

# بررسی اعتبار پرسشنامه ارزشیابی دانشجویان از اساتید

ایمان... بیگدلی \*

## چکیده:

جهت بررسی میزان اعتبار پرسشنامه ارزشیابی دانشجویان از اساتید که یکی از متداولترین و پرکاربردترین شیوهها در ارزشیابی فعالیت اعضای هیأت علمی دانشگاههای سراسر کشور می باشد ۹۵۶۲ پرسشنامه تکمیل شده توسط همه دانشجویان (دختر و پسر) دانشگاه سمنان در نیمسال اول ۷۹-۷۸ به دو شکل مورد تحلیل قرار گرفت: ۱- ابتدا همه پرسشنامهها ۲- یک درس به طور تصادفی در هر گروه آموزشی. جهت ارزیابی میزان اعتبار نتایج بدست آمده از روش آلفای کرونباخ استفاده شد. نتایج حاکی از آن است که اعتبار و دقت پرسشنامه ارزشیابی به طور جدی مورد تردید می باشد. گامهای اساسی در اصلاح افزایش اعتبار پیشنهاد شده است.

**کلید واژهها:** اعتبار، روش آلفای کرونباخ، اصلاح اعتبار



## مقدمه:

یکی از ارزشیابیهای مهم در همه نظامهای آموزشی ارزشیابی از فعالیتهای مدرسین می باشد. ارزشیابی آموزشی فرایندی است که از طریق آن دادههای مرتبط جمع آوری و در

---

\* عضو هیأت علمی دانشگاه سمنان

جهت اتخاذ تصمیم، تبدیل به اطلاعات مفید می‌شود (ویلیام کولی، ۱۹۷۶). سیف (۱۳۷۶) ارزشیابی از مدرسین را این گونه تعریف می‌کند: «تعیین میزان موفقیت معلمان و رسیدن به هدفهای آموزشی خود» و انجام چنین کاری را مستلزم جمع‌آوری اطلاعات لازم درباره فعالیتهای آموزشی معلم و انتخاب معیارهایی برای مقایسه اطلاعات به دست آمده با آن معیارها و سپس قضاوت در این باره که مدرسین تا چه میزانی به هدفهای از پیش تعیین شده دست یافته‌اند، می‌داند. بنابراین در ارزشیابی از فعالیتهای آموزشی مدرسین دو دسته اطلاعات ضروری می‌باشند:

۱- ملاکها یا معیارهای موفقیت آموزشی ۲- شاخصی از میزان توفیق مدرسین در رسیدن به آن ملاکها. ارزشیابی از فعالیتهای و فرآیندهای اساتید باروشهای گوناگون امکان‌پذیر است. سیف به نقل از مک‌نیل<sup>۱</sup> و پوپام<sup>۲</sup> (۱۳۷۰) این روشها را به هفت دسته زیر تقسیم کرده است:

۱- ارزشیابی معلم به وسیله مدیران و راهنمایان تعلیماتی ۲- ارزشیابی معلم از سوی همکاران ۳- ارزشیابی معلم از کار خود ۴- ارزشیابی از طریق مشاهدات منظم از نحوه آموزش معلم ۵- روش قراردادی با استفاده از مقدار یادگیری فراگیران ۶- روش آزمون‌های عملکردی ۷- ارزشیابی معلمان بوسیله فراگیران. در پژوهش حاضر روش آخر این فهرست، از لحاظ اعتباریابی و اعتمادپذیری مورد بررسی قرار گرفته است. استفاده از این روش به عنوان وسیله‌ای برای ارزشیابی فعالیت اساتید دانشگاه یکی از روشهای متداول در دانشگاههای سراسر کشور می‌باشد. در این روش معمولاً یک پرسشنامه نظرسنجی شامل ۱۵ تا ۲۰ سؤال درباره فرآیندها و فعالیتهای آموزشی عضو هیأت علمی در اختیار دانشجویان قرار می‌گیرد تا به این ترتیب استاد درس خود را مورد ارزشیابی قرار دهند. بدلیل اینکه این شیوه کاربرد بسیار زیادی داشته و در زمینه تصمیم‌گیری در مورد سرنوشت شغلی اساتید، مانند ارتقاء یا ترفیع و استخدام مورد استفاده قرار می‌گیرد، لازم است تا قابلیت اعتماد آن مورد ارزیابی واقع شود. یعنی این که قابلیت اعتماد، دقت و صحت نتایج این ارزشیابی مشخص گردد.

