مقایسه کیفیت سؤالات آزمون ارتقای دستیاری قبل و بعد از مداخله آموزشی در دانشکده پزشکی گیلان طی سال‌های ۱۳۸۹ و ۱۳۹۰

حسین خوشرنگ، ماهدخت طاهری*، عطاا... اسدی، آبتین حیدرزاده

چکیده

مقدمه: آزمون ارتقاء، یکی از مهم‌ترین آزمون‌های تراکمی است که در مقطع دستیاری انجام می‌شود و منجر به صدور مجوز ورود دستیاران به سال بالاتر می‌گردد، آزمون‌های چندگزینه‌ای از عینی‌ترین و رایج‌ترین انواع آزمون در آموزش علوم پزشکی هستند، بنابر این یافتن راه‌کارهایی برای بهبود کیفیت آزمون‌ها به خصوص در مقاطع تخصصی اهمیت دارد. هدف این مطالعه بررسی تأثیر کارگاه آموزشی طراحی سؤالات چندگزینه‌ای بر بهبود کیفیت طراحی سؤالات آزمون ارتقای دستیاری دانشکده پزشکی دانشگاه علوم پزشکی گیلان است.

روش‌ها: در این مطالعه نیمه تجربی (قبل و بعد)، کلیه سؤالات آزمون‌های ارتقای دستیاری دانشکده پزشکی دانشگاه علوم پزشکی گیلان، مربوط به ۱۲ رشته تخصصی در سال ۸۹ از نظر ساختار طراحی و سطح تاکسونومی با استفاده از چک لیستی روا و پایا مبتنی بر اصول میلمن بررسی شد (۱۵۰ سؤال از هر رشته و در مجموع ۱۸۰۰ سؤال). بعد از برگزاری ۱۲ دوره کارگاه آموزشی یک روزه طراحی سؤالات چندگزینه‌ای (MCQ) جهت طراحان آزمون (به تفکیک گروه‌های آموزشی) در طی یک سال، سؤالات آزمون ارتقای سال ۹۰ نیز مورد بررسی قرار گرفت. داده‌های قبل و بعد از مداخله مربوط به کلیه سؤالاتی که توسط افراد شرکت‌کننده در کارگاه مداخله‌ای طراحی شده بود با استفاده از آزمون آماری Wilcoxon signed-rank test مقایسه گردید.

نتایج: در آزمون کتبی سال ۱۳۸۹، ۱۲۰۴سؤال (۶۶/۰۹ درصد) تاکسونومی یک، ۵۶۵ سؤال (۳۱/۴ درصد) تاکسونومی دو، و ۳۱ سؤال(۱/۷درصد) تاکسونومی سه داشتند. در سال ۱۳۹۰، ۸۶۷ سؤال (۴۸/۲٪) در تاکسونومی یک، ۷۴۱ سؤال(۴۱/۲٪) تاکسونومی دو، و ۱۹۲ سؤال(۱۰/۶٪) تاکسونومی سه طراحی شده بودند. تفاوت معناداری بین تاکسونومی سؤالات رشته‌های مختلف در سال‌های ۸۹ و ۹۰ مشاهده شد (p<۰/۰۰۰۱, z=۲۳/۸۳). در زمینه رعایت اصول ساختاری میلمن در سؤالات چندگزینه‌ای، یافته‌ها نشان داد که نسبت سؤالات بدون اشکال در آزمون سال ۱۳۸۹، برابر ۶۶/۷۷٪ و در سال ۱۳۹۰، ۷۵/۲۲٪ بود. از نظر مراعات اصول میلمن بین کیفیت سوالات آزمون سال‌های ۸۹ و ۹۰ تفاوت معنادار آماری وجود دارد (z=۱۶/۷۰۳، p<۰/۰۰۰۱).

نتیجه‌گیری: مطالعه حاضر پیشرفت معناداری را در کاربرد سؤالات مربوط به سطوح عمیق دانش (تاکسونومی ۲ و ۳) در آزمون ارتقای دستیاری نشان داد که بیانگر تأثیر مثبت مداخله آموزشی انجام شده (کارگاه آموزشی MCQ) بود.

واژه‌های کلیدی: آزمون‌های چندگزینه‌ای، آزمون ارتقای دستیاری، تاکسونومی، آموزش استاد، کارگاه‌های آموزشی

مجله ایرانی آموزش در علوم پزشکی/ مهر ۱۳۹۲؛ ۱۳(۷): ۵۵۱ تا ۵۶۰

مقدمه

آموزش تخصصی و فوق تخصصی پزشکی بنابر نقش ویژه دستیاران در نظام آموزشی-درمانی دانشگاه‌ها و همچنین نقش بسیار مهم دانش‌آموختگان در نظام

*نویسنده مسؤول: دکتر ماهدخت طاهری (استادیار)، مرکز مطالعات و توسعه آموزش علوم پزشکی، دانشگاه علوم پزشکی گیلان، رشت، ایران. m_taheri@gums.ac.ir

دکتر حسین خوشرنگ (دانشیار)، گروه بیهوشی، دانشکده پزشکی، دانشگاه علوم پزشکی گیلان، رشت، ایران. (khoshrang@gums.ac.ir)؛ عطاالله اسدی (مربی)، گروه هوشبری، دانشکده پرستاری و مامایی شهید بهشتی، دانشگاه علوم پزشکی گیلان، رشت، ایران. (aa.louyeh@yahoo.com)؛ دکتر آبتین حیدرزاده، (دانشیار)، گروه پزشکی اجتماعی، دانشکده پزشکی، دانشگاه علوم پزشکی گیلان، رشت، ایران. (heidarzadeh@gums.ac.ir)

تاریخ دریافت مقاله: ۹۱/۷/۱۷، تاریخ اصلاحیه: ۹۲/۲/۱۵، تاریخ پذیرش: ۹۲/۴/۱۹

[Downloaded from ijme.mui.ac.ir on 2026-11-26]

