


A Longitudinal Study of the Relationship between Students' Final Grade and their Evaluation of Professors' Teaching in the Next Semesters

Seyed Ahmad. Madani^{1*} 

¹ Assistant Professor, Department of Educational Sciences, Faculty of Humanities, University of Kashan, Kashan, Iran

* Corresponding author email address: madani@kashanu.ac.ir

Article Info

Article type:

Original Research

How to cite this article:

Madani, S. A. (2023). A Longitudinal Study of the Relationship between Students' Final Grade and their Evaluation of Professors' Teaching in the Next Semesters. *Quarterly Journal of Research and Planning in Higher Education*, 29(4), 83-101.



© 2023 the authors. Published by Institute for Research and Planning in Higher Education (IRPHE), Tehran, Iran. This is an open access article under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License.

ABSTRACT

Student evaluation of professors is one of the common tools of universities to improve the quality of teaching in classrooms. Since in Iran's higher education system, student evaluations are used for annual promotion, contract extension or change of status, and the promotion of professors' scientific rank, the correct analysis and interpretation of the results of these evaluations is of foremost importance. This longitudinal research, which is a panel study, has analyzed the relationship between students' grades and their teacher's evaluation in the following semesters. The statistical population of the research was made up of more than 200,000 pairs of "Student's Final Grade-Teacher's Evaluation Grade" which were obtained from students studying at the University of Kashan in the conditions of face-to-face education before the pandemic. After conditioning and sifting the data, the final analysis has been done on the grades-evaluations of 3098 students during four semesters as a research sample. The regression analysis of the random effects model showed that with a one-unit increase in the final grade of the student, his/her evaluation grade of the teacher's teaching in the next semesters increases by about five percent. Additional analyzes showed that the grade can create a special mental attitude in the student, which can affect the student's evaluation of the professor even three semesters later. It seems that university administrators, educational councils, and university admissions boards should revise the guidelines for calculating and interpreting data from student evaluations, taking into account the various consequences of grade inflation at the university level.

Keywords: higher education, student evaluation of teaching, quality of teaching, professor evaluation, academic achievement.



مطالعه طولی ارتباط نمره پایانی دانشجویان با ارزشیابی آن‌ها از تدریس مدرسان در نیمسال‌های بعدی

سید احمد مدنی^۱

۱. استادیار گروه علوم تربیتی، دانشکده علوم انسانی، دانشگاه کاشان، کاشان، ایران

*ایمیل نویسنده مسئول: madani@kashanu.ac.ir

اطلاعات مقاله

چکیده

نوع مقاله

پژوهشی اصلی

نحوه استناد به این مقاله:

مدنی، سید احمد. (۱۴۰۲). مطالعه طولی ارتباط نمره پایانی دانشجویان با ارزشیابی آن‌ها از تدریس مدرسان در نیمسال‌های بعدی. *فصلنامه پژوهش و برنامه ریزی در آموزش عالی*، ۲۹(۴)، ۱۰۱-۸۳.



© ۱۴۰۲ تمامی حقوق انتشار این مقاله متعلق به نویسنده است. انتشار این مقاله به صورت دسترسی آزاد مطابق با گواهی (CC BY 4.0) صورت گرفته است.

ارزشیابی دانشجویان از اساتید یکی از ابزارهای رایج دانشگاه‌ها برای بهبود کیفیت تدریس در کلاس‌های درس است. از آنجا که در نظام آموزش عالی ایران از ارزشیابی‌های دانشجویی برای ترفیع سالیانه، تمدید قرارداد یا تبدیل وضعیت و ارتقای مرتبه علمی اساتید نیز بهره‌برداری می‌شود تحلیل و تفسیر صحیح نتایج این ارزشیابی‌ها اهمیت مضاعف دارد. این پژوهش طولی که از نوع مطالعه پانلی است ارتباط میان نمرات دانشجویان را با ارزشیابی آن‌ها از استاد در نیمسال‌های بعد مورد واکاوی قرار داده است. جامعه آماری پژوهش را بیش از ۲۰۰ هزار جفت «نمره پایان‌ترم دانشجوی-نمره ارزشیابی استاد» تشکیل داده که از دانشجویان شاغل به تحصیل دانشگاه کاشان در شرایط آموزش حضوری پیش از کرونا به دست آمده‌اند. پس از شرط‌گذاری و غربال داده‌ها، تحلیل نهایی روی نمره-ارزشیابی‌های ۳۰۹۸ نفر دانشجوی طی چهار نیمسال به عنوان نمونه تحقیق انجام شده است. تحلیل رگرسیونی مدل اثرات تصادفی نشان داد که با یک واحد افزایش در نمره پایان‌ترم دانشجوی، نمره ارزشیابی او از تدریس استاد در نیمسال بعدی حدود پنج درصد افزایش می‌یابد. تحلیل‌های تکمیلی نشان داد که نمره می‌تواند آمایه ذهنی ویژه‌ای در دانشجوی ایجاد کند که این آمایه می‌تواند حتی تا سه نیمسال بعد ارزشیابی دانشجوی از استاد را تحت تأثیر قرار دهد. به نظر می‌رسد مدیران دانشگاهی، شوراهای آموزشی و هیأت‌های جذب دانشگاهی باید دستورالعمل‌های محاسبه و تفسیر داده‌های ارزشیابی‌های دانشجویی را با در نظر گرفتن پیامدهای مختلف تورم نمرات در سطح دانشگاه بازنگری کنند.

کلیدواژگان: آموزش عالی، کیفیت آموزش، تدریس، ارزشیابی دانشجویان از تدریس، ارزشیابی اساتید، پیشرفت تحصیلی.

مقدمه

ارزشیابی یکی از اساسی‌ترین ابزارها برای بهبود عملکرد شغلی در حرفه‌های مختلف است (Marshall, 2009). در اکثر نظام‌های آموزش عالی، ارزشیابی دانشجویان از اساتید یکی از راهبردهای بهبود کیفیت آموزش و یادگیری تلقی می‌شود (Danielson & McGreal, 2000; Lakeman et al., 2022; Marshall, 2022; Radmacher & Martin, 2001; Zibaei & Azizi 2012). داده‌های برآمده از ارزشیابی دانشجویان، همچنین کاربردهای ویژه‌ای برای مسیر حرفه‌ای خود اساتید دارند (Berezvai et al., 2021). اگر چه دغدغه‌هایی که درباره اعتبار این ارزشیابی‌ها وجود دارد اما هنگام تصمیم‌گیری درباره تمدید قرارداد^۱، تبدیل وضعیت استخدامی به «رسمی»^۲، و ارتقای مرتبه علمی به جد مورد استناد قرار می‌گیرند (Linse, 2017). این در حالی است که استادان صرف نظر از ابهامات مربوط به چگونگی اجرای این ارزشیابی‌ها، نگرانی‌هایی نیز درباره چگونگی بهره‌برداری و تفسیر نتایج آن‌ها دارند (Boysen et al., 2014; Dargahi & Mohammadzadeh 2013; Ziaee et al., 2007). علی‌رغم وجود این نگرانی‌ها مدیران مؤسسات آموزشی به دلیل الزام قانونی برای تکمیل پرونده اساتید موظفانند امتیازات آموزشی آن‌ها را محاسبه و تفسیر کنند. بر اساس آئین‌نامه ارتقا امتیازات آموزشی دارای بخش‌های مختلفی نظیر فعالیت‌های برنامه‌ریزی درسی، نظم و شؤون مدرسی و کمیت تدریس است. طبق گزارش حقیقی، اکبری، چراپین و کریمی (Haghighi et al., 2021) از میان این مؤلفه‌ها به ارزشیابی‌های دانشجویی بین ۴۰ تا ۵۰ درصد وزن داده می‌شود. نکته اساسی اما اینست که اهداف ارزشیابی‌های دانشجویان از اساتید فقط به تکمیل پرونده استادان محدود نیست.

اهداف ارزشیابی دانشجویان از اساتید در سه محور اساسی خلاصه می‌شود: (۱) ارائه بازخورد تشخیصی به استادان برای بهبود آموزش (هدف تکوینی)؛ (۲) تصمیم‌گیری‌های پرسنلی نظیر تبدیل وضعیت استخدامی و ارتقای مرتبه علمی (هدف پایانی)؛ و (۳) کمک به دانشجویان در انتخاب دوره‌ها (Schmelkin et al., 1997). مارش و روکه (Marsh & Roche, 1993) به کارکرد چهارمی نیز تحت عنوان بهره‌برداری از ارزشیابی‌های برای بهبود تحقیقات در زمینه تدریس و یادگیری اشاره کرده‌اند. در پژوهش‌های ایرانی نیز به امکان بهره‌گیری از داده‌های ارزشیابی اساتید برای بهبود کیفیت خدمات آموزشی دانشگاه (Beheshtirad et al., 2014) و نیز برای بهبود قدرت بیان، روش تدریس و تسلط علمی استاد (Ganbari, 2018) اشاره شده است. اما واقعیت این است که از میان این چهار کارکرد اساسی، تصمیم‌گیری‌های پرسنلی بیشترین مناقشه و چالش را ایجاد کرده است. اخیراً گرایش به کسب افتخارات آموزشی نیز به ابعاد مسأله را پیچیده‌تر کرده است (Griffin, 2020). یکی از دلایل استمرار مناقشه‌ها وجود شواهد ضد و نقیض در زمینه ارزشیابی‌های دانشجویی است.

نتایج پژوهش‌های کلاسیک دهه‌های هفتاد و هشتاد نشان‌گر پایایی و روایی ارزشیابی‌های دانشجویی بود (Marsh, 1977; Marsh, 1982). اما به مرور اتهامات عمده‌ای به ارزشیابی‌های وارد شد از جمله اینکه پایایی و روایی ندارند؛ بیش از اندازه با نمرات همبستگی دارند، محبوبیت ندارند و از عواملی نظیر ویژگی‌های درس، اندازه کلاس، مقطع درس، الزامی یا اختیاری بودن درس و ساعت کلاس در طول روز تأثیر می‌پذیرند. همچنین این ارزیابی‌ها از ویژگی‌های استادان و نیز ویژگی‌های خود دانشجویان (جنسیت، مقطع تحصیلی، سنوات تحصیل و میزان علاقه آن‌ها به درس) تأثیر می‌پذیرند (Basirat et al., 2019; Schmelkin et al., 1997). از سوی دیگر، نقش ویژگی‌های مؤسسه آموزشی و کیفیت خدمات آموزشی آن بر روی روش تدریس استادان نیز روشن شده است (Beheshtirad et al., 2014). با ملاحظه مجموع این شواهد می‌توان ادعا کرد که عوامل مهمی نتایج ارزشیابی‌های دانشجویان را تحت تأثیر قرار می‌دهند که دانشگاه‌ها می‌توانند بسیاری از آن‌ها را مدیریت کنند.

