها اعتقاد دارند اگر اثر این فاکتورها کنترل نگردد ارزیابی منصفانه نیست. یکی از مکررترین شکایات مربوط به اجرای ابزار ارزیابی در پایان ترم بود. یکی از شرکت کننده ها این مشکل را به صورت زیر شرح می دهد:

دو گروه دانشجو در هر کلاس وجود دارند. آنهایی که در طول ترم حاضر بودند در اواخر ترم ممکن است جهت آماده شدن برای امتحانات پایان ترم کلاس در زمان ارزیابی غایب باشند. از سوی دیگر دانشجویانی که در طول ترم غایب هستند برای جبران غیبتشان در پایان ترم در کلاس شرکت می کنند. از آنجا که ابزار ارزیابی تقریباً در پایان ترم برگزار می گردد، آنهایی که در طول ترم غیبت داشته اند در ارزیابی استاد حضور دارند. در یک کلاس ۴۸ نفره فقط ۱۰ نفر عملکرد مرا مورد ارزیابی قرار دادند. در کلاس دیگر فقط ۵ دانشجو این کلاس را مورد ارزیابی قرار دادند. آیا می توانیم نظر تعداد محدودی از دانشجویان را نمایانگر نظر کل کلاس در نظر بگیریم؟

مشکل دیگر این است که دانشجویان ابزار را تحت شرایط یکسان پر نمی کنند. حضور استاد در جلسه ارزیابی می تواند جواب دانشجویان را به شدت شکل دهد. مشکل این است که برخی از اساتید کلاس را در هنگام برگزاری ابزار ترک می کنند در

حالی که برخی دیگر در کلاس می مانند. این کار نه تنها سطح بالایی از استرس را در دانشجویان به وجود می آورد بلکه پاسخهایشان را نیز شکل می دهد. نقطه نظرات شرکت کنندگان این شرایط را بهتر توصیف می کند:

کارشناس برگزاری ابزارسنجش به من گفت که بعضی اساتید در کلاس می مانند و پاسخ های دانشجویان را زیر نظر می گیرند. اساتید با حضورشان ، دانشجویان را مجبور می کنند آنگونه که آنها می خواهند ارزیابی شان کنند. از سوی دیگر در کلاسی که استاد آن را ترک می کند، دانشجویان به راحتی می توانند نظرات خود را بیان کنند. من اعتقاد دارم که نمره ارزیابی استادی که کلاس را ترک نمی کند بیشتر نشانگر نظرات سانسور شده است تا نظرات واقعی دانشجویان. اجرا باید محرمانه بودن نظرات دانشجویان را تضمین نماید.

سومین مشکل در ارتباط با اجرا این است که برگه های ارزیابی در معرض دید بسیاری افراد دیگر قرار می گیرند چرا که بلافاصله بعد از اجرا امضاء و پلمب نمی شوند. این می تواند مشکلاتی را هم برای دانشجویان و هم برای اساتید بوجود آورد. دانشجویان به این شرط جواب می دهند که دیدگاه آنها محرمانه باقی بماند و به همین ترتیب اساتید می خواهند نتایج ارزیابی شان کاملاً محرمانه باشد. شرکت کنندگان شکایت دارند که نتایج ارزیابی محرمانه نمی ماند. یکی از شرکت کننده گان می گوید:

کارشناس ارزیابی می آید و داده های ارزیابی را جمع آوری می کند. سپس آنها را در دفترش برای یک یا چند هفته نگهداری می کند . در طول این مدت این داده ها ممکن است توسط افراد ذینفع دستکاری شود و همینطور ممکن است توسط گروه های دیگری شامل خود اساتید دیده شود. اگر استاد دیدگاه های دانشجویان را مشاهده کند ممکن است این نتایج را بر علیه آنها استفاده کند. بنا بر این من اعتقاد دارم که برگه های ارزیابی باید همان لحظه که جمع آوری می شوند امضاء و پلمب شوند. داده ها باید محرمانه بمانند به طوری که اساتید و دانشجویان به سیستم ارزیابی اعتماد کنند. ما باید افرادی را که باید به داده ها دسترسی داشته باشند مشخص نموده و دلیل دسترسی آنها را هم اعلام نماییم.

