



Institute for Research
& Planning in Higher Education



National Organization
for Educational Testing

There is No Revenge! A Longitudinal Study of How Failed Students Evaluate the Teaching Quality of University Professors

Seyed Ahmad Madani¹

1. Assistant Professor, Department of Educational Sciences, Faculty of Humanities, University of Kashan, Kashan, Iran.
(Corresponding Author), Email: madani@kashanu.ac.ir

Article Info

Article Type:
Research Article

Received: 2023.06.28
Received in revised form: 2023.10.03
Accepted: 2023.11.27
Published Online: 2023.12.23

ABSTRACT

Objective: Do students, consciously or unconsciously, show biased attitudes in their evaluation of the teacher's teaching in next semesters? The purpose of the present panel study was to test the hypothesis that failed students will give a lower evaluation score to their professors in the subsequent semesters.

Methods: In order to test this assumption, the data of academic progress and evaluation of students of a specific entry during two consecutive semesters was checked and after conditioning the data according to the design and purpose of the research, a sample of 268 failed students was obtained.

Results: The analyzes showed that, contrary to the initial hypothesis, the average evaluation scores of failed students in the next semester are significantly higher than the previous semester. At the same time, after applying the random effects parameters model and controlling the variance caused by the teacher factor; on the one hand, it was observed that with a decrease of one unit in student's GPA as the main indicator of academic ability and talent, there is a decrease of thirteen percent in the evaluation of the student. On the other hand, with one-unit increase in the score of failed students, there is a significant increase of one percent in their evaluation of professors.

Conclusion: From the sum of these findings, it can be concluded that although there is a relationship between the academic ability and the grade of failed students with their evaluation of their professors, these students' evaluations is more influenced by impression and understanding of the professor's educational and professional activities in the next semester rather than their poor grade in the previous semester.

Keywords: Evaluation, Teaching, Student Evaluation of Teaching (SET), Professors.

Cite this article: Madani, Seyed Ahmad(2023). There is No Revenge! A Longitudinal Study of how Failed Students Evaluate the Teaching Quality of University Professors. *Higher Education Letter*, 16 (64): 24-33 pages.



DOI: 10.22034/HEL.2024.709931

Publisher: Institute for Research & Planning in Higher Education & National Organization of Educational Testing



مطالعه طولی نحوه ارزشیابی دانشجویان مردود از کیفیت تدریس استادان دانشگاه: انتقामी در کار نیست!

سید احمد مدنی^۱

۱. استادیار گروه علوم تربیتی، دانشکده علوم انسانی، دانشگاه کاشان، کاشان، ایران؛ (نویسنده مسئول)، رایانامه: madani@kashanu.ac.ir

اطلاعات مقاله	چکیده
نوع مقاله: مقاله پژوهشی	هدف: یکی از پرسش‌های حساسیت‌برانگیز در نظام ارزشیابی دانشجویان از تدریس اساتید این است که آیا دانشجویان ضعیفی که در درس مردود می‌شوند به صورت نظام‌مند استادان خویش را ضعیف‌تر ارزشیابی می‌کنند؟ به بیان دیگر، آیا این دسته از دانشجویان، هشیارانه یا ناهشیارانه در آینده ارزشیابی خویش از تدریس استاد را سوگیرانه انجام می‌دهند؟ از آنجایی که نمرات ارزشیابی دانشجویان از اساتید، بر اساس آئین‌نامه ارتقای مرتبه علمی و بر اساس دیگر دستورالعمل‌های داخلی دانشگاه‌ها، تصمیم‌گیری‌های دانشگاهی درباره اعضای هیأت علمی را تحت تأثیر قرار می‌دهند این احتمال وجود دارد که سنجش و اندازه‌گیری پیشرفت تحصیلی دانشجویان توسط اساتید، هشیارانه یا ناهشیارانه دستخوش ملاحظات غیرآکادمیک قرار گیرد. هدف مطالعه پنلی حاضر آزمودن این فرضیه است که دانشجویان مردود در آینده نمره ارزشیابی کمتری به استاد خویش می‌دهند.
دریافت: ۱۴۰۲/۰۴/۰۷	روش پژوهش: برای آزمودن فرضیه تحقیق، داده‌های پیشرفت تحصیلی و نیز ارزشیابی دانشجویان یک ورودی مشخص استخراج شد و پس از شرط‌گذاری متناسب با طرح و هدف تحقیق، نمونه‌ای به اندازه ۲۶۸ دانشجوی مردودی به دست آمد که اطلاعات نمرات درسی و نیز ارزشیابی آنها از استاد درس، طی دو نیمسال متوالی به دست آمد. تحقیق و ماهیت داده‌ها از نوع طولی و مبتنی بر منطق مطالعه پنلی است.
اصلاح: ۱۴۰۲/۰۷/۱۱	یافته‌ها: تحلیل‌ها نشان داد که بر خلاف فرضیه اولیه، میانگین نمرات ارزشیابی دانشجویان مردود از استاد خویش در نیمسال بعد به طور معناداری بالاتر از نیمسال قبلی است. در عین حال، پس از کاربرد مدل پارامترهای اثرات تصادفی و کنترل واریانس ناشی از عامل استاد، از یک سو ملاحظه شد که با کاهش یک واحد در «معدل دانشجو» به عنوان شاخص اصلی توان و استعداد تحصیلی، کاهش به اندازه سیزده صدم در ارزشیابی دانشجو از تدریس استاد قابل مشاهده است. از سوی دیگر، با یک واحد افزایش در نمره دانشجویان مردود، افزایش معناداری به اندازه یک درصد در ارزشیابی آنها از اساتید به وجود می‌آید.
پذیرش: ۱۴۰۲/۰۹/۰۶	نتیجه‌گیری: از مجموع این یافته‌ها می‌توان نتیجه گرفت که اگر چه بین توان و استعداد تحصیلی و نیز نمره دانشجویان مردود با ارزشیابی آنها از استاد درس ارتباط وجود دارد اما ارزشیابی این دانشجویان از تدریس استاد بیش از اینکه متأثر از نمره مردودی آنها در ترم قبل باشد مبتنی بر برداشت و ادراک آنها از فعالیت‌های آموزشی و حرفه‌ای استاد در ترم بعد است.
انتشار: ۱۴۰۲/۱۰/۰۲	کلیدواژه‌ها: ارزشیابی، تدریس، ارزشیابی دانشجویان از تدریس، استادان.

استناد: مدنی، سیداحمد. (۱۴۰۲). مطالعه طولی نحوه ارزشیابی دانشجویان مردود از کیفیت تدریس استادان دانشگاه: انتقामी در کار نیست! *نامه آموزش عالی*، ۱۶ (۶۴)،

DOI: 10.22034/HEL.2024.709931

صفحه ۳۳-۳۴



حقوق مؤلف © نویسندگان.