1- retention

2- tenure

البته چالش‌هایی ماهوی نیز در مورد نظرسنجی‌های دانشجویی مطرح شده است. از جمله اینکه هدف غایی آن‌ها بهبود یادگیری است اما در عمل، بین نمرات ارزشیابی تدریس و مقدار یادگیری واقعی دانشجویان همبستگی ناچیزی وجود دارد (Clayson, 2009). در واقع، نتایج ارزشیابی‌ها به میزان یادگیری واقعی دانشجو در درس ربطی ندارند (Uttl et al., 2017). از نظر برخی محققان این ارزشیابی‌ها بیشتر تحت تأثیر سطح توانایی و ویژگی‌های شخصیتی دانشجویان می‌باشند (McMillan et al., 2021). حتی ادراک دانشجو از نظام ارزشیابی اساتید به عنوان یک عامل اثرگذار مورد توجه واقع شده است. بر اساس پژوهش اسپورن و کریستیانز (Spooren & Christiaens, 2017) دانشجویانی که نظام ارزشیابی تدریس در مؤسسه را کارآمد و ارزشمند می‌پندارند نمرات بالاتری به استادان می‌دهند. حتی تعداد درجات مقیاس یا همان گزینه‌های نمره‌دهی دانشجو به استاد نیز می‌تواند اثرگذار باشد (Landrum & Braitman, 2008). سرانجام، آنگونه که لیکمن و همکاران (Lakeman et al., 2022) اشاره کرده‌اند گمنامی¹ و عدم امکان تشخیص هویت ارزیابان نیز نتایج ارزشیابی را تحت تأثیر قرار می‌دهد. باید توجه کرد که در برخی نظام‌های آموزش عالی، نتایج ارزشیابی‌های دانشجویی برای دیگر دانشجویان قابل مشاهده است و آن‌ها می‌توانند درس‌های مورد پسند و علاقه خویش را بر اساس همین ارزشیابی‌ها شناسایی کنند (He et al., 2022). بنابراین، بر اساس پژوهش‌های موجود تردیدی نیست که منظومه‌ای از عوامل تصادفی یا غیرتصادفی می‌توانند داده‌های ارزشیابانه را در معرض سوگیری قرار دهند.

نکته مهم اینست که مدیران دانشگاهی نباید فرض را بر این بگذارند که سوگیری‌های مزبور پیامدهای مخرب ندارند یا اینکه آثار منفی آن چندان قابل توجه نیست. واقعیت اینست که این سوگیری‌ها در درجه اول تأثیرات مختلفی بر سبک کاری استادان دارند. به عنوان مثال، استادان انتظار دارند دانشجویان اهمیت درس را درک کنند و این قدرشناسی در ارزشیابی آن‌ها از درس انعکاس داشته باشد. اما در واقع، ارزشیابی دانشجویان از اساتید، ربطی به سطح تعهد دانشجویان به موضوع و درس ندارد (Beran & Violato, 2010) و بیشتر انعکاسی از تصورات دانشجویان درباره میزان دشواری درس است (Addison et al., 2006). واقعیت اینست که استادان این تعارض‌ها را با تمام وجود درک می‌کنند. آن‌ها در ابتدا بر اساس ارزشیابی‌های دانشجویی تغییراتی را در سبک و رویه تدریس خود ایجاد می‌کنند. اما چنین نیست که همیشه نتیجه مثبتی حاصل شود. یکی از دلایل این است که بسیاری از فرم‌های ارزیابی بازخوردهای مفیدی را به مدرسان نمی‌دهند (Marshall, 2009). شاید به همین علت است که برخی از استادان به مرور زمان یاد می‌گیرند که ارزشیابی‌های دانشجویی را نادیده بگیرند! متخصصان نیز به سردرگمی ناشی از شواهد پژوهشی ضد و نقیض اشاره کرده‌اند. بولر (Buller, 2012) از متخصصان ارزشیابی اساتید در این باره می‌گوید: «بیان این حرف اغراق‌آمیز نیست که اگر شما جستجوی جدی و عمیقی انجام دهید می‌توانید مطالعاتی را در تأیید/عدم‌تأیید هر فرضیه‌ای درباره عوامل اثرگذار بر ارزیابی‌های دانشجویان از اساتید پیدا کنید» (ص، ۴۰). نظر بولر اینست که از ارزشیابی‌های دانشجویی اطلاعات چندان به دست نمی‌آید زیرا شاید تعداد دانشجویان ارزیابی‌کننده کم باشد، یا اینکه دانشجویان به دلایلی به استاد علاقه ویژه داشته باشند. بنابراین، اگر قرار است مدیران دانشگاهی از نتایج ارزشیابی‌هایی که در معرض انواع سوگیری‌ها می‌باشند در تصمیم‌گیری‌های مهم استفاده نمایند شناسایی منشأ این سوگیری‌ها و اتخاذ تدابیر مناسب در این زمینه بسیار اهمیت دارد.

یکی از عوامل اثرگذار بر سوگیری‌ها رویکرد استادان برای سنجش آموخته‌ها و تعیین نمره پایانی دانشجویان است. از منظر آموزشی «نمره» باید بازتابی از شایستگی‌های تحصیلی و یادگیری‌های واقعی دانشجویان در پایان دوره آموزشی باشد. در عمل اما مدرسان به هنگام ارزشیابی و تعیین نمره پایانی فراگیران به ابعاد دیگری غیر از یادگیری (نظیر میزان تلاش، میزان پیشرفت، تکمیل و تحویل کار، مشارکت، میزان توجه در کلاس، و تا حدی به شخصیت و رفتارهای دانشجو) امتیاز می‌دهند. گذشته از این، نمره‌دهی آن‌ها در هر دوره یا مقطع آموزشی و حتی در درون یک کلاس نیز برای افراد متفاوت است (Brookhart et al., 2016). البته، این تفاوت‌ها نه به دلیل سلیقه شخصی استاد، بلکه

¹- anonymous evaluation

به دلیل تشخیص آکادمیک او لحاظ می‌شوند و چنین کاری حتی مورد تأکید متخصصان سنجش و اندازه‌گیری نیز هست. به عنوان مثال، مارک ویلسون (Wilson, 2019) موافق است که فهرستی از ویژگی‌های شخصی که برای دستاوردها و موفقیت کلی در زندگی مهم‌اند و خود دانشجویان، والدین، مدارس و کارفرمایان آینده نیز برای آن‌ها ارزش قائل‌اند در نمره‌دهی مورد توجه قرار گیرند. پیامد منطقی و طبیعی توجه به تفاوت‌ها اینست که دانشجویان به مرور متوجه این موضوع می‌شوند. همان‌گونه که استادان در مورد کاهش غیرمنصفانه ارزشیابی‌های خود حساس‌اند دانشجویان نیز ادراکاتی ولو شهودی از دلایل افت یا رشد نمرات خویش دارند. اما در اینجا یک مسأله روان‌شناختی مهم وجود دارد. طبق نظریه اسناد^۱ (Nevid, 2018) دانشجویانی که «اسناد بیرونی» دارند علت کاهش نمره خود را در عوامل بیرونی جستجو می‌کنند. در چنین وضعیتی، آمایه ذهنی دانشجو ارزشیابی وی از تدریس استاد را تحت تأثیر قرار دهد. حتی احتمال دارد در گذر نیمسال نیز دانشجو رفتارها و اقدامات استاد را متناسب با آمایه قبلی خویش تفسیر و برداشت کند.

عوامل متعددی می‌توانند آمایه ذهنی دانشجو نسبت به استاد را تحت تأثیر قرار دهند. نمره یکی از این عوامل است، به ویژه اگر تابعی از یادگیری واقعی دانشجو باشد و خود دانشجو نیز نسبت میان نمره و یادگیری‌های خود را درک کرده باشد. در چنین حالتی، خودپنداره تحصیلی دانشجو نیز تقویت می‌شود. نتایج تحقیقات نیز مؤید آن است که تدریس اثربخش استاد باعث بهبود حس خودکارآمدی دانشجویان می‌شود (Ezadi et al., 2015). دانشجویان این اثرگذاری را درک کرده و برای آن ارزش قائل می‌شوند. اما در مورد طیف‌های ویژه‌ای از دانشجویان، نمره به طور ویژه‌ای ادراکات آن‌ها را تحت الشعاع قرار می‌دهد. در چنین شرایطی، ارزشیابی دانشجویی استاد در پایان ترم نه تابعی از میزان تلاش و اثربخشی تدریس استاد، بلکه تابعی از پنداره‌ها و آمایه ذهنی دانشجو خواهد بود. با در نظر گرفتن متغیرهایی نظیر افت تحصیلی، مشروطی و فقدان علاقه دانشجویان به درس‌ها که در حال حاضر در کلاس‌های درس دانشگاهی تجربه می‌شود و بر مبنای نظریه اسناد احتمال می‌رود که دانشجویان علت افت نمره خویش را به استاد یا تدریس استاد نسبت دهند. طبیعی است که اگر چنین وضعیتی بروز کند در صورتی که نمره دانشجو از سطح مورد انتظار وی کمتر بود شاید به اسناد موقعیتی^۲ روی آورد و تدریس استاد را ضعیف‌تر ارزیابی کند. افزون بر سوگیری ناهشیارانه، برخی شواهد نشان‌گر آن‌اند که ممکن است دانشجو به طور عامدانه و آگاهانه نیز استاد را در سطحی کمتر از استحقاق وی ارزیابی کنند. به عنوان مثال، (Lakeman et al., 2022) مشاهده کردند که پنهان بودن هویت دانشجویان به هنگام ارزیابی می‌تواند به بیان نظرات تهاجمی، تنبیهی و آبروریزی استاد منجر شود. حال، طرح پژوهش طولی حاضر می‌تواند به عنوان شواهدی برای تأیید یا رد فرضیه اخیر مورد تفسیر قرار گیرد. به طور مشخص این سؤال مطرح است که نمره‌ای که دانشجو در ترم قبل دریافت کرده تا چه اندازه می‌تواند ارزشیابی او از استاد در نیمسال‌های آتی را پیش‌بینی کند؟

اما چرا کاوش در ماهیت ارتباط این متغیرها اهمیت دارد؟ چرا تعیین اثرات احتمالی نمره دانشجویان بر ارزشیابی آن‌ها از تدریس استادان اهمیت دارد؟ اصلی‌ترین پاسخ به تغییر هشیارانه یا ناهشیارانه سبک تدریس استادان دانشگاهی مربوط می‌شود. به نظر مارش (Marsh, 1982) این پدیده خطرناک است که استادان بخواهند با دادن نمرات بالاتر و خواستن کار کمتر، در ارزشیابی‌های تدریس نمرات بهتری بگیرند. مارش (Marsh, 2001) به این باور رایج استادان اشاره می‌کند که می‌پندارند یکی از راه‌های به دست آوردن نمرات بالا در نظرسنجی‌های دانشجویی، تدریس ساده و آرام محتوایی آسان است که یادگیری آن به تلاش کمی نیاز دارد و دانشجویان را به چالش نمی‌کشد. بروز چنین وضعیتی در دانشگاه‌های ایران باید زنگ خطر را برای مدیران دانشگاهی به صدا در آورد. یادآوری مجدد این نکته اهمیت دارد که در شرایط فعلی آموزش عالی ایران نتایج ارزشیابی‌های دانشجویی در انواعی از فرایندهای ارتقاء، تبدیل وضعیت استخدامی، ترفیع سالیانه و اخیراً اعطای پایه تشویقی

1- attribution theory

2- situational attribution

مورد استفاده قرار می‌گیرند. بنابراین، تأکید بیشتر بر نتایج ارزشیابی‌ها می‌تواند در گذر زمان بر کیفیت واقعی فرایندهای یاددهی-یادگیری اثر منفی بگذارد.

طرح مسأله پژوهش حاضر به شکل فوق نباید این تصور را ایجاد کند که نمره‌ها بیش از هر چیز از ارزشیابی استادان اثر می‌پذیرند. بر عکس، واقعیت اینست که پیشرفت تحصیلی دانشجویان و «نمره» و «معدل» به عنوان شاخص‌های اصلی آن تحت تأثیر عوامل بی‌نهایت متغیری است. از کم و کیف خوابگاه (McEwan & Soderberg, 2006) تا میزان انگیزش شخصی دانشجویان (Yousefi et al., 2018) همگی بر نمره اثر می‌گذارند. به قول جان هتی (Hattie, 2008) تقریباً هر عاملی بر پیشرفت تحصیلی تأثیر دارد. اما به نظر می‌رسد تأکید بیش از حد بر نتایج ارزشیابی‌های دانشجویی می‌تواند به شکل نظام‌مند و گسترده‌تر نمرات را تحت تأثیر قرار دهد. در مقابل، باید استادان را به حفظ معیارهای آکادمیک و رفتارهای آموزشی اصیل دانشگاهی تشویق کرد. نولان و هوور (Nolan & Hoover, 2007) از متخصصان ارزشیابی آموزشی هشدار می‌دهند که هدف ارزشیابی از مدرس باید بهبود تدریس باشد. طبیعتاً برای بهبود تدریس، مدرسان باید خطرپذیری کرده و رفتارها و روش‌های تازه‌ای را امتحان کنند. اما یکی از کارکردهای پنهان فرم‌های ارزشیابی دانشجویی اینست که مدرسان در گذر زمان به مجموعه‌ای از رفتارهای یکسان که در فرم ارزیابی آمده توجه می‌کنند. از سوی دیگر، روش تدریس باید بسته به محتوا و اهداف هر درس متفاوت باشد. این در حالی است که نظام‌های دانشگاهی قابلیت طراحی فرم ارزشیابی متفاوت برای دروس را ندارند. در بهترین حالت، فرم ارزشیابی متفاوتی را برای دروس عملی و آزمایشگاهی طراحی می‌کنند.

