



میزان اشاعه استانداردهای معیار دقت در ارزشیابی‌های درونی گروه‌های آموزشی دانشگاه‌ها و مؤسسات آموزش عالی

معصومه قائمی پور^۱، دکتر مرتضی کرمی^۲، دکتر مقصود امین خندقی^۳ و غلامعلی سیفی^۴

چکیده

کیفیت آموزش یکی از دغدغه‌های مهم نظام‌های دانشگاهی و مؤسسات آموزش عالی است؛ که نمود آن را در پروژه‌های متعدد ارزشیابی درونی می‌توان دید. هدف پژوهش حاضر میزان اشاعه استانداردهای معیار دقت در ارزشیابی‌های درونی گروه‌های آموزشی دانشگاه‌ها و مؤسسات آموزش عالی است. این پژوهش از نظر روش، پژوهش ارزشیابی است. بر این اساس جامعه آماری مقاله‌هایی با موضوع ارزشیابی درونی گروه‌های آموزشی طی ۵ سال اخیر یعنی سال‌های ۹۲-۱۳۸۸ مدنظر قرار گرفت. بدین منظور به سایت‌های مگیران، اس‌آ‌دی، نورمگز و انسانی مراجعه شد و حجم جامعه حدود ۴۵ به دست آمد. به دلیل محدود بودن جامعه از شیوه سرشماری استفاده شد، بدین نحو که تمامی مقالات مورد بررسی قرار گرفت. در این پژوهش، پژوهشگر از ابزار چک‌لیست برای جمع‌آوری داده‌ها استفاده کرده است. پژوهشگر چک‌لیست پژوهش خود را بر اساس استانداردهای ارزشیابی برنامه که در سال ۲۰۱۱، نسخه سوم و بر اساس بررسی‌های دو چک‌لیست کوتاه و بلند که توسط استافل بیم (۱۹۹۹، ۲۰۱۲) برای این منظور تهیه شده، تدوین کرده است. این چک‌لیست استانداردهای معیار دقت ارزشیابی را در مقالات ذکر شده می‌سنجد. یافته‌ها نشان داد که پژوهشگران کمتر توجهی به مستند کردن فرایند کارشان داشته‌اند. شاید استانداردهای دقت را در ارزشیابی‌شان در نظر گرفته باشند ولی در گزارش کمتر اشاره‌ای به آن‌ها کرده‌اند. همچنین بیشتر استاندارد اطلاعات معتبر در گزارش اشاره و مستند شده است.

کلید واژگان: فرا-ارزشیابی، استانداردهای معیار دقت، کیفیت، آموزش عالی

مقدمه

کیفیت آموزش یکی از دغدغه‌های مهم نظام‌های دانشگاهی و مؤسسات آموزش عالی است؛ در جهت ارتقای کیفیت آن‌ها تلاش‌های قابل توجهی انجام شده است که نمود آن را می‌توان در پروژه‌های متعدد ارزشیابی درونی دید. در سال‌های اخیر در آموزش عالی ایران رویکرد اعتبارسنجی به منظور ارزشیابی و اعتبارسنجی نظام آموزش عالی بیشتر به چشم می‌خورد؛ مرحله اول آن یعنی ارزشیابی درونی در اکثر گروه‌های آموزشی دانشگاه‌های سراسر کشور اجرا شده است (عباسپور و شرفی، ۱۳۹۰). گروه‌های آموزشی نقطه شروع مناسبی جهت ارزشیابی کیفی و تضمین کیفیت در نظام آموزش عالی می‌باشند و در این میان گروه‌های مطرح و با قدمت زیاد به جهت اینکه با انجام ارزشیابی رهنمودهای مفیدی را برای سایر گروه‌های مشابه فراهم می‌سازند موقعیت حساس و ویژه‌ای دارند (زین‌آبادی، کیامنش و فرزاد، ۱۳۸۴). استفاده از ارزشیابی درونی در گروه‌های آموزشی باعث می‌شود تا توانایی‌های بالقوه گروه شناسایی شود و شرایط لازم برای خود ترمیمی امور گروه فراهم شود (عابدینی، خرمی‌راد، آخوندزاده و حیدری، ۱۳۹۲). به عبارت دیگر، بر اساس نتایج ارزشیابی درونی پیشنهادهایی برای بهبود کیفیت داده می‌شود (بازرگان، ۱۳۸۸). رویکرد ارزشیابی درونی یک فرایند سیستماتیک است که در آن وضعیت موجود درون‌دادها، فرایندها و برون‌دادهای گروه به تصویر کشیده شده و با وضعیت مطلوب مقایسه می‌شود (بازرگان، ۱۳۸۰). اکثر تعاریف ارائه شده در این حوزه بر این امر تأکید دارند که ارزشیابی درونی وسیله‌ای برای شناسایی نقاط قوت و ضعف و همچنین شناسایی

