

مقاله پژوهشی

شناسایی شاخص‌های مطلوب ارزشیابی عملکرد دانشجویان در رشته‌های علوم انسانی دانشگاه‌ها و میزان توجه به آنها

مصطفی باقریان فر^{۱*}، احمدرضا نصرافهانی^۲، محمد جوادی‌پور^۳، اصغر زمانی^۴

۱. نویسنده مسئول: استادیار، گروه نوآوری آموزشی و درسی، موسسه پژوهش و برنامه‌ریزی آموزش عالی، تهران، ایران رایانامه: m.bagherian@irphe.ac.ir
۲. استاد گروه علوم تربیتی، دانشگاه اصفهان، دانشکده علوم تربیتی و روان‌شناسی، اصفهان، ایران.
۳. دانشیار گروه برنامه‌ریزی درسی، دانشگاه تهران، دانشکده علوم تربیتی و روان‌شناسی، تهران، ایران.
۴. دانشیار گروه مدیریت آموزش عالی، موسسه پژوهش و برنامه‌ریزی آموزش عالی.

تاریخ پذیرش: ۱۴۰۱/۰۹/۲۰

تاریخ دریافت: ۱۴۰۰/۰۹/۰۴

چکیده:

ارزشیابی به‌عنوان یکی از ابزارهای بسیار ارزشمند ارتقای کیفیت تدریس و یادگیری محسوب می‌شود، زیرا می‌تواند بازخورد مفیدی برای استادان و دانشگاه در مورد کیفیت عناصر برنامه درسی فراهم سازد. پژوهش حاضر با هدف شناسایی شاخص‌های مطلوب ارزشیابی عملکرد دانشجویان در رشته‌های علوم انسانی دانشگاه‌ها و میزان توجه به آنها انجام شد. در پژوهش حاضر، از طرح آمیخته اکتشافی متوالی از نوع ابزارسازی استفاده شد. در بخش کیفی، از روش مطالعه موردی کیفی و در بخش کمی، از روش توصیفی - پیمایشی بهره گرفته شده است. در بخش کیفی با ۲۰ نفر از متخصصان و اعضای هیئت علمی دانشگاه‌های کشور که در حوزه ارزشیابی اهل نظر هستند، مصاحبه نیمه‌ساختاریافته به عمل آمد. جامعه آماری بخش کمی دانشجویان کارشناسی دانشگاه‌ها بودند که به روش خوشه‌ای چندمرحله‌ای، تعداد ۳۶۰ نفر به‌عنوان نمونه انتخاب شدند. داده‌های پژوهش در بخش کمی از طریق پرسشنامه محقق ساخته مستخرج از بخش کیفی جمع‌آوری شد. برای تعیین روایی سؤال‌های مصاحبه و پرسشنامه از روایی صوری و محتوایی و برای برآورد پایایی پرسشنامه از ضریب آلفای کرونباخ استفاده شد. برای تحلیل داده‌های پژوهش در بخش کیفی از روش‌های ساختاری و تفسیری و در بخش کمی از آمار توصیفی و استنباطی بهره گرفته شد. نتایج نشان داد که ده شاخص از منظر مصاحبه‌شوندگان مهم‌ترین شاخص‌های ارزشیابی بودند و می‌توان از شاخص‌های شناسایی شده برای ارزشیابی و بازنگری ارزشیابی عملکرد دانشجویان در دروس رشته‌های علوم انسانی دانشگاه‌ها استفاده نمود. یافته‌های کمی نشان داد میانگین توجه به شاخص‌های ارزشیابی عملکرد دانشجویان در دروس رشته‌های علوم انسانی دانشگاه‌ها نامطلوب بوده و نیاز به توجه بیشتری دارد.

کلیدواژه‌ها: ارزشیابی، برنامه درسی، دانشگاه‌ها، رشته‌های علوم انسانی.

استاد به این مقاله:

باقریان فر، مصطفی؛ نصرافهانی، احمدرضا؛ جوادی‌پور، محمد؛ زمانی، اصغر. (۱۴۰۳). شناسایی شاخص‌های مطلوب ارزشیابی عملکرد دانشجویان در رشته‌های علوم انسانی دانشگاه‌ها و میزان توجه به آنها. *اندیشه‌های نوین تربیتی*، ۲۰(۳): ۸۱-۹۸. doi: 10.22051/jontoe.2022.37656.3414

مقدمه

یکی از عناصر برنامه درسی که صاحب‌نظران مذکور به اتفاق نظر به آن اشاره داشته‌اند و در این پژوهش مورد بررسی قرار گرفته است ارزشیابی است. همچنین علی‌رغم تفاوت نظر در تعداد عناصر برنامه درسی، عنصر مشترک تمامی این طبقه‌بندی‌ها، ارزشیابی است (هریتیج و چن، ۲۰۰۵، ص. ۱۰). ارزشیابی آموخته‌های دانشجویان عنصری تأثیرگذار در برنامه درسی معرفی شده است، به طوری که هرگونه ضعف در آن، که عمدتاً در آخر چرخه برنامه درسی قرار می‌گیرد، می‌تواند به شکست یک برنامه درسی بینجامد. ارزشیابی به عنوان قضاوت و داوری مبتنی بر ارزش تعریف شده است. این در منطق کلی ارزشیابی بیان شده و در چهار مرحله تعیین معیارها، تعیین استانداردهای عملکرد، اندازه‌گیری عملکرد و سپس ترکیب و ادغام در یک قضاوت توصیف شده است (روردا و گالکسن، ۲۰۱۹، ص. ۳).

اصطلاح ارزشیابی در برخی از کشورها از جمله ایالات متحده آمریکا، برای اشاره به دستاوردهای فردی دانشجویان یا پیشرفت تحصیلی آنان به کار می‌رود در حالی که در کشور انگلستان به عنوان ارزیابی^۱ توصیف می‌شود (هارلن، ۲۰۰۷، ص. ۴). ارزشیابی شامل دو مؤلفه اصلی است: (۱) گردآوری داده‌ها با استفاده از ابزار اندازه‌گیری که معیارهای مربوطه در آن گنجانده شده است؛ (۲) استفاده از نتایج برای قضاوت یا تصمیم‌گیری اهداف با استناد به استانداردهای توافق شده (آلدمن و همکاران، ۲۰۱۴، ص. ۳). چنگ، لیو و جیا (۲۰۲۴)، ارزشیابی را به عنوان یک فرایند حیاتی می‌داند که دستیابی به اهداف آموزشی را هم برای مدرسان و هم برای دانشجویان فراهم می‌کند و هدف از ارزشیابی را تعیین تصمیم‌گیری‌هایی می‌داند که بر اساس اطلاعات جمع‌آوری شده، گرفته می‌شود. استافل بیم و شینک‌فیلد (۲۰۰۷، ص. ۵) استدلال می‌کنند یکی از اهداف اصلی ارزشیابی، ارائه بازخورد به دانشجویان و استادان در مورد میزان انجام فعالیت‌ها و کارایی آنهاست. هدف دیگر ارزشیابی به اعتقاد آنان، هدایت دانشجویان برای تعیین مراحل ارزشیابی و اهداف مورد نیاز برای بهبود برنامه‌های کاربردی است. در همین راستا، براون (۲۰۰۴، ۲۰۰۲، ۲۰۰۶) سه مفهوم اصلی اهداف ارزشیابی در کلاس درس را بهبود آموزش و یادگیری، ارزشیابی به عنوان پاسخگویی دانشگاه و ارزشیابی به عنوان پاسخگویی دانشجویان اشاره کرده است (دلوکا، کومب و لاپوینت-مک‌اوان، ۲۰۱۹، ص. ۹).

به گفته لاتوکا و استارک (۲۰۰۹، ص. ۳) با در نظر گرفتن برنامه درسی، اصطلاح سنجش به عنوان جزئی از فرایندهای ارزشیابی است که بر پیشرفت تحصیلی دانشجویان تأثیر دارد. ارزشیابی آموزشی با توجه به اهدافی که دارد می‌توان آن را به چند شکل فرض کرد. عمومی‌ترین حالت‌های ارزشیابی دو مقیاس بیرونی (خارجی) و ارزشیابی کلاسی (داخلی) هستند. مطابق با اهداف ارزشیابی، اسکریون (۱۹۶۷، ص. ۸) بین ارزشیابی برای اهداف سازنده (بهبود یادگیری) و ارزشیابی برای اهداف نهایی (تصمیم‌گیری) تمایز قائل شد. مطالعات فراتحلیل حوزه ارزشیابی سازنده، نشان می‌دهند که ضریب تأثیر این ارزشیابی در بهبود یادگیری بسیار زیاد است (هاتی، ۲۰۰۹، ص. ۳). همچنین ارزشیابی پایانی، به ارزشیابی که در پایان فرایند تدریس یا ترم تحصیلی توسط استادان برای تصمیم‌گیری نهایی گرفته می‌شود (ابوالحیجا، ۲۰۰۷، ص. ۶). صرف نظر از چیزی که مورد ارزشیابی قرار می‌گیرد، فرایند ارزشیابی یکی است. در انواع مختلف ارزشیابی‌ها چیزی که تغییر می‌کند؛ موضوع مورد ارزشیابی، نحوه به کارگیری فرایند ارزشیابی و انواع تصمیماتی است که گرفته می‌شوند.

در ارتباط با ارزشیابی الگوهای متفاوتی بوده که می‌توان به مواردی چون الگوی ارزشیابی هدف-آزاد؛ الگوی سیپ، الگوی ارزشیابی هدف‌محور، خبرگی و نقادی، کرک‌پاتریک، پاسخ‌ده، الگوی تدافعی، مشارکتی، اعتبارگذاری اشاره کرد (عزیزی محمودآباد و نیلی، ۱۳۹۸). تأثیرات ارزشیابی بر یادگیری دانشجویان به طور گسترده‌ای مورد بحث قرار گرفته است. به طوری که کوهن و سامپسون (۱۹۹۹، ص. ۷) اظهار داشتند ارزشیابی تأثیر زیادی بر یادگیری در دوره‌های رسمی دارد. این امر نقش مهمی در فرایندها (چیستی محتوا) و رویکردهای (میزان و چگونگی) یادگیری دانشجویان ایفا می‌کند (دای، متیس و رئیسی، ۲۰۲۰، ص. ۸). پترز و همکاران (۲۰۱۷، ص. ۵) بر این باور



بودند که برای ارزشیابی منصفانه و عادلانه، لازم است: الف) به‌عنوان یک ابزار تشخیصی در نظر گرفته شود، نه به‌عنوان مکانسیم طبقه‌بندی دانشجویان؛ ب) برای تقویت آموزش و یادگیری دانشجویان استفاده شود، نه به‌عنوان ابزاری برای اندازه‌گیری عملکرد آنان؛ و ج) به‌عنوان ارزشیابی کلی دانشجویان استفاده شود، نه به‌عنوان تنبیه دانشجویانی که نمره لازم را کسب نکرده‌اند.