چالش دیگر در مورد ارزشیابی دانشجویان از تدریس استادان این است که نمی‌توان ویژگی‌های شخصیتی استاد را مطلقاً از مهارت‌های تدریس او جدا کرد. رادماکر و مارتین (Radmacher & Martin, 2001) دریافتند که از میان ویژگی‌های شخصیتی استاد، متغیر برون‌گرایی قدرت پیش‌بینی ارزشیابی دانشجویان از اثربخشی تدریس را دارد. این محققان بر اساس این یافته خود نگرانی‌ها و ترس‌های استادان درباره اعتبار ارزشیابی‌ها را بر حق می‌دانند زیرا گویی نتایج ارزشیابی‌های دانشجویی بیش از هر چیز به شخصیت کلی استاد بستگی دارد. این محققان در ادامه می‌افزایند: «البته، شاید مدرسان برون‌گرایی که ارتباط‌گرا، مشتاق و شوخ‌طبع هستند معلمان اثربخش‌تری باشند» (ص، ۲۶۶) و همچنین با استناد به تحقیقات تصریح می‌کنند که دانشجویان در هر صورت تدریس اثربخش را از تدریس غیراثربخش تشخیص می‌دهند. اما به هر حال، بعید نیست سطح ارزشیابی دانشجویی استادی که به لحاظ شخصیتی آرام و درون‌گرا است کاهش پیدا کند. این مسأله به جنسیت استادان نیز ارتباط پیدا می‌کند. در پژوهشی کلاسیک (Bachen et al., 1999) این یافته ارزشمند به دست آمد که دانشجویان بر اساس انتظاراتی که متناسب با جنسیت استاد از وی دارند او را ارزشیابی می‌کنند. در واقع، دانشجویان از استاد زن انتظار یک سبک زنانه و از استاد مرد انتظار یک سبک مردانه دارند و طبق همین آمایه جنسیتی^۱ استاد را ارزشیابی می‌کنند. البته، محققان تذکر می‌دهند که این ارتباط ماهیت ساده‌ای ندارد و آمایه ذهنی جنسیتی، حداقل ارزیابی‌های برخی از دانشجویان را تحت تأثیر قرار نمی‌دهد و عاملی مثل تخصص استاد در موضوع مورد نظر اثرگذاری بیشتری دارد.

مسأله مرتبط دیگر با چگونگی بهره‌برداری از ارزشیابی‌های دانشجویی، پدیده تورم نمرات^۲ در سطح دانشگاه‌ها است. نمره همیشه برای دانشجویان جذاب بوده و دانشگاه‌ها نیز همواره کوشیده‌اند دانشجوی عالی را از خوب، متوسط و ضعیف متمایز کنند (Mackinnon & Kashif, 2022). دانشجویان همیشه خواهان کسب مدارک بیشتری هستند که شایستگی‌های تخصصی آن‌ها را به نمایش بگذارد (Morrison, 2021) و نمره یکی از این مدارک است. حال اگر استادان نیز به کسب نمره عالی در ارزشیابی‌های دانشجویی بیاندیشند یا اگر دانشگاه‌ها رویه‌های

¹- gender schema

²- grade inflation



سختگیرانه‌ای را در مورد چگونگی استفاده از نتایج ارزشیابی‌های دانشجویی تصویب و اجرا کنند چه پیامدهایی قابل تصور است؟ واقعیت اینست که هم استادان به ارزیابی‌های بالاتر از تدریس خویش نیاز دارند و هم دانشجویان به نمره‌عالی علاقه دارند (Hall, 2023). لذا بعید نیست که به مرور زمان معامله‌هشیارانه یا ناهشیارانه‌ای میان اساتید و دانشجویان شکل بگیرد: «من به شما نمره بالا می‌دهم، شما هم به من نمره ارزشیابی بالا بدهید!». مدیران دانشگاهی هرگز نباید فراموش کنند که «نمره» برای اکثر دانشجویان اهمیت بسیاری دارد. دغدغه دانشجویان برای گرفتن نمره بالا به دلایل متفاوتی است اما هر دلیلی که داشته باشد، برآیند مطالبات دانشجویان در این زمینه می‌تواند به مخدوش شدن نتایج سنجش‌ها منجر شود. اگر چه در دانشگاه‌های تراز اول تورم نمره کمتری وجود دارد (Jephcote et al., 2021) اما اکثر نظام‌های دانشگاهی این پدیده را به طور مداوم رصد می‌کنند (Müller-Benedict & Gaens, 2020). برخی محققان، مدرسان مدعو یا پاره‌وقت را عامل وخیم‌تر شدن تورم نمرات در مؤسسه دانسته‌اند (Chowdhury, 2018). اما مسأله این است که تورم نمرات نه برای دروس استادان رسمی مؤسسه و نه برای مدعوین آن قابل توجیه نیست. مدیران مؤسسه باید به نحوی از کلیه مدرسان حمایت اجرایی و روان‌شناختی کنند که مدرس با آسودگی خاطر برای نمره‌دهی دقیق اقدام نماید. در عین حال، سیاست‌ها و خط‌مشی‌های نمره‌دهی مشخصی نیز برای کلیه مدرسان دانشگاه وجود داشته باشد. باید دقت کرد که در تورم نمره، دانشجویانی که شایستگی کمتری دارند نمرات بالاتری می‌گیرند و این امر می‌تواند انگیزگی دانشجویان اخلاق‌مدار را به دنبال داشته باشد و رفتارهای انتخاب درس و استاد را نیز در دانشجویان تحت تاثیر قرار دهد به گونه‌ای که به سوی استادان خوش‌نمره گرایش پیدا کنند.

مجموعه دغدغه‌ها و اشکالات فوق، برخی از مدرسان، گروه‌های آموزشی و دانشگاه‌ها را به اجرای روش‌های جایگزین نمره‌دهی سوق داده است. چانگ و کاتینو (Chang & de Lemos Coutinho, 2022) نمره‌دهی دانشجویان مهندسی به تکالیف خود را روش نسبتاً کارآمدی ارزیابی کرده‌اند که مزایایی از لحاظ بازخورد دهی به دانشجویان و افزایش شناخت استاد از دانشجویان دارد. دانشگاه کرنل^۱ نیز در برابر پدیده تورم نمرات دو خط‌مشی اتخاذ کرد: (۱) انتشار میانگین نمرات هر درس بر روی اینترنت؛ (۲) درج میانگین نمرات دانشجویان در درس بر روی کارنامه هر دانشجو. این کار صرفاً در مورد کلاس‌های بالای ۱۰ نفر در مقطع کارشناسی و بدون انتشار نام استاد انجام شد. البته، پیامد طرح مذکور استقبال دانشجویان از دوره‌های خوش‌نمره بود (Bar et al., 2009) که این امر به خوبی نشان می‌داد اجرای تدابیر تازه باید توأم با احتیاط بیشتری باشد.

نکته مهم دیگر درباره نمره این است که بر خلاف تصور رایج، اولویت نخست و ارزشمندترین هدف دانشجویان نیست. در پژوهشی (Lastusaari et al., 2019) محققان دریافتند که قوی‌ترین عامل پیش‌بین پایداری دانشجویان در یکی از گرایش‌های شیمی و عدم تغییر رشته، رویکرد عمیق به یادگیری بود. وقتی محققان داده‌های مربوط به دانشجویان اخراجی، تغییر رشته‌ای و غیره را نیز در تحلیل‌ها راه دادند باز هم «نمره‌ها عامل پیش‌بین و قدرتمندی» نبودند. دومین عامل قدتمند، تلفیق دانشجویان در ساختار سیستم بود به این معنی که دانشجویان در سیستم حضور داشته باشند و پیروی از مقررات ساختاری و الزامات آموزشی را تجربه کنند. سومین عامل گرایش به تجربیات عملی در زمینه آموزش بود. این کار باعث می‌شد دانشجو به جای اینکه گرایش به یادگیری سطحی داشته باشد به یادگیری عمیق روی بیاورد. این یافته‌ها به خوبی نشان می‌داد که بیش از خود نمره، ایجاد جو و فضای آموزشی حمایت‌گرایانه بر روی یادگیری‌های دانشجویان تأثیر می‌گذارد. البته، در ترم نخست تحصیل، حمایت و مساعدت استادان در زمینه نمره می‌تواند آثار مثبتی داشته باشد. بنابراین، نمرات بالا حداقل در طولانی‌مدت، برای دانشجویان قانع‌کننده و رضایت‌بخش نیستند.

¹- Cornell



مسأله «نمره» از این دیدگاه نیز قابل تحلیل است که عده‌ای از استادان در زمینه سنجش و ارزشیابی دانشجویان به انعطاف‌پذیری و سنجش مجزا^۱ قائل نیستند و انتظار دارند دانشجویان محتوای تخصصی دروس را به طور کامل یاد بگیرند و یادگیری‌های خود را در امتحانات نیز نشان بدهند. برخی از این استادان متعهد، نمره خالصی را که دانشجو دریافت کرده در کارنامه او درج می‌کنند و معتقدند که به عنوان مثال، بی هیچ ارفاقی نباید نمره ۷ را در کارنامه به ۹ تبدیل کرد. معدودی از استادان نیز وجود دارند که گمان می‌کنند توزیع نمرات دانشجویان کلاس باید «نرمال» باشد. این در حالی است که توزیع نمرات قبول‌شدگان باید نرمال باشد (Yorke, 2007) نه اینکه استاد بکوشد اجباراً کل نمرات دانشجویان کلاس را به صورت نرمال در آورد. در هر صورت، پیامدهای این نمرات برای سبک و روش تدریس استادان می‌تواند مخرب باشد. باید توجه کرد عنصر اساسی انصاف در سنجش نهایی دانشجویان این است که استاد فرصت‌های جایگزین را برای افرادی که شرایط خاص دارند تدارک ببیند به نحوی که استاندارد یا هدف محتوایی مخدوش نشود، همه به سمت آن هدف تلاش کنند و این هدف/استاندارد برای همه یکسان باشد (Herman & Cook, 2019). بنابراین، ملاحظه می‌گردد که ارتباط تنگاتنگی میان نمره‌دهی استادان، واکنش عاطفی و روان‌شناختی دانشجویان به نمره پایان‌ترم خود، و نیز نحوه ارزشیابی دانشجویان از اساتید وجود دارد. هدف اصلی پژوهش طولی حاضر نیز پاسخ‌گویی به این سؤال است که نمره دانشجو در درس، تا چه اندازه ارزشیابی او از استاد را در نیمسال یا نیمسال‌های بعدی تحت تأثیر قرار می‌دهد.