تقریباً همه کشورهای که درگیر مسائل آموزشی هستند، کمابیش در زمینه «ارزشیابی از کار اساتید» فعالیت دارند. استفاده از نظر دانشجویان برای ارزشیابی تدریس و فعالیتهای آموزشی اساتید یکی از روشهای بسیار معمول و متداول در دانشگاههای جهان و منجمله کشور ماست. میلر (۱۹۸۷) روشهای گوناگونی برای ارزشیابی از فعالیت مدرسین پیشنهاد کرده است، که

عبارتند از: ۱- ارزشیابی مدرس به وسیله دانشجو ۲- ارزشیابی مدرس همکار از مدرس ۳- فرم ارزشیابی از عملکرد مدرس. ۴- آزمونهای پیشرفت تحصیلی ۵- مصاحبه‌های ارزشیابی و ۶- ارزشیابی از طریق دانش‌آموختگان. چون ارزشیابی باید براساس اطلاعات و داده‌های منظم، به اتخاذ تصمیم منجر گردد، لذا این اطلاعات باید عینی باشد و براساس معیارهای قابل قبول و ملموس حاصل گردد، (مارش و همکاران، ۱۹۷۹). مدیران و مسئولان مؤسسات آموزش عالی به طور رسمی یا غیررسمی نظرات دانشجویان را درباره ویژگیهای مختلف شخصی و حرفه‌ای اعضاء هیأت علمی مؤسسه خود جویا می‌شوند و براساس نظر آنها به قضاوت درباره میزان شایستگی اساتید می‌پردازند (سیف، ۱۳۷۶). به رغم کاربرد فراوان این شیوه، اختلاف نظرهای بسیار زیادی در باره استفاده از این شیوه وجود دارد. برای مثال سیف (به نقل از تورن<sup>۴</sup>، و همکاران ۱۹۷۶ و اکانل<sup>۵</sup> و اسمارت<sup>۶</sup> ۱۹۷۹) می‌گوید: «استفاده از نظر دانشجویان برای ارزشیابی کار استادان تنها منبع اطلاعات ملموس در دانشگاههای جهان است» از طرفی شواهد پژوهشی دیگر نشان می‌دهند که اظهار نظر دانشجویان درباره کار اساتید متأثر از عواملی است که با موضوع ارزشیابی چندان ارتباطی ندارند. برای مثال دونکین<sup>۷</sup> (۱۹۸۶، به نقل از سیف، ۱۳۷۶) اظهار می‌دارد که صفات شخصی و ویژگیهای محیطی، ادراکها و قضاوتهای افراد را تحت تأثیر قرار می‌دهد و دلیلی وجود ندارد که دانشجویان در ارزشیابی از اساتید خود از این نوع تأثیرات مصون باشند. از جمله پژوهشهای دیگری که به شدت این روش را مورد انتقاد قرار داده است، مطالعات معروف به اثر دکتر فاکس<sup>۸</sup> یا اغواگری آموزشی<sup>۹</sup> است. این مجموعه پژوهشها نخستین بار در ۱۹۷۲ به وسیله نفتولین<sup>۱۰</sup> و همکاران انجام شد. در این پژوهش نشان داده شد که یک بازیگر حرفه‌ای سینما در نقش یک استاد مدعو می‌تواند بخوبی شنوندگان را در جهت اظهار نظرهای مثبت اغوا نماید. نتایج فراتحلیلی که توسط آبرامی<sup>۱۱</sup>، و همکاران در ۱۹۸۲ انجام گرفته است چنین نقل شده است: «آنان تأثیرات اصلی و متقابل روش آموزش معلم و محتوای درس را در رابطه با اظهار نظر شاگردان و پیشرفت آنان به عنوان متغیرهای ملاک، مورد بررسی قرار دادند. وقتی که اظهار نظر شاگردان به عنوان ملاک برگزیده شد نتایج این بررسی از قرار زیر بود: در تمام ۱۲ پژوهش مورد بررسی تأثیر روش آموزشی یا طرز ارائه درس بر اظهار نظر شاگردان نسبت به درس و روش معلم نسبتاً قابل توجه و معنی دار بود، در حالی که تأثیر محتوای درس یا مطالب آموزش داده شده بر اظهار نظر شاگردان ناهماهنگ و بسیار کوچکتر بود.» (به نقل از سیف ۱۳۷۰)