سلامت جامعه، از اهمیت کلیدی در نظام آموزش عالی کشور برخوردار است و از این رو تحول و ارتقای کیفیت آن‌ها از دغدغه‌های جدی مسئولین اجرایی و صاحب‌نظران است. در این بین همانند هر سیستم آموزشی، ارزیابی فراگیران جزء مهمی از آموزش است ودر میان اجزاء مختلف آموزش تخصصی، آزمون‌ها به عنوان یکی از بخش‌های تأثیرگذار و جهت‌دهنده به روند کلی آموزش از جایگاه ویژه‌ای برخوردار است(۱). آزمون‌ها، مهم‌ترین وسیله‌اندازه‌گیری پیشرفت تحصیلی دانشجویان محسوب می‌شود(۲)؛ به طوری که هم باعث ایجاد انگیزه یادگیری می‌شوند و هم ابزاری هستند برای حصول اطمینان از این‌که دانش‌آموختگان به سطح مورد نظر از دانش و توانمندی رسیده‌اند. بدین جهت برنامه‌های ارتقای کیفیت و استانداردسازی آزمون‌ها از اثربخشی بالایی برخوردار است(۱). مطالعه تارانت (Tarrant) و همکاران در دانشکده پزشکی دانشگاه هنگ کنگ در سال ۲۰۰۶ میلادی ، نشان داد که حدود نیمی از سؤالات چندگزینه‌ای طراحی شده توسط اعضای هیأت‌علمی دارای اشکال ساختاری بوده و عمدتاً سطوح پایین شناختی را مورد سنجش قرار داده بودند(۳).

در این زمینه شواهد متعددی از جمله مطالعه Reid و همکاران، مطالعه Larsen و همکاران و برخی مطالعات دیگر نشان داده‌اند که کیفیت آزمون‌ها، به طور اساسی کیفیت آموزش و روند یاددهی–یادگیری را تحت تاثیر قرار می‌دهند(٦تا٤). مطالعه معیاری و همکاران نشان داد که برگزاری برنامه‌های آموزش طراحی آزمون‌های چندگزینه‌ای به شرط آن که به صورتی هدف‌مند بوده و در آن‌ها بر اساس اصول فنی مربوطه با تمرکز بر تقویت نقاط قوت و رفع نقاط ضعف آزمون‌های قبلی همان طراحان مد نظر قرار گیرد، می‌تواند در ارتقای کیفیت طراحی آزمون‌های دستیاری مؤثر باشد(۷)

یکی از مهم‌ترین آزمون‌های تراکمی که در مقطع دستیاری از فراگیران به عمل می‌آید و منجر به صدور مجوز ورود دستیاران به سال بالاتر می‌گردد، آزمون ارتقا است. این آزمون که تا دو سال قبل به صورت متمرکز و در سطح ملی برگزار می‌شد، در حال حاضر به صورت درون دانشگاهی برگزار می‌گردد(۸). در ایران آزمون دانشنامه از دو بخش کتبی و شفاهی تشکیل یافته که آزمون کتبی آنرا پرسش‌های چندگزینه‌ای (Multiple Choice Question MCQ) تشکیل می‌دهد(۹). این آزمون همانند دیگر آزمون‌ها دارای معایب و محاسنی است که آشنایی با آنها و تلاش در رفع معایب آن می‌تواند تا حد زیادی این آزمون‌ها را بهبود بخشد. حق‌شناس در مطالعه‌ای به بررسی آزمون‌های چندگزینه‌ای دانشکده پزشکی دانشگاه علوم پزشکی مازندران پرداخته و به این نتیجه رسید که کیفیت آزمون‌های چندگزینه‌ای مورد مطالعه از نظر توزیع تاکسونومی و رعایت اصول ساختاری MCQ بین آزمون‌های مقاطع مختلف متفاوت بوده و در برخی موارد نیازمند اتخاذ تدابیر اصلاحی است(۱۰). منظور از تاکسونومی سؤال، طبقه‌بندی است که در آن سطح دانش آزمون شونده سنجیده می‌شود و مراحل آن عبارت است از: یادآوری ساده مطالب (تاکسونومی۱)، تفسیر اطلاعات (تاکسونومی ۲) و حل مسأله و کاربرد مطالب (تاکسونومی۳)(۷). اصول میلمن اصولی است که برای اولین بار میلمن جهت نحوه تدوین طراحی ساقه (Stem) و گزینه‌ها (Option) اعلام کرد. این اصول شامل: درج اعظم اطلاعات سؤال در ساقه سؤال، پیگیری هر هدف یادگیری با یک سؤال، استفاده از لغات ساده و روشن، عدم کاربرد گزینه منفی در ساقه سؤال، عدم کاربرد گزینه همه موارد یا هیچ‌کدام، عدم کاربرد گزینه متضاد، مشخص کردن کلمات منفی، استقلال سؤالات از یکدیگر، هم‌سنگی سؤالات از لحاظ طول و ساختار لغوی، عدم کاربرد کلمات تکراری، نداشتن غلط املایی ودرج گزینه‌ها به صورت عمودی می‌باشند(جدول ۱).

گرچه آزمون‌های چند گزینه‌ای به سبب برخورداری از

[Downloaded from ijme.mui.ac.ir on 2026-11-26]

حداکثر عیینت(Objectivity) و در نتیجه پایایی (Reliability) بالا و سهولت نسبی اجرا به‌ویژه در مقیاس‌های بزرگ، واجد امتیازات ویژه‌ای برای آزمون‌های مهم آموزش پزشکی می‌باشند، اما همان طور که Kaufman و همکاران و Van der Vlueten و همکاران نشان دادند نگرانی‌های قابل توجهی در مورد برخی شاخص‌های کیفیت آزمون‌های چند گزینه‌ای شامل سطح دانش بکار رفته (تاکسونومی)، رعایت اصول ساختاری و همچنین شاخص‌های پس آزمونی وجود دارد(۱۱و۱۲). وحیدشاهی و همکاران تأثیر مثبت برنامه ارتقای کیفیت آزمون دانشنامه و بهبود معنادار سطح دانش به کار رفته (تاکسونومی) و رعایت اصول ساختاری میلمن در سؤالات را اثبات نمودند(۱۳).

همچنین فریدمن (Friedman) از دانشگاه داندی طی مطالعه‌ای ارزش و اهمیت ارتقای کیفیت آزمون‌های چندگزینه‌ای برای ارزیابی سطح دانش به کار رفته (تاکسونومی) و نیز رعایت اصول ساختاری را مؤثر دانسته است(١٤). تعیین سطح دانش به کار گرفته و اشکال‌های ساختاری سؤال‌های چندگزینه‌ای و بررسی علت‌ها و فراوانی آن‌ها موضوع با اهمیتی است که می‌تواند به کاهش نقاط ضعف و تقویت نقاط قوت در آزمون‌های چندگزینه‌ای و در نهایت به بهبود کیفیت ارزشیابی در دانشگاه منجر شود.