استفاده نادرست از نتایج ارزیابی

به دلیل ماهیت درس و بسیاری فاکتور های دیگر ارزیابی دانشجویان از عملکرد تدریس در برخی از دروس و دانشکده ها بالا و در برخی دیگر پایین است. بنا بر این مقایسه عملکرد تدریس در یک درس از یک دانشکده با درس دیگر از دانشکده دیگر بسیار اشتباه برانگیز خواهد بود. یکی از شرکت کنندگان در این رابطه می گوید:

بر اساس نمرات خام در درس زبان عمومی عملکرد من بالا و در درس تئوریهای ارتباطات پایین بود. در ظاهر روشن است که من در درس زبان عمومی عملکرد بهتری داشته ام اما اگر شما آنها را با میانگین گروه در دو دانشکده مقایسه کنید نظر دیگری خواهید داشت. گرچه نمره عملکرد من در درس تئوریهای ارتباطات به نظر پایین تر می رسد، من در دانشکده مربوطه عملکرد خوبی داشته ام زیرا نمره عملکرد من بسیار بالاتر از میانگین آن دانشکده است. از سوی دیگر، وقتی نمره درس زبان عمومی با میانگین گروه مقایسه می شود، عملکرد خوبی ندارم چون نمره پایین تری نسبت به میانگین گروه دارم. برای محاسبه میانگین نمره من، این دو مقدار را با هم جمع می شود که کاری کاملاً غیر منطقی است.

محدودیت دیگر سیستم ارزیابی کنونی این است که که بازخورد سازنده ای که بتواند برای اصلاح مشکلات آموزشی استفاده شود فراهم نمی آرد . با تمرکز بر روی جنبه های عمومی تدریس، سیستم ارزیابی جنبه های فنی تدریس را نادیده می گیرد. بازخوردی که چنین سیستمی فراهم می آورد سازنده نیست . برای اصلاح روش های تدریس در یک درس خاص ، بازخورد باید خاص باشد . نظرات شرکت کنندگان مشکل را بهتر توصیف می کند :

در جنبه های عمومی تدریس من می توانم از دیدگاه دانشجویان برای بهبود تدریس استفاده کنم . برای مثال ، ترم گذشته دانشجویان اعتراض کردند که من به طور منظم در ساعات مشاوره مشخص شده حضور نداشتم . این ترم من سعی کردم این مشکل را حل کنم . اما وقتی دانشجویان عنوان می کنند که از روش تدریس من راضی نیستند من نمی دانم چه باید بکنم تا

تدریس ام را بهبود بخشیم . من دوست دارم تا تدریسم بهتر شود. اما برای انجام این کار من نیاز به بازخورد شفاف در جنبه های خاص تدریس دارم ارزیاب نمی تواند روش تدریس مرا با یک پرسش کلی ارزیابی کنند . این ابزار سوالات خاص که جنبه های مختلف تدریس را پوشش دهد می طلبد .

اما محدودیت های اصلی سیاست کنونی ارزیابی این است که هیچ تاثیری بر روی کیفیت تدریس و یادگیری ندارد . بنابراین محدود است ، چون بیشتر برای استخدام ، ارتقاء و اخراج استفاده می شود . اما سیستم ارزیابی کنونی حتی از انجام این اعمال نیز قاصر است. دفتر نظارت عنوان می کند که :