براساس مطالعه مارش و همکاران (۱۹۷۹) در مورد فرمهای ارزشیابی دانشجویان از اساتید در دانشگاه کالیفرنیا جنوبی، اعتبار و قابلیت این روش به دلایل ذیل مورد تردید می‌باشد: دانشجویان تحت تأثیر عوامل متعددی چون: عوامل اجرایی مربوط به تدریس، ویژگیهای درس، شخصیت معلم، ویژگیهای دانشجو و علاقه قبلی و فعلی به موضوع درسی، به ارزشیابی مدرس خود می‌پردازد و بدیهی است که این گونه ارزشیابیها توان واقعی و کارایی معلم را نمی‌سنجد. بحث انگیزترین روش ارزشیابی معلم روشی است که در آن از شاگردان خواسته می‌شود تا او را مورد ارزشیابی قرار دهند. گروهی شدیداً از آن انتقاد کرده و گروهی دیگر قویاً از آن حمایت به عمل آورده‌اند. انتقادهای وارده به این روش ارزشیابی معلم را می‌توان در موارد زیر خلاصه نمود: الف- نظر دانشجویان نسبت به کار استاد درس به درستی معرف شایستگی او نیست بلکه صرفاً منعکس کننده رضایت خاطر یا عدم رضایت خاطر شاگردان از او است، شاگردانی که نمره بهتری از یک معلم می‌گیرند در مقایسه با شاگردانی که نمره خوبی از او دریافت نمی‌کنند در ارزشیابیهای خود نمره بیشتری به وی می‌دهند.

ب - جنس معلم و شاگرد در نحوه ارزشیابی شاگردان از معلم مؤثر است. بندینگ<sup>۱۲</sup> (۱۹۵۲؛ به نقل از سیف، ۱۳۷۶) دریافت که دانشجویان مؤنث در مقایسه با دانشجویان مذکر به معلمان مرد نمره کمتری می‌دهند. همچنین واکر<sup>۱۳</sup> (۱۹۶۹؛ به نقل از سیف، ۱۳۷۶) به این نتیجه رسید که دانشجویان دختر معلمان زن را از معلمان مرد مثبت‌تر ارزشیابی می‌کنند.

ج - تعداد شاگردان کلاس، سطح دشواری درس، نوع درس و سطح تحصیلات بر نحوه ارزشیابی دانشجویان از اساتید تأثیر می‌گذارد. دانشجویان کلاسهای شلوغ در مقایسه با دانشجویان کلاسهای کم جمعیت معمولاً به اساتید خود نمره کمتری می‌دهند. شاگردان دروس انتخابی در مقایسه با شاگردان دروس اجباری به اساتید خود نمره بهتری می‌دهند. دانشجویان سطوح بالا (دوره‌های فوق لیسانس و دکتری) در مقایسه با دانشجویان سطوح پایین‌تر (دوره لیسانس) به اساتید خود نمره بهتری می‌دهند.

د - نظر دانشجویان نسبت به استادان بیشتر تحت تأثیر رفتار دوستانه معلم، شیوه‌های سرگرم کننده و به اصطلاح اغواگرانه او، ویژگیهای کلی و شهرت و اعتبار اوست. گرجی (۱۳۷۳) در پژوهشی نشان داد که بین نمراتی که دانشجویان دانشگاه در یک فرم محبوبیت سنج استادی به

استادان خود دادند و نمراتی که به عنوان ارزشیابی از کیفیت آموزش در فرمهای نظر خواهی برای آنان در نظر گرفتند ضریب همبستگی ۰/۷۵ وجود دارد. این اثرات هاله‌ای<sup>۱۴</sup>، استفاده از فرمهای نظرسنجی دانشجویی برای ارزشیابی از آموزش اساتید را با کمبود روایی<sup>۱۵</sup> مواجه می‌سازد. اما اگر هدف استفاده کننده از آنها سنجش احساسات دانشجویان نسبت به اساتید باشد، بی‌تردید این فرمها از اعتبار یا روایی کافی برخوردارند. مورس<sup>۱۶</sup> و وایلد<sup>۱۷</sup> (۱۹۵۴)؛ به نقل از گرجی، (۱۳۷۳) در یک بررسی از پژوهشهای اولیه درباره روایی این گونه فرمهای نظرسنجی شواهدی یافتند حاکی از اینکه اگر معلم کوششهای آموزشی خود را متوجه شاگردان قوی کلاس بکند، مورد تأیید این شاگردان قرار خواهد گرفت و در نتیجه ضریب همبستگی بین نمرات درس شاگردان و نمراتی که در ارزشیابی معلم خود به او می‌دهند مثبت خواهد بود، اما اگر معلم آموزش خود را بر شاگردان ضعیف کلاس متمرکز کند، در آن صورت دیگر مورد تأیید شاگردان قوی کلاس نخواهد بود، در نتیجه ضریب همبستگی بین نمرات درس شاگردان و نمراتی که شاگردان در فرم نظر سنجی به معلم می‌دهند منفی خواهد بود.