ارزش و اهمیت این گونه مطالعات و بازخورد آن‌ها در زمینه ارتقای کیفیت آزمون‌های چندگزینه‌ای در مطالعات متعدد داخلی محققی و همکاران(۹) و خوشکام و همکاران(۱۵) و رمضانی و همکاران(۱۸) و برخی محققان خارجی (٦و١٤و١٩و٢٠) نشان داده شده است. با بررسی نتایج به دست آمده، انجام مداخلات مناسب و مقتضی در راستای ارتقای سطح کیفی سؤالات چندگزینه‌ای ضروری به نظر می‌رسد. لذا هدف از این مطالعه مقایسه کیفیت سؤالات آزمون ارتقای دستیاری قبل و بعد از مداخله آموزشی در دانشکده پزشکی گیلان طی سال‌های ۱۳۸۹ و۱۳۹۰ به منظور ارائه بازخورد به گروه‌های آموزشی و طراحان سؤالات بود.

روش‌ها

این تحقیق یک مطالعه نیمه تجربی از نوع قبل و بعد بود. در این تحقیق کلیه سؤالات آزمون‌های ارتقا گروه‌های دستیاری در سال‌های ۱۳۸۹ و ۱۳۹۰ مورد ارزیابی قرار گرفت (مجموعاً ٣٦٠٠ سؤال). این سؤال‌های چهار گزینه‌ای مربوط به آزمون‌های کتبی دانشنامه تخصصی در ۱۲ رشته تخصصی (جراحی عمومی، داخلی، رادیولوژی، گوش و حلق بینی جراحی سرو گردن، زنان، روانپزشکی، نورولوژی، پوست، قلب، کودکان، ارولوژی، بیهوشی) بود.

این پژوهش در دو گام و چند مرحله طراحی شد:

گام اول: ابتدا هر سؤال آزمون کتبی ارتقای دستیاری گروه‌های آموزشی در سال ۸۹ از نظر تاکسونومی و رعایت اصول میلمن بررسی شد و به ازای هر سؤال یک برگه و به ازای هر آزمون یک دفترچه تهیه گردید. نتایج ارزیابی در یک database وارد و تجزیه و تحلیل شد. برای انجام این گام متون مربوط به ارزیابی آزمون‌های چندگزینه‌ای مطالعه شد و ابزار ارزیابی کیفیت سؤالات آزمون طراحی گردید. تاکسونومی سؤالات با توجه به سطوح شناختی بلوم(۱۱و۲۱) در سه سطح الف) یادآوری، ب)کاربرد و تجزیه و تحلیل، و ج) مقایسه و تشخیص (حل مسأله) بررسی گردید. همچنین رعایت اصول ساختاری در ساقه و گزینه‌های سؤالات بر اساس چک‌لیست میلمن کنترل شد(۷و۱۱و۲۱) (جدول ۱). با توجه به استاندارد بودن چک لیست و بومی‌سازی آن و استفاده مکرر محققین داخلی(۷و۲۱)، روایی آزمون مورد تأیید بود. برای تأمین پایایی، ابتدا جلسات هماهنگی بین پژوهشگران و مصححین برگزار شد و توافق نظر اولیه در مورد نحوه بررسی سؤالات و تکمیل چک لیست به عمل آمد. با بررسی ۱۵ درصد از سؤالات

[Downloaded from ijme.mui.ac.ir on 2026-11-26]

سال ۸۸ (۲۷۰ سؤال) توسط پژوهشگران، میزان هم‌خوانی بین ارزیابی‌کنندگان (Inter-rater=۰/۸۲) و در اندازه‌گیری مجدد برای هر ممتحن پس از حداقل یک هفته ۰/۹۶ برآورد شد. بنابر این پایایی آن نیز مورد تأیید قرار گرفت.

۲- بررسی تاکسونومی و مراعات اصول میلمن در سؤالات، توسط گروهی متشکل از اعضای هیأت‌علمی بالینی که هم در عرصه بالین و هم در عرصه آموزش پزشکی خبره هستند انجام گرفت. هر آزمون توسط ۳ نفر از اساتید به طور مستقل بررسی و نظر غالب (توافق دو نفر از اساتید) در مورد تاکسونومی هر سؤال و آیتم‌های چک لیست میلمن به عنوان نتیجه نهایی در مورد سؤال لحاظ گردید.

جدول ۱: چک لیست میلمن جهت بررسی رعایت اصول ساختاری در سؤالات چند گزینه ای

	موضوع	بلی/ خیر
۱	آیا بخش اعظم اطلاعات در ساقه سؤال گنجانده شده است؟	
۲	آیا سؤال یک هدف اختصاصی یادگیری را مورد ارزیابی قرار می‌دهد؟	
۳	آیا لغات استفاده شده در ساقه یا گزینه‌ها، شفاف و مستقیم بیان شده‌اند؟	
٤	آیا از کاربرد گزینه منفی برای ساقه منفی خودداری شده است؟	
٥	آیا از کاربرد گزینه‌هایی نظیر" همه موارد" هیچ‌کدام "و گزینه‌های ترکیبی خودداری شده است؟	
٦	آیا ازکاربرد گزینه‌های متضاد با یکدیگر خودداری شده است؟	
۷	آیا از لغات مثبت درساقه سؤال استفاده شده است یا در صورت منفی بودن ساقه سؤال، لغات منفی مشخص شده‌اند؟	
۸	آیا هر سؤال مستقل از سؤالات دیگر می‌باشد؟	
۹	آیا گزینه‌ها از نظر طول، ساختار لغوی و سبک نگارش هم‌سنگ هستند؟	
۱۰	آیا تا حد امکان از کاربرد عبارت‌های تکراری در گزینه‌ها خودداری شده است؟	
۱۱	آیا کلمات به کاررفته در ساقه و یا گزینه‌ها از نظر املایی صحیح هستند؟	
۱۲	آیا گزینه‌ها به طور عمودی لیست شده‌اند؟	

۳- اطلاعات پس از جمع‌آوری با استفاده از آمار توصیفی و Wilcoxon signed-rank test از طریق نرم‌افزار آماری SPSS.Ver.16 تجزیه و تحلیل گردید. سطح معناداری آزمون ۰/۰۵ در نظر گرفته شد. به دلیل این که گروه طراح سؤالات هر گروه مشترک بودند، یافته‌های قبل و پس از مداخله را نمی توان مستقل از هم دانست. از این‌رو، شرط استفاده از آزمون‌های غیر پارامتریک مثل خی دو فراهم نبود. بنابراین برای مقایسه تغییرات قبل و بعد از مداخله، و نیز مقایسه بین گروه‌های مختلف، از Wilcoxon signed-rank test استفاده شد.