ما این نتایج را برای ارتقاء، ترفیع و تبدیل وضعیت استفاده می کنیم . اما بعد از رسیدن به این اهداف مدرسین نسبت به کیفیت تدریس بی تفاوت می شوند و نمره ارزیابی خیلی از آنها افت می کند . با دانشجویان می گویند که برخی از این اساتید بسیار بد اخلاق و بی نظم شده اند. مشکل دیگر این است که از نتایج ارزیابی نمی توان برای کنترل عملکرد بد استفاده کرد . سیستم این است که اگر نمره ارزیابی استاد برای ۲ ترم متوالی کمتر از ۲/۵ باشد ، ساعات تدریس وی کم گردد . اما مشکل اینجاست که در بعضی دانشکده ها و خصوصاً در بعضی دروس ، ما استاد کافی در اختیار نداریم . این استراژی در چنین مواقعی شکست می خورد . سومین مشکل این است که در حال حاضر هیچ طرح انگیزشی برای کسانی که عملکرد خوبی دارند وجود ندارد . در آینده ما قصد داریم از نمرات استاندارد ارزیابی برای تشویق کسانی که عملکرد خوبی داشتند استفاده کنیم .

تاثیر روش ارزیابی بر زندگی حرفه ای مدرسین

همانگونه که در بخش نتایج مورد بحث قرار گرفت ، یک محدودیت کلی سیستم ارزیابی مورد استفاده در دانشگاه صنعتی شاهرود این است که بر روی ارزیابی پایانی تاکید دارد . این روش دو اثر منفی دارد . اول اینکه روی عملکرد ضعیف مدرس تمرکز می کند و مدارکی دال بر نقاط ضعف مدرس جمع آوری می کند. بنابراین با قضاوتی بودن خود ، جوی شک و عدم اعتماد در بین مدرسین به وجود می آورد . ثانیاً عملکرد خوب را به رسمیت نشناخته و تشویق نمی نماید . پس حالتی بی تفاوت به وجود می آورد . به بیان دقیق تر ، وقتی ارزیابی دانشجویان از عملکرد تدریس مدرسا در گروه ضعیف قرار می دهد نتایج ارزیابی صحیح در نظر گرفته می شود حال آنکه اگر دانشجویان عملکرد او را موفق بدانند نتایج غیر قابل اعتماد و نادرست تلقی می شوند . گروه ضعیف گروه قوی را به افزایش نمرات پایانی و جدی نبودن در حضور و غیاب دانشجویان در کلاس متهم می کنند . یکی از شرکت کنندگان می گوید :

زمانی که من گزارش ارزیابی ام را دریافت کردم و رئیس دانشکده مرا متهم کرد به اینکه به دلیل درصد بالایی قبولی در امتحان پایان ترم دانشجویان به من علاقه دارند . او حتی مرا متهم به دوستی با دانشجویان و جدی نبودن در کنترل کلاس نمود . او شکایت داشت که نمره ارزیابی وی پایین است زیرا او در حضور و غیاب و نمره پایان ترم جدی است.

به علاوه ، از آنجا که این ارزیابی حاوی هیچ گونه طرح انگیزشی نمی باشد ، یک حالت بی تفاوتی در بین هر دو گروه ضعیف و قوی ایجاد می کند و همین طور دانشجویان این ارزیابی را جدی تلقی نمی کنند . یکی از شرکت کنندگان می گوید:

تقریباً همه نسبت به ارزیابی بی تفاوت هستند. دلیل این است که هیچ تفاوتی بین آنهایی که نمره پایین می گیرند و آنهایی که نمرات بالا می گیرند وجود ندارد. آنها بی تفاوتند چون تاثیر ارزیابی را در محیط نمی بینند. ما باید از نتایج ارزیابی برای بهبود عملکرد آنهایی که نمرات پائین گرفته اند استفاده کنیم. باید کارگاه های تدریس برای آنها دایر نمائیم. از آنجایی که هیچ جایی برای بهبود در نظر گرفته نشده است، گروه ضعیف بی تفاوت است. آنان که نمرات بالا گرفته اند باید مورد تشویق قرار گیرند. عملکرد خوب آنها باید به رسمیت شناخته شود و منتشر گردد. آنها باید به عنوان الگوی موفق تدریس عمل کنند. از آنجا که این امر هیچ وقت اتفاق نیافتاده است، این گروه نیز بی تفاوت است. دانشجویان نیز به همین منوال بی تفاوتند. آنها اساتید دارای نمرات ارزیابی پائین را می بینند که همان درس ها را بدون هیچ گونه بهبودی دوباره تدریس می کنند. آنها اعتقاد دارند که ارزیابی اساتید اتلاف وقت است.