از سوی دیگر، طرفداران روش ارزشیابی معلم توسط شاگردان نیز مزایایی برای آن ذکر کرده‌اند که می‌توان آنها را به نحو زیر خلاصه کرد:

الف- شاگردان منبع اصلی اطلاعات درباره توفیق معلم در رسیدن به هدفهای آموزشی، درس و نحوه برخورد او با شاگردان و روشهای ارائه درس و کلاس‌داری هستند. بنابراین، هیچ‌کس به خوبی شاگردان یک معلم از کم و کیف کار او با اطلاع نیست. لذا بهترین داوران اثر بخشی کار معلم همانا شاگردان هستند.

ب- استفاده از نظرات شاگردان یک معلم برای داوری درباره کار او از هر روش دیگر ارزشیابی امکان پذیرتر است، زیرا برخلاف روشهایی چون مشاهده کلاسی توسط مدیران یا سایر افراد خارج از کلاس، در روش ارزشیابی معلم توسط شاگردان جریان طبیعی درس و بحث معلم مختل نمی‌شود.

ج- نظر شاگردان نسبت به معلمان خود، وسیله مناسبی برای ایجاد ارتباط بین شاگردان و معلم فراهم می‌آورد و بهترین راه تأثیر گذاری متقابل به هر دو طرف ذینفع است. در نتیجه با استفاده از نظر دانشجویان، هم می‌توان روند آموزشی اساتید را اصلاح کرد و آنان را از سطح توقع و خواسته‌های دانشجویان با خبر گردانید و هم دانشجویان را نسبت به کار و کوشش معلمان حساس و

علاقه‌مند نمود. (به نقل از سیف، ۱۳۷۶)

پژوهش حاضر برای بررسی این سؤال طرح و اجرا شده است که آیا ارزشیابی دانشجویان از اساتید دارای اعتبار و اعتماد می‌باشد؟ آیا این وسیله یک وسیله قابل اعتماد و با دقت است؟

روش

الف) آزمودنیها

آزمودنیهای این پژوهش کلیه دانشجویان دانشگاه سمنان در دوره کارشناسی (۹۵۶۲ نفر) رشته‌های مهندسی برق، مواد، مکانیک، عمران، ریاضی کاربردی، فیزیک، مدیریت بازرگانی، ادبیات عرب، ادبیات انگلیسی، دبیری ریاضی، ریاضی محض، علوم آزمایشگاهی و دامپزشکی، در نیمسال اول ۷۸-۷۹ بودند. نظر خواهی در کلیه دروس در همه رشته‌های تحصیلی انجام گرفت. ابتدا از هر گروه آموزشی یک درس بصورت تصادفی جهت تحلیل نتایج پرسشنامه‌های استاد درس مربوطه انتخاب گردید. سپس کل پرسشنامه‌های پاسخ داده شده از لحاظ اعتبار پرسشنامه در همه درسها و مربوط به همه اساتید مورد بررسی قرار گرفت.

ب) ابزار

ابزار پژوهش پرسشنامه‌ای با عنوان پرسشنامه ارزشیابی دانشجویان از اساتید، شامل ۱۷ سؤال بود که در آن از دانشجو خواسته شده بود تا نظر خود را نسبت به روش آموزش و محتوای درس اساتید در امتداد یک مقیاس درجه‌بندی (بسیار خوب، خوب، متوسط و ضعیف) در مقابل هر سؤال مشخص کند.

چند نمونه از سؤالات به شرح ذیل می‌باشد:

- قدرت بیان و تفهیم مطالب درس - ارائه برنامه و بیان اهداف کلی درس در شروع کلاس - تسلط بر موضوع - حضور استاد طبق برنامه تعیین شده به منظور رفع اشکال و مشاوره - استقبال از نظرات و پیشنهادها و انتقادات دانشجویان - علاقه شما به این درس .

شیوه اجرا

پرسشنامه‌های ارزشیابی در هفته دوازدهم نیمسال تحصیلی به طور همزمان در کلیه دروس نظری و عملی در دانشگاه سمنان در ابتدای کلاس درس توسط مسئولین آموزشی دانشگاه در

دانشکده‌ها توزیع شدند و دانشجویان پس از پاسخ‌گویی آن را به مسئول آموزشی عودت دادند. کلیه دانشجویان از اینکه قرار بود نتایج ارزشیابی برای مقاصد پژوهش بکار رود بی‌اطلاع بودند و مانند نیمسالهای گذشته به طور طبیعی به پرسشنامه‌ها پاسخ دادند. پرسشنامه‌ها پس از استفاده از نرم‌افزار مخصوص این پرسشنامه خوانده و اطلاعات به صورت کمی ثبت و ضبط شدند.