٤- پس از تجزیه و تحلیل مجموعه سؤالات و مشخص شدن نقاط قوت و ضعف آزمون مزبور، نتایج ارزیابی هر سؤال آزمون کتبی دستیاری از نظر تاکسونومی ورعایت اصول میلمن، به طراحان سؤال بازخورد داده شد. پس از برگزاری جلسه مشترک با حضور طراحان سؤالات آزمون ارتقای دستیاری گروه‌های آموزشی، به تفکیک برای هر گروه آموزشی، و ارائه بازخورد سؤالات طراحی شده، به رفع اشکالات سؤالات و ارائه راه حل جهت طراحی سؤالات با رعایت اصول میلمن و تاکسونومی بالاتر پرداخته و نیز ارزیابان پمفلت آموزشی آزمون چندگزینه‌ای را تهیه و به طراحان ارائه کردند.

[Downloaded from ijme.mui.ac.ir on 2026-11-26]

گام دوم : این گام خود از چند مرحله به شرح زیر تشکیل شده بود:

۱- دوازده دوره کارگاه آموزشی یک روزه (٦ ساعته) با عنوان "نحوه نگارش سؤالات چندگزینه‌ای" با همکاری مرکز مطالعات و توسعه آموزش علوم پزشکی توسط پژوهشگران (که خود از اعضای مرکز مطالعات دانشگاه هستند) و با حضور فعال در گروه‌های آموزشی، به تفکیک رشته‌های تخصصی (هر هفته یک گروه آموزشی با هماهنگی مدیر گروه) برگزار گردید. در کارگاه در مورد آزمون‌های چندگزینه‌ای (MCQ)، ویژگی‌ها و شرایط یک آزمون چندگزینه‌ای استاندارد بحث گردید. علاوه بر سطوح یادگیری شناختی بلوم به بحث تاکسونومی و معیارها و اصول میلمن به طور مفصل پرداخته شد. همچنین سؤالات مربوط به هر یک از گروه‌های آموزشی به تفکیک در کارگاه مربوطه بررسی شد و پس از تعیین نقاط ضعف و قوت از نظر تاکسونومی و رعایت اصول میلمن و انجام کار عملی در گروه‌های کوچک و با ارائه مثال‌های عینی، بازخورد و آموزش مناسب به شرکت کنندگان داده شد.

۲- بررسی هر سؤال آزمون کتبی دانشنامه از نظر تاکسونومی، رعایت اصول میلمن (مطابق گام یک) پس از برگزاری آزمون دستیاران در تیر ماه ۱۳۹۰.

۳- ارسال بازخورد بررسی هر سؤال آزمون کتبی دانشنامه از نظر تاکسونومی، رعایت اصول میلمن به طراحان سؤال و برگزاری نشست با طراحان طبق برنامه تنظیمی بر اساس گروه‌های تخصصی.

٤- مقایسه سؤالات آزمون تخصصی دستیاران در سال‌های ۱۳۸۹ و ۱۳۹۰ از نظر"تاکسونومی، رعایت اصول میلمن" و بررسی تأثیر مداخله در نحوه طراحی سؤالات آزمون ۱۳۹۰.

نتایج

در مجموع ۳٦۰۰ سؤال (۱۸۰۰ سؤال از سال ۸۹ و ۱۸۰۰سؤال از سال ۹۰) مربوط به آزمون کتبی دستیاران ۱۲ رشته تخصصی مورد بررسی قرار گرفت. (۱۵۰ سؤال برای هر رشته تخصصی). بر اساس یافته ها، در آزمون کتبی سال ۸۹، تاکسونومی ۱۲۰٤ سؤال(٦٦/۹ درصد) در سطح یک، ٥٦٥ سؤال(۳۱/٤ درصد) سطح دو، و ۳۱ سؤال(۱/۷ درصد) سطح سه بود. در سال ۹۰، تاکسونومی ۸٦۷ سؤال (٤۸/۲ درصد) یک، ۷٤۱ سؤال(٤۱/۲ درصد) سطح دو و ۱۹۲ سؤال(۱۰/٦ درصد) در سطح سه طراحی شده‌بود. آزمون آماری Wilcoxon signed-rank test تفاوت معناداری بین سال‌های ۸۹ و ۹۰ از نظر توزیع فراوانی تاکسونومی سؤالات رشته‌های مختلف نشان داد. جدول شماره دو بیانگر بهبود چشمگیر تاکسونومی سؤالات در گروه‌های آموزشی است. در بین گروه‌های آموزشی تنها گروه آموزشی قلب تفاوت معناداری را در تاکسونومی سؤالات سال ۹۰ نسبت به سال ۸۹ نشان نداد.