ناشر: مؤسسه پژوهش و برنامه‌ریزی آموزش عالی و سازمان سنجش آموزش کشور

مقدمه

مدیران دانشگاه‌های ایران داده‌های ارزشیابی‌های دانشجویان از تدریس اساتید را برای تصمیم‌گیری‌های مختلفی از جمله تمدید قرارداد پیمانی، ارتقای مرتبه علمی و تبدیل وضعیت استخدامی آنها مورد استفاده قرار می‌دهند. دیگر نظام‌های آموزش عالی جهان نیز گاه حتی در تصمیم‌گیری‌های مختلف بر استفاده از نتایج ارزشیابی‌های دانشجویی تأکید می‌کنند (برزوای، لوکاتز و مولانتای، ۲۰۲۰). در واقع، «تدریس» به معیاری بنیادین در سنجش عملکرد کلی اعضای هیات علمی تبدیل شده است که نظام‌های دانشگاهی در مقایسه با راهنمایی دانشجویان، کار در کمیته‌ها، پژوهش و خدمات عمومی وزن بیشتری به آن می‌دهند (برک، ۲۰۰۶). البته، تبدیل شدن «تدریس» به عنوان معیار اساسی عملکرد اساتید باعث حساسیت اعضای هیات علمی درباره نظام ارزشیابی دانشجویان از تدریس اساتید شده است. حساسیت برخی از اساتید به اندازه‌ای است که معتقدند اختصاص زمان بیشتر به پژوهش به بهای تحت تأثیر قرار گرفتن ارزشیابی‌های دانشجویی تمام می‌شود (بیکر و مانینگ، ۲۰۲۱). یکی از پرسش‌های همیشگی این است که آیا ارزشیابی اساتید توسط دانشجویانی که پیشرفت تحصیلی بسیار ضعیف دارند رویه درستی است؟

واقعیت این است که تدریس باکیفیت در کلاس‌های درس دانشگاهی وابستگی و همبستگی بالایی با دانش محتوایی آموزشی دارد در واقع، اساتیدی در تدریس دانشگاهی موفق‌تر و اثربخش‌تر هستند که افزون بر داشتن تسلط علمی، روش ویژه‌ای را برای توضیح محتوا در غالب مفاهیم، مسائل، موضوعات و مثال‌ها خلق کرده‌اند (کینگ، ۲۰۲۲). با چنین دیدگاهی، بسیاری از اساتید این اشکال را مطرح می‌کنند که دانشجویان شناخت و درک کاملی برای ارزشیابی میزان «تسلط و توان علمی استاد» ندارند. عده‌ای از اساتید بر همین اساس مدعی‌اند که ارزشیابی دانشجویان از اساس فاقد وجهت است. عده‌ای معتقدند آیت‌های مربوط به تسلط علمی یا شایستگی تخصصی استاد باید از فرم‌های ارزشیابی حذف شود. سرانجام، گروهی از اساتید نیز هستند که کلیت ارزشیابی دانشجویان از تدریس اساتید را پذیرفته‌اند اما پیشنهاد می‌دهد نتایج ارزشیابی دانشجویان مشروط یا مردود در درس، از مجموع داده‌ها حذف شوند.

در زمینه ارزشیابی دانشجویان از کیفیت تدریس اساتید پژوهش‌های متعدد و گسترده‌ای انجام شده است. بحث و گفتگو درباره این پژوهش‌ها می‌تواند به چالشی علمی و مجادله‌ای بزرگ تبدیل شود زیرا گستره شواهد ضد و نقیض در این حوزه به اندازه‌ای است که به قول جفری بولر (۲۰۱۲) با جستجو در منابع علمی می‌توان هر فرضیه‌ای را در زمینه اعتبار یا عدم اعتبار ارزشیابی‌های دانشجویان از تدریس اساتید تأیید یا رد کرد. البته، در مورد برخی از عوامل اتفاق نظر نسبی وجود دارد. به عنوان مثال، کریترز و سویت‌کاشمن (۲۰۲۲) از بررسی بیش از یکصد گزارش پژوهشی به این نتیجه رسیدند که از میان انواع سوگیری‌های اثرگذار بر ارزشیابی‌های دانشجویی، برای سوگیری‌های ناشی از جنسیت استاد شواهد محکمی وجود دارد. کوک و پاتر (۲۰۲۲) نیز به برخی از این سوگیری‌ها اشاره کرده‌اند. به عنوان مثال، اینکه دانشجویان هیجان‌ها و عواطف ویژه‌ای را به اساتید زن نسبت می‌دهند از جمله اینکه آنها رسیدگی و مراقبت بیشتری مبذول می‌دارند یا اینکه شایستگی‌های پژوهشی و آموزشی آنها را کمتر از اساتید مرد برآورد می‌کنند. به نظر این محققان، ظاهر جذاب‌تر برای مدرسان زن نمرات ارزشیابی بالاتری را به ارمغان می‌آورد و «این یافته دلالت بر این دارد که ادراک دانشجویان از شایستگی یک استاد زن تا اندازه‌ای به جذاب به نظر آمدن او بستگی دارد» (ص، ۸۶). کن و کپریانی (۲۰۲۲) نیز استدلالی آماری را در مورد سوگیرانه بودن ارزشیابی‌های دانشجویی مطرح کرده‌اند. از نظر این متخصصان به لحاظ نظری اگر اثر هاله‌ای در کار نباشد آیت‌های فرم ارزشیابی لزوماً نباید با یکدیگر همبستگی داشته باشند. به عنوان مثال، نباید همواره میان آیت‌های مربوط به سازماندهی اثربخش محتوای درسی و آیت‌های مربوط به احترام گذاشتن به دانشجویان همبستگی وجود داشته باشد. در عمل اما همبستگی‌های معنادار حتی میان آیت‌های نامرتب مشاهده می‌شود.

بنابراین، در یک جمع‌بندی نهایی می‌توان با اطمینان ادعا کرد که برخی از سوگیری‌های آگاهانه یا ناآگاهانه، ارزشیابی‌های دانشجویان از تدریس اساتید را تحت تأثیر قرار می‌دهند. وجود سوگیری اما به معنای عدم اعتبار مطلق ارزشیابی‌های دانشجویی نیست. در واقع، اگر برای ارزیابی پژوهش‌های موجود معیارهای روش‌شناسی دقیقی انتخاب شود می‌توان گفت که ارزشیابی‌های دانشجویی در مجموع اعتبار دارند. هربرت مارش (۱۹۷۷، ۲۰۰۱) که در سراسر مسیر حرفه‌ای خویش درباره این موضوع پژوهش کرد ضمن پذیرش اعتبار ارزشیابی‌های دانشجویی، علت بروز چالش در این زمینه را بی‌دقتی مؤسسات آموزشی در نحوه استفاده از داده‌های ارزشیابانه می‌داند. از نظر مارش اگر دانشگاه‌ها با رویکردی منسجم و بخردانه از نتایج ارزشیابی‌ها بهره‌برداری می‌کردند شاهد بدبینی بخشی از اساتید درباره آسیب‌ها یا پیامدهای ناخواسته این ارزشیابی‌ها نبودیم. پژوهشگران متأخر نظیر ایساری و والدز (۲۰۲۰) نیز پس از بررسی شواهد به این نتیجه رسیده‌اند که اگر چه ارزشیابی‌های دانشجویی می‌تواند دارای پایایی و اعتبار بالایی باشند و نمرات آنها با کیفیت واقعی عملکرد اعضای هیات علمی همبستگی داشته باشد اما روش‌های مرسوم بهره‌برداری از نتایج آنها می‌تواند به تصمیم‌گیری‌های غیرمنصفانه منجر شود.