^۱ دانشجوی کارشناسی ارشد تحقیقات آموزشی دانشگاه فردوسی مشهد، مشهد، ایران. نویسنده مسئول: ghaemi88@yahoo.com

^۲ استادیار دانشکده علوم تربیتی و روانشناسی دانشگاه فردوسی مشهد، مشهد، ایران. mor.karami@gmail.com

^۳ دانشیار دانشکده علوم تربیتی و روانشناسی دانشگاه فردوسی مشهد، مشهد، ایران. aminkhandaghi@ferdowsi.um.ac.ir

^۴ دانشجوی دکتری برنامه ریزی درسی دانشگاه فردوسی مشهد، مشهد، ایران. ali_seifi7@yahoo.com



نواحی تغییر و دگرگونی می‌باشد (اسونسون^۵، ۲۰۰۴؛ بالباستر^۶، ۲۰۰۳؛ هیلمن^۷، ۱۹۹۴؛ فین^۸، ۱۹۹۴؛ به نقل از زین آبادی و همکاران، ۱۳۸۴).

باید توجه داشت هر ارزشیابی به هدف‌های مطلوب منجر نخواهد شد و ممکن است عواملی مانند: اجرای نامناسب، گزارش با تأخیر، یافته‌های سوگیرانه، یافته‌های مبهم، نتیجه‌گیری‌های ناموجه، تفسیرهای ناکافی و پیشنهادهای ناموجه باعث کم اعتبار شدن ارزشیابی شوند (استافل بیم^۹، ۲۰۰۱). بر این اساس اهمیت فرا-ارزشیابی‌ها دوچندان می‌شود. به عقیده‌ی اسکریون (۲۰۰۸)، فرا-ارزشیابی را ارزشیابی از ارزشیابی‌ها تعریف می‌کند. فرا-ارزشیابی می‌تواند برای سنجش کیفیت مطالعات انجام شده استفاده شود (مادزیوهاندیلا، ۲۰۱۰). استانداردهای ارزشیابی برنامه مخصوصاً در ارزشیابی از ارزشیابی‌ها یعنی فرا-ارزشیابی‌ها، قابل کاربرد هستند. ارزیاب‌ها برای قضاوت درباره اینکه تا چه حدی ارزشیابی برنامه از هر استاندارد پیروی می‌کند، اطلاعات جمع‌آوری می‌کنند. نیمرخی از این قضاوت‌ها پایه‌هایی برای قضاوت ارزشیابی به‌منظور بررسی‌های سودمندی، قابلیت کاربرد، درستی و دقت در ارتباط با هر استاندارد ارائه می‌دهد (استافل بیم، ماداس و کلاگان^{۱۰}، ۲۰۰۰).

با توجه به روند رو به رشد ارزشیابی‌های درونی گروه‌های آموزشی در پنج سال اخیر (۱۳۸۸-۱۳۹۲) و علی‌رغم انجام پژوهش‌های ارزشیابی درونی گروه‌های مختلف آموزشی مانند، عابدینی و همکاران (۱۳۹۲)، رحیمی فرد و همکاران (۱۳۹۲) و جعفری قوام‌آباد و همکاران (۱۳۹۲)، تاکنون پژوهشی جز مطالعه حکیمی (۱۳۹۰) یافت نشد که به فرا-ارزشیابی از ارزشیابی‌های درونی گروه‌های آموزشی بر اساس استانداردهای ارزشیابی برنامه پردازد. حکیمی (۱۳۹۰)، در این پژوهش باهدف ارزیابی طرح‌های ارزیابی درونی و راه‌های توسعه اثربخشی آن در دانشگاه فردوسی مشهد هرچند به آسیب‌شناسی طرح‌های ارزیابی درونی پرداخته اما به فرا-ارزشیابی ارزشیابی‌های درونی بر اساس استانداردهای ارزشیابی نپرداخته است.

با توجه به اهمیت فرا-ارزشیابی محقق در این مطالعه با استفاده از استانداردهای ارزشیابی برنامه انجمن ارزشیابی آمریکا (۲۰۱۱)، به فرا-ارزشیابی پژوهش‌های ارزشیابی آموزشی انجام شده پرداخته و میزان اشاعه استانداردهای معیار دقت را در مقاله‌هایی با موضوع ارزشیابی درونی گروه‌های آموزشی بررسی نموده است. با توجه به اینکه ممکن است ارزیابان استانداردهای معیار دقت را در پژوهش ارزشیابی خود رعایت کرده باشند اما در گزارش خود شواهدی دال بر استفاده از استانداردهای دقت در پژوهش ارزشیابی خود ذکر نکرده باشند بنابراین پژوهشگر به دنبال میزان اشاعه استانداردهای معیار دقت بوده است.