جدول۲: فراوانی مطلق (و نسبی) تاکسونومی سؤالات آزمون ارتقای دستیاری در سال‌های ۸۹ و ۹۰، به تفکیک گروه‌های آموزشی

گروه آموزشی	سطح تاکسونومی سؤالات در سال ۸۹			سطح تاکسونومی سؤالات در سال ۹۰			z	P vlue
	۳	۲	۱	۳	۲	۱		
داخلی	٦(٤)	۷۹(٥۲/۷)	٦٥(٤۳/۳)	۱۰(٦/۷)	۸۷(٥۸)	٥۳(۳٥/۳)	٤/۰۰۰	۰/۰۰۰۱
کودکان	٤(۲/۷)	۸٥(٥٦/٦)	٦۱(٤۰/۷)	۲۱(۱٤)	۹۱(٦۰/۷)	۳۸(۲٥/۳)	٦/۳۲	۰/۰۰۰۱
جراحی عمومی	۰(۰)	٥۳(۳٥/۳)	۹۷(٦٤/۷)	۳۲(۲۱/۳)	۷۳(٤۸/۷)	٤٥(۳۰)	۹/۱٦	۰/۰۰۰۱
زنان	۰(۰)	۱٥(۱۰)	۱۳٥(۹۰)	۱۱(۷/۳)	٤٤(۲۹/٤)	۸٦(٥۷/۳)	۷/۱٤	۰/۰۰۰۱

[Downloaded from ijme.mui.ac.ir on 2026-11-26]

رادیولوژی	۲(۱/۳)	۵۲(۳٤/۷)	۹٦(٦٤)	۱۹(۱۲/۷)	۹٤(٦۲/۷)	۳۷(۲٤/٦)	۱۲/۱۲	۰/۰۰۰۱
روانپزشکی	٦(٤)	۲۱(۱٤)	۱۲۳(۸۲)	۱۲(۸)	۵۹(۳۹/۳)	۷۹(۵۲/۷)	۷/۰۷	۰/۰۰۰۱
پوست	۵(۳/۳)	۳۸(۲۵/۳)	۱۰۷(۷۱/٤)	۲۲(۱٤/۷)	۵۷(۳۸)	۷۱(٤۷/۳)	۷/۲۸	۰/۰۰۰۱
گوش، گلو و بینی و جراحی سرو گردن	۰(۰)	۵۰(۳۳/۳)	۱۰۰(٦٦/۷)	۱٦(۱۰/۷)	٤۸(۳۲)	۸٦(۵۷/۳)	۵/۵٦	۰/۰۰۰۱
بیهوشی	۰(۰)	٦۵(٤۳/۳)	۸۵(۵٦/۷)	۱۸(۱۲)	۵٤(۳٦)	۷۸(۵۲)	۵/۰۰۰	۰/۰۰۰۱
قلب	۲(۱/۳)	٤۳(۲۸/۷)	۱۰۵(۷۰)	۷(٤/۷)	۳۷(۲٤/۷)	۱۰٦(۷۰/٦)	۱/۳۳٦	۰/۱
جراحی کلیه و مجاری ادراری	۳(۲)	٤۱(۲۷/٤)	۱۰٦(۷۰/٦)	۱۰(٦/۷)	٤۷(۳۱/۳)	۹۳(٦۲)	٤/٤۷۲	۰/۰۰۰۱
نورولوژی	۳(۲)	۲۳(۱۵/۳)	۱۲٤(۸۲/۷)	۱٤(۹/۳)	۵۰(۳۳/۳)	۸٦(۵۷/٤)	۷/۰۰۰	۰/۰۰۰۱
کل سؤالات	۳۱(۱/۷)	۵٦۵(۳۱/٤)	۱۲۰٤(٦٦/۹)	۱۹۲(۱۰/۲)	۷٤۱(٤۱/۲)	۸٦۷(٤۸/۲)	۲۳/۸۳۳	۰/۰۰۰۱

از نظر تاکسونومی درون گروه‌ها اختلاف معنادار آماری بین سؤال‌های طراحی شده در سال‌های ۸۹ و ۹۰ وجود داشت ولی در سوالات گروه قلب تفاوت معنادار آماری مشاهده نشد.

در زمینه رعایت اصول ساختاری میلمن در سؤالات چندگزینه‌ای، یافته‌ها نشان داد که در سال ۸۹، اغلب (٦٦/۷۷درصد) سؤالات بدون اشکال بودند. این رقم در در سال ۹۰، به ۷۵/۲۲ درصد رسید(جدول ۳).

جدول ۳: فراوانی مطلق و نسبی رعایت اصول میلمن در سؤالات آزمون ارتقای گروه‌های مختلف آموزشی در سال‌های ۸۹ و ۹۰

گروه‌های آموزشی	سال ۸۹	سال ۹۰	z	p
داخلی	۸۳ (٪۵۵/۳۳)	۲٤ (٪۱٦)	٦/٤۰۳	۰/۰۰۰۱
کودکان	۹۵ (٪٦۳/۳۳)	۱۱۹ (٪۷۹/۳۳)	٤/۸۹۹	۰/۰۰۰۱
جراحی عمومی	۹۳ (٪٦۲)	۱۱۹ (٪۷۹/۳۳)	۵/۰۹۹	۰/۰۰۰۱
زنان	۸٦ (٪۵۷/۳۳)	۹۸ (٪٦۵/۳۳)	۳/٦۰٦	۰/۰۰۰۳
رادیولوژی	۱۱۸ (٪۷۸/٦٦)	۱۳۰ (٪۸٦/٦٦)	٦/۲٤۵	۰/۰۰۰۱
روانپزشکی	۱۳۷ (٪۹۱/۳۳)	۱۳۷ (٪۹۱/۳۳)	۰	۰
پوست	۱۲۵ (٪۸۳/۳۳)	۱٤۱ (٪۱۰۰/٦٦)	٤/۰۰۰	۰/۰۰۰۱
گوش، گلو و بینی و جراحی سرو گردن	۱۰۲ (٪٦۸)	۱۱۱ (٪۷٤)	۳/۰۰۰	۰/۰۰۲۷
بیهوشی	۹۵ (٪٦۳/۳۳)	۱۲۵ (٪۸۳/۳۳)	۵/۳۸۵	۰/۰۰۰۱
قلب	۷٤ (٪٤۹/۳۳)	۱۰۰ (٪٦٦/٦٦)	۵/۰۹۹	۰/۰۰۰۱
جراحی کلیه و مجاری ادراری	۱۱۰ (٪۷۳/۳۳)	۱۲۲ (٪۸۱/۳۳)	۳/٤٦٤	=۰/۰۰۰۵
نورولوژی	۸٤ (٪۵٦)	۱۲۸ (٪۸۳/۳۳)	٦/۳۳٦	<۰/۰۰۰۱
جمع	۱۲۰۲(٪٦٦/۷۷)	۱۳۵٤(٪۷۵/۲۲)	۱٦/۷۰۳	<۰/۰۰۰۱

از نظر رعایت اصول میلمن بین سؤال‌های طراحی شده در سال‌های ۸۹ و ۹۰ تفاوت معنادار آماری وجود داشت(p<۰/۰۰۰۱ , z=۱۶/۷۰۳). از جمله گروه‌های داخلی، قلب و نورولوژی بهبود چشم‌گیری را نشان دادند. بررسی روند بهبود کیفیت سؤالات رشته‌های مختلف آزمون دستیاری، پیشرفت درصد سؤالات بدون اشکال ساختاری را در ۱۲ رشته نشان داد.