روش پژوهش

پژوهش طولی^۲ حاضر از نوع مطالعه پنلی^۳ است. در مطالعه پنلی «افراد یکسان در گذر زمان و در نقاط متوالی مورد مطالعه قرار می‌گیرند» (Johansson, 2016). انواع مدل‌های پنلی و روش‌های تحلیل داده‌های پنلی در برخی منابعی (Reed & Ye, 2011) تشریح و در پژوهش‌های مختلفی (Felix et al., 2019; Kaplan et al., 1995) به کار رفته است. جامعه آماری پژوهش حاضر را کلیه دانشجویان شاغل به تحصیل در دانشگاه کاشان در چهار نیمسال آموزش حضوری پیش از شرایط کرونا و آموزش مجازی تشکیل داده‌اند. گزارش داده‌های مربوط به ارزشیابی دانشجویی و نمرات پایان‌ترم آن‌ها از سامانه آموزش دانشگاه استخراج گردید. در فایل استخراج شده اطلاعاتی نظیر نمره‌ای که دانشجو در پایان‌ترم گرفته بود، نمره‌ای که در ارزشیابی تدریس به استاد داده بود، به همراه کد درس و چند متغیر قابل تحلیل نظیر اندازه کلاس، میانگین نمره دانشجویان کلاس، و جنسیت دانشجو موجود بود. با توجه به اینکه هر دانشجو در هر نیمسال به طور متوسط بیش از پانزده واحد درس دارد فایل داده‌های اصلی یا جامعه آماری شامل ۲۲۰ هزار سطر بود. شرط انتخاب افراد در نمونه تحقیق اینگونه تعریف شد که دانشجو در چهار نیمسال متوالی با یک استاد مشخص درس گرفته باشد. حجم نمونه مورد نیاز بر انجام تحلیل‌های اصلی با نرم‌افزار جی‌پاور^۴ برآورد شد. با توجه به اینکه داده‌ها از نوع پنل طولی^۵ بود و برای تعیین تأثیر نمره دانشجو در هر ترم نیز کاربرد مدل‌های اثرات ترکیبی چندسطحی^۶ لازم بود لذا از خانواده آزمون^۷ گزینه رگرسیون چندگانه^۸ به عنوان روش تحلیل آماری انتخاب شد. اندازه اثر^۹ متغیر مستقل (نمره دانشجو) برابر با ۰/۰۲، سطح آلفا برابر با ۰/۰۱، توان آزمون آماری برابر با ۰/۹۹، و تعداد متغیرهای پیش‌بین ۵ عدد در نظر گرفته شد. نمونه برآوردی نرم‌افزار

1- differentiated assessment

2- longitudinal

3- panel study

4- GPower

5- longitudinal/panel data

6- multilevel mixed-effects models

7- test family

8- multiple regression

9- effect size



۱۶۸۳ نفر بود. پس از شرط‌گذاری، غربال اولیه و حذف سطرهایی که حتی یکی از سلول‌های اساسی آن مفقود بود نمونه‌ای به تعداد ۳۰۹۸ نفر به دست آمد که خوشبختانه بسیار بیشتر از حداقل حجم برآوردی نرم‌افزار بود. ابزار گردآوری داده‌ها در این تحقیق، فرم رسمی ارزشیابی دانشجویان از اساتید در بازه چهار نیمسال بود که در هفده آیتم در مقیاسی نه درجه‌ای (از ۱۲ تا ۲۰) به همراه گزینه «نظری ندارم» در تمامی دانشکده‌ها اجرا شده بود. برای محاسبه نمره‌ای که دانشجویان در ارزشیابی به استاد داده بود میانگین ساده آیتم‌های هفده‌گانه نظرسنجی، بدون اعمال ضرایب در نظر گرفته شد.

یافته‌ها

آیا نمره‌ای که دانشجویان در یک نیمسال از استاد درس می‌گیرند در نیمسال یا نیمسال‌های آتی، ارزشیابی او از کیفیت تدریس استاد را تحت تأثیر قرار می‌دهد؟ پیش از انجام آزمون آماری لازم برای پاسخ به این سؤال، نمایش ساختار اولیه داده‌ها در **جدول ۱** لازم به نظر می‌رسد.

جدول ۱

ساختار داده‌های ۴ نیمسال نمرات پایان‌ترم دانشجویان و نمرات ارزشیابی همان دانشجویان از استاد

کد یکتا	نمره پایان‌ترم دانشجویان				نمره ارزشیابی دانشجویان از استاد				جنسیت استاد	جنسیت دانشجویان	اندازه کلاس	معدل کلاس	تعداد افتاده‌ها
	۹۵-۲	۹۶-۱	۹۶-۲	۹۷-۱	۹۵-۲	۹۶-۱	۹۶-۲	۹۷-۱					
کد دانشجویان* کد استاد	۱۶	۱۵	۱۷	۱۶	۱۸	۱۸	۱۹	۱۹	مرد	زن	۲۸	۱۴	۱

چنانکه در **جدول ۱** مشخص است ساختار داده‌ها از نوع طولی است. برای انجام آزمون ابتدا ساختاردهی لازم انجام گردید و داده‌ها به صورت پنلی در آمد. سپس در محیط نرم‌افزار استاتا^۱ رگرسیون مدل اثرات تصادفی^۲ اجرا شد. نمره ارزشیابی دانشجویان از اساتید به عنوان متغیر وابسته و سایر متغیرها به عنوان متغیرهای مستقل در نظر گرفته شدند. نتیجه این تحلیل در جدول شماره ۲ ارائه شده است.

جدول ۲

مدل اثرات تصادفی برای پیش‌بینی نمره ارزشیابی استاد از روی نمره پایان‌ترم دانشجویان

متغیر مستقل	Coef.	Std. Err.	z	P> z	[۹۵% Conf. Interval]
اندازه کلاس	-۰/۰۲۵	۰/۰۰۳۶	-۶/۹۵	۰/۰۰۰	-۰/۰۳۲ -۰/۰۱۸
معدل کلاس	-۰/۰۱۷	۰/۰۲۲۲	-۰/۷۹	۰/۴۲۸	-۰/۰۶۱ -۰/۰۲۵
افتاده‌ها	-۰/۰۰۰۱	۰/۰۱۲۳	-۰/۰۱	۰/۹۹۱	-۰/۰۲۴ -۰/۰۲۴
نمره دانشجویان	۰/۰۵۲	۰/۰۰۵۴	۹/۵۲	۰/۰۰۰	۰/۰۴۱ ۰/۰۶۳
اندازه ثابت	۱۸/۷۰	۰/۳۴۰۵	۹۱/۹۱	۰/۰۰۰	۱۸/۰۳ ۱۹/۳۶

ضریب همبستگی درون-رده‌ای: ۰/۴۶۴ کای اسکور والد^۳: ۱۵۰/۶۸ سطح معنی‌داری: ۰/۰۰۰۰

1- STATA

2- GLS random effects

3- Wald Chi



در جدول ۲ ضریب رگرسیونی قدرتمند «نمره دانشجو» نشان می‌دهد که با یک واحد افزایش در نمره پایان‌ترم دانشجو نمره ارزشیابی استاد در نیمسال بعد به اندازه ۰/۵۲٪ افزایش می‌یابد. بر این اساس با یک واحد افزایش در نمره پایان‌ترم دانشجو با اطمینان ۹۹ درصد می‌توان گفت که بر مبنای فرم ارزشیابی نه‌درجه‌ای^۱ افزایشی به اندازه ۰/۱۲ در نمره ارزشیابی دانشجو از استاد در نیمسال‌های بعدی به وجود آمده است. ضریب رگرسیونی معنادار دوم مربوط به متغیر «اندازه کلاس» و نشان‌گر آن است که با افزایش یک نفر در تعداد ثبت‌نام کنندگان کلاس درس، کاهش معناداری به اندازه ۰/۰۲۵ در واحد نمره ارزشیابی استاد به وجود آمده است. نکته مهم در مورد جدول شماره ۲ مقدار ضریب همبستگی درون-رده‌ای است که نشان می‌دهد ۴۶/۴ درصد از تغییرات در نمرات ارزشیابی استادان به علت تفاوت میان پنل‌ها (در اینجا نیمسال‌ها) می‌باشد. با توجه به تغییر در درس‌ها، شرایط نیمسال‌ها و دیگر تفاوت‌های میان نیمسال‌ها، این مقدار قابل پیش‌بینی بود. در تحلیل قبل مشخص شد که نمره دانشجو در ترم قبل با ارزشیابی دانشجو از استاد در ترم بعد ارتباط دارد. به لحاظ منطقی این سؤال پیش می‌آید که آیا نمره تا نیمسال‌های بعدی نیز روی ارزشیابی دانشجو از استاد تأثیر می‌گذارد؟ به بیان دیگر و به لحاظ روان‌شناختی، تأثیر آمایه‌ای که به سبب دریافت نمره‌ای خاص در ذهن دانشجو ایجاد گردیده چقدر پایدار است؟ برای پاسخ به این سؤال از رگرسیون چندسطحی اثرات ترکیبی^۲ از نوع اثرات تصادفی^۳ استفاده شده است. استفاده از این روش به این دلیل لازم بود که باید برای تعیین تأثیر نمره یک ترم، تأثیر نمرات نیمسال‌های بعدی آن کنترل می‌شد. در اولین تحلیل، نمره نیمسال دوم ۱۳۹۵ به عنوان متغیر مستقل و نمره نیمسال دوم ۱۳۹۶ به عنوان متغیر وابسته در مدل وارد شد. طبیعتاً چنانکه در جداول زیر مشخص است تأثیر نمره نیمسال اول ۱۳۹۶ کنترل شده است.

جدول ۳

پیش‌بینی ارزشیابی استاد در نیمسال دوم ۱۳۹۶ از روی نمره دانشجو در نیمسال دوم ۱۳۹۵

متغیر مستقل	Coef.	Std. Err.	z	P> z	[۹۵% Conf. Interval]
نمره نیمسال دوم ۱۳۹۵	۰/۲۸۷	۰/۰۱۷۸	۱۶/۱۵	۰/۰۰۰	۰/۲۵۲ ۰/۳۲۲
اندازه ثابت	۱۰/۹۳	۰/۳۴۷	۳۱/۵۱	۰/۰۰۰	۱۰/۲۵۵ ۱۱/۶۱۶

لگاریتم درست‌نمایی: -۷۴۳۱/۶۸ کای اسکوتر والد: ۲۶۰/۷۶ سطح معنی‌داری: ۰/۰۰۰۰

جدول ۴

برآورد، خطای معیار و فاصله اطمینان پارامترهای اثرات تصادفی

پارامترهای اثرات تصادفی	برآورد	خطای معیار	فاصله اطمینان ۹۵٪
نمره نیمسال اول ۱۳۹۶:	۶/۴۴۹	۰/۸۵۱	۴/۹۶۷ ۸/۳۷۳
واریانس سطح ۲	۶/۲۸۲	۰/۱۶۳	۵/۹۶۹ ۶/۶۱۲

مقدار آماره کای اسکوتر = ۹۵۵/۹۳ سطح معنی‌داری = ۰/۰۰۰۰

بر اساس اطلاعات فوق، ضریب همبستگی درون‌رده‌ای که از تقسیم واریانس سطح دو بر مجموع واریانس سطح دو و واریانس خطا به دست می‌آید برابر با ۰/۵۱ و نشان‌گر آن است ۵۱٪ از واریانس نمرات ارزشیابی اساتید در نیمسال دوم ۱۳۹۶ ناشی از تفاوت بین نمرات

^۱ - مقیاس فرم از ۱۲ تا ۲۰ بوده است و لذا هر واحد نمره معادل ۲/۲۲۲ می‌باشد.