روش پژوهش

پژوهش حاضر از نظر روش، پژوهش ارزشیابی است، پژوهشگر به فرا-ارزشیابی یعنی، ارزشیابی از ارزشیابی‌ها پرداخته است. جامعه آماری مقاله‌هایی با موضوع ارزشیابی درونی گروه‌های آموزشی طی ۵ سال اخیر یعنی سال‌های ۹۲-۱۳۸۸ مدنظر قرار گرفت. بدین منظور به سایت‌های مگیران، اس‌آی‌دی، نورمگز و انسانی مراجعه شد و حجم جامعه حدود ۴۵

⁵.Svenson

⁶.Balbaster

⁷.Hilman

⁸.Finne

⁹. Stufflebeam

¹⁰. Madaus & Kellaghan



به دست آمد. به دلیل محدود بودن جامعه از شیوه سرشماری استفاده شد، بدین نحو که تمامی مقالات مورد بررسی قرار گرفت. از ابزار چک‌لیست برای جمع‌آوری داده‌ها استفاده شد. این چک‌لیست استانداردهای معیار دقت را می‌سنجد. پژوهشگر چک‌لیست پژوهش خود را بر اساس استانداردهای ارزشیابی برنامه که در سال ۲۰۱۱، نسخه سوم و بر اساس چک‌لیست استافل بیم (۲۰۱۲) که برای این منظور تهیه شده است و بر اساس بررسی‌های دو چک‌لیست کوتاه و بلند استافل بیم (۱۹۹۹) که بر اساس نسخه دوم استانداردها تهیه شده است، تدوین کرد. برای هر استاندارد معیار دقت ۶ نشانگر وجود دارد. قضاوت درباره‌ی هر استاندارد به این صورت است، اگر هر ۶ نشانگر دیده شد، عالی، اگر از ۶ نشانگر، ۵ نشانگر دیده شد، خیلی خوب، اگر از ۶ نشانگر، ۴ نشانگر دیده شد، خوب، اگر از ۶ نشانگر، ۲ یا ۳ نشانگر دیده شد، متوسط و اگر از ۶ نشانگر، ۱ نشانگر دیده شد یا اصلاً دیده نشد، ضعیف است.

یافته‌های پژوهش

سؤال پژوهش: تا چه اندازه استانداردهای معیار دقت در مقاله‌های ارزشیابی درونی گروه‌های آموزشی اشاعه یافته است؟

به‌منظور پاسخگویی به سؤال پژوهش، یافته‌های حاصل از تحلیل توصیفی با استفاده از شاخص‌های آمار توصیفی ارائه شد. ابتدا جداول مربوط به یافته‌های توصیفی مربوط به اشاعه‌ی نشانگرهای استانداردهای معیار دقت و سپس جدول مربوط به یافته‌های توصیفی مربوط به اشاعه استانداردهای معیار دقت ارائه شده است.

جدول (۱): یافته‌های توصیفی مربوط به اشاعه نشانگرهای استاندارد تصمیمات و نتایج موجه

کل	عدم مشاهده شواهد	مشاهده شواهد	شاخص آماری	نشانگرها	استاندارد
۴۵	۴۵	۰	فراوانی	۱- محدود کردن نتایج به دوره‌های زمانی، بافت، اهداف و فعالیت‌های قابل اجرا	تصمیمات و نتایج موجه
۱۰۰	۱۰۰	۰	درصد		
۴۵	۰	۴۵	فراوانی	۲- مستند کردن اطلاعاتی که از هر نتیجه حمایت می‌کند	
۱۰۰	۰	۱۰۰	درصد		
۴۵	۴۵	۰	فراوانی	۳- اخطار دادن در برابر برداشت‌های غلط رایج	
۱۰۰	۱۰۰	۰	درصد		
۴۵	۰	۴۵	فراوانی	۴- تمرکز کردن نتایج به‌طور مستقیم بر روی سؤالات ارزشیابی	
۱۰۰	۰	۱۰۰	درصد		
۴۵	۶	۳۹	فراوانی	۵- به‌دقت منعکس کردن فرایند و یافته‌های ارزشیابی	
۱۰۰	۱۳/۳	۶۷/۷	درصد		
۴۵	۴۴	۱	فراوانی	۶- گزارش محدودیت‌های ارزشیابی	
۱۰۰	۹۷/۸	۲/۲	درصد		



نتایج جدول (۱) نشان داد ، همه‌ی مقاله‌ها نشانگرهای دوم و چهارم را اشاعه داده‌اند. تعداد ۳۹ مقاله یعنی ۶۷/۷ درصد نشانگر پنجم را اشاعه داده‌اند، یک مقاله یعنی ۲/۲ درصد نشانگر ششم را اشاعه داده است و هیچ کدام نشانگرهای اول و سوم را اشاعه نداده‌اند.