در جمع‌بندی نظرات متخصصانی که به اعتبار ارزشیابی‌های دانشجویی اعتقاد دارند می‌توان گفت که اگر دانشگاه‌ها از ارزشیابی‌ها به طور منصفانه و بخردانه استفاده می‌کردند واکنش استادان این بود که از ارزشیابی فقط به عنوان ابزاری برای بهبود کیفیت تدریس استفاده کنند و هرگز هنجارها یا معیارهای آکادمیک را تنزل ندهند. اکنون کسب نمره بالاتر در ارزشیابی‌های دانشجویی برای اطمینان از تمدید قرارداد پیمانی، تبدیل وضعیت استخدامی، ارتقای مرتبه علمی، یا کسب عنوان استاد سرآمد آموزشی، برای برخی از استادان دانشگاه‌های ایران به یک هدف تبدیل شده است. گریفین (۲۰۲۰) به وجود این گرایش در دیگر نظام‌های آموزش عالی نیز اشاره کرده است.

با عنایت به چنین استدلال‌هایی، اگر مدیران دانشگاه‌های ایران خواستار بهبود واقعی فرایندهای یاددهی-یادگیری و کیفیت آموزش در کلاس‌های اند افزون بر بهره‌گیری عالمانه از داده‌های ارزشیابی دانشجویی باید از منشأ سوگیری‌های مختلف در این ارزشیابی‌ها آگاه باشند. یکی از مهم‌ترین منابع سوگیری در ارزشیابی‌های دانشجویی، توان و استعداد علمی دانشجویان است. این یکی از باورهای رایج استادان است که اطلاعات بسیاری از دانشجویان درباره محتوای درس سطحی و ناچیز است. به نظر این گروه از استادان، دانشجویی که اطلاعاتی درباره درس ندارد، امید و انگیزه و پشتکار لازم برای مطالعه را ندارد و همواره با ارفاق و مساعدت استادان درس‌ها را می‌گذراند شایستگی لازم برای ارزشیابی استادان را ندارد. شاید باورها و اعتقادات این گروه از استادان با مراجعه به اثر هلمز (۲۰۱۵) راسخ‌تر شود زیرا هلمز پافراتر نهاده و شواهدی ارائه کرده که نشان می‌دهد دانشجویان عادی حتی پس از تدریس استاد نمی‌دانند یاد گرفته‌اند یا نه! در واقع، آنها خیال می‌کنند که یاد گرفته‌اند. بنابراین، به لحاظ نظری و پژوهشی می‌توان این پرسش را موجه و مهم دانست که آیا دانشجویان از مجموعه اطلاعات، نگرش‌ها و تجربیاتی برخوردار هستند که بتوانند تدریس استادان را منصفانه ارزشیابی کنند؟ به لحاظ منطقی و نظری، حساسیت و اهمیت این پرسش در مورد دانشجویان ضعیف بیشتر است. به بیان دقیق‌تر، این سؤال مطرح است که آیا توان و استعداد تحصیلی و ترجیحات یادگیری دانشجویان ضعیف باعث ادراک متفاوت آنها از اثربخشی تدریس استاد می‌شود؟ با عنایت به مبانی نظری و پژوهشی فوق، حساسیت استادان درباره چرایی و چگونگی ارزشیابی‌های دانشجویی از تدریس استادان طبیعی و منطقی به نظر می‌رسد. نتایج برخی از پژوهش‌ها همسو با دیدگاه‌های استادان متعهدی است که دغدغه بهبود واقعی فرایندهای تدریس را دارند. در ادامه به چهار مورد از جدیدترین این پژوهش‌ها اشاره می‌شود.

استروبی (۲۰۲۰) از مرور نظام‌مند مجموعه‌ای گسترده از تحقیقات به چهار نتیجه کلی درباره ارزشیابی‌های دانشجویی دست یافت: (۱) دانشجویان به نمرات خوب، با ارزشیابی مثبت از تدریس پاداش می‌دهند؛ (۲) دانشجویان به درس آسان، با ارزشیابی مثبت از تدریس پاداش می‌دهند؛ (۳) دانشجویان درس‌هایی را برمی‌گزینند که نویدبخش نمره‌ای خوب هستند؛ و (۴) استادان، نیازمند/خواستار نمرات ارزشیابی بالا هستند. این متخصص در نهایت نتیجه می‌گیرد که ارزشیابی‌های دانشجویی شاخص معتبری برای اثربخشی تدریس نیستند و لذا نباید توسط رؤسای دانشکده‌ها یا مدیران دانشگاهی برای ارزشیابی اعضای هیأت علمی مورد استفاده قرار گیرند.

بورچ، ساندول و ریزور (۲۰۲۰) از بررسی آیت‌های فرم‌های ارزشیابی دانشجویان از تدریس اساتید دریافتند که غالب این آیت‌ها با محوریت تدریس و با محتوایی کلی طراحی می‌شوند که نتیجه طبیعی آن اینست که میزان رضایت از تدریس به درستی مشخص نمی‌شود. مهم‌تر اینکه با چنین آیت‌هایی دانشجویان درباره آن ابعادی از تدریس که به فرایندهای یادگیری آنها مربوط است به استادان بازخورد خاصی نمی‌دهند. این محققان اشاره کرده‌اند که وقتی استادان با دانشجویان درباره کم و کیف تدریس گفتگو می‌کنند هم برای دانشجویان و هم برای استادان آگاهی بهتری درباره ابعاد مختلف درس و تدریس به وجود می‌آید.