^۲ - Mixed-effects ML regression

^۳ - random effects



پایان ترم دانشجویان در نیمسال اول ۱۳۹۶ است. این مقدار کاملاً طبیعی به نظر می‌رسد زیرا از یک سو به لحاظ روان‌شناختی و طبق اصل تأخر انتظار می‌رود که آخرین نمره بر آمایه ذهنی دانشجویان تأثیر قدرتمندتری داشته باشد و از یک سو خود نمره می‌تواند به متغیرهای مختلفی که روی ارزشیابی مثبت یا منفی دانشجویان از استاد تأثیر دارند ارتباط داشته باشد. آنچه مسلم اینک از ملاحظه بخش اول و دوم جدول شماره ۳ می‌توان دریافت که با یک واحد افزایش در نمره پایان ترم دانشجویان در نیمسال دوم ۱۳۹۵ پس از گذشت دو نیمسال افزایشی به اندازه ۰/۲۸۷ واحد (طبق فرم ارزشیابی نه‌درجه‌ای، معادل ۰/۶۳۷) در نمره ارزشیابی دانشجویان از استاد در نیمسال دوم ۱۳۹۶ به وجود آمده است. آیا این یافته در مورد نیمسال مشابه نیز قابل مشاهده است؟ نتیجه تحلیل در زیر ارائه شده است.

جدول ۵

پیش‌بینی ارزشیابی استاد در نیمسال اول ۱۳۹۷ از روی نمره دانشجویان در نیمسال اول ۱۳۹۶

متغیر مستقل	Coef.	St d. Err.	z	P> z	[۹۵% Conf. Interval]
نمره نیمسال اول ۱۳۹۶	۰/۳۰۴	۰/۰۱۸۴	۱۶/۴۸	۰/۰۰۰	۰/۲۶۸ - ۰/۳۴۰
اندازه ثابت	۱۰/۴۰۶	۰/۳۳۷	۳۰/۸۲	۰/۰۰۰	۹/۷۴۴ - ۱۱/۰۶۸

لگاریتم درست‌نمایی: -۷۱۰۸/۹۵ کای اسکوتر والد: ۲۷۱/۷۳ سطح معنی‌داری: ۰/۰۰۰۰

جدول ۶

برآورد، خطای معیار و فاصله اطمینان پارامترهای اثرات تصافی

پارامترهای اثرات تصادفی	برآورد	خطای معیار	فاصله اطمینان ۹۵٪
نمره نیمسال دوم ۱۳۹۶:	۵/۵۷۵	۰/۷۴۲	۴/۲۹۴ - ۷/۲۳۶
واریانس سطح ۲	۵/۰۹۳	۰/۱۳۲	۴/۸۳۹ - ۵/۳۶۰

مقدار آماره کای اسکوتر = ۹۵۹/۲۲ = سطح معنی‌داری = ۰/۰۰۰۰

طبق اطلاعات فوق، در اینجا نیز ضریب همبستگی درون‌رده‌ای معادل ۰/۵۲ است که نشان می‌دهد ۵۲ درصد از واریانس نمرات ارزشیابی اساتید در نیمسال اول ۱۳۹۷ ناشی از تفاوت بین نمرات پایان ترم دانشجویان در نیمسال دوم ۱۳۹۶ است. پس از کنترل تأثیر واریانس این متغیر، می‌توان دریافت که با یک واحد افزایش در نمره پایان ترم دانشجویان در نیمسال اول ۱۳۹۶ دو نیمسال بعد افزایشی به اندازه ۰/۳۰۴ واحد (طبق فرم ارزشیابی نه‌درجه‌ای، معادل ۰/۶۷۵) در نمره ارزشیابی دانشجویان از استاد در نیمسال اول ۱۳۹۷ به وجود آمده است.

اما آیا پس از گذشت ۳ نیمسال، همچنان آمایه ذهنی برآمده از نمره نیمسال نخست می‌تواند ارزشیابی دانشجویان از استاد را تحت تأثیر قرار دهد؟ پیش از تحلیل داده‌های مربوط به این سؤال، ذکر این نکته لازم است که بر اساس مبانی و منطق رگرسیون چندسطحی، افزون بر این که نمره دانشجویان از استاد را تحت تأثیر قرار می‌دهد انتظار می‌رود که دانشجویان درون یک کلاس نیز نظرات یکدیگر را در مورد کیفیت تدریس استاد تحت تأثیر قرار دهند. در واقع، این امکان وجود دارد که بخشی از واریانس نمرات ارزشیابی اساتید ناشی از تفاوت‌های درونی کلاس‌های درس باشد. با توجه به اینکه تنها شاخص معتبر در دسترس، در مورد کلاس درس، معدل کلاس درس مربوطه بود تأثیر متغیر ترکیبی «معدل کل کلاس * نمره دانشجویان» نیز کنترل شد. نتیجه در زیر گزارش شده است.



جدول ۷

پیش‌بینی ارزشیابی استاد در نیمسال اول ۱۳۹۷ از روی نمره دانشجو در نیمسال دوم ۱۳۹۵

متغیر مستقل	Coef.	Std. Err.	z	P> z	[۹۵% Conf. Interval]
نمره نیمسال دوم ۱۳۹۵	۰/۱۳۲	۰/۰۱۲۸	۱۰/۳۳	۰/۰۰۰	۰/۱۰۷
اندازه ثابت	۱۶/۵۵۶	۰/۲۱۵	۷۶/۹۲	۰/۰۰۰	۱۶/۱۳۴

لگاریتم درست‌نمایی: ۲۴۴۵/۶۱ - کای اسکوتر والد: ۱۰۶/۷۷ سطح معنی‌داری: ۰/۰۰۰۰

جدول ۸

برآورد، خطای معیار و فاصله اطمینان پارامترهای اثرات تصافی

پارامترهای اثرات تصادفی	برآورد	خطای معیار	فاصله اطمینان ۹۵٪
نمره نیمسال اول ۱۳۹۶ *معدل کلاس	۰/۶۶۸	۰/۵۵۵۴	۰/۱۳۱
واریانس سطح ۲	۳/۴۸۶	۰/۵۶۳	۲/۵۴۰
نمره نیمسال دوم ۱۳۹۶ *معدل کلاس	۰/۱۰۹	۰/۰۰۳	۰/۱۰۳۵
واریانس سطح ۳			
واریانس خطا			

مقدار آماره کای اسکوتر = ۸۴۲۰/۲۲ = سطح معنی‌داری = ۰/۰۰۰۰

طبق اطلاعات فوق، ضرایب همبستگی درون‌رده‌ای به ترتیب برای اثر ترکیبی «نمره نیمسال اول ۱۳۹۶ *معدل کلاس» برابر با ۰/۱۵۶ و برای «نمره نیمسال اول ۱۳۹۶ *معدل کلاس» برابر با ۰/۸۱۷ می‌باشد. این ضرایب به خوبی نشان می‌دهند که بخشی از واریانس نمرات ارزشیابی اساتید توسط اثر متقابل نمره دانشجو و سطح پیشرفت تحصیلی کل کلاس در نیمسال‌های قبلی تبیین می‌شود. اما پس از کنترل تأثیرات مربوط به نیمسال‌های قبل، ضریب رگرسیونی مربوط به نمره نیمسال دوم ۱۳۹۵ که برابر با ۰/۱۳۲ است در سطح اطمینان ۹۹ درصد معنادار می‌باشد که نشان می‌دهد آمایه ذهنی برآمده از نمره گذشته همچنان پس از گذشت سه نیمسال تأثیر پایدار دارد. بر اساس داده‌های جدول شماره ۵ می‌توان گفت که با یک واحد افزایش در نمره پایان‌ترم دانشجو در نیمسال دوم ۱۳۹۵، پس از سه نیمسال، افزایش معناداری (بر اساس فرم ارزشیابی نه‌درجه‌ای) معادل ۰/۲۹۳ در نمره ارزشیابی دانشجو از استاد به وجود آمده که قابل تأمل است و دلالت‌های آن در قسمت نتیجه‌گیری تشریح شده است.

بحث و نتیجه‌گیری

متخصصان ارزشیابی در آموزش عالی تذکر می‌دهند که نباید به ارزشیابی‌های دانشجویی به عنوان محرک و مشوق اصلی استادان برای تلاش بیشتر جهت بهبود کیفیت آموزش اشاره کرد. این در حالی است که در دانشگاه‌های ایران «امتیازات کیفیت تدریس» که بخش مهمی از آن به ارزشیابی‌های دانشجویی مربوط است به ابزار و اهرم اصلی برای ارتقای آموزش تبدیل شده‌اند. افزون بر تصمیم‌گیری‌های پرسنلی در پرونده‌های تبدیل وضعیت استخدامی و ارتقای مرتبه علمی، از امتیازات برآمده از ارزشیابی‌ها برای انتخاب استادان سرآمد یا ممتاز و اخیراً اعطای پایه تشویقی آموزشی استفاده می‌شود. در پژوهش‌های موجود در زمینه ارزشیابی دانشجویان از کیفیت تدریس و کاربردهای آن یافته‌هایی مشاهده می‌شود که ضرورت سیاست‌گذاری‌های بخردانه‌تر در این زمینه را روشن می‌سازند. یافته‌های پژوهش حاضر که از بررسی و تحلیل داده‌های چهار نیمسال متوالی به دست آمده‌اند ارتباط روشنی را میان نمره دانشجو در هر نیمسال با ارزشیابی او از تدریس استاد در نیمسال بعد نشان

می‌دهند. در مجموع، چهار ضریب رگرسیونی به عنوان چهار یافته و مدرک پژوهشی مستقل به دست آمد که همگی نشانگر وجود رابطه معنادار میان نمره دانشجو و ارزشیابی او از استاد درس می‌باشند. یافته اول و نتیجه کلی این است که پس از کنترل اثر تفاوت‌های میان پنل‌ها (نیمسال‌ها)، می‌توان ادعا کرد که با یک واحد افزایش در نمره پایان‌ترم دانشجو، ارزشیابی او از استاد در نیمسال بعدی بیش از پنج درصد افزایش می‌یابد. یافته دوم این است که با یک واحد افزایش در نمره پایان‌ترم دانشجو در نیمسال دوم ۱۳۹۵ پس از گذشت دو نیمسال، افزایشی به اندازه ۰/۲۸۷ واحد در نمره ارزشیابی دانشجو از استاد در نیمسال دوم ۱۳۹۶ به وجود آمده است. یافته سوم این است که با یک واحد افزایش در نمره پایان‌ترم دانشجو در نیمسال اول ۱۳۹۶ پس از گذشت دو نیمسال، افزایشی به اندازه ۰/۳۰۴ واحد در نمره ارزشیابی دانشجو از استاد در نیمسال اول ۱۳۹۷ به وجود آمده است. یافته چهارم و بنیادی‌ترین نتیجه این است که پس از کنترل اثر متقابل نمره دانشجو و سطح پیشرفت تحصیلی کل کلاس، همچنان یک ضریب رگرسیونی معنادار به اندازه ۰/۱۳۲ مشاهده می‌شود که به خوبی نشان می‌دهد آمایه ذهنی برآمده از نمره دانشجو در پایان‌ترم نیمسال دوم ۱۳۹۵ پس از گذشت سه نیمسال همچنان ارزشیابی دانشجو از استاد را تحت تأثیر قرار می‌دهد. به طور ملموس، بر اساس فرم ارزشیابی تدریس در دانشگاه کاشان که نُه درجه‌ای (از نمره ۱۲ تا ۲۰) می‌باشد با یک واحد افزایش در نمره پایان‌ترم دانشجو در نیمسال دوم ۱۳۹۵، پس از سه نیمسال، افزایش معناداری معادل ۰/۲۹۳ در نمره ارزشیابی دانشجو از استاد به وجود آمده است. از معدود تحقیقات داخلی موجود در این زمینه، یافته‌های پژوهش قاضی طباطبایی و یوسفی افراشته (Ghazi Tabatabaee & Yousefi, 2023) همسو با نتایج حاضر است که در آن وجود ارتباط میان رضایت دانشجو از نمره و ارزشیابی او از تدریس استاد تأیید شده است. در مقابل، بصیرت، متوسلی، میرفرهادی و طاهری (Basirat et al., 2019) در طرح طولی خویش دریافتند که تفاوت معناداری میان ارزشیابی حین و پایان‌ترم دانشجویان از استاد وجود ندارد. در تبیین روان‌شناختی ارتباط نمره و ارزشیابی استادان می‌توان به یافته‌های ارزشمند کاری و کاریفیو (Carey & Carifio, 2012) اشاره کرد که دریافتند نمراتی که فراگیران در ترم نخست می‌گیرند نقش مؤثری در بهبود نمرات بعدی و نیز بهبود وضعیت روانی و نگرشی آن‌ها در درس‌ها دارد. بعید نیست که برخی از مدرسان بر حسب تجربه به وجود این ارتباط پی ببرند و هشیارانه یا ناهشیارانه، در راستای افزایش رضایت مؤسسه و یا دانشجویان نمرات بالاتری برای دانشجویان در نظر بگیرند. در چنین شرایطی، تورم نمرات در مؤسسه نه نشان‌گر افزایش آموخته‌ها و یادگیری‌های دانشجویان، بلکه حکایت‌گر معامله‌ای پنهان در جریان آموزش است.