جدول (۲): یافته‌های توصیفی مربوط به اشاعه نشانگرهای استاندارد اطلاعات معتبر

کل	عدم مشاهده شواهد	مشاهده شواهد	شاخص آماری	نشانگرها	استاندارد
۴۵	۰	۴۵	فراوانی	۱- تمرکز ارزشیابی روی سؤالات کلیدی	اطلاعات معتبر
۱۰۰	۰	۱۰۰	درصد		
۴۵	۰	۴۵	فراوانی	۲- سنجش و گزارش اینکه هر یک از فرایندهای به کار گرفته شده به چه نوع اطلاعاتی احتیاج دارند	
۱۰۰	۰	۱۰۰	درصد		
۴۵	۱۰	۳۵	فراوانی	۳- مستند کردن اینکه چگونه اطلاعات هر رویه، نمره‌گذاری، تجزیه و تحلیل و تفسیر شدند	
۱۰۰	۲۲/۲	۷۷/۸	درصد		
۴۵	۱۱	۳۴	فراوانی	۴- استنتاج‌ها را به‌تنهایی و در ترکیب، گزارش و توجیه کنید	
۱۰۰	۲۴/۴	۷۵/۶	درصد		
۴۵	۱۱	۳۴	فراوانی	۵- در صورت مناسب بودن، به‌کارگیری اندازه‌گیری‌های چندگانه برای مخاطب قرار دادن هر سؤال	
۱۰۰	۲۴/۴	۷۵/۶	درصد		
۴۵	۴۵	۰	فراوانی	۶- مستند کردن و گزارش شرایط و فرایند جمع‌آوری داده‌ها	
۱۰۰	۱۰۰	۰	درصد		

نتایج جدول (۲) نشان داد ، همه نشانگرهای اول و دوم را اشاعه داده‌اند. تعداد ۳۵ مقاله یعنی ۷۷/۸ درصد نشانگر سوم را اشاعه داده‌اند، تعداد ۳۴ مقاله یعنی ۷۵/۶ درصد نشانگرهای چهارم و پنجم را اشاعه داده‌اند و هیچ کدام نشانگر ششم اشاعه نداده‌اند.



جدول (۳): یافته‌های توصیفی مربوط به اشاعه نشانگرهای استاندارد اطلاعات پایا

استاندارد	نشانگرها	شاخص آماری	مشاهده شواهد	عدم مشاهده شواهد	کل
اطلاعات پایا	۱- شناسایی و توجیه انواع و میزان پایایی ادعا شده	فراوانی	۲۴	۲۱	۴۵
		درصد	۵۳/۳	۴۶/۷	۱۰۰
	۲- در صورت امکان، وسیله اندازه‌گیری‌ای انتخاب کنید که در گذشته سطح قابل اطمینان پایایی برای کاربردهای قصد شده، داشته است	فراوانی	۸	۳۷	۴۵
		درصد	۱۷/۸	۸۲/۲	۱۰۰
	۳- در گزارش پایایی ابزار، سنجش و گزارش عوامل مؤثری که تأثیر گذاشته در پایایی، شامل ویژگی‌های داوطلبان، شرایط جمع‌آوری داده‌ها و سوگیری‌های ارزیاب می‌شود	فراوانی	۰	۴۵	۴۵
		درصد	۰	۱۰۰	۱۰۰
	۴- بررسی و گزارش ثبات نمره‌گذاری، طبقه‌بندی و کدگذاری	فراوانی	۳۹	۶	۴۵
		درصد	۸۶/۷	۱۳/۳	۱۰۰
	۵- آموزش دادن و درجه‌بندی کردن امتیاز دهندگان و تحلیل گران تا نتایج محکم ارائه دهند	فراوانی	۰	۴۵	۴۵
		درصد	۰	۱۰۰	۱۰۰
۶- اجرای مقدماتی ابزار جدید به منظور شناسایی و کنترل منابع خطا	فراوانی	۰	۴۵	۴۵	
	درصد	۰	۱۰۰	۱۰۰	

نتایج جدول (۳) نشان داد، تعداد ۲۴ مقاله یعنی ۵۳/۳ درصد نشانگر اول را اشاعه داده‌اند، تعداد ۸ مقاله یعنی ۱۷/۸ درصد نشانگر دوم را اشاعه داده‌اند، تعداد ۳۹ مقاله یعنی ۸۶/۷ درصد نشانگر چهارم را اشاعه داده‌اند و هیچ کدام نشانگرهای سوم، پنجم و ششم را اشاعه نداده‌اند.