نتایج پژوهش کوک، واستون و وب (۲۰۲۴) نشان داد ارزشیابی از عملکرد دانشجویان یک معیار کلیدی برای ارزیابی کیفیت تدریس در آموزش عالی بوده و در این زمینه سوگیری‌هایی آماری و شناختی صورت گرفته است. سوگیری‌ها می‌تواند در عملکرد دانشجویان تأثیرگذار باشد و عملکرد فردی آنها را تحت تأثیر قرار دهد. نتایج پژوهش‌های چنگ و همکاران (۲۰۲۴)، موریلو و هیدالگو (۲۰۲۰، ص. ۶)، رسولی، زندی و دلوکا (۲۰۱۸، ص. ۱)، ژو و براون (۲۰۱۶، ص. ۱۰)، دلوکا و همکاران (۲۰۱۶، ص. ۸) و دلوکا (۲۰۱۲، ص. ۴) نشان داد که رعایت اصل عدالت یکی از اصول مهم در نحوه ارزشیابی استادان می‌باشد. همچنین تیری (۲۰۱۴، ۲۰۱۶) رعایت انصاف در ارزشیابی را برای تضمین عدالت ضروری می‌داند و از سازگاری ارزشیابی با نیازها و خصوصیات دانشجویان پشتیبانی می‌کند. افزون بر این، پتیفور و ساکلوفسک (۲۰۱۲، ص. ۵) ادعا می‌کنند که روش‌های مختلفی وجود دارد که می‌تواند ارزشیابی آموزشی را به یک عمل عادلانه‌تر تبدیل کنند. یکی از این موارد، آشنایی دانشجویان با معیارهای ارزشیابی است که در ارزشیابی اعمال خواهد شد یا حتی، این معیارها می‌تواند به‌طور مشترک توسط استادان و دانشجویان تعریف شوند. نتایج پژوهش‌های پیرا، فلورس، سیموو و باروس (۲۰۱۶، ص. ۸)، شوت (۲۰۰۸، ص. ۶) و کورنلیوس - وایت (۲۰۰۷، ص. ۱) حاکی از آن است که بازخورد یکی از اصول ارزشیابی است و باید روی کار دانشجویان متمرکز باشد نه ویژگی‌های شخصی آنان. همچنین بازخورد نباید دانشجویان را دلسرد و با دیگران مقایسه کند و برای کارهای سخت از بازخورد فوری و برای کارهای ساده از بازخورد تأخیری استفاده نماید. استرووین و همکاران (۲۰۰۵، ص. ۷) اعتقاد دارند حالت‌های ارزشیابی سازنده، فعال و خلاقانه، مانند خودارزیابی و ارزشیابی فرایند یادگیری، دانشجویان را ترغیب می‌کند که بر ساخت دانش، درک عمیق و معنادار دانش و یادگیری متمرکز شوند در حالی که ارزشیابی سنتی مانند امتحانات و آزمون باعث می‌شود دانشجویان بیشتر به حفظ مطالب و نمرات توجه کنند. همچنین نتایج پژوهش‌های (وان دن برق، ۲۰۰۶؛ گیس، ۱۹۹۹؛ اسولر، ۱۹۹۸) نشان داد تنوع ارزشیابی تأثیر مهمی در روند یادگیری دانشجویان دارد. این بدان معناست که اگر استادان در محیط‌های یادگیری از ارزشیابی‌های متنوعی استفاده نمایند می‌توان نتایج بهتری به‌دست آورد.

جین و رون (۲۰۲۳) در پژوهشی نشان دادند زبان ارزشیابی حامل معانی نگرشی تشدید شده که ارتباط نزدیکی با تلاش و عملکرد دانشجویان دارد. سؤالات به‌عنوان یک ابزار ارزشیابی برای فرموله کردن بازخورد ارتباطی برای ارائه دانشگاهی در نظر گرفته شده است. گریتنسن ون لیوونکمپ، جوستن ده برینک و کستر (۲۰۱۹، ص. ۸) در پژوهشی دریافت که شش شاخص مرتبط با کیفیت ارزشیابی عبارت‌اند از: تأثیر ارزشیابی بر پیشرفت تحصیلی، انصاف در ارزشیابی؛ شرایط و وضعیت ارزشیابی؛ تفسیر نمرات ارزشیابی؛ صحت ارزشیابی و اعتبار ارزشیابی. اگر استادان بتوانند در شیوه‌های ارزشیابی خود درک مثبتی از این شاخص‌ها در دانشجویان ایجاد نمایند رویکرد یادگیری عمیق‌تر و استراتژیک‌تر می‌شود و نتایج بهتری محقق می‌شود. ارزشیابی و شرایط آن بر یادگیری تأثیرگذار بوده است. نتایج پژوهش‌های (داروین، ۲۰۱۱؛ بلک‌مور، ۲۰۰۹؛ آبرامی و همکاران، ۲۰۰۷؛ تیل و فلدمن، ۲۰۰۷؛ آرولا، ۲۰۰۷؛ مارش، ۲۰۰۷؛ برک، ۲۰۰۵؛ کانن، ۲۰۰۱؛ جانسون و رایان، ۲۰۰۰) حاکی از آن بود در آموزش عالی، برخی از روش‌های ارزشیابی در مؤسسات از اصول اسکریون استفاده نکرده‌اند و دارای موارد ضعفی چون نداشتن اصول و اهداف مشخص، نادیده گرفتن تغییرات در ماهیت یادگیری و تدریس و محیطی که در آن اتفاق می‌افتد، تمرکز تقریبی منحصراً بر رتبه دانشجویان که در محیط کلاس با استادان انفرادی کار می‌کنند؛ عدم اعتبار و روایی و استفاده از داده‌ها برای اهداف نامناسب هستند.

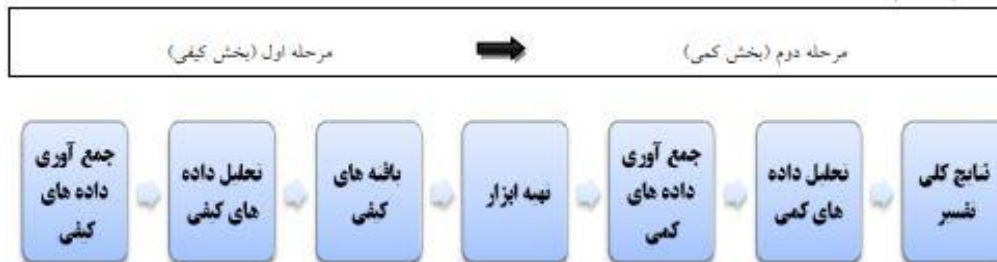
ارزشیابی از چگونگی یادگیری دانشجویان به‌عنوان جزء ضروری در آموزش مؤثر و به‌عنوان یک مسیر کلیدی در دستیابی به پیشرفت‌های مهم در توانایی دانشجویان تلقی می‌شود؛ اما شاخص‌های مرتبط برای تعیین کیفیت ارزشیابی چیست؟. با توجه به کمبود پژوهش در این زمینه، همچنین محدودیت‌هایی در برخی از مطالعات صورت گرفته، ضرورت انجام چنین پژوهشی روشن می‌شود؛ لذا پژوهش حاضر

درصد است تا شاخص‌های ارزشیابی مطلوب در آموزش عالی را شناسایی و میزان توجه به آنها در رشته‌های علوم انسانی دانشگاه‌ها را بررسی نماید و از این طریق به سؤالات پژوهش پاسخ دهد.

۱. شاخص‌های مطلوب برای ارزشیابی عملکرد دانشجویان در درس رشته‌های علوم انسانی دانشگاه‌ها چه مواردی است؟
۲. میزان توجه به شاخص‌های مطلوب ارزشیابی عملکرد دانشجویان در درس رشته‌های علوم انسانی در دانشگاه مورد مطالعه چه میزان است؟
۳. میزان توجه به شاخص‌های مطلوب ارزشیابی عملکرد دانشجویان در رشته‌های علوم انسانی در دانشگاه مورد مطالعه چه میزان است؟
۴. میزان توجه به شاخص‌های مطلوب ارزشیابی عملکرد دانشجویان در دانشگاه مورد مطالعه چه میزان است؟
۵. آیا بین نظرات دانشجویان در مورد میزان توجه به شاخص‌های مطلوب ارزشیابی عملکرد بر حسب رشته‌های علوم انسانی در دانشگاه مورد مطالعه تفاوت معناداری وجود دارد؟
۶. آیا بین نظرات دانشجویان در مورد میزان توجه به شاخص‌های مطلوب ارزشیابی عملکرد بر حسب جنسیت در دانشگاه مورد مطالعه تفاوت معناداری وجود دارد؟

روش شناسی پژوهش

پژوهش حاضر از نظر هدف کاربردی و به لحاظ روش، ترکیبی اکتشافی متوالی از نوع ساخت ابزار است که در آن داده‌های کیفی به منظور کمک به تنظیم پرسشنامه گردآوری و تحلیل شد. در بخش کیفی دیدگاه استادان دانشگاه‌های کشور و متخصصان حوزه تعلیم و تربیت درباره ارزشیابی و در بخش کمی دیدگاه دانشجویان سال سوم و چهارم دانشگاه اصفهان در سال تحصیلی ۱۳۹۸-۱۳۹۹ گردآوری شده است. بخش کیفی به روش مطالعه موردی کیفی و بخش کمی به روش توصیفی از نوع پیمایشی اجرا گردیده است.



شکل شماره ۱. فرآیند تحقیق ترکیبی اکتشافی متوالی از نوع ابزارسازی (اقتباس از کرسول و پلانوکلاک، ۲۰۰۷)

افراد متخصص در حوزه ارزشیابی به عنوان مشارکت‌کنندگان بالقوه، با استفاده از رویکرد هدفمند و روش نمونه‌گیری مطلعان کلیدی متناسب با موضوع پژوهش (صاحب‌نظران کلیدی^۱) و استفاده از معیار کفایت «اشباع نظری^۲ داده‌ها» انتخاب شدند. از این رو، ۲۰ نفر از استادان مجرب و اهل نظر کشور در حوزه ارزشیابی در این پژوهش مشارکت داده شدند. با توجه به اینکه اعتبار یافته‌ها در این قبیل پژوهش‌ها به توانمندی، دانش و تجارب مصاحبه‌شوندگان بستگی دارد، بدین گونه سعی شد افرادی انتخاب گردند که بالقوه می‌توانستند پاسخگوی سؤالات پژوهش باشند و دارای سوابق اجرایی و عملی، تسلط علمی و تجربه غنی از پدیده مورد بررسی و توانایی و تمایل به بیان روشن آن داشته و دارای تألیف و در آموزش و پژوهش شاخص باشند. از میان مصاحبه‌شوندگان ۶ نفر زن و ۱۴ نفر مرد و از نظر مرتبه علمی ۱۳ نفر در مرتبه دانشجویی و ۷ نفر استاد بودند. مشخصات دموگرافیک مصاحبه‌شونده‌های بخش کیفی در جدول شماره (۱) ارائه شده است:

جدول شماره ۱. مشخصات دموگرافیک مصاحبه‌شوندگان

کد شرکت‌کننده	جنسیت	سابقه کار	مرتبه علمی	کد شرکت‌کننده	جنسیت	سابقه کار	مرتبه علمی
۱	مرد	۲۶	استاد	۱۱	مرد	۲۸	استاد
۲	مرد	۲۷	استاد	۱۲	زن	۱۲	دانشیار
۳	زن	۱۲	دانشیار	۱۳	مرد	۲۶	استاد
۴	مرد	۱۴	دانشیار	۱۴	زن	۱۱	دانشیار
۵	زن	۱۵	دانشیار	۱۵	مرد	۹	دانشیار
۶	مرد	۲۶	استاد	۱۶	مرد	۱۰	دانشیار
۷	مرد	۲۳	استاد	۱۷	زن	۱۱	دانشیار
۸	مرد	۱۳	دانشیار	۱۸	مرد	۲۵	استاد
۹	زن	۱۴	دانشیار	۱۹	مرد	۱۲	دانشیار
۱۰	مرد	۱۰	دانشیار	۲۰	مرد	۱۴	دانشیار

جامعه آماری در بخش کمی پژوهش شامل تمامی دانشجویان سال سوم و چهارم دوره کارشناسی دانشگاه اصفهان در سال تحصیلی ۹۹ - ۱۳۹۸ می‌باشد. روش نمونه‌گیری در بخش کمی خوشه‌ای چندمرحله‌ای و بر اساس جدول نمونه‌گیری کرجسی و مورگان (۱۹۷۰) بوده است؛ بدین صورت که از بین دانشکده‌های حوزه علوم انسانی، دانشکده‌های علوم تربیتی و روانشناسی، علوم جغرافیا و برنامه‌ریزی و علوم اداری و اقتصاد انتخاب شدند. سپس، در هر دانشکده، دو گروه که دارای رشته کارشناسی^۱ بودند، انتخاب و سپس دو درس از درس‌های تخصصی^۲ هر رشته انتخاب و پرسشنامه‌ها در بین دانشجویان آنها به صورت سرشماری توزیع شده است. به طور کلی ۳۶۰ نفر از دانشجویان به عنوان نمونه انتخاب شدند. از این تعداد ۲۹۶ نفر معادل ۸۲/۲ درصد دانشجویان زن و ۶۴ نفر معادل ۱۷/۸ درصد دانشجویان مرد بودند. همچنین ۳۳/۳۳ درصد دانشجویان دانشکده علوم تربیتی و روانشناسی، ۳۳/۳۳ درصد دانشجویان دانشکده علوم جغرافیا و برنامه‌ریزی و ۳۳/۳۳ درصد دانشجویان دانشکده علوم اداری و اقتصاد بودند.