### نتایج

به منظور بررسی میزان اعتبار پرسشنامه ارزشیابی دانشجویان از اساتید، ابتدا پاسخهای آزمودنی‌ها به صورت رتبه‌های یک تا چهار، نمره‌گذاری شد و بعد پرسشنامه‌ها با استفاده از روش آلفای کرونباخ مورد تحلیل قرار گرفت.

بدلیل اینکه در تمامی زمینه‌های علوم رفتاری اندازه‌گیری‌های بدون اعتبار و اعتماد از نظرات مردم از پیشگویی صحیح رفتار آنها جلوگیری می‌کند، مسئله دقت اندازه‌گیری، زمانی که متغیرها به سختی مشاهده می‌شوند، مشکلاتی را در زمینه تحقیقات کاربردی ایجاد می‌کند. تحلیل اعتبار و اعتمادپذیری در کلیه مسائلی که به نحوی با اندازه‌گیری‌ها سر و کار دارند، با بهبود مقیاسهای موجود و ارزیابی مقیاسهای قبلی، مقیاسهای اندازه‌گیری قابل اعتمادی را ایجاد می‌کند.

ارزیابی اعتبار مقیاسهای اندازه‌گیری مثل مقیاس ارزشیابی دانشجویان از اساتید بر میزان همبستگی‌ها بین ماده‌ها یا اندازه‌هایی که مقیاس را تشکیل می‌دهند نسبت به واریانس آنها، مبتنی است. در تحلیل اعتبار مقیاس یا پرسشنامه پاسخ به هر سؤال پرسشنامه از دو جنبه ارزیابی می‌شود:

#### ۱- پیش‌داوریه‌های افراد نسبت به مسئله ۲- محتوای درونی سؤال مورد نظر

برای مثال در رابطه با سؤال میزان تسلط استاد به موضوع درس، پاسخ دانشجویان به این سؤال از یک سو به نظر کلی او نسبت به تسلط استادش مربوط می‌شود و از سوی دیگر نوع سؤال و ترکیب کلمات آن نیز در پاسخگویی بی‌تأثیر نیست. برای مثال ممکن است دانشجویی مورد نظر تسلط را به معنی بیان خوب در نظر بگیرد. اعتباریابی یا اعتمادپذیری تعریفی صریح دارد، اندازه‌ای دارای اعتبار است که نسبت به یک میزان خطای معین منعکس‌کننده نظرات حقیقی باشد. بنابراین به عنوان اندازه‌ای برای اعتباریابی می‌توان گفت: اعتبار یا اعتمادپذیری برابر است با نسبت واریانس امتیازات واقعی به واریانس کل امتیازات.

در این بررسی به هر پرسشنامه که دانشجو آن را پر کرده بود به نسبت میزان ارزیابی او از

بسیار خوب تا ضعیف عددی بین ۱ تا ۴ داده شد.

یکی از تفاسیر اعتبار، همسانی درونی<sup>۱۸</sup> آزمون است: داده‌های آزمون همگون هستند. این تفسیر در حقیقت همان مفهومی است که در مورد سایر تفسیرها مطرح می‌شود، یعنی دقت. در این تفسیر از ماده‌های یک آزمون، یک نمونه تصادفی انتخاب می‌شود و سپس نمونه تصادفی دیگری از ماده‌های آن انتخاب می‌شود. هر یک از نمونه‌ها را به عنوان یک آزمون فرضی درباره گروهی از افراد اجرا می‌کنیم. در این صورت هر فرد دو نمره خواهد داشت: یک  $x_t$  در یک نمونه فرضی و یک  $x_t$  دیگر در نمونه فرضی دوم. همبستگی بین این دو مجموعه نمره را محاسبه کرده و این فرایند را بی‌نهایت تکرار می‌کنیم. میانگین، همبستگی‌های متقابل بین نمونه‌های فرضی همسانی درونی آزمون را نشان می‌دهد.