جدول (۴): یافته‌های توصیفی مربوط به اشاعه نشانگرهای استاندارد برنامه روشن و توصیف موقعیت

کل	عدم مشاهده شواهد	مشاهده شواهد	شاخص آماری	نشانگرها	استاندارد
۴۵	۴۵	۰	فراوانی	۱- جمع‌آوری توصیفات برنامه‌ی قصد شده از منابع کتبی مختلف و جمع‌آوری توصیفات برنامه‌ی قصد شده از مشتری و دیگر ذی‌نفعان کلیدی	برنامه روشن و توصیف موقعیت
۱۰۰	۱۰۰	۰	درصد		
۴۵	۴۵	۰	فراوانی	۲- حفظ سوابق از منابع مختلفی که برنامه چطور عمل کرده است	
۱۰۰	۱۰۰	۰	درصد		
۴۵	۰	۴۵	فراوانی	۳- توصیف اینکه برنامه چطور واقعاً عمل کرده است	
۱۰۰	۰	۱۰۰	درصد		
۴۵	۴۴	۱	فراوانی	۴- تجزیه و تحلیل بین اینکه برنامه قصد داشت تا چگونه عمل کند و اینکه برنامه واقعاً چطور عمل کرد	
۱۰۰	۹۷/۸	۲/۲	درصد		
۴۵	۴۵	۰	فراوانی	۵- ثبت اینکه تا چه حد اهداف برنامه در طول زمان تغییر کرده است	
۱۰۰	۱۰۰	۰	درصد		
۴۵	۴۵	۰	فراوانی	۶- پرسیدن از مشتریان و ذی‌نفعان مختلف برای سنجش دقت توصیف‌های ثبت شده از هر دو، هم برنامه قصد شده و هم برنامه واقعی (که اجرا شده)	
۱۰۰	۱۰۰	۰	درصد		

نتایج جدول (۴) نشان داد، همه نشانگر سوم را اشاعه داده‌اند، یک مقاله یعنی ۲/۲ درصد نشانگر چهارم را اشاعه داده است و هیچ کدام نشانگرهای اول، دوم، پنجم و ششم را اشاعه نداده‌اند.



جدول (۵): یافته‌های توصیفی مربوط به اشاعه نشانگرهای استاندارد مدیریت اطلاعات

استاندارد	نشانگرها	شاخص آماری	مشاهده شواهد	عدم مشاهده شواهد	کل
مدیریت اطلاعات	۱- مستندسازی، توجیه و گزارش ابزار استفاده شده برای به دست آوردن اطلاعات از هر منبع	فراوانی	۲۶	۱۹	۴۵
		درصد	۵۷/۸	۴۲/۲	۱۰۰
	۲- آوردن ابزار جمع‌آوری داده‌ها در پیوست فنی در گزارش	فراوانی	۰	۴۵	۴۵
		درصد	۰	۱۰۰	۱۰۰
	۳- مستندسازی و گزارش هر ویژگی سوگیرانه در اطلاعات به دست آمده	فراوانی	۰	۴۵	۴۵
		درصد	۰	۱۰۰	۱۰۰
	۴- نظام‌دار کردن و کنترل ذخیره‌سازی اطلاعات ارزشیابی	فراوانی	۴۵	۰	۴۵
		درصد	۱۰۰	۰	۱۰۰
	۵- استفاده کردن از اطلاعات مناسب که قبلاً جمع‌آوری شده بعد از تأیید اعتبار	فراوانی	۰	۴۵	۴۵
		درصد	۰	۱۰۰	۱۰۰
۶- در صورت مناسب بودن، به کار گرفتن روش‌ها و منابع مختلف برای جمع‌آوری داده‌ها	فراوانی	۳۴	۱۱	۴۵	
	درصد	۷۵/۶	۲۴/۴	۱۰۰	

نتایج جدول (۵) نشان داد، همه نشانگر چهارم را اشاعه داده‌اند، تعداد ۲۶ مقاله یعنی ۵۷/۸ درصد نشانگر اول را اشاعه داده‌اند، تعداد ۳۴ مقاله یعنی ۷۵/۶ درصد نشانگر ششم را اشاعه داده‌اند و هیچ کدام نشانگرهای دوم، سوم و پنجم را اشاعه نداده‌اند.

جدول (۶): یافته‌های توصیفی مربوط به اشاعه نشانگرهای استاندارد تجزیه و تحلیل و طراحی‌های درست

استاندارد	نشانگرها	شاخص آماری	مشاهده شواهد	عدم مشاهده شواهد	کل
تجزیه و تحلیل و طراحی‌های درست	۱- ویرایش و بررسی جداول تولید شده از خروجی کامپیوتر یا هر وسیله‌ی دیگر	فراوانی	۴۵	۰	۴۵
		درصد	۱۰۰	۰	۱۰۰
	۲- آیا ارائه‌دهندگان داده، اطلاعاتی که ارائه دادند را بررسی کرده‌اند	فراوانی	۰	۴۵	۴۵
		درصد	۰	۱۰۰	۱۰۰