برای گردآوری داده‌ها در بخش کیفی از مصاحبه نیمه‌ساختاریافته استفاده شد. با اجازه و رضایت مشارکت‌کنندگان و با اطمینان از محرمانه ماندن و همچنین بالا بردن اعتبار داده‌های مصاحبه، مصاحبه‌ها با بهره‌گیری از ابزارهای دیجیتال ضبط شد. کوتاه‌ترین مصاحبه ۴۳ دقیقه و طولانی‌ترین آن ۵۶ دقیقه به طول انجامید. متوسط زمان مصاحبه‌ها ۴۹ دقیقه بوده است. مصاحبه فردی توسط پژوهشگر در شرایطی غیررسمی انجام شد. پژوهشگران برای انجام مصاحبه‌ها، نامه‌ای را با ذکر اهداف پژوهش، نقش مصاحبه‌شونده در انجام پژوهش و سؤالات مصاحبه تنظیم و به صورت حضوری و پست الکترونیک تقدیم مصاحبه‌شوندگان نمودند. پژوهشگران با کوشش زیاد سعی نمودند رضایت مصاحبه‌شونده را جلب کنند. زمان و مکان مصاحبه توسط مصاحبه‌شوندگان تنظیم شد. ارسال این اطلاعات باعث شد تا مصاحبه‌شوندگان به اهمیت پژوهش پی ببرند و با آمادگی قبلی به سؤال‌ها پاسخ دهند.

برای تجزیه و تحلیل داده‌های کیفی از روش‌های ساختاری^۳ و تفسیری^۴ استفاده شد. بدین منظور، پژوهشگر برای تجزیه و تحلیل داده‌های حاصل از مصاحبه، پس از گردآوری آنها، چندین بار به دقت به متن آنها گوش سپرد و سپس آنها را کلمه به کلمه به صورت مکتوب درآورد و سطر به سطر مورد بررسی قرار داد و چندین بار به طور دقیق خوانده شدند تا تحلیل‌گر اطلاعات کلی نسبت به

۱. رشته‌های علوم تربیتی و روانشناسی از دانشکده علوم تربیتی و روانشناسی، رشته‌های برنامه‌ریزی روستایی و برنامه‌ریزی شهری از دانشکده علوم جغرافیا و برنامه‌ریزی و رشته‌های حقوق و مدیریت دولتی از دانشکده علوم اداری و اقتصاد انتخاب شدند.

۲. دروس تخصصی بر مبنای اینکه واقعاً از دروس اساسی رشته باشند و با مشورت با استادان و متخصصان هر رشته و گروه انتخاب گردیدند که عناوین آن در قسمت یافته‌ها آمده است.

۳. در تحلیل ساختاری متن پیاده شده، واژه‌ها، مفاهیم و ارتباط میان آنها بر حسب میزان تکرار، تعداد واژه‌ها، الفاظ، کنایه‌ها و اصطلاحات به کار رفته در جمله‌ها و میزان تکرارشان شمارش و بررسی می‌گردد تا الگوهای موجود در گفته‌ها کشف شود (کریمی و نصر، ۱۳۹۱)

۴. در این روش تحلیل، پژوهشگر درصدد است تا حد امکان، به آشکار نمودن پیام‌های نهفته در متن نوشتاری مصاحبه بپردازد. طبق توصیه گیلهام، برای دستیابی به این هدف، متن مصاحبه باید به طور کامل و با شرح جزئیات نوشته شود (گیلهام، ۲۰۰۰، ص. ۶۳).

داده‌ها کسب کند. سرانجام این مراحل دنبال شد: (۱) ابتدا به هر مصاحبه‌شونده یک کد اختصاص داده شد؛ (۲) در زیر هر سؤال پژوهش، جمله‌هایی با معنای مشابه در تمام مصاحبه‌ها کنار یکدیگر قرار گرفتند و مطابق با مفهوم مستتر در آنها برایشان نام و کد در نظر گرفته شد، به عبارتی دیگر، جملات معنادار مشخص و به صورت کد تعریف شدند. سپس کدهای استخراج شده دسته‌بندی شده و برای هر دسته عنوان مناسبی انتخاب شد؛ (۳) با مرور مجدد، کدهای مشابه در کنار یکدیگر قرار گرفتند و مقوله‌هایی بزرگ‌تر تشکیل شد. این کار روی تمام سؤال‌های مصاحبه انجام شد- و ادغام و دسته‌بندی مقوله‌ها تا رسیدن به یک نظام مقوله‌بندی مناسب ادامه یافت و سپس، بر اساس مشابهت‌ها، مقوله‌های اصلی و فرعی استخراج و الگوی کلی ارائه شد، به عبارت دیگر، در جریان تحلیل داده‌ها، مقایسه بین داده‌ها صورت گرفته و بخش‌هایی که دارای محتوای مشابه بودند ادغام یا بازنگری شده و سعی شد مقوله‌های کلی استخراج شوند (کرسول، ۲۰۱۱، ص. ۴). این مقوله‌ها را متخصصان و استادان بازنگری نمودند و اصلاحات لازم در مقوله‌های اصلی و فرعی صورت گرفت. برای افزایش دقت در تحلیل داده‌ها در بخش کیفی از نرم‌افزار مکس کیودا^۱ استفاده شده است. برای اعتباریابی کیفی از تکنیک قابل قبول و معتبر بودن استفاده شد. برای رسیدن به این باور، از روش همسوسازی استفاده شد و سعی گردید با گردآوری داده‌های کافی از منابع چندگانه، این باورپذیری را ایجاد نمود. به علاوه، از تکنیک کنترل توسط اعضا از طریق ارائه نتایج تحلیل داده‌ها به مشارکت‌کنندگان برای چک کردن و بررسی نتایج نیز استفاده گردید. برای افزایش پایایی پژوهش، مصاحبه‌ها با یک برنامه قبلی در یک فضای مناسب و رعایت شرایط مصاحبه با راهنمایی‌های لازم و به دور از سوگیری و اعمال نظر شخصی و با استفاده از دستگاه ضبط صوت انجام می‌گرفت. همزمان با گردآوری داده‌ها، تجزیه و تحلیل آنها با دو هدف بازخورد برای مصاحبه‌های فردی و اطمینان از اشباع داده‌ها آغاز شد. افزون بر این، از روش مسیر ممیزی برای تأیید صحت و درستی داده‌ها استفاده شد. در مسیر ممیزی، افراد از خارج از پژوهش که به مطالعات کیفی آشنا هستند به‌عنوان چک‌کننده و ناظر عمل می‌نمایند. در پژوهش حاضر از یک ناظر خارجی باتجربه در تحقیق کیفی استفاده شد.

برای گردآوری داده‌ها در بخش کمی از پرسشنامه محقق ساخته مبتنی بر شاخص‌های استخراج شده از بخش کیفی و متون تخصصی استفاده شده است. پرسشنامه ارزشیابی شامل ۱۰ گویه است. هر کدام از گویه‌ها بر اساس طیف پنج ارزشی لیکرت از خیلی کم (۱) تا خیلی زیاد (۵) نمره‌گذاری شدند. به‌منظور تعیین روایی پرسشنامه هشت نفر از اعضای هیئت علمی و متخصصان تعلیم و تربیت در مورد آن اظهارنظر کردند و اصلاحات لازم اعمال شد. برای برآورد پایایی و همبستگی درونی سؤالات پرسشنامه از آلفای کرونباخ استفاده شد. پس از اجرای آزمایشی پرسشنامه بین ۳۰ نفر از افراد نمونه، پایایی آن ۰/۹۰۶ برآورد گردید. برای تجزیه و تحلیل ویژگی‌های جمعیت‌شناختی گروه‌های مورد مطالعه از آمار توصیفی و به‌منظور پاسخگویی به پرسش‌های تحقیق، از روش‌های آماری استنباطی (آزمون t تک‌گروهی و آزمون مانوا) با رعایت پیش‌فرض‌ها استفاده شده است. محرمانه بودن اطلاعات شخصی مصاحبه‌شوندگان، مختار بودن برای شرکت در مصاحبه و نتایج حاصله از مصاحبه‌ها از رعایت نکات اخلاقی در هر پژوهش است که به مصاحبه‌شونده‌های پژوهش حاضر در خصوص رعایت این موارد اطمینان خاطر داده شد.

یافته‌های پژوهش

یافته‌های حاصل از مصاحبه‌ها بر اساس نمونه متن مصاحبه و پرسشنامه‌های توزیع شده در بین نمونه پژوهش با توجه به سؤال‌های پژوهش به تفکیک ارائه می‌گردند:

سؤال اول: شاخص‌های مطلوب برای ارزشیابی عملکرد دانشجویان در دروس رشته‌های علوم انسانی دانشگاه‌ها چه مواردی است؟

با توجه به نظر صاحب‌نظران و افراد اهل نظر، ده مقوله به شرح زیر تدوین شده است:

۱. تناسب تکالیف ارزشیابی با اهداف درس: ۸۵ درصد از مصاحبه‌شوندگان بر این مؤلفه تأکید داشتند و در این راستا: پاسخگوی کد ۹:

«تکالیف ارزشیابی از طراحی محتوای درس، سؤالات میان‌ترم و پایان‌ترم درس باید دقیق و هدف‌دار باشد. اگر تکالیف ارزشیابی بدون هدف و بی‌برنامه باشد هیچ مزایا و سودی برای دانشجو و استاد ندارد؛ لذا استادان باید به‌گونه‌ای تکالیف ارزشیابی خود را طراحی کنند که هم برای استاد و هم برای دانشجو مفید باشد. هدف از تکالیف ارزشیابی در جهت رشد و ارتقای یادگیری باشد، نه رفع تکلیف».

مصاحبه‌شونده شماره ۱۰ بیان داشت:

«تعدادی از دانشجویان اظهار داشتند که برخی از استادان تکالیفی را بر عهده ما می‌گذارند که هیچ سنخیت و سازگاری با اهداف درس ندارد. به‌طور مثال، بخشی از نمره پایانی شما، تهیه مقاله و ارسال آن به مجلات یا ترجمه کردن مقالات یا کتب انگلیسی است. در حالی که موضوع مقاله یا کتب با اهداف درس تناسبی ندارد».

۲. رعایت اصول علمی و حرفه‌ای در هنگام طراحی سؤال و برگزاری آزمون: تعداد ۱۷ نفر از مصاحبه‌شوندگان (۸۵ درصد) صریحاً به این شاخص اشاره و بر آن تأکید داشتند.

پاسخگوی کد ۲ اظهار داشت:

«ارزشیابی باید از دقت و صحت بالایی برخوردار باشد. برای نمونه، من ناظر یکی از امتحانات بودم. مهمان بودم و استاد بیست تا سؤال تستی داده بود بیست نمره. در یک ساعت و این بوک (Open Book) بود. این ارزشیابی از نظر من ارزشیابی موثق و معتبری نیست؛ یعنی درجه اعتبارش برای ما مهم هست. افزون بر این، مواقعی برخی از دانشجویان تعریف می‌کنند که فلان استاد سؤالات ارزشیابی درس خود را به دانشجویان مقاطع بالاتر می‌دهد تا آنها را طراحی کند. هم کلاسی‌های من هم‌اتاق آنهاست و سؤالات را به هم کلاسی ما می‌دهد و با اینکه در درس مشارکت فعالی نداشت، ولی بهترین نمره را گرفت».

همچنین به عقیده مصاحبه‌شونده کد ۳:

«استادان در هنگام ارزشیابی باید اصول طراحی سؤالات از جمله طراحی سؤالات از تمامی صفحات درس، درصد دشواری سؤالات، توجه به سطوح یادگیری به‌خصوص در بخش حیطه شناختی، شرایط عمومی دانشجویان، مستقل بودن هر سؤال و سایر اصول را رعایت نمایند. اگر اصول صحیح رعایت نشود، پیشرفت تحصیلی اعتبار نخواهد داشت».

۳. دادن بازخورد به تکالیف دانشجویان و راهنمایی آنها: از دیدگاه تمامی مصاحبه‌شوندگان (۲۰ نفر) یکی از شاخص‌هایی که باید

در ارزشیابی مورد توجه قرار گیرد دادن بازخورد به تکالیف دانشجویان و راهنمایی آنهاست. بنابر اظهارات مصاحبه‌شونده کد ۱۷:

«وقتی می‌خواهم از دانشجویان کوئیز بگیرم همان جلسه یا جلسه بعد بازخورد آن را ارائه می‌دهم. دانشجویانی که مشکل داشته باشند آنها را راهنمایی می‌کنم. در واقع همکاران بعد از اجرای ارزشیابی باید به دانشجویان خود بگویند چه عملکردی داشتند. کجای کارشون ضعیف بوده و نقاط قوتشون چی بوده است. وقتی نقاط قوت و ضعف آنان گفته شود دانشجویان خواهند فهمید که در چه زمینه‌هایی نیاز به تقویت دارند و کجای کار باید بهبود یابد».