برزوی، لوکاتز و مولانتای (۲۰۲۰) در بررسی ارتباط میان نمرات دانشجویان و ارزشیابی آنها از تدریس استادان به این نتیجه رسیدند که با افزایش یک درجه در نمره پایانی دانشجو به‌طور تقریبی افزایشی به اندازه ۰/۲ تا ۰/۴ در ارزشیابی او از آن استاد منجر خواهد شد. «این بدان معناست که با دست و دل‌بازی در نمره‌دهی می‌توان ارزشیابی‌های بهتری کسب کرد. البته، این یافته دلالت بر این ندارد که استادان به‌طور هدفمند و برای رسیدن به نمرات ارزشیابی دانشجویی بالاتر، دست و دلبازانه نمره می‌دهند» (ص، ۱۴). محققان بر مبنای این یافته، اتکای صرف دانشگاه‌ها بر نمرات ارزشیابی دانشجویی را خطا محسوب کرده‌اند.

گوندوز و فوکوب (۲۰۲۱) از تحلیل نتایج ارزشیابی‌های دانشجویی در بافت دانشگاهی در ترکیه به این نتیجه رسیدند که دانشجویان متعهد نسبت به کسانی که درس را جدی نمی‌گیرند رضایت بیشتری از تجارب یادگیری خود ابراز می‌کنند. محدودیت این پژوهش به وجود تعداد بالای پاسخ‌های با واریانس صفر مربوط می‌شود که محققان دلیل آن را تعداد بالای آیت‌های پرسشنامه (۲۸ عدد) عنوان کرده‌اند. این محققان توصیه کرده‌اند که داده‌های ارزشیابی‌های دانشجویی به طور عمیق‌تر تحلیل شوند.

سرانجام، کتن و کپریانی (۲۰۲۲) از پژوهش خویش به این نتیجه رسیدند که ارزشیابی‌های دانشجویی از تدریس استاد متأثر از اثر هاله‌ای است. بر اساس تبیین این محققان، دانشجو به برداشتی کلی درباره استاد می‌رسد اما به هنگام ارزشیابی وی نمی‌تواند میان قضاوت‌های کلی خود و قضاوت‌های مربوط به تدریس اثربخش استاد تمایز قابل شود. یکی از دلایل این محققان برای طرح این ادعا، وجود همبستگی بالا میان آیت‌های فرم ارزشیابی است. در واقع، به لحاظ نظری و در مورد گروه‌هایی از استادان، انتظار می‌رود که میان آیت‌های ارزشیابی آنها همبستگی نباشد اما در عمل این همبستگی وجود دارد. پژوهشگران برای آزمودن این فرضیه، یک آیت بی‌ربط در فرم ارزشیابی قرار دادند که مربوط به کیفیت فضای فیزیکی کلاس بود. با توجه به اینکه مشخص شد نمرات مربوط به این آیت با سایر آیت‌ها همبسته‌اند محققان مذکور معتقدند این ارزشیابی‌ها متأثر از اثر هاله‌ای است و بازخوردهای به دست آمده از آنها دقت کامل ندارد و نمی‌تواند به بهبود دقیق ابعاد تدریس استادان کمک کند.

هم‌راستا با مجموعه تحقیقات و مبانی پژوهشی موجود در زمینه ارزشیابی دانشجویان از تدریس استادان، هدف پژوهش حاضر پاسخ‌گویی به یکی از پرسش‌های رایج و مهم در این حوزه است. آیا مردودی دانشجویان در درس به افت ارزشیابی آنها از استاد در نیمسال بعد منجر می‌شود؟ البته، فرضیه اولیه به‌طور ویژه بر اساس نظریه اسناد بیرونی بود که طبق آن وقتی افراد نتایج نامطلوب (شکست) را تجربه می‌کنند آن را به عاملی غیر از خویش نسبت می‌دهند (اسمیت و فرت، ۲۰۱۸). به هر حال، فرضیه تحقیق به بیان دقیق‌تر این است که مردودی دانشجو در درس به ایجاد مجموعه‌ای از سوگیری‌های هشیارانه یا ناهشیارانه منجر می‌شود که ارزشیابی دانشجو از تدریس استاد را در نیمسال بعدی تحت تأثیر قرار می‌دهد.

روش پژوهش

پژوهش توصیفی حاضر از نوع طولی^۱ و مبتنی بر منطق مطالعه پنلی^۲ است که در آن «افراد یکسان در گذر زمان و در نقاط متوالی مورد مطالعه قرار می‌گیرند» (جانسون و کریستنسن^۳، ۲۰۱۶، ص ۸۴۱). با توجه به اینکه در این تحقیق نمرات پیشرفت تحصیلی و ارزشیابی دانشجویان مردود از کیفیت تدریس استادان در دو نیمسال متوالی بررسی گردیده می‌توان آن را از نوع مطالعه پنلی محسوب کرد. جامعه آماری پژوهش را کل دانشجویان یک ورودی دانشگاه به تعداد ۱۲۸۸ نفر تشکیل داده‌اند. نمونه‌گیری در پژوهش‌های پنلی متناسب با هدف و طرح تحقیق است. در پژوهش حاضر پس از استخراج داده‌های مربوط به ورودی مورد نظر از سامانه آموزش، شرط‌گذاری به این صورت انجام شد که در مرحله اول، دانشجویان مردودی ورودی از بقیه تفکیک و کسانی که نمره ۰/۲۵ و یا کمتر از آن را گرفته بودند از فهرست خارج شدند. در مرحله دوم داده‌های آن دسته از دانشجویان مردود که در نیمسال بعدی، همان درس را با همان استاد گرفته بودند استخراج شد. تعداد این افراد ۲۶۸ نفر بود که حجم نمونه پژوهش را تشکیل دادند^۴. ساختار داده‌ها و اطلاعات دیگر متغیرهای در دسترس در جدول شماره ۱ نشان داده شده است.

جدول ۱: ساختار داده‌های دو نیمسال نمرات پایان ترم دانشجویان مردودی و نمرات ارزشیابی این دانشجویان از استاد

کدیکتا	نمره پایان ترم دانشجو		نمره ارزشیابی دانشجو از استاد		جنسیت دانشجو	اندازه کلاس	معدل کلاس	معدل دانشجو
	نیمسال ۱	نیمسال ۲	نیمسال ۱	نیمسال ۲				
کد دانشجو* کد استاد	۸/۵	۱۱/۵	۱۸	۱۹	زن	۵۱	۱۴	۱۳/۴۷

داده‌های ارزشیابی دانشجویان از تدریس استادان بر اساس نظرسنجی الکترونیکی به دست آمده‌اند. رویه این نظرسنجی در دانشگاه به این صورت است که حدود دو هفته پیش از امتحانات پایان ترم، دانشجویان فرم‌های ارزشیابی الکترونیکی را برای یکایک درس‌های خود به‌طور جداگانه تکمیل می‌کنند. فرم ارزشیابی دارای مقیاس نُه‌درجه‌ای است که دانشجو می‌تواند در هر آیت نمره‌ای بین ۱۲ تا ۲۰ برای استاد درس در نظر بگیرد و گزینه «نظری ندارم» نیز برای هر آیت وجود دارد. در گزارش این داده‌ها نمره گزینه «نظری ندارم» به صورت (۰) درج شده بود به همین دلیل در کلیه مراحل تحلیل، در محاسبه میانگین ارزشیابی دانشجو از استاد این صفرها از تحلیل خارج شده‌اند. برای انجام تحلیل‌های عاملی نیز به جای نمره صفر، میانگین کلی نمره ارزشیابی دانشجو از استاد در فایل داده‌ها جایگزین شد.