از معدود پژوهش‌های دارای طرح تحقیقی دقیق که به ارتباط بین نمره، سوگیری شخصی^۱ دانشجو و ارزشیابی وی از استاد پرداخته‌اند می‌توان به پژوهش پارک و چو (Park & Cho, 2023) اشاره کرد که دریافتند دانشجویان ضعیف در ارزشیابی‌ها نمرات کمتری به استاد می‌دهند. این محققان نتیجه گرفته‌اند که اگر دانشجو به دلایلی، میان نمره ایدآل و نمره واقعی خود، شکاف یا اختلافی را حس کنند استاد را تنبیه کرده یا پاداش می‌دهند. البته این محققان گزارش کرده‌اند که در صورت اختلاف مثبت یا همان تورم نمره، پاداش کاهش می‌یابد اما تنبیه کاهش نخواهد داشت. تبیین محققان این است که تورم نمرات، انتظارات دانشجویان را تغییر خواهد داد. ایراد روش‌شناختی پژوهش مزبور این است که معدل کلی دانشجو را به عنوان متغیر مستقل در نظر گرفته است. در واقع، در پژوهش مذکور بیشتر به نقش سطح و توان تحصیلی دانشجو توجه شده است. در یکی از معدود پژوهش‌های ایرانی در زمینه رابطه نمره دانشجو و ارزشیابی اساتید نیز این ایراد روش‌شناختی مشاهده می‌شود؛ شکورنیا، ملایری، تراب‌پور و الهام‌پور (Shakurnia et al., 2006) همبستگی ضعیفی را بین معدل دانشجویان و نمره ارزشیابی اساتید مشاهده کرده و نتیجه گرفته‌اند که این عامل «نمی‌تواند در ارزشیابی اساتید تأثیر چندانی داشته باشد» (ص، ۵۱). از آنجا که در طرح تحقیقی

¹- self-directed bias



حاضر ارتباط نمره دانشجویان در درس و ارزشیابی آن دانشجویان از کیفیت تدریس استاد آن درس در نیمسال‌های بعدی واکاوی شد می‌توان یافته‌ها را تأییدی بر این فرضیه دانست که آمایه ذهنی برآمده از نمره می‌تواند تأثیری طولانی‌مدت بر ارزیابی‌های دانشجویان داشته باشد.

در تبیین تکمیلی یافته اصلی پژوهش حاضر، اشاره به یافته‌های جدید مارشال (Marshall, 2022) نیز روشن‌گر خواهد بود. این محقق از تحلیل بار معنایی مثبت و منفی نظرات کیفی دانشجویان به این نتیجه رسید که این نظرات با نتیجه ارزشیابی کمی دانشجویان از استاد همبستگی معناداری دارند. این یافته از این نظر اهمیت دارد که نشان می‌دهد دانشجویان بر مبنای تصور مثبت یا منفی ویژه‌ای که از استاد دارد او را ارزیابی می‌کند. چنانکه تحلیل‌های مقاله حاضر نیز نشان داد تبیین محتمل این است که تصور مثبت یا منفی دانشجویان که برآمده از نمره بالا یا پایین او است در گذر زمان ثابت مانده و لذا حتی تا چند نیمسال بعد روی ارزیابی او از استاد تأثیر می‌گذارد. سرانجام، باید به نقش ویژگی‌های فردی و شخصیتی استادان نیز توجه کرد. در تحقیق یعقوبی و سلیمی (Yaghoubi & Salimi 2018) معلوم شد که ویژگی‌های فردی استادان نظیر «آراستگی ظاهر» و «انعطاف‌پذیری در برابر دانشجویان» بالاترین رتبه و جایگاه را نزد دانشجویان دارند. بنابراین، بعید نیست حس نخستین یا آمایه ذهنی ویژه‌ای که به واسطه این ویژگی‌ها در دانشجویان ایجاد گردیده در گذر زمان ثابت بماند و ارزیابی‌های او را تحت تأثیر قرار دهد.

اما دلالت‌های عملی یافته‌های پژوهش حاضر کدامند؟ تردیدی نیست که مدیران دانشگاهی به دلایل مختلف طرفدار وجود نظام ارزشیابی تدریس هستند. اکثر استادان دانشگاه، به ویژه کسانی که امتیازات بالایی در این زمینه می‌گیرند (Shakurnia, 2012) نیز با وجود این نظام موافق‌اند و اطلاعات برآمده از آن را منبع مهمی برای بهبود کیفیت آموزش می‌دانند. در عین حال، این انتظار وجود دارد که نقش «عوامل مداخله‌گر» در نظر گرفته شوند (Dargahi & Mohammadzadeh 2013). به عنوان مثال، حدود یک‌سوم از استادان اعتقاد دارند که نتایج ارزشیابی‌ها متأثر از عواملی نظیر حضور و غیاب و سیستم نمره‌دهی استاد است (Ziaee et al., 2007). نتایج تازه‌ترین پژوهش‌ها نیز نشان‌گر حدود ۱۰ درصد کاهش در میزان رضایت دانشجویان از تدریس اساتید بعد از تغییر شرایط از آموزش حضوری به مجازی است (Fry, ۲۰۲۳). مدیران دانشگاهی باید توجه کنند که این عوامل اثرگذار خارج از اختیار و کنترل استادان هستند. اما اساسی‌ترین عامل مداخله‌گر که مغفول مانده است کیفیت و ویژگی‌های برنامه‌دروسی هر دوره است. هال (Hal I, ۲۰۲۲) به این نکته بنیادین اشاره کرده که حدود بیست درصد از واریانس پیشرفت تحصیلی دانشجویان توسط عوامل مربوط به ویژگی‌های دوره/برنامه‌دروسی تبیین می‌شود در حالی که فقط حدود ۳ درصد از واریانس توسط ویژگی‌های مدرسان تبیین می‌شود. با عنایت به چنین یافته‌های پژوهشی مهمی، تمرکز بیش از حد مدیران دانشگاهی بر نتایج ارزشیابی‌های دانشجویی می‌تواند پیامدهای ناخواسته به دنبال داشته باشد.

اولین دلالت عملی یافته‌های این تحقیق، ضرورت سیاست‌گذاری برای پیشگیری از تورم نمرات در مؤسسه است. تحلیل‌های پژوهش حاضر به روشنی، وجود ارتباط میان نمره دانشجویان و ارزشیابی استاد را نشان می‌دهند. چنانکه پارک و چو (Park & Cho, 2023) نیز تصریح کرده‌اند بعید نیست که استادان با دادن نمرات بالاتر به دانشجویان پاداش (نمره ارزشیابی بالاتر) دریافت کنند. در واقع، بعید نیست حداقل در مورد طیف خاصی از کلاس‌ها، معامله‌ای هشیارانه یا ناهشیارانه میان استاد و دانشجویان جریان پیدا کند. این در حالی است که طبق قواعد صریح سنجش و اندازه‌گیری پیشرفت تحصیلی، اگر چه در نظر گرفتن ویژگی‌هایی به غیر از پیشرفت تحصیلی قابل توجیه است (Wilson, 2019) اما نمره باید بیش از هر چیز انعکاسی از یادگیری‌های واقعی دانشجویان در آن درس باشد. متأسفانه در دانشگاه‌های ایران، به هنگام بررسی پرونده تمدید پیمانی، تبدیل وضعیت استخدامی یا ارتقای مرتبه علمی، صرفاً به خود نمرات ارزشیابی تدریس توجه می‌شود و همین امر می‌تواند به ایجاد رفتارهای آموزشی غیرسازنده منجر شود. صرف نظر از این مسأله، همان‌طور که تحلیل‌های ناماری و ناکاچی (Maamari & Naccache, 2022) نشان می‌دهد تأکید بیش از اندازه مؤسسه روی نمرات ارزشیابی دانشجویان از اساتید می‌تواند به تشدید مسأله تورم نمرات منجر شود. تورم نمرات می‌تواند آفت بسیار مهلک و مخربی برای نظام دانشگاهی ایران باشد. به ویژه با در نظر گرفتن اینکه در کارنامه‌های دانشگاهی (ریز



نمرات دانشجو) رتبه یا جایگاه وی در کلاس درج نمی‌شود. در واقع، معلوم نیست فردی که در یک درس تخصصی نمره ۱۶ گرفته است رتبه اول کلاس است یا اینکه ضعیف‌ترین دانشجوی کلاس بوده است. برای روشن تر شدن این مسأله ذکر یک مثال واقعی مقتضی به نظر می‌رسد. از بررسی داده‌های ده نیمسال دانشگاه که برای تحقیق حاضر انجام شد معلوم گردید که از حدود ۱۲ هزار درس در مقطع کارشناسی با تعداد ثبت‌نامی بیش از ۱۰ نفر، معدل نمرات پایان‌ترم ۱۵۸۵ کلاس بیشتر از ۱۸ بود! این مقدار معادل ۱۳ درصد از دروس کارشناسی است. اطلاعات تازه‌ای در مورد وضعیت تورم نمره در دیگر دانشگاه‌های ایران موجود نیست. اما بر اساس نمونه‌ای که گفته شد احتمال می‌رود که مسأله تورم نمره در دیگر دانشگاه‌ها نیز وجود دارد و به هر حال، این مسأله به ارزشیابی دانشجویان از اساتید نیز مربوط می‌شود. بنابراین، مدیران دانشگاهی باید تدابیر ویژه‌ای را در این زمینه اتخاذ کنند.