روش اعتباریابی مورد استفاده در این پژوهش روش آلفای کرونباخ بود که یکی از متداول‌ترین روش‌های اعتباریابی پرسشنامه‌هاست. این روش بر اساس سازگاری یا هماهنگی درونی پرسشنامه شکل گرفته است، اگر انحراف معیار سؤالها یک باشد این روش مقایسه را بر مبنای متوسط همبستگی میان سؤالهای پرسشنامه انجام می‌دهد. در غیر این صورت بر مبنای متوسط کوواریانس بین سؤالها تصمیم‌گیری می‌شود.

با استفاده از نرم‌افزار آماری SPSS هر دو روش انجام می‌شود. فرض زیربنایی این روش این است که اگر سؤالهای موجود با یکدیگر همبستگی مثبتی نداشته باشند، انتظار نداریم در صورت تکرار پرسشنامه بر روی افراد نمونه دیگری نتایج مشابهی به دست بیاید و در واقع اعتبار آزمون (پرسشنامه) کاهش می‌یابد.

در این روش ابتدا واریانس هر سؤال و واریانس مجموع مقیاس اندازه‌گیری محاسبه می‌شود. اگر تغییرپذیری بین افراد تقریباً نزدیک بهم باشد، واریانس مجموع اندازه‌ها از مجموع واریانس سؤالات مختلف کمتر خواهد بود. اگر به سؤالات امتیاز واقعی داده نشود و پاسخهای افراد کاملاً به یکدیگر بی‌ارتباطی باشند، ضریب آلفا به سمت صفر میل می‌کند. اگر تمام سؤالها قابل اعتماد باشند و یک نتیجه را نشان دهند، ضریب آلفا یک خواهد بود.

با توجه به تحلیل اعتبار آزمون و نتایج حاصله، ضریب آلفای کرونباخ هم در مورد کل آزمودنیهایی که به پرسشنامه پاسخ داده‌اند ( $N = 9562$ ) و هم در مورد یکی از دروس که بصورت تصادفی از یکی از گروههای آموزشی انتخاب شده است ( $n = 116$ ) در حدود بالایی است



( $\alpha = 0/91$ ) لیکن تحلیل واریانس پاسخها، که تفاوت پاسخگویان و پاسخهای سؤالات (اندازه‌ها) را آزمون می‌کند و در بحث تحلیل اعتبار باید مورد توجه قرار گیرد نشان از عدم همگنی و یکدستی پرسشها داشته و اعتبار پرسشنامه را مورد تردید قرار می‌دهد.

همانطور که در جدول ۱ و ۲ تحلیل واریانس به تفکیک برای هر دو نمونه ملاحظه می‌شود منبع تغییر بین آزمودنی‌ها<sup>۱۹</sup> به تفاوت‌هایی بین پاسخگویان مختلف اشاره می‌کند، که در این تحلیل واریانس آزمون نشده است، اما می‌توان با محاسبه آماری  $F = \frac{MS_B}{MS_R} = \frac{5/4402}{0/4633}$  و مقایسه با نقطه بحرانی در توزیع F با درجات آزمودنی ۱۵۲۹۹۲ و ۹۵۶۱ آزمون را اجرا کرد.

در سطر درون آزمودنی‌ها<sup>۲۰</sup> جدول ۱ تحلیل واریانس، تفاوت بین پاسخهای مختلف یک فرد به سؤالات سنجیده می‌شود. منبع تغییر دیگر منبع تغییر بین اندازه‌ها<sup>۲۱</sup> است و به تغییرات بین اندازه‌های مختلفی که به سؤالات نسبت داده شد، اشاره می‌کند. در این سطر مقدار آماره F معنی‌دار ( $p < 0/0001$ ) است و فرض یکسان بودن اندازه‌های منتسب به سؤالات رد می‌شود. در اینجا باید به این نکته مهم توجه کرد که تغییرات بین اندازه‌ها در یک فرد یا ناشی از تفاوت سؤالات است یا ناشی از خطاهای تصادفی، در واقع مجموع مربعات تفاوت‌های درون فردی برابر است با مجموع مربعات تفاوت‌های بین سؤالات و خطاهای تصادفی. در یک پرسشنامه مناسب باید تفاوت بین افراد به خوبی تشخیص داده شود و تفاوت‌های درون فردی تنها ناشی از خطاهای تصادفی باشد.