بحث

در این مطالعه دو شاخص مهم کیفیت سؤالات چندگزینه‌ای یعنی، تاکسونومی و رعایت اصول میلمن، در سؤالات آزمون دستیاری قبل و بعد از یک مداخله

[Downloaded from ijme.mui.ac.ir on 2026-11-26]

آموزشی مورد بررسی قرار گرفت.

نتایج حاصل از این مطالعه نشان داد که برگزاری کارگاه در زمینه طراحی سؤالات چندگزینه‌ای مؤثر بوده است که شاید دلیل آن طراحی مناسب و منطبق بر اهداف کارگاه و همچنین استفاده عینی، شفاف وکاربردی از سؤالات طراحی شده در آزمون دستیاری گروه‌های آموزشی مربوط به هر رشته تخصصی و بازخورد مناسب در حین برگزاری کارگاه باشد. بدون اشکال بودن ساختاری بیش از دوسوم سؤالات آزمون‌های کتبی تأثیر مثبت برنامه ارتقای کیفیت سؤالات بر این شاخص بود که با مطالعه محققی و همکاران همخوانی دارد(۹). نتایج تحقیق معیاری و همکاران نشان داد که برگزاری کارگاه در زمینه آموزش طراحی مناسب سؤالات به همراه ارائه مثال‌های شفاف و عینی و کاربردی مؤثر بوده است(۷).

نتایج حاکی از آن بود که توزیع سطح تاکسونومی به کار گرفته در سال ۹۰ نسبت به سال ۸۹ بهبود چشم‌گیری را نشان می‌دهد که مؤید تأثیر مثبت مداخله آموزشی بوده است. اما از سوی دیگر نشان می‌دهد که اکثریت سؤالات آزمون‌های مزبور همچنان به ارزیابی سطوح پایین دانش (محفوظات) بسنده نموده است که با مطالعات معیاری، محققی و همکاران، Bulter و Tarrant همخوانی دارد(٦و۹و۲۰و۲۲). نتایج مطالعه شکورنیا نشان داد که بیش از دو سوم سؤال‌ها در سطح تاکسونومی یک (دانش و فهمیدن) تهیه شده‌اند(۱۷). غلبه سؤالات مبتنی بر حافظه (تاکسونومی یک) در آزمون‌ها، معضلی ریشه‌دار در سیستم آموزش پزشکی بوده و می‌توان گفت تقریباً در آزمون‌های تمامی سطوح تحصیلی این مشکل وجود دارد. چنین وضعیتی موجب سوق دادن دانشجویان به کسب دانش سطحی و محفوظات می‌گردد. یکی از دلایل غلبه سوؤالات سطح دانش، می‌تواند ساده تر بودن طراحی این گونه سؤالات نسبت به سطوح بالاتر باشد.

نگاهی اجمالی به مطالعات انجام شده در دانشکده‌های پزشکی به خصوص آزمون‌های ارتقای دستیاری نشان می‌دهد که گرچه سطح تاکسونومی سؤالات مزبور نیز آنچه را قابل انتظاراست برآورده ننموده است، اما در سال‌های اخیر با توجه به تأکید دبیرخانه شورای آموزش تخصصی پزشکی وزارت متبوع، تعداد قابل توجهی از سؤالات در سطوح پیچیده‌تر شناختی طراحی شده است. به نظر می‌رسد اهتمام سیستم مدیریت آموزشی در ترغیب طراحان به تدوین سؤالات با تاکسونومی بالا، در این زمینه نقش مؤثری داشته است(۷).

تحقیقات نشان داده است که در آزمون‌های پیشرفت تحصیلی سال‌های اخیر، سهم بیش‌تری به سطوح عمیق‌تر دانش (تاکسونومی ۲ و۳) اختصاص یافته-است(۷و۱٤و۲۱). از جملهه محققی و همکاران نشان دادند که پیشرفت معناداری در فراوانی سؤالات مربوط به سطوح بالاتر دانش (تاکسونومی ۲ و ۳) در آزمون دانشنامه سال ۸۸ نسبت به سال‌های ۸٦ و ۸۷ حاصل شده‌است. این امر می‌تواند بیانگر تأثیر مثبت مداخله آموزشی انجام شده (برنامه ارتقای کیفیت آزمون کتبی دانشنامه) باشد(۹). بر اساس گزارش دبیرخانه شورای آموزش پزشکی و تخصصی وزارت بهداشت از تحلیل نتایج و ساختار آزمون دانشنامه تخصصی سال ۸۸، ٥٦±۱٥/٥٦ درصد سؤالات در تاکسونومی‌های ۲ و ۳ طراحی شده است که در مقایسه با سال گذشته آن (٤٥±۱۹/۳ درصد)، تفاوت آماری معناداری را نشان می‌دهد(۹).

نتایج این مطالعه افزایش درصد سؤالاتی که بر اساس رعایت اصول ساختاری (اصول میلمن) طراحی شده بود را پس از برگزاری کارگاه آموزشی طراحی سؤال ، نشان داد. در سال۸۹، ٦٦/۷۷ درصد سؤالات بدون اشکال ساختاری بودند که این شاخص در سال ۹۰ به ۷٥/۲۲ درصد ارتقا پیدا کرد. مطالعه معیاری و همکاران نیز نشان داد که مداخله‌های آموزشی در زمینه طراحی آزمون‌های چندگزینه‌ای می‌تواند سبب ارتقای کیفیت آزمون‌های دستیاری گردد(۷).