دومین دلالت عملی یافته‌های این تحقیق ضرورت تدوین و تصویب دستورالعمل‌های دقیق‌تر در شورای آموزشی دانشگاه‌ها می‌باشد. آنگونه که مارتینز، شویگ و گولداشمیت (Martínez et al., 2016) می‌گویند، برای ارزشیابی صحیح استادان بهره‌گیری از اندازه‌گیری‌های چندگانه‌ای ضرورت دارد که بر مبنای منطقی تدوین و بر اساس توافق ذینفعان، وزن‌دهی شده باشند. در هر صورت، چنانکه بویسن، کلی، ریزلی و کاسنر (Boysen et al., 2014) تذکر داده‌اند تفاوت‌های ناچیز در میانگین ارزشیابی دانشجویی اساتید به قضاوت‌های جدی و سرنوشت‌ساز منجر گردیده و به همین دلیل، برای برداشت‌های صحیح از این داده‌ها باید تحلیل‌های آماری دقیقی انجام داد. این ابعاد مسأله باید از منظر احتمالی منفی ارزشیابی‌ها در تخریب حس امنیت شغلی اساتید و فرهنگ سازمانی دانشگاه در شوراهای آموزشی و هیأت‌های جذب مد نظر قرار گیرد. در پژوهش علی‌بیگی، بارانی و کرمی دهکردی (Alibeygi et al., 2019) به ناخرسندی استادان از این نتایج منفی اشاره شده است. مسأله آثار ارزشیابی‌ها بر امنیت روانی اساتید به قدری اهمیت دارد که بسیاری از دانشگاه‌های جهان حتی نظرات کیفی (نوشتاری) دانشجویان را قبل از ارسال برای استاد درس، هم با استفاده از فناوری هوش مصنوعی و هم به صورت دستی، غربال و پالایش می‌کنند و حتی در صورت نیاز برای اصلاح به دانشجو باز می‌گردانند (Cunningham et al., 2023). البته، در مطالعه اخیر کمتر از پنج درصد نظرات، غیرسازنده و نیازمند اصلاح بودند. اما به هر حال، باید توجه کرد که حتی یک نظر منفی غیرمنصفانه می‌تواند آثار مخرب جبران‌ناپذیر داشته باشد. در پایان مجدداً به نظر هربرت مارش (Marsh, 1977; Marsh, 2001; Marsh & Overall, 1979) اشاره می‌شود که بیش از سه دهه به پژوهش در این زمینه پرداخته است. او معتقد است نظام‌های ارزشیابی نباید به گونه‌ای طراحی شوند که استادان به این نتیجه برسند که بهترین روش، ارائه نمرات بالا و کاستن از بار کاری دانشجویان است. مورین و مارش (Morin & Marsh, 2015) وجود الگوی مناسب در این زمینه را ضروری می‌دانند. مولر (Müller-Benedict & Gaens, 2020) با عنایت به در هم تنیدگی عوامل اثرگذار بر ارزشیابی‌های دانشجویی تأکید می‌کند که کارگروه‌های بررسی صلاحیت باید فقط بر ابعادی از نمرات ارزشیابی‌های دانشجویی تمرکز کنند که دانشجو در بهترین موقعیت قضاوت قرار دارد و هرگز از این نمرات به عنوان تنها روش تعیین کیفیت تدریس استفاده نکنند. سرانجام، بولر توصیه می‌کند که نمرات ارزشیابی دانشجویی با در نظر گرفتن و مقایسه عواملی نظیر گروه آموزشی، اندازه کلاس، درس‌های مشابه، و روند تغییرات در گذر زمان تفسیر کنند. علی‌رغم بلوغ نظام دانشگاهی ایران، هنوز اقدامات بسیاری در این زمینه لازم است.

پیشنهادها

۱) مدیران دانشگاهی می‌توانند وضعیت تورم نمرات را در سطح دانشگاه، دانشکده، گروه آموزشی و عضو هیأت علمی تحلیل و نتایج آن را در اختیار گروه‌های آموزشی خویش قرار دهند.



- ۲) در دستورالعمل‌های دانشگاهی برای محاسبه امتیازات مربوط به کیفیت آموزش، سیاست‌هایی برای پیشگیری از تورم بیش از حد نمرات پیش‌بینی شود. به عنوان مثال، میانگین نمرات پیشرفت تحصیلی دانشجویان کارشناسی استاد طی هشت یا ده نیمسال نباید بیش از ۲ انحراف معیار از متوسط دانشکده یا گروه بالاتر باشد.
- ۳) ماهیت دروس اثر قدرتمندی بر نمرات دانشجویان در پایان دوره و نیز ارزشیابی آن‌ها از کیفیت تدریس دارد. در دستورالعمل‌های محاسبه امتیازات آموزشی باید به میزان دشواری دروس پایه و تخصصی توجه شود به گونه‌ای که استادان این دروس در مورد افت نتایج ارزشیابی‌های خویش نگرانی نداشته باشند و نمرات پایانی دانشجویان را با تمرکز بیشتر بر میزان آموخته‌های واقعی آن‌ها تعیین نمایند.
- ۴) تشویق یک استاد به شکل‌هایی نظیر اعطای پایه تشویقی آموزشی یا انتخاب به عنوان استاد ممتاز آموزشی نباید به بهای نادیده گرفتن تلاش‌های سایر استادان متعهد تمام شود. رویکرد اکثر دانشگاه‌ها برای اقدامات تشویقی مذکور، اتکا بر نتایج ارزشیابی‌های کمی است. مدیران دانشگاهی باید توجه کنند که برخی از اعضای هیأت علمی، متعهدانه در صدد افزایش یادگیری‌های دانشجویان هستند و در این مسیر همواره انتظارات سطح بالای خویش را حفظ می‌کنند.
- ۵) تصور موجود این است که هر درسی باید توسط دانشجویان مورد ارزشیابی قرار بگیرد. شاید دلیل این رویکرد، سهولت اجرای ارزشیابی الکترونیکی در سامانه‌های آموزش دانشگاه باشد. اما مدیران دانشگاهی می‌توانند این رویکرد را نیز مد نظر قرار دهند که حداقل در مورد برخی از دروس بنیادین دوره، ارزشیابی با نظر گروه آموزشی انجام شود.
- ۶) فرم‌های ارزشیابی تدریس باید با تمرکز بر رفتارهای تدریس اثربخش طراحی و تدوین شوند. کاربرد واژه‌های انتزاعی و کلی در این فرم‌ها می‌تواند سوگیری در نمره‌دهی را افزایش دهد. در واقع، دانشجو در مواجهه با چنین آیتم‌هایی بر اساس برداشت ذهنی و حس خویش استاد را ارزشیابی می‌کند.
- ۷) چنانکه تحلیل‌ها نشان داد اندازه کلاس‌ها عامل اثرگذاری بر نتایج ارزشیابی‌های دانشجویی است. از آنجا که بسیاری از قضاوت‌ها در مورد نتایج ارزشیابی‌ها مبتنی بر مقایسه با دیگر استادان است به لحاظ حرفه‌ای و اخلاقی انتظار می‌رود که شرایط برابر و عادلانه باشد. در دانشکده‌ها مشاهده می‌شود که به عنوان مثال، یک عضو هیأت علمی، در یک نیمسال ۲۰۰ تا ۳۰۰ دانشجوی کارشناسی دارد در حالی که موظفی آموزشی عضو دیگر با یک درس دکتری و یک درس کارشناسی تکمیل می‌شود. از مدیران دانشگاهی انتظار می‌رود در دستورالعمل‌های داخلی برای تصمیم‌گیری‌های پرسنلی و نیز اجرای برنامه‌های تشویقی به این عامل مهم توجه نمایند.

تعارض منافع

در انجام مطالعه حاضر، هیچ‌گونه تضاد منافی وجود ندارد.

References

- Addison, W. E., Best, J., & Warrington, J. D. (2006). STUDENTS' PERCEPTIONS OF COURSE DIFFICULTY AND THEIR RATINGS OF THE INSTRUCTOR. *College Student Journal*, 40(2). <https://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jml=01463934&asa=Y&AN=21375566&h=LaZpnVMcIjk5s%2B1mB9ccw0U2y8l6KEDyU0yLwMKy99snXsOIwIhFWHh1zukT3LZ%2FxlZ2ZTJqarmzYHBLan aig%3D%3D&url=c>
- Alibeygi, A. H., Barani, S., & Karami Dehkordi, M. (2019). Designing a comprehensive model for teaching quality evaluation of faculty members: the case of Razi University. *Research in Curriculum Planning*, 16(63), 21-34. https://jsr-e.isfahan.iau.ir/article_669703.html?lang=en
- Bachen, C. M., McLoughlin, M. M., & Garcia, S. S. (1999). Assessing the role of gender in college students' evaluations of faculty. *Communication Education*, 48(3), 193-210. <https://doi.org/10.1080/03634529909379169>

- Bar, T., Kadiyali, V., & Zussman, A. (2009). Grade Information and Grade Inflation: The Cornell Experiment. *Journal of Economic Perspectives*, 23(3), 93-108. <https://doi.org/10.1257/jep.23.3.93>
- Basirat, M., Motevasseli, S., Mirfarhadi, N., & Taheri, M. (2019). Dentistry Students' Viewpoints about Evaluation of Faculty members during and at the end of Semester in 2016 [Research]. *Research in Medical Education*, 11(1), 3-9. <https://doi.org/10.29252/rme.11.1.3>
- Beheshtirad, R., Ghalee, A., & Galavandi, H. (2014). Investigation of the Student's Perceptions and Expectations of Educational Service Quality in Urmia University and Its Relationship with the Evaluation of Faculty Member's Performance. *Quarterly Journal of New Thoughts on Education*, 9(4), 119-145. <https://www.magiran.com/paper/1243495>
- Beran, T., & Violato, C. (2010). Student Ratings of Teaching Effectiveness: Student Engagement and Course Characteristics. *Canadian Journal of Higher Education*, 39(1). <https://doi.org/10.47678/cjhe.v39i1.491>
- Berezvai, Z., Lukáts, G. D., & Molontay, R. (2021). Can professors buy better evaluation with lenient grading? The effect of grade inflation on student evaluation of teaching. *Assessment & Evaluation in Higher Education*, 46(5), 793-808. <https://doi.org/10.1080/02602938.2020.1821866>
- Boysen, G. A., Kelly, T. J., Raesly, H. N., & Casner, R. W. (2014). The (mis)interpretation of teaching evaluations by college faculty and administrators. *Assessment & Evaluation in Higher Education*, 39(6), 641-656. <https://doi.org/10.1080/02602938.2013.860950>
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T., & Welsh, M. E. (2016). A Century of Grading Research: Meaning and Value in the Most Common Educational Measure. *Review of Educational Research*, 86(4), 803-848. <https://doi.org/10.3102/0034654316672069>
- Buller, J. L. (2012). *Best Practices in Faculty Evaluation: A Practical Guide for Academic Leaders*. Wiley. <https://books.google.com/books?id=WzHGZj8K6gC>
- Carey, T., & Carifio, J. (2012). The Minimum Grading Controversy: Results of a Quantitative Study of Seven Years of Grading Data From an Urban High School. *Educational Researcher*, 41(6), 201-208. <https://doi.org/10.3102/0013189x12453309>
- Chang, K., & de Lemos Coutinho, L. (2022). Administering Homework Self-Grading for an Engineering Course. *College Teaching*, 70(3), 387-395. <https://doi.org/10.1080/87567555.2021.1957666>
- Chowdhury, F. (2018). Grade inflation: Causes, consequences and cure. *Journal of Education and Learning*, 7(6), 86-92. <https://eric.ed.gov/?id=EJ1191199>
- Clayson, D. E. (2009). Student Evaluations of Teaching: Are They Related to What Students Learn?: A Meta-Analysis and Review of the Literature. *Journal of Marketing Education*, 31(1), 16-30. <https://doi.org/10.1177/0273475308324086>
- Cunningham, S., Laundon, M., Cathcart, A., Bashar, M. A., & Nayak, R. (2023). First, do no harm: automated detection of abusive comments in student evaluation of teaching surveys. *Assessment & Evaluation in Higher Education*, 48(3), 377-389. <https://doi.org/10.1080/02602938.2022.2081668>
- Danielson, C., & McGreal, T. L. (2000). *Teacher Evaluation to Enhance Professional Practice*. Association for Supervision and Curriculum Development. <https://books.google.com/books?id=GrLEtfxBEbYC>
- Dargahi, H., & Mohammadzadeh, N. (2013). Faculty Member's Evaluation by Students: Valid or Invalid. *Iranian Journal of Medical Education*, 13(1), 39-48. <https://www.magiran.com/paper/1109932>
- Ezadi, S., Alavi Langerudi, S. K., Pakmehr, H., & Tajari, M. (2015). Predict Efficacy Beliefs in Human Science Students through Quality of Teaching. *Journal of higher education curriculum studies*, 6(11), 47-70. https://www.icsajournal.ir/article_15682_bd875276d98a2e20df4f3623718f1e66.pdf
- Felix, E. D., Binmoeller, C., Sharkey, J. D., Dowdy, E., Furlong, M. J., & Latham, N. (2019). The influence of different longitudinal patterns of peer victimization on psychosocial adjustment. *Journal of School Violence*, 18(4), 483-497. <https://doi.org/10.1080/15388220.2018.1528552>
- Ganbari, S., & Soltanzadeh, V. (2018). Improving the Quality of Teaching in the Light of the Evaluation of Professors: Reflectively on Students' Perspective. *Research in Teaching*, 6(2), 15-31. https://trj.uok.ac.ir/article_60838_4836097e945ca6de5d0bc41c6f263297.pdf
- Ghazi Tabatabaee, M., & Yousefi Afrashteh, M. (2023). Relationship Analysis of some of the Variables Associated with Teaching Evaluation by Students: An Application of Structural Equation Modeling. *Quarterly Journal of Research and Planning in Higher Education*, 18(2), 83-107. https://journal.irphe.ac.ir/article_702784_937551382cc48c6fbb0ad295d58ecb1c.pdf
- Griffin, K. A. (2020). Institutional Barriers, Strategies, and Benefits to Increasing the Representation of Women and Men of Color in the Professoriate. In L. W. Perna (Ed.), *Higher Education: Handbook of Theory and Research: Volume 35* (pp. 277-349). Springer International Publishing. https://doi.org/10.1007/978-3-030-31365-4_4
- Haghighi, M., Akbari, A., Cherabin, M., & Karimi, M. (2021). A comparative study of the educational evaluation system of faculty members in Iranian universities of medical sciences [Original research article]. *Iranian Journal of Medical Education*, 21(0), 237-246. <http://ijme.mui.ac.ir/article-1-5259-en.html>
- <http://ijme.mui.ac.ir/article-1-5259-en.pdf>
- Hall, R. E. (2023). Grade inflation investigation: ethical challenges to the social work academy. *Social Work Education*, 42(8), 1134-1144. <https://doi.org/10.1080/02615479.2022.2063829>
- Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. routledge. <https://www.taylorfrancis.com/books/mono/10.4324/9780203887332/visible-learning-john-hattie>
- He, J., Zheng, X., Liu, M., Du, Y., Liu, G., Cui, J., & Su, Y. (2022). Reciprocity in college teaching: a big data study based on online student evaluation of 919,750 professors. *Assessment & Evaluation in Higher Education*, 47(8), 1401-1415. <https://doi.org/10.1080/02602938.2022.2067980>
- Herman, J., & Cook, L. (2019). Fairness in Classroom Assessment. In (pp. 243-264). <https://doi.org/10.4324/9780429507533-14>
- Jephcote, C., Medland, E., & Lygo-Baker, S. (2021). Grade inflation versus grade improvement: Are our students getting more intelligent? *Assessment & Evaluation in Higher Education*, 46(4), 547-571. <https://doi.org/10.1080/02602938.2020.1795617>
- Johansson, S. (2016). International large-scale assessments: what uses, what consequences? *Educational Research*, 58(2), 139-148. <https://doi.org/10.1080/00131881.2016.1165559>