۴۵	۴۵	۰	فراوانی	۳- آموزش کارکنان ارزشیابی برای پابندی به پروتکل های داده
۱۰۰	۱۰۰		درصد	
۴۵	۲۰	۲۵	فراوانی	۴- به طور نظام دار بررسی دقت کدگذاری و نمره گذاری
۱۰۰	۴۴/۴	۵۵/۶	درصد	
۴۵	۰	۴۵	فراوانی	۵- انتخاب فرایند مناسب برای سؤال های ارزشیابی و ماهیت داده ها
۱۰۰	۰	۱۰۰	درصد	
۴۵	۴۵	۰	فراوانی	۶- بررسی دقت یافته ها با به دست آوردن شواهد تأییدی از منابع متعدد از جمله ذی نفعان
۱۰۰	۱۰۰	۰	درصد	

نتایج جدول (۶) نشان داد ، همه نشانگرهای اول و پنجم را اشاعه داده اند، تعداد ۲۵ مقاله یعنی ۵۵/۶ درصد نشانگر چهارم را اشاعه داده اند و هیچ کدام نشانگرهای دوم، سوم و ششم را اشاعه نداده اند.

جدول (۷): یافته های توصیفی مربوط به اشاعه نشانگرهای استاندارد استدلال واضح (روشن) ارزشیابی

کل	عدم مشاهده شواهد	مشاهده شواهد	شاخص آماری	نشانگرها	استاندارد
۴۵	۱۷	۲۸	فراوانی	۱- به کارگیری فرایند تحلیلی متعدد برای بررسی ثبات و تکرار یافته ها	استدلال واضح (روشن) ارزشیابی
۱۰۰	۳۷/۸	۶۲/۲	درصد		
۴۵	۴۵	۰	فراوانی	۲- بررسی هم بی ثباتی هم گرایش های مرکزی (آمار توصیفی)	
۱۰۰	۱۰۰	۰	درصد		
۴۵	۰	۴۵	فراوانی	۳- شناسایی و تجزیه و تحلیل روابط متقابل آماری	
۱۰۰	۰	۱۰۰	درصد		
۴۵	۰	۴۵	فراوانی	۴- مشخص کردن اینکه چگونه به فرضیات کلیدی هر فرایند پاسخ داده شده است	
۱۰۰	۰	۱۰۰	درصد		
۴۵	۱۰	۳۵	فراوانی	۵- استخراج نتایج و پیشنهادات و نشان دادن معنی هایشان	
۱۰۰	۲۲/۲	۷۷/۸	درصد		
۴۵	۴۴	۱	فراوانی	۶- گزارش محدودیت های اطلاعات، تجزیه و تحلیل و استنتاج های اشاره شده	
۱۰۰	۹۷/۸	۲/۲	درصد		



نتایج جدول (۷) نشان داد ، همه نشانگرهای سوم و چهارم را اشاعه داده‌اند، تعداد ۲۸ مقاله یعنی ۶۲/۲ درصد نشانگر اول را اشاعه داده‌اند، تعداد ۳۵ مقاله یعنی ۷۷/۸ درصد نشانگر پنجم را اشاعه داده‌اند، یک مقاله یعنی ۲/۲ درصد نشانگر پنجم را اشاعه داده است و هیچ کدام نشانگر دوم را اشاعه نداده‌اند.

جدول (۸): یافته‌های توصیفی مربوط به اشاعه نشانگرهای استاندارد ارتباط و گزارش

استاندارد	نشانگرها	شاخص آماری	مشاهده شواهد	عدم مشاهده شواهد	کل
ارتباط و گزارش	۱- درگیر کردن مشتریان به منظور تعیین مراحل برای حصول اطمینان از گزارش‌های عادلانه و بی طرفانه	فراوانی	۰	۴۵	۴۵
		درصد	۰	۱۰۰	۱۰۰
	۲- حفاظت گزارش‌ها از تحریفات عمدی یا غیرعمدی	فراوانی	۰	۴۵	۴۵
		درصد	۰	۱۰۰	۱۰۰
	۳- گزارش دیدگاه‌های همه گروه‌های ذی نفعان	فراوانی	۴۵	۰	۴۵
		درصد	۱۰۰	۰	۱۰۰
	۴- شرح اقدامات انجام گرفته شده برای کنترل سوگیری	فراوانی	۰	۴۵	۴۵
		درصد	۰	۱۰۰	۱۰۰
	۵- تدوین گویا و مناسب گزارش	فراوانی	۴۵	۰	۴۵
		درصد	۱۰۰	۰	۱۰۰
	۶- تعیین مخاطبانی که به درستی شناخته شده‌اند	فراوانی	۴۵	۰	۴۵
		درصد	۱۰۰	۰	۱۰۰

نتایج جدول (۸) نشان داد ، همه نشانگرهای سوم، پنجم و ششم را اشاعه داده‌اند و هیچ کدام نشانگرهای اول، دوم و چهارم را اشاعه نداده‌اند.