مصاحبه‌شونده شماره ۱ اذعان داشت:

«بازخورد می‌تواند این آگاهی را به دانشجویان دهد که از پیشرفتشان در درس با خبر شوند و بتوانند نقاطی که نیاز به بازنگری و بهبود دارند آنها را مشخص و رفع کنند. افزون بر این دیدگاه و نظر دانشجویان را نسبت به موضوع و درس بازتر خواهد کرد. بیشتر درس‌هایی را که تدریس می‌کنم از دانشجویان کارهای پژوهشی می‌خواهم. تمام جلسات که کارهای آنها را ارزشیابی می‌کنم با بازخوردهایی که به آنها می‌دهم آنها را راهنمایی می‌کنم».

۴. متنوع بودن روش‌های ارزشیابی: در این راستا تمامی مصاحبه‌شونده‌ها بر این مؤلفه تأکید داشتند. به‌زعم مصاحبه‌شونده‌ی کد ۴:

«در حال حاضر بسیاری از همکاران فقط از دانشجویان می‌خواهند که در پایان ترم مقاله‌ای بنویسند و به مجلات ارسال کنند و پذیرش بگیرند و نمره بیست برای آنها در نظر بگیرند. در حالی که این تنها روش مفیدی برای ارزشیابی درس نیست و به‌طور بدیهی روش منطقی نیست. استادان باید از روش‌های ارزشیابی متنوعی برای بررسی میزان آموخته‌های دانشجویان استفاده کنند. باید به دانشجویان مسئله داده شود، سؤال داده شود. فعالیتی تحویل داده شود که عیب‌یابی کنند».

همچنین بنابر گفته‌های مشارکت‌کننده‌ی کد ۲۰:

«استادان جهت ارزشیابی باید از روش‌های متنوعی از جمله تکالیف، نوشتن مقاله، شرکت در کلاس، ارائه گزارش‌ها، آزمون‌ها و امتحان نهایی استفاده کند تا دانشجویان بتوانند در همه جوانب پیشرفت نمایند. در طول یک ترم نباید همه چیز به یک برگه امتحان ختم شود. هر قدر هم خوب طراحی شده باشد، یعنی فرض کنیم ترکیبی از انواع سؤالات باشد، یک نوع ارزشیابی، آن هم در یک مقطع خاص کفایت نمی‌کند. این ضعف استاد هست، چرا؟ چون دانشجو ممکنه به هر دلیلی در آن مقطع خاص مشکلی براش پیش بیاد».

افزون بر این، به‌زعم پاسخ‌گوی کد ۱۱:

«استادان می‌توانند از پوشه کار الکترونیکی که با بهره‌گیری از فناوری‌های روز فراهم می‌شود به‌عنوان یکی از روش‌های ارزشیابی استفاده کنند؛ چراکه این روش در کنار توجه به فرایند و فرآورده یادگیری، امکان ارائه بازخوردهای مستمر و استفاده از تکالیف و فعالیت‌های مناسب و کیفی را فراهم می‌کند».

۵. تناسب و هماهنگی ارزشیابی با محتوای درس: تعداد ۱۷ نفر از مصاحبه‌شوندگان (۸۵ درصد) صریحاً به این شاخص اشاره و بر آن تأکید داشتند؛ لذا به‌زعم مصاحبه‌شونده‌ی کد ۵:

«برخی اوقات تعدادی از دانشجویان پیش من می‌آیند و از دیگر استادان گلابه و شکایت دارند. برای نمونه، یکی از دانشجویان می‌گفت با اینکه استادی تعداد زیادی منابع معرفی کرده بود و آنها را مطالعه کردیم ولی متأسفانه سؤالات از منابع معرفی شده نبودند؛ یعنی هیچ ربطی به محتوای ارائه شده نداشت. به‌گونه‌ای ما را ارزشیابی کرده که هنوز برای ما گنگ هست که این سؤالات کجا بودند. احساس می‌کنیم که سؤالات درس دیگری را به ما جهت پاسخ دادن به آنها داده‌اند و شاید سؤالات مقطع دیگری باشد. سؤالاتی که دانشجویان با آن هیچ‌گونه آشنایی و شناختی نداشته باشند بعضی مواقع منجر به فشار روانی، اضطراب و افت نمرات دانشجویان می‌شود».

همچنین پاسخگوی کد ۶ اظهار داشت:

«ارزشیابی باید مبتنی بر محتوای ارائه‌شده در طول ترم تحصیلی باشد. بدین‌گونه که ارزشیابی استاد باید مبتنی بر کاری باشد که در کلاس درس در یک ترم انجام داده است. آنچه که در محیط آموزشی انجام و اهدافی که زمینه یادگیریشون فراهم شده، دقیقاً باید مورد ارزشیابی قرار بگیرند».

۶. آگاه ساختن دانشجویان از معیارها و روش‌های ارزشیابی: تعداد ۱۸ نفر از مصاحبه‌شوندگان (۹۰ درصد) صریحاً به این شاخص اشاره و بر آن تأکید داشتند. در همین راستا، مشارکت‌کننده‌ی کد ۸ بیان داشت:

«قبل از آنکه تدریس را در جلسه اول شروع کنم تمامی روش‌های ارزشیابی مانند ارائه پژوهش، فعال بودن، آزمون‌های کوتاه پاسخ، حضور و غیاب، ارزشیابی پایانی و معیارها را به اطلاع آنها می‌رسانم. همچنین به اطلاع دانشجویان می‌رسانم که هر یک از آنها چند نمره دارد. اگر این چهارچوب، معیارها و روش‌ها مشخص نباشد ضعف استاد هست. استادان باید انتظارات خود را بر اساس معیارها تعریف کنند. بدین صورت که بارم‌بندهای حضور در کلاس، پروژه آخر کلاس، ارزشیابی پایان‌ترم و میان‌ترم را برای دانشجویان مشخص کرده و آنها را در قالب جدول به آگاهی دانشجویان برساند».

همچنین به عقیده مصاحبه‌شونده شماره ۱۵:

«معیارهای ارزشیابی از درس باید از پیش تعیین شده باشند و قضاوت در مورد کیفیت یادگیری دانشجویان باید بر اساس معیارها و شاخص‌های صریح تعیین شده باشد، نه بر اساس پیشرفت دانشجویان دیگر و یا استناد به نمرات دیگر دانشجویان».

۷. رعایت عدالت و انصاف در طراحی سؤال و نمره دادن: منظور از رعایت عدالت و انصاف این است که ارزشیابی باید در مورد تمام گروه‌های دانشجویان بدون در نظر گرفتن سن، جنس، زمینه‌های اجتماعی و موقعیت جغرافیایی یکسان باشد. تعداد ۱۹ نفر از مصاحبه‌شوندگان بر این شاخص تأکید داشتند. برای نمونه، مصاحبه‌شونده کد ۱۱ اظهار داشت:

«استادان باید در طراحی سؤال و تصحیح کردن امتحانات، عدالت و انصاف را رعایت کند. مثلاً در طراحی سؤالات، از سؤال‌های دشوار، متوسط و آسان استفاده کند یا اینکه تمامی سؤالات را از یک نوع سؤال طراحی نکند؛ مثلاً بیست سؤال تستی و بیست نمره. محتوا و قالب ارزشیابی باید به گونه‌ای طراحی شده باشد که هیچ یک از دانشجویان احساس بی‌عدالتی نکنند».

همچنین بنا بر اظهارات مصاحبه‌شونده کد ۱۲:

«استادان باید در هنگام تصحیح اوراق امتحانی یا آزمون عملی عدالت و انصاف را رعایت نمایند. به طور مثال، دانشجویانی که در کلاس نسبت به سایر دوستان خود فعالیت‌های بیشتری را انجام می‌داند یا فعال بودند باید مورد توجه قرار بگیرند».

۸. توجه به ارزشیابی‌های مستمر و تکوینی: پاسخگوی کد ۷ اظهار داشت:

«توجه به ارزشیابی‌های مستمر منجر به این می‌شود که دانشجویان محتوای جدید را مطالعه کنند و ایده‌های جدید در ذهن آنها تراوش کند و به درس ارزش بیشتری بدهند. استادان باید مکرراً ارزشیابی تکوینی را در کلاس خود اجرا کنند تا درک کنند پیشرفت یادگیری دانشجویان در درس به چه میزان بوده است».

همچنین به‌زعم مصاحبه‌شونده کد ۱۹:

«انجام ارزشیابی‌های مستمر و تکوینی منجر به شناسایی نقاط قوت و ضعف دانشجویان، ایجاد انگیزه برای تداوم یادگیری، تقویت اعتماد به نفس، پرورش روحیه تفکر، تلاش، خلاقیت، توجه به جنبه‌های مهم درس، مشخص کردن نتایج یادگیری، کمک به پیشرفت تحصیلی، توسعه مهارت‌ها، مطلع کردن دانشجویان از عملکرد خویش، شناخت توانایی‌ها و زمینه‌های علمی، بهبود و اصلاح فعالیت‌های آموزشی شود».

۹. ارزشیابی از دانش، نگرش و مهارت‌های دانشجویان در خصوص درس: مقصود از جامعیت در ارزشیابی، توجه همزمان به ابعاد دانش، مهارت و نگرش است. تعداد ۱۴ نفر از مصاحبه‌شوندگان (۷۰ درصد) صریحاً به این شاخص اشاره و بر آن تأکید داشتند. در همین راستا، مصاحبه‌شونده کد ۱۸ اظهار داشت:

«برای اینکه ارزشیابی جامع داشته باشیم هر سه بعد دانش، نگرش و مهارت را مورد ارزشیابی قرار دهیم. اگر نتوانیم یکی از ابعاد را ارزشیابی کنیم، به فرایند ارزشیابی لطمه زده‌ایم و نمی‌توانیم ارزشیابی مؤثری داشته باشیم. ارزشیابی از دانش و مطالبی که دانشجویان در حافظه خود جهت امتحان پایان ترم نگهداری می‌کنند نمی‌تواند به‌تنهایی ارزشیابی تأثیرگذاری برای دانشجویان و استادان باشد».

همچنین بنابر گفته‌های مصاحبه‌شونده شماره ۱۳:

«در برخی از رشته‌ها بیشتر دروس عملی هستند یعنی واحدهای عملی بیشتری دارند. چنین دروسی را نمی‌توانیم فقط به حیطة دانش آن توجه کنیم؛ چراکه در این دروس حیطة مهارت بسیار مهم بوده و استادان باید آزمون عملی از دانشجویان به عمل آورند».

۱۰. فعالیت‌محور بودن ارزشیابی: ۶۵ درصد از مصاحبه‌شوندگان بر این مقوله تأکید داشتند و در این رابطه مصاحبه‌شونده کد ۱۶ اظهار داشت:

«یکی از تکنیک‌های که همکاران باید در ارزشیابی مد نظر قرار دهند فعالیت محور بودن ارزشیابی است؛ بدین گونه که استادان باید به فعالیت‌های دانشجویان ارزش و بهای زیادی بدهند نه به آن محتوایی که حفظ کرده و به سرعت آنها را فراموش می‌کنند. بخش

خیلی کمی از نمره درس باید به ارزشیابی پایانی اختصاص داده شود. ارزشیابی باید در قالب فعالیت‌های متنوع از جمله پژوهش، سمینار و تکالیف هفتگی صورت گیرد. توجه به فعالیت محور بودن ارزشیابی منجر می‌شود که توجه به خلاقیت و نوآوری‌های دانشجویان در فرایند ارزشیابی دوچندان شود».

مصاحبه‌شونده کد ۱۴ اظهار داشت:

«ارزشیابی که توسط همکاران انجام می‌شود باید فعالیت محور باشد؛ بدین گونه که کار تیمی، فرایند انجام فعالیت‌ها، مشارکت و تعامل با دیگر دوستان و زمان انجام را در نظر گرفت».

سؤال دوم: میزان توجه به شاخص‌های مطلوب ارزشیابی عملکرد دانشجویان در دروس رشته‌های علوم انسانی در دانشگاه مورد مطالعه چه میزان است؟

نظرات دانشجویان در مورد میزان توجه به شاخص‌های مطلوب ارزشیابی عملکرد دانشجویان در دروس رشته‌های علوم انسانی در دانشگاه مورد مطالعه در جداول ۱ تا ۶ ارائه گردیده است. عنوان درس‌ها در برخی از جداول به صورت کد ارائه شده است. روش‌های تدریس (کد ۱)، درس روش تحقیق (کد ۲)، نظریه‌های مشاوره (کد ۳)، اصول راهنمایی و مشاوره (کد ۴)، برنامه‌ریزی شهری (کد ۵)، مدیریت شهری (کد ۶)، برنامه‌ریزی روستایی (کد ۷)، مدیریت روستایی (کد ۸)، حقوق اساسی (کد ۹)، روش تحقیق در حقوق (کد ۱۰)، مدیریت توسعه (کد ۱۱) و مدیریت تحول سازمانی (کد ۱۲).