- Kaplan, D. S., Liu, X., & Kaplan, H. B. (1995). Explaining Dropout Behavior: A Longitudinal Study of Direct and Indirect Psychosocial Effects of Seventh and Eighth Grade Experiences. *Educational Research and Evaluation*, 1(3), 205-225. <https://doi.org/10.1080/1380361950010301>
- Lakeman, R., Coutts, R., Hutchinson, M., Lee, M., Massey, D., Nasrawi, D., & Fielden, J. (2022). Appearance, insults, allegations, blame and threats: an analysis of anonymous non-constructive student evaluation of teaching in Australia. *Assessment & Evaluation in Higher Education*, 47(8), 1245-1258. <https://doi.org/10.1080/02602938.2021.2012643>
- Landrum, R. E., & Braitman, K. A. (2008). The Effect of Decreasing Response Options on Students' Evaluation of Instruction. *College Teaching*, 56(4), 215-218. <https://doi.org/10.3200/CTCH.56.4.215-218>
- Lastusaari, M., Laakkonen, E., & Murtonen, M. (2019). Persistence in studies in relation to learning approaches and first-year grades: a study of university chemistry students in Finland [10.1039/C8RP00244D]. *Chemistry Education Research and Practice*, 20(3), 452-467. <https://doi.org/10.1039/C8RP00244D>
- Linse, A. R. (2017). Interpreting and using student ratings data: Guidance for faculty serving as administrators and on evaluation committees. *Studies in Educational Evaluation*, 54, 94-106. <https://doi.org/10.1016/j.stueduc.2016.12.004>
- Maamari, B. E., & Naccache, H. S. (2022). The Impact of Grade Inflation on Teachers' Evaluation: A Quantitative Study Conducted in the Context of Five Lebanese Universities. *Journal of Global Education and Research*, 6(2), 192-205. <https://eric.ed.gov/?id=EJ1347283>
- Mackinnon, S. P., & Kashif, S. (2022). A Short History of Grading Practices at Dalhousie University (1901-2021). *Online Submission*. <https://eric.ed.gov/?id=ED620603>
- Marsh, H. W. (1977). The Validity of Students' Evaluations: Classroom Evaluations of Instructors Independently Nominated As Best and Worst Teachers by Graduating Seniors. *American Educational Research Journal*, 14(4), 441-447. <https://doi.org/10.3102/00028312014004441>
- Marsh, H. W. (1982). Factors Affecting Students' Evaluations of the Same Course Taught by the Same Instructor on Different Occasions. *American Educational Research Journal*, 19(4), 485-497. <https://doi.org/10.3102/00028312019004485>
- Marsh, H. W. (2001). Distinguishing Between Good (Useful) and Bad Workloads on Students' Evaluations of Teaching. *American Educational Research Journal*, 38(1), 183-212. <https://doi.org/10.3102/00028312038001183>
- Marsh, H. W., & Overall, J. U. (1979). Long-term stability of students' evaluations: A note on Feldman's "consistency and variability among college students in rating their teachers and courses". *Research in Higher Education*, 10(2), 139-147. <https://doi.org/10.1007/BF00976226>
- Marsh, H. W., & Roche, L. (1993). The Use of Students' Evaluations and an Individually Structured Intervention to Enhance University Teaching Effectiveness. *American Educational Research Journal*, 30(1), 217-251. <https://doi.org/10.3102/00028312030001217>
- Marshall, K. (2009). *Rethinking Teacher Supervision and Evaluation: How to Work Smart, Build Collaboration, and Close the Achievement Gap*. Wiley. <https://books.google.com/books?id=qxIY3BP2fCAC>
- Marshall, P. (2022). Contribution of open-ended questions in student evaluation of teaching. *Higher Education Research & Development*, 41(6), 1992-2005. <https://doi.org/10.1080/07294360.2021.1967887>
- Martínez, J. F., Schweig, J., & Goldschmidt, P. (2016). Approaches for Combining Multiple Measures of Teacher Performance: Reliability, Validity, and Implications for Evaluation Policy. *Educational Evaluation and Policy Analysis*, 38(4), 738-756. <https://doi.org/10.3102/0162373716666166>
- McEwan, P. J., & Soderberg, K. A. (2006). Roommate Effects on Grades: Evidence from First-Year Housing Assignments. *Research in Higher Education*, 47(3), 347-370. <https://doi.org/10.1007/s11162-005-9392-2>
- McMillan, A., Solanelles, P., & Rogers, B. (2021). Bias in student evaluations: Are my peers out to get me? *Studies in Educational Evaluation*, 70, 101032. <https://doi.org/10.1016/j.stueduc.2021.101032>
- Morin, A. J. S., & Marsh, H. W. (2015). Disentangling Shape from Level Effects in Person-Centered Analyses: An Illustration Based on University Teachers' Multidimensional Profiles of Effectiveness. *Structural Equation Modeling: A Multidisciplinary Journal*, 22(1), 39-59. <https://doi.org/10.1080/10705511.2014.919825>
- Morrison, A. (2021). JUST DESERTS? GRADE INFLATION AND DESERT-BASED JUSTICE IN ENGLISH HIGHER EDUCATION. *British Journal of Educational Studies*, 69(4), 437-451. <https://doi.org/10.1080/00071005.2020.1832957>
- Müller-Benedict, V., & Gaens, T. (2020). A new explanation for grade inflation: the long-term development of German university grades. *European Journal of Higher Education*, 10(2), 181-201. <https://doi.org/10.1080/21568235.2020.1718516>
- Nevid, J. S. (2018). *Essentials of psychology: Concepts and applications*. Cengage Learning. <https://thuvienso.hoasen.edu.vn/handle/123456789/7441>
- Nolan, J., & Hoover, L. A. (2007). *Teacher Supervision and Evaluation: Theory into Practice*. Wiley. <https://books.google.com/books?id=zRlbAAAACAAJ>
- Park, B., & Cho, J. (2023). How does grade inflation affect student evaluation of teaching? *Assessment & Evaluation in Higher Education*, 48(5), 723-735. <https://doi.org/10.1080/02602938.2022.2126429>
- Radmacher, S. A., & Martin, D. J. (2001). Identifying Significant Predictors of Student Evaluations of Faculty Through Hierarchical Regression Analysis. *The Journal of psychology*, 135(3), 259-268. <https://doi.org/10.1080/00223980109603696>
- Reed, W. R., & Ye, H. (2011). Which panel data estimator should I use? *Applied Economics*, 43(8), 985-1000. <https://doi.org/10.1080/00036840802600087>
- Schmelkin, L. P., Spencer, K. J., & Gellman, E. S. (1997). Faculty Perspectives on Course and Teacher Evaluations. *Research in Higher Education*, 38(5), 575-592. <https://doi.org/10.1023/A:1024996413417>
- Shakurnia, A. (2012). Results of student ratings: Does faculty attitude matter? [Short Communication]. *Iranian Journal of Medical Education*, 12(1), 33-35. <http://ijme.mui.ac.ir/article-1-2014-en.html>
- Shakurnia, A., Malayeri, A. R., Torabpour, M., & Elhampour, H. (2006). Correlation between Student Evaluation of Teaching and Students' Grades. *Iranian Journal of Medical Education*, 51-58.
- Spooren, P., & Christiaens, W. (2017). I liked your course because I believe in (the power of) student evaluations of teaching (SET). Students' perceptions of a teaching evaluation process and their relationships with SET scores. *Studies in Educational Evaluation*, 54, 43-49. <https://doi.org/10.1016/j.stueduc.2016.12.003>



- Uttl, B., White, C. A., & Gonzalez, D. W. (2017). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation*, 54, 22-42. <https://doi.org/10.1016/j.stueduc.2016.08.007>
- Wilson, M. (2019). Emerging issues in classroom assessment. *Classroom Assessment and Educational Measurement*, 265. <https://library.oapen.org/bitstream/handle/20.500.12657/54181/9780429017612.pdf?sequence=1#page=276>
- Yaghoubi, M., & Salimi, M. (2018). Determining the Factors affecting Facultie's educational evaluation in a Military University of Medical Sciences in Tehran, Iran. *Journal of Military Medicine*, 20(1), 73-82. <https://www.magiran.com/paper/1811220>
- Yorke, M. (2007). *Grading student achievement in higher education: Signals and shortcomings*. Routledge. <https://www.taylorfrancis.com/books/mono/10.4324/9780203939413/grading-student-achievement-higher-education-mantz-yorke>
- Yousefi, A., Jamshid nejad, A., & Moatamed, N. (2018). Evaluation of Relationship between Academic Motivation and Achievement in Students of Zanjan University of Medical Sciences [Orginal Research]. *Journal of Medical Education Development*, 11(29), 54-62. <https://doi.org/10.29252/edcj.11.29.54>
- Ziaee, M., Miri, M., Haji-Abadi, M., Azarkar, G., & Eshbak, P. (2007). Academic staff and student's impressions on academic evaluation of students in Birjand University of Medical Sciences and Health Services. *Birjand University of Medical Sciences*, 13(4), 61. <https://www.magiran.com/paper/414915>
- Zibaei, M., & Azizi, K. (2012). Designing and Evaluating Teaching Quality Assessment from the Viewpoint of Lorestan University of Medical Sciences Students in 2010. *Journal of Medical Education Development*, 5(8), 58-68. <https://www.magiran.com/paper/1025502>