بحث و نتیجه گیری

با مقایسه ی الگو های نظری ارزیابی و آنچه در عمل در دانشگاه صنعتی شاهرود به عنوان ارزیابی روی می دهد، در می یابیم که شکاف عظیمی بین تئوری و عمل وجود دارد. الگوهای نظری امیدوارکننده است ولی در عمل ارزیابی از سنت های محلی تبعیت می کند تا یافته های تحقیقاتی. با مقایسه ی ارزیابی در جنبه های نظری و عملی می توانیم به دو نقطه ضعف اصلی در سیستم ارزیابی دانشگاه صنعتی شاهرود و دیگر دانشگاهها پی ببریم: استناد به نظر دانشجویان به عنوان اصلی ترین منبع و بسنده کردن به ارزیابی پایانی برای ترفیع، ارتقا، و تبدیل وضعیت و نادیده گرفتن نقش ارزیابی در بهبود کیفیت تدریس و رشد مدرس. برای مشخص و روشن نمودن این مشکلات سیستم ارزیابی دانشگاه صنعتی شاهرود موارد زیر را مد نظر قرار دهد:

- بجای استناد به نظر دانشجویان به عنوان اصلی ترین منبع، تصمیمات بر پایه منابع چندگانه از قبیل بررسی طرح درس، مشاهده ی اجرای طرح درس در عمل و همچنین بررسی پورت فوی استاد توسط یک کار گروه تخصصی.
- بجای استناد به ارزیابی پایانی باید بر هر دو نقش ارزیابی تاکید شود. وقتی این دو با هم تلفیق شوند شرایط لازم برای رشد مدرس را به وجود می آورند (نولانو هوور، ۲۰۰۵ و برانت، ۲۰۰۷).
- استفاده از نتایج این تحقیق نیز می تواند کمک خوبی به بر طرف کردن نقاط ضعف استفاده از نظرات دانشجویان کند.

به طور خلاصه، نتیجه تحلیل دیدگاه های مدرسین راهکارهای زیادی را برای بهبود ارزیابی دانشجویان از عملکرد تدریس ارائه می دهد. بدین منظور سیستم ارزیابی بهتر است:

- از روایی و پایایی ابزار ارزیابی اطمینان حاصل کند.
- از قابل قبول بودن ضریب تشخیص هر سوال اطمینان حاصل کند. یعنی شواهدی ارائه دهد که هر سوال گروه های قوی را از گروه های ضعیف تمیز می دهد.
- نمرات خام را به نمرات استاندارد تبدیل نموده و عملکرد را با توجه به نمره استاندارد بسنجد نه با استناد به نمرات خام.
- واریانس های مربوط به محدودیت های محیطی از جمله نا همگن بودن کلاس، اجرای غیر استاندارد ابزار و واریانس های غیر مرتبط دیگر را از واریانس مربوط به کیفیت تدریس جدا نماید. در حال حاضر، ارزیابی تمام واریانس نمره ی ارزیابی را مرتبط با تدریس میدانند. از نظر منطقی مدرسان مسئول محدودیت های فراگیران و محیط هستند.
- تمام سوالات به صورت تعریف عملیاتی ارائه شوند تا به طور عینی قابل سنجش باشند.
- اجرای ابزار تحت شرایط یکسان صورت پذیرد و از محرمانه بودن نتایج آن اطمینان حاصل شود.
- از ارزیابی به عنوان مکانیسمی علمی برای ایجاد رشد مدرس و بهبود کیفیت تدریس استفاده شود.