^۱. longitudinal

^۲. panel study

^۳. Johnson & Christensen

^۴. حجم نمونه برآوری نرم‌افزار تخصصی تعیین حجم نمونه (GPower) با تنظیم پارامترها روی رگرسیون چندمتغیره، اندازه اثر = ۰/۱، توان آزمون = ۰/۹۵، و تعداد متغیرهای پیش‌بین = ۵، برابر با ۱۹۱ نفر است.

یافته‌ها

پس از شرط‌گذاری روی داده‌ها مشخص شد دانشجویان مردودی که شرایط قرار گرفتن در نمونه آماری تحقیق را دارند از دانشجویان دانشکده‌های مهندسی، ریاضی، فیزیک و شیمی شیمی بوده‌اند. باید دقت کرد که تعداد کل دانشجویان مردودی در سطح دانشگاه به مراتب بیشتر بوده است. اما تعداد دانشجویانی که پس از مردودی، بلافاصله همان درس را با همان استاد گرفته بودند ۲۶۸ نفر بوده است. این دانشجویان در ۳۸ کلاس متفاوت بوده‌اند که متوسط تعداد ثبت‌نامی کلاس حدود ۴۳ نفر بوده است. غالب این کلاس‌ها مربوط به دروس پایه یا سرویس بودند که دانشجویان به طور معمول آنها را در نیمسال‌های نخست تحصیل خود می‌گذرانند.

ارزشیابی دانشجویان از تدریس اساتید بر اساس فرم رسمی مؤسسه انجام شده بود. آیتم‌های هفده‌گانه این فرم توسط مدیریت دانشگاه دسته‌بندی یا مؤلفه‌بندی نشده بودند و به همین دلیل، برای تحلیل دقیق‌تر ارزشیابی‌های دانشجویان، ابتدا مؤلفه‌های مکنون در فرم ارزشیابی دانشگاه با استفاده از تحلیل عامل اکتشافی شناسایی شدند.

جدول ۲: نتایج تحلیل‌های عامل اکتشافی بر روی آیتم‌های پرسشنامه ارزشیابی تدریس دانشجویان

مؤلفه‌های ارزشیابی دانشجویان از تدریس	تعداد آیتم‌ها	پایایی	شاخص‌های تحلیل عامل اکتشافی	
			ESSL ^۱	KMO ^۲
تسلط و توان علمی	۳	۰/۹۱	۸۵/۱۱۵	۰/۷۵۹
انگیزش و احترام	۳	۰/۸۹	۸۲/۵۰۲	۰/۷۲۵
مدیریت کلاس و ارتباطات	۳	۰/۹۰	۸۴/۲۸۷	۰/۷۵۷
رفتارهای تدریس اثربخش	۴	۰/۹۳	۸۳/۱۰۸	۰/۸۶۰
اخلاق حرفه‌ای	۴	۰/۹۱	۸۰/۳۸۳	۰/۸۲۳
کل فرم ارزشیابی	۱۷	۰/۹۸	۷۴/۷۳۲	۰/۹۶۳

چنانکه اطلاعات جدول ۲ نشان می‌دهد، تحلیل عامل اکتشافی به شناسایی پنج مؤلفه برای کیفیت تدریس منجر شد. تفکیک این مؤلفه از آنجا اهمیت دارد که تحلیل نمره کلی ارزشیابی دانشجویان از اساتید می‌تواند گمراه‌کننده باشد. در واقع، از تحلیل نمره کلی مشخص نمی‌شود که دانشجویان هر بخش از رفتارهای حرفه‌ای استادان را چگونه ارزیابی کرده‌اند.

پرسش اصلی در پژوهش حاضر این بود که آیا دانشجویانی که در درس مردود می‌شوند در نیمسال بعد که درس را با همان استاد می‌گیرند کیفیت تدریس او را آگاهانه یا ناآگاهانه به شکلی سوگیرانه ارزشیابی می‌کنند؟ طبیعتاً اگر پاسخ این سؤال مثبت باشد باید کاهش معناداری در ارزشیابی این دسته از دانشجویان از اساتید مشاهده شود. با توجه به اینکه در اینجا ارزشیابی‌های یک دانشجو از استادی مشخص و در درسی مشخص در دو نیمسال متوالی استخراج شده‌اند برای مقایسه ارزشیابی‌های دو نیمسال از آزمون مقایسه زوجی استفاده شده است.

جدول ۳: نتایج مقایسه زوجی میانگین ارزشیابی دانشجویانی مردودی از استاد درس در دو نیمسال متوالی

نیمسال‌ها	n	میانگین	انحراف معیار	تفاوت میانگین	t	Sig.
۱	۲۶۸	۱۷/۸۷۶	۲/۲۰۱			
۲	۲۶۸	۱۸/۶۵۹	۱/۸۹۹	۰/۷۸۲	۵/۱۱۶	۰/۰۰۰

چنانکه اطلاعات جدول ۳ نشان می‌دهد در نیمسال دوم افزایش معناداری به اندازه ۰/۷۸ در ارزشیابی نیمسال دوم ایجاد شده است. همبستگی نمرات دو نیمسال برابر با ۰/۲۶۰ و در سطح اطمینان ۹۹ درصد معنادار بود. این یافته را می‌توان مدرکی برای ابطال این فرضیه دانست که اگر دانشجو در درس مردود شود در نیمسال آتی نمره ارزشیابی کمتری به استاد خواهد داد. حداقل، در مجموع می‌توان اطمینان داشت که وقتی نمرات ارزشیابی دانشجویان مردودی تجمیع شوند میانگین کلی آن نسبت به نیمسال قبلی بالاتر است. چنانکه پیشتر اشاره شد آیتم‌های فرم ارزشیابی

1. extraction sums of squared loadings

2. Kaiser-Meyer-Olkin Measure of Sampling Adequacy

دانشگاه متنوع بودند و ابعاد مختلفی از عملکرد حرفه‌ای استادان در آن لحاظ شده بود. در مرحله دوم تحلیل، مقایسه دقیق‌تر نمرات ارزشیابی دانشجویان از کیفیت تدریس اساتید دو نیمسال بر حسب مؤلفه‌های به دست آمده از تحلیل اکتشافی انجام شد که نتایج آن در جدول شماره ۳ گزارش شده است.