[Downloaded from ijme.mui.ac.ir on 2026-11-26]

بر اساس نتایج مطالعه حاضر، در هر دو سال مورد بررسی، بیش‌تر از دوسوم سؤالات منطبق با اصول ساختاری طراحی سؤالات چندگزینه‌ای (اصول میلمن) بود. شکورنیا و همکاران در اهواز، و معیاری در همدان نشان دادند که بیش از نیمی از سؤالات فاقد اشکالات ساختاری است(۷و۱۷). تحقیقات مک کوربی (McCourbie) و همکاران، هاموند (Hammond) و همکاران، وجود اشکالات در درصد قابل توجهی از سؤالات را نشان دادند(۲۳و۲٤). عدم رعایت اصول ساختاری به هر میزان، روایی سازه را در آزمون مختل نموده و در این صورت کارکردهای آموزشی را تحت تأثیر قرار می‌دهد.

شایع‌ترین اشکالات در ساختار سؤالات در کل آزمون‌های مورد بررسی، شامل «هم‌سنگ نبودن گزینه‌ها»، و «منفی بودن ساقه سؤال» بود. در این زمینه مطالعه حق شناس و همکاران در دانشگاه علوم پزشکی مازندران نشان داد، شایع‌ترین اشکالات در آیتم‌ها، «عدم چیدمان عمودی گزینه‌ها»، «هم‌سنگ نبودن گزینه‌ها» و «نیاوردن بخش اعظم اطلاعات در ساقه سؤال» بود(۱۰). همچنین مطالعه جوادی و همکاران در بم حاکی از آن بود که شایع‌ترین اشکالات در آیتم‌ها، «درج مطالب تکراری در گزینه‌ها» و «وجود لغات منفی در ساقه سؤال» بود(۲۵).

این مطالعه با مشکلاتی نظیر حضور تمامی اعضای هیأت‌علمی هر یک از گروه‌های آموزشی در روز کارگاه مربوطه و مقاومت اعضای هیأت‌علمی در پذیرش نقاط ضعف در طراحی سؤالات مواجه بود. از نقاط قوت مداخله می‌توان تشریک مساعی و هماهنگی هر چه بیش‌تر اعضای هیأت‌علمی در طراحی سؤالات و تلاش در جهت ارتقای کیفیت سؤالات را نام برد.

نتایج حاصل از این مطالعه لزوم توجه بیش‌تر به اصول آزمون‌سازی را نشان می‌دهد. برنامه‌ریزی برای افزایش آگاهی، مهارت و توجه بیش‌تر استادان در این زمینه می‌تواند به اصلاح این رویه و بهبود طراحی سؤال‌های چندگزینه‌ای کمک نماید. به منظور بهبود وضعیت کنونی، توصیه می‌شود ضمن برگزاری دوره‌های آموزشی مناسب، با ارزیابی مستمر سؤال‌های آزمون دستیاری و دادن بازخورد مناسب، آنان را در ارتقای این آزمون‌ها یاری دهیم(۱۷).

نتیجه‌گیری

مطالعه حاضر پیشرفت معناداری را در کاربرد سؤالات مربوط به سطوح عمیق دانش (تاکسونومی ۲ و ۳) و ساختار سؤالات (اصول میلمن) در آزمون سال ۹۰ نسبت به سال ۸۹ و تأثیر مثبت مداخله آموزشی انجام شده (کارگاه‌های آموزشی MCQ) نشان داد. لذا می توان گفت برگزاری کارگاه‌های طراحی سؤالات چندگزینه‌ای، تحلیل سؤالات و همچنین حضور فعال در هر گروه آموزشی می‌تواند راه‌گشایی برای بهبود آزمون‌های مشابه در دانشگاه باشد.

قدردانی

پژوهشگران لازم می‌دانند از همکاری کلیه مدیران گروه‌های آموزشی و همچنین اعضای هیأت‌علمی ۱۲ گروه آموزشی تربیت کننده دستیار تخصصی (طراحان سؤال) دانشکده پزشکی دانشگاه علوم پزشکی گیلان تشکر و قدردانی نمایند.

منابع

1. Vahidshahi K, Mohagheghi A, Shakeri S. [Comparison of quality indicators question multiple-choice written exam Board in 2008 & 2009]. [Cited 2010 Oct 19]. Available from: http://www.Cgme.behdasht.gov.ir/uploads/264_781_N4_sakhtarDastyari.pdf.[Persian]
2.Kave Tabatabaei M, Bahraini Toosi MH, Modaber Azizi MJ, Ebrahimzadeh S. [Barrasye tahliye

[Downloaded from ijme.mui.ac.ir on 2026-11-26]