جدول (۹): یافته‌های توصیفی مربوط به اشاعه استانداردهای معیار دقت

معیار	استانداردها	شاخص آماری	عالی	خیلی خوب	خوب	متوسط	ضعیف	کل
۳	۱- تصمیمات و نتایج موجه	فراوانی	۰	۰	۱	۴۴	۰	۴۵
		درصد	۰	۰	۲/۲	۹۷/۸	۰	۱۰۰
	۲- اطلاعات معتبر	فراوانی	۰	۱۹	۲۰	۶	۰	۴۵
		درصد	۰	۴۲/۲	۴۴/۴	۱۳/۳	۰	۱۰۰
	۳- اطلاعات پایا	فراوانی	۰	۰	۰	۲۵	۲۰	۴۵
		درصد	۰	۰	۰	۵۵/۶	۴۴/۴	۱۰۰
	۴- برنامه روشن و توصیف موقعیت	فراوانی	۰	۰	۰	۱	۴۴	۴۵
		درصد	۰	۰	۰	۲/۲	۹۷/۸	۱۰۰
	۵- مدیریت اطلاعات	فراوانی	۰	۰	۰	۴۰	۵	۴۵
		درصد	۰	۰	۰	۸۸/۹	۱۱/۱	۱۰۰
	۶- تجزیه و تحلیل و طراحی‌های درست	فراوانی	۰	۰	۰	۴۵	۰	۴۵
		درصد	۰	۰	۰	۱۰۰	۰	۱۰۰
	۷- استدلال واضح (روشن) ارزشیابی	فراوانی	۰	۰	۰	۲۲	۰	۴۵
		درصد	۰	۰	۰	۴۸/۹	۵۱/۱	۱۰۰
	۸- ارتباط و گزارش	فراوانی	۰	۰	۰	۴۵	۰	۴۵
		درصد	۰	۰	۰	۱۰۰	۰	۱۰۰

نتایج جدول (۹) نشان داد، در رابطه با استاندارد اول یک مقاله یعنی ۲/۲ درصد در سطح خوب و تعداد ۴۴ مقاله یعنی ۹۷/۸ درصد در سطح متوسط ارزشیابی شدند. در رابطه با استاندارد دوم تعداد ۱۹ مقاله یعنی ۴۲/۲ درصد در سطح بسیار خوب، تعداد ۲۰ مقاله یعنی ۴۴/۴ درصد در سطح خوب و تعداد ۶ مقاله یعنی ۱۳/۳ درصد در سطح متوسط ارزشیابی شدند. در رابطه با استاندارد سوم تعداد ۲۵ مقاله یعنی ۵۵/۶ درصد در سطح متوسط و تعداد ۲۰ مقاله یعنی ۴۴/۴ درصد در سطح ضعیف ارزشیابی شدند. در رابطه با استاندارد چهارم یک مقاله یعنی ۲/۲ درصد در سطح متوسط و تعداد ۴۴ مقاله یعنی ۹۷/۸ درصد در سطح ضعیف ارزشیابی شدند. در رابطه با استاندارد پنجم تعداد ۴۰ مقاله یعنی ۸۸/۹ درصد در سطح متوسط و تعداد ۵ مقاله یعنی ۱۱/۱ درصد در سطح ضعیف ارزشیابی شدند. در رابطه با استاندارد هفتم تعداد ۲۳ مقاله یعنی ۵۱/۱ درصد در سطح خوب و تعداد ۲۲ مقاله یعنی ۴۸/۹ درصد در سطح متوسط ارزشیابی شدند. در رابطه با استاندارد هشتم همه در سطح متوسط ارزشیابی شدند.



بحث و نتیجه گیری

هدف این مطالعه میزان اشاعه استانداردهای معیار دقت در ارزشیابی های درونی گروه های آموزشی دانشگاه ها و مؤسسات آموزش عالی بود. با عنایت به اینکه رعایت استانداردهای ارزشیابی در پژوهش ها یکی از عوامل مهم بالا بردن کیفیت پژوهشی می باشد و مراحل ارزشیابی شامل: درگیر کردن ذی نفعان، توصیف برنامه، تمرکز روی طرح ارزشیابی، جمع آوری شواهد معتبر، نتیجه گیری های موجه، اطمینان از استفاده و به اشتراک گذاری (کاپلان، میلستین و وترحال^{۱۱})، می باشد رعایت استانداردهای دقت در هر کدام از مراحل از اهمیت بسزایی برخوردار است باین حال یافته ها نشان داد که پژوهشگران کمتر توجهی به مستند کردن فرایند کارشان داشته اند. شاید استانداردهای دقت را در ارزشیابی شان در نظر گرفته باشند ولی در گزارش کمتر اشاره ای به آن ها کرده اند. با توجه به اینکه استانداردهای معیار دقت به راست گویی، اعتبار، قابلیت اعتماد بودن یافته های ارزشیابی، تفسیر و قضاوت درباره کیفیت اشاره دارند و همچنین به طرح ها، روش ها، جمع آوری داده ها، تجزیه و تحلیل ها، پایایی، اعتبار، خطاها و سوگیری ها و منطق نتیجه گیری ها می پردازند (روهه و بودریو، ۲۰۱۳) از اهمیت بالایی برخوردار هستند و با در نظر گرفتن این استانداردها در پژوهش های ارزشیابی درونی گروه های آموزشی اعتبار آن ها افزایش می یابد تا بتوان به یافته ها و نتایج اعتماد کرد و در تصمیم گیری ها از آن ها بهره برد.