جدول ۴: نتایج مقایسه زوجی میانگین ارزشیابی دانشجویانی مردودی از اساتید در دو نیمسال متوالی

مؤلفه‌ها	نیمسال	n	میانگین	انحراف معیار	تفاوت میانگین	t	Sig.
تسلط و توان علمی	۱	۲۶۸	۱۷/۵۹۲	۲/۵۱۴	۱/۰۶۱	۶/۰۲۷	۰/۰۰۰
	۲	۲۶۸	۱۸/۶۵۳	۲/۰۴۰			
انگیزش و احترام	۱	۲۶۸	۱۸/۰۶۱	۲/۱۷۶	۰/۵۶۱	۴/۰۳۱	۰/۰۰۰
	۲	۲۶۸	۱۸/۶۲۳	۱/۶۱۶			
مدیریت کلاس و ارتباطات	۱	۲۶۸	۱۷/۷۵۲	۲/۳۵۳	۰/۷۸۱	۴/۶۵۷	۰/۰۰۰
	۲	۲۶۸	۱۸/۵۳۳	۲/۰۷۴			
رفتارهای تدریس اثربخش	۱	۲۶۸	۱۷/۶۳۳	۲/۴۰۸	۰/۹۳۰	۵/۴۱۹	۰/۰۰۰
	۲	۲۶۸	۱۸/۵۶۴	۲/۰۸۷			
اخلاق حرفه‌ای	۱	۲۶۸	۱۸/۳۰۵	۲/۰۷۵	۰/۴۶۵	۳/۱۷۰	۰/۰۰۲
	۲	۲۶۸	۱۸/۷۷۰	۱/۸۷۵			

چنانکه اطلاعات جدول ۴ نشان می‌دهد در نیمسال دوم در تمام مؤلفه‌های فرم ارزشیابی دانشجویان از تدریس، در سطح اطمینان ۹۹ درصد افزایش معناداری حدوداً بین ۰/۵ تا ۱ نمره قابل مشاهده است. البته، باید افزایش میانگین نمرات ارزشیابی دانشجویان مردودی در نیمسال دوم را از مسأله رابطه میان نمرات دانشجویان و ارزشیابی آنها از تدریس اساتید تفکیک کرد. اطلاعات تکمیلی در این مورد در جدول شماره ۵ ارائه شده است.

جدول ۵: پیش‌بینی ارزشیابی کلی استاد از روی معدل، نمره، جنسیت، وضعیت تأهل دانشجو، و تعداد واحد درس

متغیرهای مستقل	Coef.	Std. Err.	z	P > z	[95% Conf. Interval]
معدل دانشجو	-۰/۱۳۰	۰/۰۴۸	-۲/۷۰	۰/۰۰۷	-۰/۲۲۵ - ۰/۰۳۵
نمره دانشجو	۰/۰۹۹	۰/۰۲۴	۴/۰۴	۰/۰۰۰	۰/۰۵۱ - ۰/۱۴۸
جنسیت دانشجو	-۰/۳۲۲	۰/۱۹۹	-۱/۶۲	۰/۱۰۵	-۰/۷۱۲ - ۰/۰۶۷
وضعیت تأهل	۰/۳۰۰	۰/۴۰۹	۰/۷۳	۰/۴۶۳	-۰/۵۰۱ - ۱/۱۰۲
تعداد واحد درس	۰/۴۳۷	۰/۳۶۶	۱/۱۹	۰/۲۳۲	-۰/۲۸۰ - ۱/۱۵۵
اندازه ثابت	۱۷/۴۶	۱/۳۸۲	۱۲/۶۳	۰/۰۰۰	۱۴/۷۵۳ - ۲۰/۱۷۴

مقدار آماره لگاریتم درست‌نمایی: -۱۱۳۱/۸۸ کای اسکوتر والد: ۲۳/۶۷ سطح معنی‌داری: ۰/۰۰۰۳

پارامترهای اثرات تصادفی	برآورد	خطای معیار	فاصله اطمینان ۹۵٪
استاد درس: واریانس سطح ۲	۰/۴۷۹	۰/۲۱۱	۰/۲۰۲ - ۱/۱۳۶
واریانس خطا	۳/۷۵۷	۰/۲۳۸	۳/۳۱۸ - ۴/۲۵۳

مقدار آماره کای اسکوتر = ۲۱/۲۶ سطح معنی‌داری = ۰/۰۰۰۰

طبق اطلاعات جدول شماره ۵ در اینجا نیز ضریب همبستگی درون‌رده‌ای (حاصل تقسیم واریانس سطح دو بر مجموع واریانس سطح دو و واریانس خطا) معادل ۰/۱۱ است به این معنا که یازده درصد از واریانس نمرات ارزشیابی ناشی از تفاوت‌های میان اساتید است که در تحلیل فوق به عنوان پارامترهای اثرات تصادفی^۱ کنترل شده است. شاید بتوان مهم‌ترین ضریب را در تحلیل فوق به متغیر «معدل دانشجو» مربوط دانست که با کاهش یک واحد در آن، کاهش به اندازه سیزده صدم در نمرات ارزشیابی استادان قابل مشاهده است. به بیان دقیق‌تر، با کاهش معدل دانشجویان مردودی که می‌توان آن را ملاک معتبری برای استعداد و توان تحصیلی آنها محسوب کرد کاهش معناداری نیز در نمرات ارزشیابی اساتید قابل ملاحظه است. در مقابل، ضریب رگرسیون «نمره دانشجو» نیز به خوبی نشان می‌دهد که با یک واحد افزایش در نمرات دانشجویان مردودی، افزایش معناداری به اندازه یک درصد در ارزشیابی آنها از اساتید مشاهده می‌شود.

بحث و نتیجه‌گیری

شور و شوق دانشجویان و اندیشه‌های آنها می‌تواند الهام‌بخش تغییراتی اساسی در تدریس استادان باشد (کودی^۲ و همکاران، ۲۰۲۲). ارزشیابی‌های آنها اما می‌تواند فرایندهای تدریس و تجربه‌های یادگیری را در جهاتی پیش‌بینی‌ناپذیر دگرگون کند. در واقع، انبوه^۳ ارزشیابی‌هایی که به مدد ابزارهای فناوریانه میسر شده‌اند بر ارتباطات میان مدرسان با دانشگاه و نیز با دانشجویان اثرات متعددی می‌گذارند (لیکمن و همکاران، ۲۰۲۲). البته باید توجه کرد که در برخی نظام‌های آموزش عالی، نتایج ارزشیابی‌های دانشجویی برای خود دانشجویان نیز قابل مشاهده است و آنها می‌توانند درس‌های مورد پسند و علاقه خویش را بر اساس همین ارزشیابی‌ها شناسایی کنند (هی^۴ و همکاران، ۲۰۲۲). در دانشگاه‌های ایران اما نتایج ارزشیابی‌ها به طور عمده در اختیار مدیران دانشگاهی است و ماهیتی محرمانه دارند و لذا می‌توان امیدوار بود که برخی از پیامدهای منفی از جمله رقابت‌های ناشی از انتشار نتایج ارزشیابی‌ها در بافت دانشگاه‌های ایران کمتر باشد. در این مقاله به انتقادات و حساسیت‌های مختلف استادان دانشگاه در زمینه ارزشیابی‌های دانشجویی اشاره شد. به نظر می‌رسد مدیران دانشگاهی می‌توانند با بهره‌گیری از یافته‌های پژوهشی موجود و یا سفارش پژوهش‌های مورد نیاز به دغدغه‌ها و پرسش‌های استادان پاسخ دهند و ارزشیابی‌های دانشجویان از کیفیت تدریس را به ابزاری کارآمدتر برای بهبود فرایندهای یاددهی-یادگیری تبدیل کنند.