در ادامه، نتایج توصیفی نظرات دانشجویان نسبت به شاخص‌های ارزشیابی در هر یک از دروس در جدول شماره ۱ ارائه گردیده است.

جدول شماره ۱. جدول توصیفی شاخص‌های ارزشیابی بر حسب دروس مختلف^۱

ردیف	شاخص‌های ارزشیابی	کد ۱	کد ۲	کد ۳	کد ۴	کد ۵	کد ۶	کد ۷	کد ۸	کد ۹	کد ۱۰	کد ۱۱	کد ۱۲
۱	تناسب تکالیف ارزشیابی با اهداف درس	۳.۴۰	۲.۸۳	۲.۵۳	۲.۴۰	۲.۷۷	۲.۳۷	۲.۹۰	۲.۳۳	۳.۰۳	۳.۳۰	۳.۱۷	۲.۸۳
	رعایت اصول علمی و حرفه‌ای در هنگام طراحی سؤال و برگزاری آزمون	۳.۲۳	۲.۸۳	۲.۲۷	۲.۲۳	۲.۶۰	۲.۳۷	۲.۷۰	۲.۰۷	۳.۱۳	۳.۰۰	۳.۴۷	۲.۷۷
	دادن بازخورد به تکالیف دانشجویان و راهنمایی آنها	۰.۹۳	۱.۰۵	۰.۹۷	۰.۹۳	۰.۹۷	۰.۹۹	۱.۰۶	۱.۱۲	۱.۱۲	۰.۹۹	۰.۸۷	۱.۱۱
۲	متنوع بودن روش‌های ارزشیابی	۳.۵۷	۲.۷۷	۲.۱۰	۲.۳۰	۲.۴۷	۲.۳۷	۲.۳۷	۲.۱۰	۲.۷۳	۲.۹۷	۳.۱۷	۲.۹۰
	تناسب و هماهنگی ارزشیابی با محتوای درس	۳.۴۷	۲.۷۳	۲.۲۷	۲.۳۳	۲.۵۷	۲.۳۰	۲.۸۷	۲.۳۰	۲.۶۳	۳.۱۳	۳.۳۷	۲.۶۳
	آگاه ساختن دانشجویان از معیارها و روش‌های ارزشیابی در آغاز ترم	۰.۸۱	۰.۹۷	۱.۰۰	۰.۷۲	۱.۰۶	۰.۸۸	۱.۲۷	۱.۰۷	۱.۳۱	۱.۱۷	۱.۲۰	۱.۲۰
۳	رعایت انصاف در طراحی سؤال و نمره دادن	۳.۴۷	۲.۷۳	۲.۶۳	۲.۱۰	۲.۳۳	۲.۲۳	۲.۹۰	۲.۵۰	۲.۷۷	۳.۰۳	۳.۲۷	۲.۶۰
	رعایت انصاف در طراحی سؤال و نمره دادن	۰.۸۶	۱.۰۱	۰.۹۴	۰.۸۴	۱.۰۷	۰.۸۷	۱.۲۲	۱.۲۶	۱.۱۲	۱.۱۳	۱.۱۵	۱.۰۶
	آگاه ساختن دانشجویان از معیارها و روش‌های ارزشیابی در آغاز ترم	۳.۳۳	۲.۷۳	۲.۸۷	۲.۶۰	۲.۶۰	۲.۰۳	۲.۸۰	۲.۳۳	۲.۷۷	۳.۳۰	۳.۴۳	۲.۹۳
۴	رعایت انصاف در طراحی سؤال و نمره دادن	۳.۴۷	۲.۷۳	۲.۶۳	۲.۱۰	۲.۳۳	۲.۲۳	۲.۹۰	۲.۵۰	۲.۷۷	۳.۰۳	۳.۲۷	۲.۶۰
	رعایت انصاف در طراحی سؤال و نمره دادن	۰.۸۶	۱.۱۵	۱.۱۰	۰.۹۵	۱.۲۱	۰.۸۵	۱.۲۶	۱.۳۳	۱.۲۷	۱.۲۴	۱.۰۴	۱.۱۳
	آگاه ساختن دانشجویان از معیارها و روش‌های ارزشیابی در آغاز ترم	۳.۳۳	۲.۵۳	۲.۵۷	۲.۳۰	۲.۶۷	۲.۱۷	۲.۵۰	۲.۱۶	۲.۹۰	۲.۸۳	۲.۲۳	۲.۸۲

۱. در جدول، سطر اول مربوط به میانگین و سطر دوم مربوط به انحراف معیار می‌باشد.

۰.۹۴	۱.۱۹	۰.۹۸	۱.۲۹	۱.۱۴	۱.۳۰	۱.۰۱	۱.۲۶	۰.۹۱	۰.۸۹	۱.۱۰	۰.۹۵	توجه به ارزشیابی‌های مستمر و تکوینی	۸
۲.۸۳	۳.۰۳	۲.۹۷	۲.۹۰	۲.۲۳	۲.۴۳	۲.۳۷	۲.۵۳	۲.۴۰	۲.۵۰	۲.۶۰	۳.۴۷	ارزشیابی از دانش، نگرش و مهارت‌های	۹
۰.۸۷	۱.۲۹	۱.۲۱	۱.۲۵	۱.۱۹	۱.۰۷	۱.۰۹	۱.۰۴	۰.۸۹	۰.۹۳	۱.۲۲	۰.۸۱	دانشجویان در خصوص درس	۱۰
۲.۷۳	۲.۸۰	۳.۲۰	۲.۸۰	۲.۳۷	۲.۴۳	۲.۱۷	۲.۶۳	۲.۱۰	۲.۲۰	۲.۸۰	۳.۴۰	فعالیت‌محور بودن ارزشیابی	
۱.۰۴	۱.۳۹	۱.۲۱	۱.۲۴	۱.۰۹	۱.۲۵	۰.۹۸	۰.۹۳	۰.۹۲	۰.۹۶	۱.۲۱	۰.۸۹		
۲.۷۹	۳.۲۲	۳.۰۸	۲.۸۵	۲.۲۹	۲.۶۶	۲.۲۷	۲.۶۰	۲.۳۲	۲.۴۴	۲.۷۲	۳.۴۳		
۰.۷۰	۰.۹۳	۰.۷۷	۰.۹۲	۰.۹۶	۰.۹۱	۰.۶۹	۰.۸۶	۰.۴۸	۰.۶۱	۰.۷۰	۰.۷۳	میانگین	

داده‌های مندرج در جدول شماره ۱ نشان می‌دهند که به‌طور کلی میانگین شاخص‌های ارزشیابی در دروس روش تدریس، مدیریت توسعه و روش تحقیق در حقوق بالاتر از نمره معیار و در سایر دروس پایین‌تر از نمره ملاک (۳) می‌باشد. افزون بر این شاخص‌های ردیف‌های ۱، ۳، ۴، ۵ و ۷ تا ۱۰ بیشترین میانگین را در روش‌های تدریس، شاخص‌های ردیف‌های ۱ و ۶ بیشترین میانگین را در مدیریت توسعه دارا می‌باشد. همچنین شاخص‌های ردیف‌های ۱، ۲، ۸ و ۹ کمترین میانگین را در مدیریت روستایی، شاخص‌های ردیف‌های ۳، ۷ و ۱۰ کمترین میانگین را در اصول راهنمایی و مشاوره و شاخص ردیف ۶ کمترین میانگین را در مدیریت شهری و شاخص ردیف ۵ کمترین میانگین را در دروس نظریه‌های مشاوره و شاخص ردیف ۴ کمترین میانگین را در دو درس نظریه‌های مشاوره و مدیریت روستایی به‌طور مشترک دارا هستند.

سؤال سوم: میزان توجه به شاخص‌های مطلوب ارزشیابی عملکرد دانشجویان در رشته‌های علوم انسانی در دانشگاه مورد مطالعه چه میزان است؟

نظرات دانشجویان نسبت به شاخص‌های مطلوب ارزشیابی عملکرد دانشجویان در رشته‌های مختلف در جدول شماره ۲ ارائه گردیده است.

جدول شماره ۲. جدول توصیفی شاخص‌های ارزشیابی بر حسب رشته‌های مختلف

ردیف	شاخص‌های ارزشیابی	علوم تربیتی	روان‌شناسی	برنامه‌ریزی شهری	برنامه‌ریزی روستایی	حقوق دولتی	مدیریت دولتی
		M ¹ (SD) ²	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
۱	تناسب تکالیف ارزشیابی با اهداف درس	۳.۱۲ (۱.۰۲)	۲.۴۷ (۰.۹۴)	۰.۹۹ (۰.۹۹)	۲.۶۲ (۱.۱۲)	۱.۰۴ (۱.۰۴)	۳.۰۰ (۱.۰۴)
۲	رعایت اصول علمی و حرفه‌ای در هنگام طراحی سؤال و برگزاری آزمون	۳.۰۳ (۱.۰۸)	۲.۲۵ (۱.۰۱)	۰.۹۸ (۰.۹۸)	۲.۳۸ (۱.۱۷)	۱.۱۶ (۱.۱۶)	۳.۱۱ (۱.۰۹)
۳	دادن بازخورد به تکالیف دانشجویان و راهنمایی آنها	۳.۱۸ (۰.۹۸)	۲.۴۷ (۰.۸۷)	۰.۹۹ (۰.۹۹)	۲.۵۳ (۱.۱۷)	۱.۲۳ (۱.۲۳)	۳.۰۵ (۱.۱۲)
۴	متنوع بودن روش‌های ارزشیابی	۳.۱۶ (۱.۱۱)	۲.۲۰ (۰.۹۸)	۰.۹۹ (۰.۹۹)	۲.۲۳ (۱.۰۹)	۱.۲۴ (۱.۲۴)	۳.۰۳ (۱.۲۲)
۵	تناسب و هماهنگی ارزشیابی با محتوای درس	۳.۱۰ (۱.۰۰)	۲.۳۰ (۰.۸۸)	۰.۹۸ (۰.۹۸)	۲.۵۸ (۱.۲۷)	۱.۱۵ (۱.۱۵)	۳.۰۰ (۱.۱۶)

1. مخفف میانگین mean

2. مخفف standard deviation (انحراف معیار): نشان می‌دهد به‌طور میانگین داده‌ها چه مقدار از مقدار متوسط فاصله دارند.

۳.۱۸ (۱.۰۴)	(۱.۱۶)	۲.۵۷ (۱.۲۵)	(۱.۰۹)	۲.۷۳ (۱.۰۵)	۳.۰۳ (۱.۱۱)	آگاه ساختن دانشجویان از معیارها و روش‌های ارزشیابی در آغاز ترم	۶
۲.۹۳ (۱.۹۳)	(۱.۲۵)	۲.۷۰ (۱.۳۰)	(۱.۰۴)	۲.۳۳ (۱.۰۵)	۳.۰۵ (۱.۰۹)	رعایت انصاف در طراحی سؤال و نمره دادن	۷
۳.۰۳ (۱.۱۸)	(۱.۱۴)	۲.۳۳ (۱.۲۳)	(۱.۱۷)	۲.۴۳ (۰.۹۱)	۲.۹۳ (۱.۱۰)	توجه به ارزشیابی‌های مستمر و تکوینی	۸
۲.۹۳ (۱.۱۰)	(۱.۲۱)	۲.۳۳ (۱.۱۲)	(۱.۰۶)	۲.۴۵ (۰.۹۰)	۳.۰۳ (۱.۱۱)	ارزشیابی از دانش، نگرش و مهارت‌های دانشجویان در خصوص درس	۹
۲.۷۷ (۱.۲۲)	(۱.۲۳)	۲.۴۰ (۱.۱۷)	(۰.۹۷)	۲.۱۵ (۰.۹۳)	۳.۱۰ (۱.۱۰)	فعالیت‌محور بودن ارزشیابی	۱۰
۳.۰۰ (۰.۸۵)	۲.۹۶ (۰.۸۶)	۲.۴۷ (۱.۹۵)	(۰.۷۹)	۲.۳۸ (۰.۵۵)	۳.۰۷ (۰.۸۰)	میانگین	
			۲.۴۳				

یافته‌ها جدول شماره ۲ نشان می‌دهد که میانگین رشته علوم تربیتی بالاتر از نمره ملاک (۳)، میانگین رشته مدیریت دولتی برابر با نمره ملاک و دیگر رشته‌ها پایین‌تر از نمره ملاک بوده است. افزون بر این شاخص‌های ردیف‌های ۲، ۶ و ۸ بیشترین میانگین را در رشته مدیریت دولتی، شاخص‌های ردیف‌های ۳ تا ۵، ۷، ۹ و ۱۰ بیشترین میانگین را در رشته علوم تربیتی را دارند. همچنین شاخص‌های ۱ تا ۵ و ۱۰ کمترین میانگین را در رشته روان‌شناسی، شاخص‌های ردیف‌های ۶ و ۷ کمترین میانگین را در رشته برنامه‌ریزی شهری و شاخص‌های ۸ و ۹ کمترین برنامه‌ریزی روستایی را دارا بودند.