azmonhaye chahar gozineye daneshkadeye pezeshkye daneshgahe oloom pezeshkye mashhad dar nimsale dovome 79-80]. Journal of Medical Faculty of Mashhad University of Medical Sciences. 2002; 45(2): 89-95.[Persian]
3. Tarrant M, Ware J. Impact of item-writing flaws in multiple – choice questions on student achievement in high – stakes nursing assessment. Med Educ. 2008; 42(2): 198-206.
4. Reid WA, Duvall E, Evans P. Relationship between assessment results and approaches to learning and studying in year Two medical students. Med Educ. 2007; 41(8): 754-62.
5. Larsen DP, Butler AC, Ruediger HL. Test- enhanced learning in medical education. Med Educ. 2008 : 42(10): 959-66.
7. Meyari A, Biglarkhani M, Zandi M, Vahedi M, Miresmaeili A. [The Effect of Education on Improvement of Multiple Choice Questions' Designing in Annual Residency Exams of Dental School]. Iranian Journal of Medical Education. 2012; 12(1): 36-45. [Persian]
6. Brady AM. Assessment of Learning with multiple choice questions. Nurse Educ pract. 2005; 5(4): 238-42.
8. Jalili M, Mirzadeh A, Tavangar SM, Meysami AP, azizian, afzali. [Arzyabi keifiyat soalat azmonhay erteghay dastyari va eraeh bazkhord be gorohay amozeshi daneshkadeh pezeshkey daneshgah olome pezeshkye Tehran]. Royesh. 2009; 5:12-13. [Persian]
9. Mohagheghi MA, Vahidshahi K, Shakeri S,Sabouri M,Razavi SM,Mohammadi M,et al. [Moghayese barkhi shakheshay keyfiyat porseshhay chand gozinei azmoon katbi daneshname takhasosi 86-88]. Ministry of Health. [citd 2013 July 20]. available from: http://cgme.behdasht.gov.ir/uploads/264_975_Maghale_ShakhesKeifiyat.pdf [Persian]
10. Haghshenas MR, Vahidshahi K,Mahmudi M,Shahbaznejad L, Parvinnejad N, Emadi A. [Evaluation of multiple choice questions in the school of medicine, Mazandaran University of Medical Sciences, the first Semester of 2007]. Strides in Development of Medical Education. 2008; 5(2): 119.126. [Persian]
11. Kaufman DM. Assessing medical students hit or miss. Student BMJ. 2001; 9(3): 87-88.
12. Van Der Vleuten CPM. The assessment of professional competence: developments, research and practical implications. Adv Health Sci Educ. 1996; 1(1): 41-67.
13. Vahidshahi K, Mohagheghi MA, Shakeri S,Sabouri M,Razavi SM,Mohammadi M,et al. [Moghayese barkhi shakheshay keyfiyat porseshhay chandgozinei azmon katbi daneshname takhasosi 86-87]. [citd 2013 July 20]. available from: http://cgme.behdasht.gov.ir/uploads/264_781_N4_Moghayese.html [Persian]
14. Friedman Ben David M. Standard setting in student assessment. Med Teach. 2000; 22(2): 120-30.
15. Khoshkam M, Shajari Zh. Trend in processing of structural designing general science exam in the study in 1999 in Tehran. Journal of Research in Medical science. 1998; 3(1): 160.
16. Arab M. editor. [Barrasye naghayese soalate 4 gozineyei emtehanate daneshkadeye pezeshkiye daneshgahe oloom pezeshkye Hamedan dar nimeye dovome sale 1378]. 4th National Congress of Medical Education; 2000; Tehran University of Medical Sciences and Health Services, Iran. [Persian]
17. Shakournia AH, Mozafari A,Khosravi borojeni A. [Survey on Structural of MCQs of residency exam in AJUMS]. Jundishapur Scientific Medical Journal. 2009; 8(4): 491-502. [Persian]
18. Ramezani Badr F, Shaban M. [Survey on effect of evaluation and Analysis of summative examination on test formation of faculties]. 8th National Congress of Medical Education; 2007; Kerman University of Medical Sciences and Health Services, Iran. [Persian]
19. Epstein RM. Assessment in medical education. N Engl J Med. 2007; 356(4): 387-96.
20. Tarrant M, Knierim A, Hayes SK, Ware J.The frequency of item writing flaws in multiple-choice questions used in high stakes nursing assessments. Nurse Educ Today. 2006; 26(8): 662-71.
21. Saif AA. [Educational measurement, assessment, and evaluation]. 2 th ed.Tehran: Doran; 2003. [Persian]
22. Butler AC, Karpicke JD, Roediger HL. Correcting metacognitive error: feedback increases retention of low.confidence correct response. J Exp Psychol Learn Mem Cogn. 2008 ; 34(4): 418-28.
23. McCoubrie P. Improving the fairness of multiple choiece questions: A literature review. Med Teach. 2004: 26(8): 709-12.
24. Hammond EJ,McIndoe AK,Sansome AJ,Spargo PM. Multiple choice examination: Adapting an evidence-based approach to exam technique. Anesthesia. 1998; 53(11): 1105-8.
25. Javadi M, Abbas zadeh A, Borhani F, Abdoli R. [Analytical survey of summative MCQs of Bam Nursing faculty]. 8th National Congress of Medical Education; 2007; Kerman University of Medical Sciences and Health Services, Iran. [Persian]

[Downloaded from ijme.mui.ac.ir on 2026-11-26]

Quality of Residents' Promotion Exams before & after Educational Intervention in 2010-2011

Hossein Khoshrang[1], Mahdokht Taheri[2], Ataollah Asadi[3], Abtin Heidarzadeh[4]

Abstract

Introduction: *Resident promotion exams are considered one of the most important summative exams in residency period that allow residents to proceed to a higher grade. Multiple choice questions are the most common and objective test in medical science education. Therefore, finding methods to improve quality of exams, particularly in specialty programs, is important. The aim of this study is to determine effectiveness of training workshops on improving quality of residency promotion exams in medical school of Guilan University of Medical Sciences (GUMS).*

Methods: *In this Quasi-experimental study (before & after), all questions of resident promotion exams of GUMS for 12 specialty programs in year 2010 (150 questions for each specialty-1800 questions in total) were structurally and taxonomically evaluated using a valid and reliable checklist based on Millman principles. After12 one–day educational workshops (one for each department) on multiple choice questions), questions of residents' promotion exam in year 2011 were evaluated. Data related to questions designed by faculties who had participated in the interventional workshop were analyzed using statistical Wilcoxon signed-rank test.*

Results: *Taxonomy of questions of written promotion exam in years 2010 and 2011 were as follows: 1204(%66.09) questions were of first level taxonomy, 565 (%31.4) of 2nd level, and 31(% 1.7) of 3rd level in the 2010 exam, and 867 questions (%48.2) were of first level taxonomy, 741 (%41.2) of 2nd level, and 192 (%10.6) of 3rd level in the 2011. Findings showed significant differences between taxonomy of questions of different programs in 2010 and 2011 (p=0.0001, Z=23.83). Also, according to Millman's structural principles, the rate of questions with no structural problem was %66.77 in the year 2010 and %75.22 in 2011. There is a statistically significant difference based on Millman principles between 2010 and 2011 (p=0.0001, Z=16.703).*

Conclusion: *Current study showed that training workshops are effective in improving the rate of questions targeted at higher levels of knowledge (Taxonomies 2 & 3) in residency promotion exams.*

Keywords: Multiple Choice Questions, Residency Promotion Exam, Taxonomy, Teacher Training, Educational Workshops

Addresses:

[1] Associate Professor, Department of Anesthesiology, School of Medicine, Guilan University of Medical Sciences, Rasht, Iran. Email: khoshrang@gums.ac.ir

[2](✉)Assistant Professor, Education Development Center, Guilan University of Medical Sciences, Rasht, Iran. Email: m_taheri@gums.ac.ir

[3] Instructor, Department of Anesthesia, School of Nursing and Midwifery, Guilan University of Medical Sciences, Rasht, Iran. Email: aa.louyeh@yahoo.com

[4] Associate Professor, Department of Community Medicine, School of Medicine, Guilan University of Medical Sciences, Rasht, Iran. Email: heidarzadeh@gums.ac.ir

[Downloaded from ijme.mui.ac.ir on 2026-11-26]