یافته‌های پژوهش حاضر نشان داد که بر خلاف فرضیه اولیه تحقیق و نیز تصور بسیاری از استادان، دانشجویانی که در درس مردود می‌شوند و دوباره درس مورد نظر را با استاد اخذ می‌کنند در نیمسال بعد ارزشیابی بهتری از تدریس استاد خود دارند. باید توجه کرد که وقتی داده‌ها به صورت مردودی بررسی می‌شوند انواع رفتارها و سبک‌های نمره‌دهی در دانشجویان قابل مشاهده است. برخی دانشجویان در نیمسال نخست نمره بالاتری به استاد داده بودند. برخی دانشجویان در هر دو نیمسال، حداقل نمره و برخی حداکثر نمره را به استاد داده بودند. عده‌ای از دانشجویان نیز در نیمسال دوم، بر خلاف نیمسال اول، استاد خویش را بسیار ضعیف ارزشیابی کرده بودند. نکته مهم اما اینست که اگر میانگین کلی ملاک باشد این گروه از دانشجویان در نیمسال بعدی، مجموعاً نمرات ارزشیابی بالاتری به استادان خود می‌دهند. جالب اینکه بسیاری از این دانشجویان در نیمسال دوم نیز در درس مردود شده بودند و فقط ۳۵ درصد از این دانشجویان نمرات بالاتر از ۱۲ گرفته بودند. به لحاظ منطقی غالب این دانشجویان می‌دانسته‌اند که برای بار دوم نیز نمره چندان بالایی در انتظار آنها نیست^۵.

یافته دیگر پژوهش این بود که افزایش معنادار نه فقط در نمره کلی ارزشیابی، بلکه در تمام مؤلفه‌ها یعنی تسلط و توان علمی، انگیزش و احترام، مدیریت کلاس و ارتباطات، رفتارهای تدریس اثربخش، و اخلاق حرفه‌ای مشاهده شد. این یافته را می‌توان همسو با نتیجه‌گیری کزن و کپریانی (۲۰۲۲) محسوب کرد زیرا این محققان دریافتند که میان نمرات آیتم‌های فرم ارزشیابی همیشه همبستگی بالایی وجود دارد. اما تبیین منطقی‌تر به ترجیحات یادگیری دانشجویان ضعیف مربوط است. در یک تبیین نظری جالب، جهانگیری و موسیولو^۶ (۲۰۱۲) ترجیحات یادگیری دانشجویان

1. Random-effects Parameters

2. Code

3. massification

4. He

۵. یادآوری مجدد این نکته اهمیت دارد که دانشجویان ارزشیابی تدریس را قبل از ایام امتحانات انجام می‌دهند.

6. Jahangiri & Mucciolo

را به ادراک آنها از اثربخشی استادان ارتباط داده‌اند. این متخصصان دانشجویان را به دو دسته «فراگیران دانشجو»^۱ و «فراگیران حرفه‌ای»^۲ تقسیم کرده‌اند. برای فراگیران دانشجو مهم‌ترین مهارت‌های تدریس استادان به طراحی، سازماندهی و تدوین محتوا مربوط می‌شود. برای فراگیران حرفه‌ای اما اعتماد به نفس، زبان بدن، و انرژی استاد اهمیت بیشتری دارد. برای فراگیران دانشجو بسیار مهم است که استاد موضوع را خلاصه و موجز، سازماندهی شده و گام به گام تدریس کند. برای فراگیران حرفه‌ای و خود استادان اما خصوصیات نظیر اشتیاق، ارتباط‌پذیری^۳، احترام، دلسوزی و رسیدگی و قدری شوخ‌طبعی بر اثربخشی استاد دلالت می‌کنند. وجوه مشترک دو گروه اینست که دوست دارند استاد بتواند موضوعات دشوار و چالش‌برانگیز را برای آنها قابل فهم کند، شفاف سخن بگوید، از موضوع منحرف نشود، آزادانه با دانشجویان ارتباط برقرار کند. با عنایت به مبانی نظری ترجیحات یادگیری، این احتمال هست که دانشجویان ضعیف استاد را بر اساس میزان کمک وی به یادگیری آنها ارزشیابی کرده باشند. در واقع، می‌توان گفت که نمرات ارزشیابی دانشجویان ضعیف بیش از اینکه متأثر از نمره مردودی آنها در ترم قبل باشد مبتنی بر برداشت و ادراک آنها از فعالیت‌های آموزشی و حرفه‌ای استاد در ترم بعد است.

سرانجام، یکی از یافته‌ها مهم در پژوهش حاضر به ارتباط میان توان و استعداد تحصیلی و ارزشیابی دانشجو از کیفیت تدریس استاد مربوط بود. اگر بتوان معدل را ملاک معتبری برای استعداد و توان تحصیلی دانشجویان محسوب کرد می‌توان گفت دانشجویان ضعیف‌تر نمرات ارزشیابی کمتری به استادان می‌دهند. در عین حال، با یک واحد افزایش در نمرات همین دانشجویان مردود افزایش معناداری به اندازه یک درصد در ارزشیابی آنها از اساتید به وجود می‌آید که همسو با یافته‌های برزوی، لوکاتز و مولانتای (۲۰۲۰) و استروبی (۲۰۲۰) است.

با عنایت به این یافته‌ها به نظر نمی‌رسد که برآیند کلی ارزشیابی دانشجویان مردود از تدریس استاد سوگیرانه یا مغرضانه باشد. از سوی دیگر، بر فرض وجود سوگیری در ارزشیابی‌های این دانشجویان، تا وقتی که مردودی‌ها درصد محدودی از دانشجویان هر کلاس را تشکیل دهند جایی برای نگرانی شدید نیست زیرا بر اساس آئین‌نامه ارتقا، برای محاسبه امتیاز کیفیت تدریس استاد میانگین نمرات ارزشیابی کل دانشجویان در چندین نیمسال مد نظر قرار گیرد. افزون بر این، باید توجه کرد که اولین اصل پیشنهادی پرت^۴ (۱۹۹۷) برای ارزشیابی تدریس در آموزش عالی اینست که در ارزشیابی باید به گوناگونی باورها، نیات و اقدامات دانشجویان احترام گذاشته شود.