سؤال چهارم: میزان توجه به شاخص‌های مطلوب ارزشیابی عملکرد دانشجویان در دانشگاه مورد مطالعه چه میزان است؟

نتایج توصیفی آزمون t تک نمونه‌ای مقایسه میانگین نمرات شاخص‌های ارزشیابی بر حسب دانشگاه با توجه به نمره ملاک (۳) در جدول شماره ۳ ارائه شده است.

جدول شماره ۳. نتایج آزمون t تک نمونه‌ای مقایسه میانگین نمرات شاخص‌های ارزشیابی

ردیف	شاخص‌های ارزشیابی	میانگین	انحراف معیار	T	Sig
۱	تناسب تکالیف ارزشیابی با اهداف درس	۲.۸۲	۱.۰۵	-۳.۱۵	۰.۰۰۲**
۲	رعایت اصول علمی و حرفه‌ای در هنگام طراحی سؤال و برگزاری آزمون	۲.۷۲	۱.۱۴	-۴.۵۶	۰.۰۰۱***
۳	دادن بازخورد به تکالیف دانشجویان و راهنمایی آنها	۲.۷۹	۱.۱۱	-۳.۵۷	۰.۰۰۱***
۴	متنوع بودن روش‌های ارزشیابی	۲.۶۵	۱.۱۸	-۵.۶۰	۰.۰۰۱***
۵	تناسب و هماهنگی ارزشیابی با محتوای درس	۲.۷۲	۱.۱۱	-۴.۸۲	۰.۰۰۱***
۶	آگاه ساختن دانشجویان از معیارها و روش‌های ارزشیابی در آغاز ترم	۲.۸۱	۱.۱۹	-۳.۰۰	۰.۰۰۱***
۷	رعایت انصاف در طراحی سؤال و نمره دادن	۲.۷۰	۱.۱۸	۴.۸۱	۰.۰۰۱***
۸	توجه به ارزشیابی‌های مستمر و تکوینی	۲.۶۷	۱.۱۴	۵.۵۱	۰.۰۰۱***
۹	ارزشیابی از دانش، نگرش و مهارت‌های دانشجویان در خصوص درس	۲.۶۹	۱.۱۲	-۵.۲۵	۰.۰۰۱***
۱۰	فعالیت‌محور بودن ارزشیابی	۲.۶۴	۱.۱۶	-۵.۹۶	۰.۰۰۱***
	میانگین	۲.۷۲	۰.۸۶	-۶.۱۵	۰.۰۰۰***

نتایج مندرج در جدول شماره ۳ نشان می‌دهد که شاخص ردیف ۱ دارای بالاترین میانگین و شاخص ردیف ۱۰ دارای کمترین میانگین بوده است. یادآور می‌شود که میانگین تمامی شاخص‌های ارزشیابی پایین‌تر از میانگین فرضی جامعه (۳) می‌باشد. همچنین شاخص ردیف ۱ در سطح ۹۹ درصد و سایر شاخص‌ها در سطح اطمینان $p < 0.001$ معنادار بوده است. نتایج نشان‌دهنده آن است که میزان توجه به شاخص‌های ارزشیابی در دانشگاه نامطلوب بوده و نیاز به توجه جدی دارد.

سؤال پنجم: آیا بین نظرات دانشجویان در مورد میزان توجه به شاخص‌های مطلوب ارزشیابی عملکرد بر حسب رشته‌های علوم انسانی در دانشگاه مورد مطالعه تفاوت معناداری وجود دارد؟

نتایج آزمون تحلیل واریانس چند متغیری مانوا برای بررسی معنی‌داری اثر متغیر رشته‌ها در میزان کاربرد شاخص‌های ارزشیابی حاکی از آن است که با توجه به میزان لامبدا ویلکز (۰/۷۵۲)، (F=۱/۸۴)، (Eta2 = ۰/۰۵۵) و سطح معناداری به دست آمده (P= ۰/۰۰۰)، تفاوت معناداری بین رشته‌ها وجود دارد. برای بررسی اختلاف معناداری بین رشته‌های مختلف از آزمون تحلیل واریانس یک‌راهه در متن مانوا استفاده گردیده شده که نتایج آن در جدول شماره ۴ قابل مشاهده است.

جدول شماره ۴. نتایج تحلیل واریانس یک‌راهه در متن مانوا

ردیف	شاخص‌های ارزشیابی	مجموع مجذورات	درجه آزادی	F	سطح معناداری
۱	تناسب تکالیف ارزشیابی با اهداف درس	۲۸.۲۵۶	۵	۵.۴۹۰	۰.۰۰۱***
۲	رعایت اصول علمی و حرفه‌ای در هنگام طراحی سؤال و برگزاری آزمون	۴۵.۹۷	۵	۷.۷۷۹	۰.۰۰۱***
۳	دادن بازخورد به تکالیف دانشجویان و راهنمایی آنها	۲۷.۰۸۹	۵	۴.۷۱۴	۰.۰۰۱***
۴	متنوع بودن روش‌های ارزشیابی	۵۳.۰۶۷	۵	۸.۵۶۲	۰.۰۰۱***
۵	تناسب و هماهنگی ارزشیابی با محتوای درس	۳۱.۶۰	۵	۵.۳۵۸	۰.۰۰۱***
۶	آگاه ساختن دانشجویان از معیارها و روش‌های ارزشیابی در آغاز ترم	۳۲.۸۵۶	۵	۴.۸۶۳	۰.۰۰۱***
۷	رعایت انصاف در طراحی سؤال و نمره دادن	۳۱.۵۰	۵	۴.۷۴۴	۰.۰۰۱***
۸	توجه به ارزشیابی‌های مستمر و تکوینی	۲۸.۴۱۴	۵	۴.۶۰۱	۰.۰۰۱***
۹	ارزشیابی از دانش، نگرش و مهارت‌های دانشجویان در خصوص درس	۲۸.۷۲۲	۵	۴.۷۹۱	۰.۰۰۱***
۱۰	فعالیت‌محور بودن ارزشیابی	۴۲.۷۴۷	۵	۶.۹۰۱	۰.۰۰۱***
	میانگین	۳۱.۶۱۳	۵	۹.۶۴۳	۰.۰۰۱***

یافته‌های حاصل از جدول شماره ۴ گویای آن است که تفاوت میانگین تمامی شاخص‌های ارزشیابی در شش گروه مذکور در سطح ۹۹ درصد معنادار بوده است.

در ادامه برای آنکه مشخص گردد در کدام رشته‌ها تفاوت معناداری وجود دارد از آزمون تعقیبی گابریل استفاده و نتایج آن در جدول شماره ۵ ارائه شده است.

جدول شماره ۵. آزمون تعقیبی گابریل در شاخص‌های ارزشیابی با توجه به رشته تحصیلی

ردیف	شاخص‌های ارزشیابی	تفاوت میانگین گروه‌ها و سطح معناداری
۱	تناسب تکالیف ارزشیابی با اهداف درس	علوم تربیتی _ روان‌شناسی (۰.۶۵*); روان‌شناسی _ حقوق (۰.۷* -)
۲	رعایت اصول علمی و حرفه‌ای در هنگام طراحی سؤال و برگزاری آزمون	علوم تربیتی _ روان‌شناسی (۰.۷۸**); روان‌شناسی _ حقوق (۰.۸۲** -); مدیریت دولتی (۰.۸۷** -); برنامه‌ریزی روستایی _ حقوق (۰.۶۸* -); برنامه‌ریزی روستایی _ مدیریت دولتی (۰.۷۲* -)
۳	دادن بازخورد به تکالیف دانشجویان و راهنمایی آنها	علوم تربیتی _ روان‌شناسی (۰.۷۲*)
۴	متنوع بودن روش‌های ارزشیابی	علوم تربیتی _ روان‌شناسی (۰.۹۷**); علوم تربیتی _ برنامه‌ریزی شهری (۰.۷۵*); علوم تربیتی _ برنامه‌ریزی روستایی (۰.۹۳**); روان‌شناسی _ مدیریت دولتی (۰.۸۳** -); برنامه‌ریزی روستایی - مدیریت دولتی (۰.۸** -)
۵	تناسب و هماهنگی ارزشیابی با محتوای درس	علوم تربیتی _ روان‌شناسی (۰.۸**); علوم تربیتی _ برنامه‌ریزی شهری (۰.۶۸*); روان‌شناسی _ مدیریت دولتی (۰.۷۲* -)
۶	آگاه ساختن دانشجویان از معیارها و روش‌های ارزشیابی در آغاز ترم	علوم تربیتی _ برنامه‌ریزی شهری * برنامه‌ریزی شهری - حقوق (۰.۷۲* -); برنامه‌ریزی شهری _ مدیریت دولتی (۰.۸۷** -)

۷	رعایت انصاف در طراحی سؤال و نمره دادن	علوم تربیتی _ روان‌شناسی ($0/72^*$)؛ علوم تربیتی- برنامه‌ریزی شهری ($0/77^*$)
۸	توجه به ارزشیابی‌های مستمر و تکوینی	برنامه‌ریزی روستایی- مدیریت دولتی ($0/7^*$)
۹	ارزشیابی از دانش، نگرش و مهارت‌ها	علوم تربیتی- برنامه‌ریزی روستایی ($0/7^*$)
۱۰	فعالیت‌محور بودن ارزشیابی	علوم تربیتی _ روان‌شناسی ($0/95^{**}$)؛ علوم تربیتی- برنامه‌ریزی شهری ($0/7^*$)؛ علوم تربیتی- برنامه‌ریزی روستایی ($0/7^*$)؛ روان‌شناسی _ حقوق ($0/85^{**}$)
		علوم تربیتی- روان‌شناسی ($0/7^{**}$)؛ علوم تربیتی- برنامه‌ریزی شهری ($0/64^{**}$)؛ علوم تربیتی- برنامه‌ریزی روستایی ($0/6^{**}$)؛ روان‌شناسی _ حقوق ($0/59^{**}$)؛ روان‌شناسی _ مدیریت دولتی ($0/63^{**}$)؛ برنامه‌ریزی شهری _ حقوق ($0/53^*$)؛ برنامه‌ریزی شهری _ مدیریت دولتی ($0/57^*$)؛ برنامه‌ریزی روستایی- مدیریت دولتی ($0/53^*$)

میانگین

یافته‌های مندرج در جدول شماره ۵ نشان می‌دهد که بیشترین تفاوت میانگین بین رشته‌های علوم تربیتی و روان‌شناسی در شاخص ردیف ۴ و کمترین تفاوت میانگین بین رشته‌های علوم تربیتی و روان‌شناسی در شاخص ردیف ۱ وجود دارد.

سؤال ششم: آیا بین نظرات دانشجویان در مورد میزان توجه به شاخص‌های مطلوب ارزشیابی عملکرد بر حسب جنسیت در دانشگاه مورد مطالعه تفاوت معناداری وجود دارد؟

نتایج آزمون تحلیل واریانس چند متغیری مانوا برای بررسی معنی‌داری اثر متغیر جنسیت در میزان کاربرد شاخص‌های ارزشیابی حاکی از آن است که با توجه به میزان لامبدا و یلکز ($0/920$)، ($F = 2/734$)، ($Eta2 = 0/080$) و سطح معناداری به دست آمده ($P = 0/002$)، تفاوت معناداری بین جنسیت وجود دارد. برای بررسی اختلاف معناداری بین دانشجویان زن و مرد از آزمون تحلیل واریانس یک‌راهه در متن مانوا استفاده گردیده شده که نتایج آن در جدول شماره ۶ قابل مشاهده است.