جدول ۱- تحلیل واریانس نتایج کل پرسشنامه‌های ارزشیابی دانشجویان از اساتید در همه درسها

و همه گروههای آموزشی در نیمسال اول ۷۸-۷۹

منبع تغییر	مجموع مجذورات	درجه آزادی	میانگین مجذورات	نسبت F	سطح معناداری
بین آزمودنیها	۵۲۰۱۳/۸۴۳۷	۹۵۶۱	۵/۴۴۰۲		
درون آزمودنیها	۷۹۹۱۳/۱۷۶۵	۱۵۲۹۹۲	۰/۵۲۲۳		
بین اندازه‌ها	۹۰۲۱/۷۶۹۷	۱۶	۵۶۳/۸۶۰۶	۱۲۱۶/۷۵۰۳	۰/۰۰۰۰
باقیمانده	۷۰۸۹۱/۴۰۶۸	۱۵۲۹۹۲	۰/۴۶۳۴		
کل	۱۳۱۹۲۷/۰۲۰۲	۱۶۲۵۵۳	۰/۸۱۱۶		

ضرایب اعتبار ۱۷ ماده:  $\alpha = 0/9148$  آلفای استاندارد شده =  $0/9130$

بنابراین معنی دار شدن آماره F در سطر بین اندازه‌ها نشانه عدم همگنی و یکدستی پرسشها بوده و اعتبار پرسشنامه را بطور معنی داری مورد تردید قرار می‌دهد.

جدول ۲- تحلیل واریانس نتایج پرسشنامه‌های ارزشیابی دانشجویان از اساتید در یک درس

منبع تغییر	مجموع مجذورات	درجه آزادی	میانگین مجذورات	نسبت F	سطح معناداری
بین آزمودنیها	۵۲۷/۹۴۳۲	۱۱۵	۴/۵۹۰۸		
درون آزمودنیها	۹۴۹/۲۹۴۱	۱۸۵۶	۰/۵۱۱۵		
بین اندازه‌ها	۲۰۳/۹۷۸۷	۱۶	۱۲/۷۴۸۷	۳۱/۴۷۳۳	۰/۰۰۰۰
باقیمانده	۷۴۵/۳۱۵۴	۱۸۴۰	۰/۴۰۵۱		
کل	۱۴۷۷/۲۳۷۳	۱۹۷۱	۰/۷۴۶۵		

ضرایب اعتبار ۱۷ ماده: آلفا = ۰/۹۱۱۸ آلفای استاندارد شده = ۰/۹۱۴۶

## بحث

همانطور که از تحلیل واریانس نتایج برمی‌آید پرسشنامه ارزشیابی دانشجویان از اساتید به لحاظ خطاهای درون آزمودنی از لحاظ اعتبار و دقت لازم کفایت مطلوب را نداشته و باید برای استفاده مناسب از آن به اصلاح اعتبار پرسشنامه پرداخت. بر اساس یافته‌های تحقیق به نظر می‌رسد که یکی از عوامل مؤثر در کاهش دقت پرسشنامه بی‌دقتی دانشجویان در پاسخگویی به سؤالات پرسشنامه است. این یافته تحقیق با تحقیقات مشابه که در دانشگاه‌های مختلف ایران و جهان صورت گرفته است مطابقت دارد (به عنوان مثال: سیف، ۱۳۷۰). دانشجویان در اواخر نیمسال در یک فاصله زمانی بسیار کوتاه (ده دقیقه‌ای) ناچار به پاسخگویی بوده و دستورالعمل مناسب، روشن و یکسانی در تکمیل پرسشنامه‌ها از سوی مسئولین اجرا دریافت نمی‌نمایند. دانشجویان معمولاً بصورت سریع و بی‌دقت و بدون انگیزه لازم به سؤالات پاسخ داده و این امر یکی از دلایل بی‌اعتباری پرسشنامه‌ها می‌باشد. در اصلاح افزایش اعتبار می‌توان به موارد زیر اشاره نمود:

● اصلی که اصلاح اعتبار را در پی دارد چیزی است که به آن اصل نظارت حداکثر - حداقل می‌گویند و مفهوم آن چنین است: واریانس تفاوت‌های فردی را به حداکثر و واریانس خطا را به

حداقل برسانید. اولین قدم در اصلاح افزایش اعتبار نوشتن ماده‌های ابزار اندازه‌گیری روانی و آموزشی بدور از ابهام و بوضوح و روشن است. زیرا یک رویداد مبهم به صورتهای مختلف تفسیر می‌شود. ماده مبهم موجب می‌شود که واریانس خطا وارد اندازه‌گیری شود زیرا افراد ممکن است آن را به صور مختلف تفسیر کنند، چنین تفسیرهایی معمولاً تصادفی هستند، بنابراین واریانس خطا را افزایش و اعتبار را کاهش می‌دهند.