منابع و مأخذ

- Attinello, J. R., Lare, D. W., & Source, F. (2006). The value of teacher portfolios for evaluation and growth. *NASSP Bulletin*, 90(2), 132-152.
- Berry, B. (2007, March). *Connecting teacher and student data: Benefits, challenges, and lessons learned*. Presentation for the Data Quality Issue meeting, Washington, DC. Retrieved February 25, 2008, from http://www.dataqualitycampaign.org/files/Meetings-DQC_Quarterly_Flyer_031207.pdf
- Brandt, C., Mathers, C., Oliva, M., Brown-Sims, M., & Hess, J. (2007). *Examining district guidance to schools on teacher evaluation policies in the Midwest Region* (Issues & Answers Report, REL, 2007-No.030). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional



- Educational Laboratory Midwest. Retrieved February 25, 2008, from http://ies.ed.gov/ncee/edlabs/regions/midwest/pdf/REL_2007030_sum.pdf
- Brown, S. C. (1999). *Learning across the campus: How college facilitates the development of wisdom*. Unpublished doctoral dissertation, University of Maryland, College Park.
- Clifford, J. (1986). Introduction: Partial truths. In J. Clifford & G. Marcus (Eds.), *Writing culture: The poetics and politics of ethnography* (pp. 1-26). Berkeley, CA: University of California Press.
- Crew, L. A., Everitt, T. J., & Nunez, R. W. (1984). *Improving teacher performance through systematic teacher evaluation*. Virginia Beach, VA: American Association of School Personnel Administrators.
- Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional learning*. Princeton, NJ: Educational Testing Service.
- Glaser, B. (1978). *Theoretical sensitivity*. Mill Valley, CA: Sociology Press.
- Goffman, E. (1959). *The presentation of self in everyday life*. New York, NY: Doubleday Anchor Books.
- Haraway, D. (1988). Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist Studies*, 14(3), 575-599.
- Kulik, J. A. (2001). Student ratings: Validity, utility and controversy. *New Directions for Institutional Research*, 27(5), 9-25.
- Marshall, K. (2005). It's time to rethink teacher supervision and evaluation. *Phi Delta Kappan*, 86(10), 727-735.
- Mujis, D. (2006). Measuring teacher effectiveness: Some methodological reflections. *Educational Research and Evaluation*, 12(1), 53-74.
- Nolan, J., Jr., & Hoover, L. A. (2005). *Teacher supervision and evaluation: Theory into practice*. New York, NY: Wiley.
- Peterson, K. (2000). *Teacher evaluation: A comprehensive guide to new directions and practices* (2nd ed.). Thousand Oaks, CA: Corwin Press, Inc. (ERIC Document Reproduction Service No. ED445087)
- Peterson, P. L., & Comeaux, M. A. (1990). Evaluating the systems: Teachers' perspectives on teacher evaluation. *Educational Evaluation and Policy Analysis*, 12(1), 3-24.
- Price, J., & Schwabacher, S. (1993). *The multiple forms of evidence study: Assessing reading through student work samples, teacher observations and tests*. New York, NY: National Center for Restructuring Education, Schools and Teaching.
- Seidman, I. E. (1991). *Interviewing as qualitative research: A guide for researchers in education and the social sciences*. New York, NY: Teachers College Press.
- Shannon, D. M. (1991, February). *Teacher evaluation: A functional approach*. Paper presented at the annual meeting of the Eastern Educational Research Association, Boston, MA.
- Sproudle, R. (2002). The un-determination of instructor performance by data from the student evaluation of teaching. *Economics of Education Review*, 21, 287-294.
- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Thousand Oaks, CA: Sage.
- Stronge, J. H. (2007). Planning and organizing for instruction. In J. H. Stronge (Ed.), *Qualities of effective teachers* (pp. 212-243). Alexandria, VA: Association for Supervision and Curriculum Development.
- Tucker, P. D., Stronge, J. H., & Gareis, C. R. (2002). *Handbook on teacher portfolios for evaluation and professional development*. Larchmont, NY: Eye on Education.
- Uhlenbeck, A. M., Verloop, N., & Beijaard, D. (2002). Requirements for an assessment procedure for beginning teachers: Implications from recent theories on teaching and assessment. *Teachers College Record*, 104(2), 242-256.