جدول شماره ۶. نتایج تحلیل واریانس یک‌راهه در متن مانوا

ردیف	شاخص‌های ارزشیابی	میانگین مردها	میانگین زن‌ها	مجموع مجدورات	df	F	سطح معناداری
۱	تناسب تکالیف ارزشیابی با اهداف درس	۳۰۰۹	۲۰۷۶	۵۰۷۳۹	۱	۳۰۳۱	۰۰۰۲۲*
۲	رعایت اصول علمی و حرفه‌ای در هنگام طراحی سؤال و برگزاری آزمون	۳۰۱۶	۲۰۶۳	۱۴۰۶۶۳	۱	۱۱۰۶۷۷	۰۰۰۳۲*
۳	دادن بازخورد به تکالیف دانشجویان و راهنمایی آنها	۲۰۸۶	۲۰۷۷	۰۰۳۸۷	۱	۰۰۳۱۹	۰۰۰۱۱***
۴	متنوع بودن روش‌های ارزشیابی	۳۰۰۹	۲۰۵۵	۱۵۰۳۲۷	۱	۱۱۰۵۱۴	۰۰۰۳۱*
۵	تناسب و هماهنگی ارزشیابی با محتوای درس	۳۰۰۳	۲۰۶۵	۷۰۷۰۳	۱	۶۰۲۷۶	۰۰۰۱۷*
۶	آگاه ساختن دانشجویان از معیارها و روش‌های ارزشیابی در آغاز ترم	۳۰۲۳	۲۰۶۵	۱۳۰۹۴۵	۱	۱۰۰۰۴۱	۰۰۰۲۷*
۷	رعایت انصاف در طراحی سؤال و نمره دادن	۲۰۹۸	۲۰۶۴	۶۰۲۹۵	۱	۴۰۵۵	۰۰۰۱۳*
۸	توجه به ارزشیابی‌های مستمر و تکوینی	۳۰۰۰	۲۰۵۹	۸۰۵۰۵	۱	۶۰۶۶	۰۰۰۱۸*
۹	ارزشیابی از دانش، نگرش و مهارت‌های دانشجویان در خصوص درس	۳۰۰۰	۲۰۶۲	۷۰۵۳۴	۱	۶۰۰۵۳	۰۰۰۱۷*
۱۰	فعالیت‌محور بودن ارزشیابی	۳۰۰۸	۲۰۵۴	۱۵۰۲۰۸	۱	۱۱۰۶۸	۰۰۰۳۲*
	میانگین	۳۰۰۵	۲۰۶۵	۸۰۵۵۲	۱	۱۱۰۹۹۷	۰۰۰۳۲*

یافته‌های حاصل از جدول شماره ۶ گویای آن است که میانگین دو گروه مرد و زن در میزان کاربرد شاخص‌های ارزشیابی درس در سطح اطمینان ۹۵ درصد معنادار بوده است. یافته‌های حاکی از آن است که میانگین تمامی شاخص‌های ارزشیابی در گروه زن‌ها از

میانگین فرضی جامعه (۳) پایین‌تر است. افزون بر این، میانگین تمامی شاخص‌های ارزشیابی در گروه مردان از گروه زنان بیشتر است؛ به عبارت دیگر، مردان میزان توجه به شاخص‌ها را بیشتر از زنان دانسته‌اند.

نتیجه‌گیری و بحث

در راستای پاسخ به سؤال‌های پژوهش، ابتدا داده‌های کیفی و در ادامه داده‌های کمی مورد نیاز گردآوری گردید. نتایج سؤال اول پژوهش نشان داد از منظر مصاحبه‌شوندگان مهم‌ترین شاخص‌های ارزشیابی در درس شامل رعایت اصول علمی و حرفه‌ای در هنگام طراحی سؤال و برگزاری آزمون، دادن بازخورد به تکالیف دانشجویان و راهنمایی آنها، متنوع بودن روش‌های ارزشیابی و رعایت انصاف در طراحی سؤال و نمره دادن می‌باشد. این موارد با نتایج پژوهش‌های (کوک و همکاران (۲۰۲۴)، چنگ و همکاران (۲۰۲۴)، جین و رون، (۲۰۲۳)؛ موریلو و هیدالگو، ۲۰۲۰؛ گرتسن ون یوونکمپ، ۲۰۱۹؛ رسولی و همکاران، ۲۰۱۸؛ پترز و همکاران، ۲۰۱۷؛ تیری، ۲۰۱۶؛ پیرا و همکاران، ۲۰۱۶؛ ژو و براون، ۲۰۱۶؛ دلوکا و همکاران، ۲۰۱۶؛ تیری، ۲۰۱۴؛ دلوکا، ۲۰۱۲؛ پتیفور و ساکلوفسک، ۲۰۱۲؛ داروین، ۲۰۱۱) همخوانی دارد. روندهای جدید ارزشیابی بر استفاده متنوع از روش‌های ارزشیابی، شیوه‌های متمرکز بر یادگیرنده و بازخورد مداوم تأکید می‌کند. تنوع در ارزشیابی موجب می‌شود تا دانشجویان از ارزشیابی استادان، برخی از مفاهیم را یاد بگیرند، از یادگیری آنان ارزشیابی به عمل آورند و ارزشیابی را به‌عنوان یادگیری در نظر بگیرند. در همین راستا، براون و همکاران (۱۹۹۷) اذعان داشتند: اگر می‌خواهید یادگیری دانشجویان را تغییر دهید، از ارزشیابی‌های متنوع جهت بررسی عملکرد و سطح شایستگی دانشجویان استفاده نمایید. به اعتقاد آلتون _ لی (۲۰۰۳)، شیوه‌های ارزشیابی مناسب، ضمن ایجاد انگیزه در دانشجویان، یادگیری را بهبود می‌بخشند. افزون بر این، شیوه‌های متنوع ارزشیابی منجر به تفکر مستقلانه دانشجویان، کاربرد دانش، خلق ایده‌های جدید، ارتباط دانش با مسائل زندگی، ادراک محتوای ارائه شده؛ جست‌وجوی اطلاعات از منابع مختلف، توسعه پیشرفت‌ها و رشد تفکر دانشجویان، کاهش پدیده مطالعه شب امتحان، کاهش استرس دانشجویان، فرایندی بودن ارزشیابی و افزایش کیفیت یادگیری دانشجویان می‌شود.

از دیگر شاخص‌های ارزشیابی مطلوب، ارائه بازخورد توسط استادان است. بازخورد یکی از ویژگی‌های اصلی ارزشیابی است که به ارتقاء کیفیت یادگیری و تغییرات مهم کلاس درس کمک می‌کند. بازخورد به موقع، منظم و مرتبط منجر به مؤثر بودن آن خواهد شد. بازخورد از اهمیت ویژه‌ای برخوردار است؛ زیرا ارتباط بین استاد و دانشجویان را تقویت می‌کند و به‌عنوان فرصتی برای یادگیری و تقویت فرایند یادگیری شناخته می‌شود. نتایج پژوهش‌ها (پیرا و همکاران، ۲۰۱۶؛ شوت، ۲۰۰۸) حاکی از آن است که بازخورد نقش مهمی در ارزشیابی دانشجویان ایفا می‌کند و برای افزایش قابلیت‌های دانشجویان و نظارت بر یادگیری آنها، تسلط بر مشکلات و تمرکز بر روی فرایند یادگیری مورد استفاده قرار می‌گیرد. افزون بر این، ارائه بازخورد به دانشجویان کمک می‌کند تا دلایل فاصله بین سطح موجود و مطلوب خود را شناسایی و برای پر کردن شکاف به وجود آمده، اقدام نمایند و به جست‌وجوی راه‌حل‌های جایگزین بگردند. همچنین پاسخ سؤالات خود از جمله؛ من به کجا می‌روم؟ چگونه می‌روم؟ و به کجا بروم؟ دست پیدا کنند.

رعایت عدالت و انصاف از دیگر شاخص‌هایی است که از نظر مصاحبه‌شوندگان باید در ارزشیابی درس در نظر گرفته شود. این نتایج با یافته‌های پژوهش‌های (موریلو و هیدالگو، ۲۰۲۰؛ رسولی و همکاران، ۲۰۱۸؛ ژو و براون، ۲۰۱۶؛ دلوکا و همکاران، ۲۰۱۶، دلوکا، ۲۰۱۲) همخوانی دارد. رعایت عدالت منجر به کاهش تعصبات دانشجویان و افزایش رضایت‌مندی و مشارکت آنان می‌شود و یادگیری معنادارتری به وجود می‌آید. همچنین در دانشجویان، انگیزه ایجاد می‌کند تا برای رسیدن به اهداف، اصلاح و بهبود عملکرد خود، تلاش دوچندان نمایند و به این باور برسند که عملکرد و تلاش‌شان مورد توجه استادان قرار گرفته است. رعایت نکردن انصاف در ارزشیابی، آسیب‌های جبران‌ناپذیری بر باورها، ذهنیت و عملکرد دانشجویان، کاهش انگیزه آنان بر جای خواهد گذاشت و عملکرد و خشم و رفتار دانشجویان را تحت تأثیر قرار می‌دهد. به عبارتی، می‌توان گفت که به هر میزان در دانشجویان باور به عدم رعایت

انصاف در ارزشیابی‌های استادان بیشتر شود، به همان اندازه استرس بر عملکرد دانشجویان افزایش می‌یابد. در ارزشیابی باید شفافیت در اطلاعات، عینیت آزمون‌ها، ماهیت کمی بودن نتیجه نهایی رعایت شود و دانشجویان باید در مورد زمان، محتوا، ابزار ارزشیابی و معیارهایی که استادان به‌عنوان بخشی از ارزشیابی از آنها استفاده می‌کند، مطلع شوند و در ارائه نمرات وضوح کاملی وجود داشته باشد.

نتایج بخش کمی پژوهش نشان داد بر اساس جداول، میانگین کلی دروس روش‌های تدریس، روش تحقیق در حقوق و مدیریت توسعه بالاتر از نمره ملاک بوده است؛ اگرچه روح کلی نتایج حاکی از آن بود که میانگین شاخص‌های ارزشیابی در دانشگاه مطلوب نبوده و نیاز به توجه بیشتری دارد. در تفسیر نتایج می‌توان گفت بین وضعیت موجود ارزشیابی عملکرد دانشجویان با وضعیت مطلوب آن از دیدگاه دانشجویان فاصله و شکاف وجود دارد. این نتایج با نتایج پژوهش‌های رسولی و همکاران (۲۰۱۸)، دلوکا و همکاران (۲۰۱۶)، پیرا و همکاران (۲۰۱۶) و دلوکا (۲۰۱۲) همخوانی دارد. شاید بتوان گفت از دیگر دلایل وضعیت نامطلوب شاخص‌های ارزشیابی، می‌توان به محتوای ثابتی که استادان ارائه می‌دهند اشاره کرد. استفاده از محتواهای مختلف از جمله کتب، فیلم، عکس، کارگاه، فایل صوتی، اسلاید باعث می‌شود تا استادان از روش‌های مختلفی برای ارزشیابی دانشجویان بهره ببرند و عدم تنوع محتوایی منجر می‌شود، استادان تنها از یک روش استفاده کنند و استفاده از یک ابزار فقط بخشی از پیشرفت تحصیلی دانشجویان را نشان می‌دهد. شاید بتوان گفت از دلایل عدم استفاده از روش‌های مختلف ارزشیابی، بی‌انگیزگی برخی از اعضای هیئت علمی دانشگاه‌ها باشد. از دلایل بی‌انگیزگی استادان می‌توان به مسائلی چون وجود استادان کم‌کار، ناآشنایی آنان با روش‌های ارزشیابی، اشتغال زیاد به کارهای پژوهشی و پست‌های اجرایی اشاره کرد. از دیگر عواملی که می‌توان اشاره کرد این است که در کلاس درس به صورت نظری تدریس می‌شود و به همان صورت هم ارزشیابی می‌شود؛ بدین گونه که اگر دانشجویان در برگه امتحانی خوب بنویسد و تحویل دهند اظهار می‌دارند درس را یاد گرفته‌اند. در حالی که مطالعات این را رد می‌کنند و این نوع آموخته‌هایی که بر روی برگه می‌آیند پس از مدتی به فراموشی سپرده می‌شوند. به نظر می‌رسد از دلایل اینکه استادان کمتر به تنوع ارزشیابی توجه دارند وقت‌گیر بودن آن، بی‌انگیزه بودن، حوصله کم و دغدغه‌های پژوهشی آنان است. دغدغه‌های پژوهشی استادان باعث می‌شود که فرصت مطالعه و کسب دانش و اطلاعات بروز را از آنان ربوده و تلاش آنها در راستای افزایش و تقویت امتیازات پژوهشی خود باشد. در حالی که آموزش یک فرایند پیچیده است که ممکن است بی‌توجهی به آن یا هرگونه سهل‌انگاری در مورد آن منجر به هدر رفتن سرمایه انسانی شود.