با عنایت به یافته‌ها، استدلال‌ها و نظرات متخصصان، تجمیع و محاسبه میانگین نمرات ارزشیابی کل دانشجویان، صرف نظر از وضعیت پیشرفت تحصیلی آنها، منطقی و قابل توجیه به نظر می‌رسد. در عین حال، نباید ارتباط میان نمرات دانشجویان در درس‌ها و ارزشیابی آنها از تدریس اساتید مغفول واقع شود. یکی از پیامدهای بی‌توجهی به این ارتباط، بروز تورم نمرات^۵ در سطح مؤسسه است. مدیران دانشگاهی باید همیشه به هدف ارزشیابی دانشجویان از کیفیت تدریس توجه کنند (لوید و رایت‌براف^۶، ۲۰۲۲). هدف هزاران ارزشیابی دانشجویی از تدریس استادان، بهبود فرایندهای یاددهی-یادگیری در کلاس‌های دانشگاهی است. مدیران دانشگاهی باید با در نظر داشتن این هدف غایی و اساسی، ابهامات مدرسان را درباره ارزشیابی‌های دانشجویی برطرف کنند به گونه‌ای که مدرسان به جای سهل‌گیری، ارفاق و یا تنزل معیارهای آکادمیک، دانشجویان را بر اساس شایستگی‌های تحصیلی و تلاش آنها برای رسیدن به اهداف آموزشی، سنجش و ارزشیابی کنند.

1. student learners

2. professional learners

3. approachable

4. Pratt

5. grade inflation

6. Lloyd & Wright-Brough

References

- Baker, Vicki. L. & Manning, Caroline. F. N. (2021). A Mid-Career Faculty Agenda: A Review of Four Decades of Research and Practice. In: Laura W. Perna, *Higher Education: Handbook of Theory and Research*. Switzerland: Springer.
- Berk, Ronald. A. (2006). *Thirteen Strategies to Measure College Teaching*. Virginia: Stylus.
- Berezvai, Zombor., Lukats, Gergely, D., & Molontav, Roland. (2020). Can professors buy better evaluation with lenient grading? The effect of grade inflation on student evaluation of teaching. *Assessment & Evaluation in Higher Education*. DOI: 10.1080/02602938.2020.1821866
- Borch, I., Sandvoll, R. & Risor, T. (2020). Discrepancies in purposes of student course evaluations: what does it mean to be "satisfied"? *Educational Assessment, Evaluation and Accountability*. 32 (1): 83-102.
- Buller, Jeffrey. L. (2012). *Best practices in faculty evaluation: A Practical guide for academic leaders*. San Francisco, Jossey-Bass.
- Cannon, Edmund., & Cipriani, Giam, Pietro. (2022). Quantifying halo effects in students' evaluation of teaching. *Assessment & Evaluation in Higher Education*. 47 (1): 1-14.
- Code, Warren., et al. (2022). Collaboration and mentoring to enhance professional learning in higher education. In: Helen King, *Developing Expertise for Teaching in Higher Education*. London: Routledge.
- Esarey, Justin., & Valdes, Natalie. (2020). Unbiased, reliable, and valid student evaluations can still be unfair. *Assessment & Evaluation in Higher Education*. 45:8. 1106-1120.
- Griffin, Kimberly. A. (2020). Institutional Barriers, Strategies, and Benefits to Increasing the Representation of Women and Men of Color in the Professoriate In: Laura W. Perna, *Higher Education: Handbook of Theory and Research* (Volume 35). Philadelphia: Springer.
- Gunduz, Necla., & Fokoueb, Ernest. (2021). Understanding students' evaluations of professors using non-negative matrix factorization. *Journal of Applied Statistics*. 48 (13-15): 2961-2981.
- He, Jibo., & et al. (2022). Reciprocity in college teaching: a big data study based on online student evaluation of 919,750 professors. *Assessment & Evaluation in Higher Education*, DOI:10.1080/02602938.2022.2067980.
- Holmes, Jeffrey. D. (2016). *Great myths of education and learning*. Hoboken: Wiley-Blackwell.
- Jahangiri, Leila., & Mucciolo, Tom. (2012). *A Guide to Better Teaching: Skills, Advice, and Evaluation for College and University Professors*. Lanham: Rowman & Littlefield Publishers, Inc.
- Johnson, R. Burke., & Christensen, Larrv. (2016). *Educational research: Quantitative, qualitative, and mixed approaches*. Los Angeles: Sage.
- King, Helen. (2022). The characteristics of expertise for teaching in higher education. In: Helen King, *Developing Expertise for Teaching in Higher Education*. London: Routledge.
- Kwok, Kathryn & Potter, Jackie. (2022). Supportive woman, engaging man Gendered differences in student perceptions of teaching excellence. In: Helen King, *Developing Expertise for Teaching in Higher Education*. London: Routledge.
- Kreitzer, R.J., Sweet-Cushman, J. (2022). Evaluating Student Evaluations of Teaching: A Review of Measurement and Equity Bias in SETs and Recommendations for Ethical Reform. *Journal of Academic Ethics*. 20 (1): 73-84.
- Lakeman, Richard., & et al. (2022). Playing the SET game: how teachers view the impact of student evaluation on the experience of teaching and learning. *Assessment & Evaluation in Higher Education*, DOI: 10.1080/02602938.2022.2126430
- Lloyd, Margaret., & Wright-Brough, Freva. (2022). Setting out SET: a situational mapping of student evaluation of teaching in Australian higher education, *Assessment & Evaluation in Higher Education*, DOI: 10.1080/02602938.2022.2130169
- Marsh, Herbert. 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.
- Marsh, Herbert. W. (2001). Distinguishing between Good (Useful) and Bad Workloads on Students' Evaluations of Teaching. *American Educational Research Journal*. 38 (1): 183-212.
- Miller, Michael T., & Gearhart, David. (2021). *Handbook of Research on the Changing Role of College and University Leadership*. Hershey PA, USA: IGI Global.
- Pratt, Daniel. D. (1997). Reconceptualizing the Evaluation of Teaching in Higher Education. *Higher Education*. 34 (1): 23-44.
- Smith, Marc., & Firth, Jonathan. (2018). *Psychology in the Classroom: A Teacher's Guide to What Works*. New York: Routledge.
- Stroebe, Wolfgang. (2020). Student Evaluations of Teaching Encourages Poor Teaching and Contributes to Grade Inflation: A Theoretical and Empirical Analysis. *Basic and Applied Social Psychology*, 42 (4): 276-294.