● دومین قدم اینست که اگر ابزاری به اندازه کافی معتبر نباشد به آن ماده‌های دیگری از همان نوع و با همان کیفیت افزوده شود. این امر معمولاً، لیکن نه لزوماً، اعتبار را به مقداری که قابل پیش بینی است افزایش می‌دهد. افزایش تعداد بیشتری از داده‌ها این احتمال را افزایش می‌دهد که نمره  $x_i$  هر فرد به نمره  $x_{i+1}$  او نزدیک باشد. این امر به سبب نمونه‌گیری ویژگی یا فضای ماده‌هاست، اگر تعداد ماده‌ها کم باشد، خطای تصادفی بیشتر می‌شود. چنانچه تعداد ماده‌ها زیادتر باشد، خطای تصادفی کمتر افزایش می‌یابد. وقتی که تعداد ماده‌ها زیاد باشد احتمال اینکه خطاهای تصادفی در مجموع یکدیگر را خنثی کنند بیشتر است. به همین ترتیب تعداد ماده‌های بیشتر احتمال صحت اندازه‌گیری را افزایش می‌دهد.

● سومین قدم و مهم‌ترین کار دستور کارهای روشن و استاندارد در تکمیل پرسشنامه است. این کار معمولاً خطای اندازه‌گیری را کاهش می‌دهد. در نوشتن دستور کارها همیشه باید دقت زیاد به عمل آید که دستور کارها به گونه‌ای روشن بیان شوند. دستور کار مبهم افزایش واریانس خطا را موجب می‌شود. به علاوه ابزارهای اندازه‌گیری همواره باید تحت شرایط استاندارد، کنترل شده و مشابه اجرا گردد. اگر موقعیتهای اجرا متفاوت باشند، بار دیگر واریانس خطا وارد عمل می‌شود.

● به رغم اختلاف نظرهای فراوان در باره استفاده از نظر دانشجویان برای ارزشیابی کار اساتید اگر پرسشنامه‌های نظر سنجی مطابق اصول صحیح تهیه و اجرا گردند و نتایج آنها برای اصلاح و بهسازی آموزش معلمان بکار روند و آنجا که قرار است تصمیم‌های در باره معلم گرفته شود، این نتایج با نتایج حاصل از روشهای دیگر ارزشیابی تکمیل شوند، استفاده از روش ارزشیابی معلم توسط دانشجویان روش مفید و قابل استفاده‌ای است.

## یادداشت‌ها

- |                          |                          |
|--------------------------|--------------------------|
| 1- McNeil                | 2- Popham                |
| 3- Reliability           | 4- Thorne                |
| 5- O'Connel              | 6- Smartt                |
| 7- Dunkin                | 8- Dr.Fox effect         |
| 9- Educational seduction | 10- Naftulin             |
| 11- Abrami               | 12- Bending              |
| 13- Walker               | 14- Halo effects         |
| 15- Validity             | 16- Morsh                |
| 17- Wildes               | 18- Internal consistency |
| 19- Between people       | 20- Within people        |
| 21- Between measures     |                          |

## منابع:

- سیف، علی اکبر (۱۳۷۰). ارزشیابی دانشجویان از استادان: تا چه اندازه می‌توان به آن اعتماد کرد؟ پژوهش‌های روان‌شناختی. دوره اول شماره ۱ و ۲.
- سیف، علی اکبر (۱۳۷۶). روش‌های اندازه‌گیری و ارزشیابی آموزشی. ویرایش دوم. تهران، دوران.
- شرکت آمار پردازان (۱۳۷۷). راهنمای کاربران نرم‌افزار SPSS.6. ج دوم. تهران، حامی.
- کرلینجر، فردان. (۱۳۷۶). مبانی پژوهشی در علوم رفتاری. ترجمه: حسن پاشا شریفی و جعفر نجفی زند. تهران، آوای نور. (تاریخ انتشار اثر به زبان اصلی ۱۹۸۶).
- گرجی، یوسف (۱۳۷۳). رابطه بین محبوبیت اجتماعی استادان از نظر دانشجویان و نحوه ارزشیابی دانشجویان از فعالیتهای آموزشی آنان. پایان‌نامه کارشناسی ارشد چاپ نشده. دانشگاه علامه طباطبائی.
- Cooley, W. (1976). *Evaluation research in education*, NewYork: Irvington Publisher.
- Marsh, H. W. , Overall, J. V. & Kesler, S. P. (1979). Validity of instructional effectiveness: A comparison of faculty self-evaluation and evaluation by their students. *Journal of Educational Psychology*, Vol. 71, No. 2.
- Miller, R. (1987), *Evaluating teaching. The role of students rating evaluating faculty for promotion and tenure*, NewYork: Academic Press.