درنهایت، نتایج حاکی از آن است که موفقیت یا عدم موفقیت در دروس نباید تنها مبتنی بر امتحان پایان ترم دانشجویان باشد، بلکه لازم است از روش‌های ارزشیابی مختلف برای سنجش آموخته‌های دانشجویان استفاده کرد تا پیشرفت تحصیلی به‌گونه‌ای دقیق مشخص و قابلیت‌ها و توانمندی‌های خود را ارائه دهند. استفاده از محتواهای مختلف توسط استادان منجر می‌شود تا ارزشیابی از نگاه بروندادمحور و نتیجه‌محور به سمت رویکردهای مستمرمحور و فرایندمحور برود. اگر ارزشیابی برای دانشجویان چالش‌برانگیز و انگیزشی باشد، می‌توان شاهد یادگیری عمیق و معنادار بود. افزون بر این نگاه فرایندمحور به ارزشیابی توانایی حل مسئله، تفکر انتقادی و کاربرد دانش را در موقعیت‌های واقعی ارتقا می‌دهد. نگاه سنتی به فرایند ارزشیابی منجر به استفاده خیلی کم از فناوری‌های روز، یادگیری‌های سطحی، تعامل ضعیف با دانشجویان و تأکید بر حفظ طوطی‌وار می‌شود. در نهایت پیشنهاد می‌شود استادان فقط از ارزشیابی کتبی برای ارزشیابی هر درس استفاده نکنند، بلکه باید از انواع روش‌های نوین ارزشیابی استفاده کنند؛ جنبه‌های کیفی مانند تلاش، پیشرفت و مشارکت دانشجویان را جهت ارزشیابی منصفانه در نظر بگیرند و از شاخص‌های شناسایی شده در پژوهش حاضر در راستای افزایش مطلوبیت ارزشیابی خود بهره ببرند.

References

Abrami, PC, Rosenfeld, S, & Dedic, H. (2007). The dimensionality of student ratings of instruction: an update on what we know, do not know, and need to do, in RP Perry & JC Smart (eds). *The scholarship of teaching and learning*

- in higher education: an evidence-based approach*, Springer, Dordrecht, The Netherlands, 446-456. https://doi.org/10.1007/1-4020-5742-3_10
- Azizi Mahmoodabad, M, Nili, M R. (2019). Evaluating elementary school math curriculum: providing a suggested model. *Journal of New thoughts on Education*, 15(2): 123 – 146. [10.22051/JONTOE.2019.17959.2006](https://doi.org/10.22051/JONTOE.2019.17959.2006) (Text in Persian).
- Abu-Alhija, F. N. (2007). Large-scale testing: Benefits and pitfalls. *Studies in educational evaluation*, 33(1), 50-68. <https://doi.org/10.1016/j.stueduc.2007.01.005>
- Akker, J, V, D. (2003). *Curriculum perspectives: An Introduction: In J. van den Akker, U. Hameyer, & W. Kuiper (Eds.), Curriculum landscapes and trends* (pp. 1-10). Dordrecht, the Netherlands: Kluwer. https://doi.org/10.1007/978-94-017-1205-7_1
- Alderman, L, Towers, SJ & Bannah, S. (2012). Student feedback systems in higher education: a focused literature review and environmental scan. *Quality in Higher Education*, 18(3), 261-280. <https://doi.org/10.1080/13538322.2012.730714>
- Alderman, L., Towers, S., Bannah, S., & Phan, L. H. (2014). Reframing evaluation of learning and teaching: An approach to change. *Evaluation Journal of Australasia*, 14(1), 24-34. <https://doi.org/10.1177/1035719X1401400104>
- Alton-Lee, A. (2003). *Quality teaching for diverse students in schooling: best evidence synthesis*. Wellington: Ministry of Education.
- Arreola, RA. (2007). *Developing a comprehensive faculty evaluation system: a guide to designing, building, and operating large-scale faculty evaluation systems*. 3rd edn, Anker, San Francisco.
- Beauchamp, G. (1981). *Curriculum theory (4th Ed.)*. Itasca, Ill.: F.E. Peacock Publishers.
- Berk, RA. (2005). Survey of 12 strategies to measure teaching effectiveness. *International Journal of Teaching and Learning in Higher Education*, 17(1), 48-62.
- Blackmore, J. (2009). Academic pedagogies, quality logics and per formative universities: evaluating teaching and what students want. *Studies in Higher Education*, 34 (8), 857-872. <https://doi.org/10.1080/03075070902898664>
- Brooman, S., Darwent, S., & Pimor, A. (2015). The student voice in higher education curriculum design: is there value in listening?. *Innovations in Education and Teaching International*, 52(6), 663-674. <https://doi.org/10.1080/14703297.2014.910128>
- Brown, G, Bull, J, & Pendlebury, M. (1997). *Assessing student learning in higher education*, London, Routledge.
- Brown, G. T. L. (2002). *Teachers' conceptions of assessment*. Unpublished dissertation, New Zealand, University of Auckland.
- Brown, G. T. L. (2006). Teachers' conceptions of assessment: Validation of an abridged version. *Psychological reports*, 99, 166-170. <https://doi.org/10.2466/pr0.99.1.166-170>
- Cannon, R. (2001). Broadening the context for teaching evaluation. *New Directions for Teaching and Learning*, 88, 87-97.
- Cheng, B., Liu, Y., & Jia, Y. (2024). Evaluation of students' performance during the academic period using the XG-Boost Classifier-Enhanced AEO hybrid model. *Expert Systems with Applications*, 238, 122136. <https://doi.org/10.1016/j.eswa.2023.122136>
- Cook, S., Watson, D., & Webb, R. (2024). Performance evaluation in teaching: Dissecting student evaluations in higher education. *Studies in Educational Evaluation*, 81, 101342. <https://doi.org/10.1016/j.stueduc.2024.101342>
- Cornelius-White, J. (2007). Learner-Centered teacher-student relationships are effective: a meta-analysis. *Review of Educational Research*, 77,113-143. <https://doi.org/10.3102/003465430298563>
- Creswell, J., W. and V. L Plano Clark, (2007). *Designing and Conducting Mixed Methods Research*, London: Sage Publication Inc.
- Creswell, W. (2011). *Educational research: planning, conducting, and evaluating quantitative and qualitative research*. (4th ed). Boston: Pearson pub
- Dai, K., Matthews, K. E., & Reyes, V. (2020). Chinese students' assessment and learning experiences in a transnational higher education programme. *Assessment & Evaluation in Higher Education*, 45(1), 70-81. <https://doi.org/10.1080/02602938.2019.1608907>
- Darwin, S. (2011). Moving beyond face value: re-envisioning higher education evaluation as a generator of professional knowledge. *Assessment & Evaluation in Higher Education*, 37 (6),733-745. <https://doi.org/10.1080/02602938.2011.565114>
- DeLuca, C. (2012). Preparing teachers for the age of accountability: Toward a framework for assessment education. *Action in Teacher Education*, 34, 576-591. <https://doi.org/10.1080/01626620.2012.730347>
- DeLuca, C., Coombs, A., & LaPointe-McEwan, D. (2019). Assessment mindset: Exploring the relationship between teacher mindset and approaches to classroom assessment. *Studies in Educational Evaluation*, 61, 159-169. <https://doi.org/10.1016/j.stueduc.2019.03.012>
- DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016). Teacher assessment literacy: A review of international standards and measures. *Educational Assessment, Evaluation and Accountability*, 1-22. <https://doi.org/10.1007/s11092-015-9233-6>

- European Students' Union (2015). *Overview on student-centred learning in higher education in Europe: Research study*. Retrieved from Brussels, Belgium: European Students' Union.
- Gerritsen-van Leeuwenkamp, K. J., Joosten-ten Brinke, D., & Kester, L. (2019). Students' perceptions of assessment quality related to their learning approaches and learning outcomes. *Studies in Educational Evaluation*, 63, 72-82. <https://doi.org/10.1016/j.stueduc.2019.07.005>
- Gibbs, G. (1999). Using assessment strategically to change the way students learn. In S. Brown & A. Glasner (Eds.), *Assessment matters in higher education: Choosing and using diverse approaches* (pp. 41- 53). Buckingham: Open University Press.
- Harlen, W. (2007). Criteria for evaluating systems for student assessment. *Studies in Educational Evaluation*, 33(1), 15-28. <https://doi.org/10.1016/j.stueduc.2007.01.003>
- Johnson, TD & Ryan, KE. (2000). A comprehensive approach to the evaluation of college teaching. *New Directions for Teaching and Learning*, 83(9), 109–123.
- Keinänen, M., Ursin, J., & Nissinen, K. (2018). How to measure students' innovation competences in higher education: Evaluation of an assessment tool in authentic learning environments. *Studies in Educational Evaluation*, 58, 30-36. <https://doi.org/10.1016/j.stueduc.2018.05.007>
- Lattuca, L., & Stark, J. (2009). *Shaping the college curriculum: Academic plans in context*, San Francisco: Jossey Bass.
- Marsh, HW. (2007). Students' evaluations of university teaching: dimensionality, reliability, validity, potential biases and usefulness', in RP Perry & JC Smart (eds). *The scholarship of teaching and learning in higher education: an evidencebased approach*, Springer, Dordrecht, The Netherlands.
- Pereira, D., Flores, M. A., Simão, A. M. V., & Barros, A. (2016). Effectiveness and relevance of feedback in Higher Education: A study of undergraduate students. *Studies in Educational Evaluation*, 49, 7-14.
- Peters, R., Kruse, J., Buckmiller, T., & Townsley, M. (2017). It's Just Not Fair! making sense of secondary students' resistance to a standards-based grading. *American Secondary Education*, 45(3), 9.
- Pettifor, J. L., & Saklofske, D. H. (2012). Fair and ethical student assessment practices. In C. F. Webber, & J. L. Lupart (Eds.). *Leading student assessment* (pp. 87–106). Dordrecht, The Netherlands: Springer.
- Rasooli, A., Zandi, H., & DeLuca, C. (2018). Re-conceptualizing classroom assessment fairness: A systematic meta-ethnography of assessment literature and beyond. *Studies in Educational Evaluation*, 56, 164-181. <https://doi.org/10.1016/j.stueduc.2017.12.008>
- Roorde, M., & Gullickson, A. M. (2019). Developing evaluation criteria using an ethical lens. *Evaluation Journal of Australasia*, 19(4), 179-194. <https://doi.org/10.1177/1035719X19891991>
- Rust, C., O'Donovan, B., & Price, M. (2005). A social constructivist assessment process model: how the research literature shows us this could be best practice. *Assessment & Evaluation in Higher Education*, 30, 231–240. <https://doi.org/10.1080/02602930500063819>
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78, 153–189. <https://doi.org/10.3102/0034654307313795>
- Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: A review. *Assessment and Evaluation in Higher Education*, 30, 325–341. <https://doi.org/10.1080/02602930500099102>
- Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation theory, models & applications*. San Francisco: Jossey-Bass.
- Tierney, R. D. (2014). Fairness as a multifaceted quality in classroom assessment. *Studies in Educational Evaluation*, 43, 55–69. <https://doi.org/10.1016/j.stueduc.2013.12.003>
- Tierney, R. D. (2016). *Fairness in educational assessment*. In M. A. Peters (Ed.). *Encyclopedia of educational philosophy and theory* (pp. 1–6). Singapore: Springer Science.
- Urillo, F. J., & Hidalgo, N. (2020). Fair student assessment: A phenomenographic study on teachers' conceptions. *Studies in Educational Evaluation*, 65, 100860. <https://doi.org/10.1016/j.stueduc.2020.100860>
- Van den Bergh, V., Mortelmans, D., Spooren, P., Van Petegem, P., Gijbels, D., & Vanthournout, G. (2006). New assessment modes within project-based education-the stakeholders. *Studies in educational evaluation*, 32(4), 345-368. <https://doi.org/10.1016/j.stueduc.2006.10.005>
- Xu, Y., & Brown, G. T. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149–162. <https://doi.org/10.1016/j.tate.2016.05.010>
- Jin, X., & Ruan, Z. (2023). University Students' Perceptions of Their Lecturer's Use of Evaluative Language in Oral Feedback. *Linguistics and Education*, 78, 101233. <https://doi.org/10.1016/j.linged.2023.101233>