بر اساس یافته های حاصل از این پژوهش پیشنهاد می شود که نظام آموزشی برای مقابله با این مشکل و ارتقای عملکرد کیفی پژوهش های ارزشیابی، پژوهشگران را نسبت به استانداردهای ارزشیابی آشنا کند. بر این اساس، آگاه کردن پژوهشگران جهت ارتقای عملکرد کیفی پژوهشی و چگونگی استفاده از آن ها در موقعیت های گوناگون ضروری است.

منابع

- جعفری قوام آباد، ا؛ خامی، م؛ یزدانی، ر؛ نصیبی، م؛ محمدی، م (۱۳۹۲). ارزیابی درونی گروه سلامت دهان و دندان پزشکی اجتماعی؛ دانشگاه علوم پزشکی تهران، راهبردهای آموزش در علوم پزشکی، ۶(۲): ۱۱۲-۱۰۸.
- حکیمی، و (۱۳۹۰). آسیب شناسی طرح های ارزیابی درونی و راه های توسعه اثربخشی آن، پایان نامه کارشناسی ارشد، دانشکده علوم تربیتی و روانشناسی، دانشگاه فردوسی مشهد.
- رحیمی فرد، ه؛ کوهپایی، ع؛ ارست، ی؛ بهنامی پور؛ س؛ مهدی نیا؛ م (۱۳۹۲). ارزیابی درونی گروه آموزشی حرفه ای دانشگاه علوم پزشکی قم، سال تحصیلی ۱۳۹۰-۱۳۸۹، مجله دانشگاه علوم پزشکی قم، ۷(۲): ۹۱-۸۶.
- زین آبادی، ح؛ کیامنش، ع؛ فرزاد، و.ا (۱۳۸۴). ارزیابی درونی کیفیت گروه مشاوره و راهنمایی دانشگاه تربیت معلم تهران به منظور پیشنهاد الگویی جهت بهبود کیفیت و حرکت در جهت اعتباربخشی گروه های مشاوره و راهنمایی کشور، پژوهش های مشاوره، ۴(۱۵): ۹۲-۶۹.
- عابدینی، ز؛ خرمی راد، ا؛ آخوندزاده، ک؛ حیدری، س (۱۳۹۲). ارزیابی درونی گروه آموزشی پرستاری دانشکده پرستاری و مامایی، دانشگاه علوم پزشکی قم، مجله دانشگاه علوم پزشکی قم، ۷(۲): ۷۱-۶۴.
- عباس پور، ع؛ شرفی، م (۱۳۹۰). ارزیابی درونی در گروه مدیریت و برنامه ریزی آموزشی دانشگاه علامه طباطبایی، فصلنامه اندازه گیری تربیتی، ۵(۲): ۱۵۹-۱۳۱.

Koplan, J. P., Milstein, R., & Wetterhall, S. (1999). Framework for program evaluation in public health. MMWR: Recommendations and Reports, 48, 1-40.

¹¹ . Koplan, Milstein & Wetterhall



- Madzivhandila, T., Griffith, G. R., Fleming, E. M., & Nesamvuni, A. (2010), Meta-evaluations in government and government institutions: A case study example from the Australian Centre for International Agricultural Research. Paper presented at the Conference (54th), February 10-12, Adelaide, Australia.
- Ruhe, V., & Boudreau, J. D. (2013). The 2011 Program Evaluation Standards: a framework for quality in medical education programme evaluations. *Journal of evaluation in clinical practice*, 19(5), 925-932.
- Scriven, M. (2008). Meta-evaluation revisited. *Journal of MultiDisciplinary Evaluation*, 6(11), iii-viii.
- Stufflebeam, D. L. (2001). The metaevaluation imperative. *The American Journal of Evaluation*, 22(2), 183-209.
- Stufflebeam, D. L., Madaus, G. F., & Kellaghan, T. (2000). *Evaluation models: Viewpoints on educational and human services evaluation (Vol. 49)*: Springer.