

# تأثیر کارگاه‌های آموزشی بر تدوین سوالات چند گزینه‌ای امتحانات ارتقاء دستیاری بر اساس اصول میلمن

فریبا درخشان<sup>۱</sup>، عباس علامی<sup>۲</sup>، سلیمان احمدی<sup>۳\*</sup>

تاریخ پذیرش: ۹۳/۱۰/۷

تاریخ دریافت مقاله: ۹۳/۵/۱۳

## چکیده

**مقدمه:** سوالات چندگزینه‌ای در آزمون‌های ارتقاء دستیاری یکی از متداول‌ترین روش‌های ارزشیابی دانشجویان است. هدف از این پژوهش بررسی تأثیر برنامه توانمندسازی اساتید بر تعداد اشکالات ساختاری سوالات آزمون ارتقاء دستیاری در دانشگاه علوم پزشکی قزوین در سال‌های ۹۱ و ۹۲ بود.

**روش‌ها:** این پژوهش مطالعه‌ای نیمه تجربی از نوع قبل و بعد مداخله است. برنامه توانمندسازی اساتید در سه مرحله صورت پذیرفت. ابتدا کتاب اصول طراحی و ارزشیابی سوالات چند گزینه‌ای در اختیار کلیه گروه‌های آموزشی قرار گرفت. کلیه سوالات آزمون ارتقای دستیاری دانشکده پزشکی دانشگاه علوم پزشکی قزوین، در سال ۹۱ از نظر ساختار سوالات با استفاده از چک لیستی مشتمل بر ۱۹ آیتم مبتنی بر اصول میلمن مورد بررسی قرار گرفت. ۱۵۰ سوال از هر رشته بررسی شد. سپس نتایج بررسی سوالات هر رشته در اختیار گروه‌ها قرار گرفت. در مرحله سوم کارگاه‌های آموزشی به تفکیک گروه‌های آموزشی جهت بررسی سوالات آن گروه، قبل از برگزاری آزمون ارتقای سال ۹۲ تشکیل گردید. سپس سوالات آزمون ارتقای سال ۹۲ مورد بررسی قرار گرفت. داده‌ها با استفاده از آزمون‌های t و ANOVA از طریق نرم‌افزار آماری SPSS مورد تجزیه و تحلیل قرار گرفت.

**یافته‌ها:** ۲۱۰۰ سوال در ۷ رشته تخصصی اطفال، جراحی، داخلی، قلب، روانپزشکی، زنان، بیهوشی مورد تجزیه و تحلیل قرار گرفت. یافته‌ها نشان داد اشکال ساختاری هر سوال در سال ۹۱،  $0/58 \pm 0/02$  و در سال ۹۲،  $0/44 \pm 0/02$  بدست آمد که با استفاده از آزمون تی گروه‌های مستقل اختلاف معنی‌داری قبل و بعد از آموزش دیده شد ( $P < 0/001$ ).

**نتیجه گیری:** برگزاری برنامه‌های توانمندسازی اساتید به طور موثری می‌تواند باعث کاهش تعداد اشکالات ساختاری سوالات گردد و لزوم برگزاری و گسترش این نوع برنامه‌ها را در سطوح آموزش پزشکی نشان می‌دهد.

**کلید واژه‌ها:** کارگاه، سوالات امتحانی، دستیاری، توانمندسازی، پزشکی

## مقدمه

گواهینامه برگزار می‌گردد. در آزمون‌های ارتقا دستیاری نیز، سوالات چندگزینه‌ای به‌عنوان یکی از متداول‌ترین روش‌های ارزشیابی دستیاران مورد استفاده قرار می‌گیرد. آزمون دانشنامه از دو بخش کتبی و شفاهی تشکیل یافته که آزمون کتبی آن را MCQ تشکیل می‌دهد. گرچه استفاده از این نوع آزمون در بررسی سطح دانش گروه‌های بزرگی از آزمون شونده‌گان مفید است ولی عدم آشنایی با فنون طراحی آن‌ها به‌طور عمده‌ای از قدرت و عملکرد این نوع آزمون می‌کاهد. سوالات MCQ اگر به‌خوبی طراحی نشود در تفسیر دقیق و معنی‌دار نمرات آزمون اختلال ایجاد کرده و در میزان موفقیت دانشجویان تأثیر منفی می‌گذارد (۲). همچنین اگر در طراحی این آزمون‌ها تجربه کافی وجود نداشته باشد نوع سؤال ممکن است هدایت‌کننده دانشجو به پاسخ صحیح باشد (۳) لذا آنالیز پس آزمون از اهمیت خاصی برخوردار است. مطالعات Quinn نشان داده که در سال‌های اخیر نمره دهی، نرم‌افزار تجزیه و تحلیل کامپیوتر،

در طول دهه گذشته تعداد زیاد دانشجویان، کاهش منابع و افزایش استفاده از فن‌آوری‌های جدید منجر به افزایش استفاده از سوالات چندگزینه‌ای (MCQ [Multiple choice question]) به‌عنوان یک روش ارزیابی برای پذیرش در دوره‌های آموزش عالی تقریباً در هر بخشی از دنیا شده است (۱). در ایران نیز آزمون‌های چند گزینه‌ای متداول‌ترین آزمون‌های کتبی در ارزشیابی دانشجویان گروه پزشکی است. آزمون‌های دوره تخصصی در رشته‌های بالینی در ایران در سه سطح پذیرش، ارتقاء و

\* نویسنده مسئول: دکتر سلیمان احمدی (استادیار)، گروه آموزش پزشکی، دانشکده آموزش علوم پزشکی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران.

soleiman.ahmady@gmail.com

فریبا درخشان (دانشجوی کارشناسی ارشد رشته آموزش پزشکی)، دانشکده آموزش علوم پزشکی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران.

دکتر عباس علامی (دانشیار)، گروه عفونی، دانشکده پزشکی، دانشگاه علوم پزشکی قزوین، ایران.

مرحله نخست، کتاب اصول طراحی و ارزشیابی سوالات چندگزینه‌ای (MCQ) در ابتدای سال ۹۱ در اختیار کلیه گروه‌های آموزشی قرار گرفت (۱۰). در مرحله دوم، پس از برگزاری آزمون دستیاران در سال ۹۱ کلیه سوالات آزمون ارتقای دستیاری دانشکده پزشکی دانشگاه علوم پزشکی قزوین، از لحاظ رعایت اصول ساختاری در ساقه و گزینه‌های سوالات بررسی شد. جهت بررسی رعایت اصول طراحی سوالات از چک لیست میلمن (Milman) استفاده گردید. چک لیست مزبور مشتمل بر ۱۹ آیتم در خصوص رعایت قواعد ساختاری بود. در چک لیست مقابل هر آیتم، شماره سوالاتی که آن مورد رعایت نشده است، مشخص شد. با توجه به استاندارد بودن چک لیست و بومی سازی آن و استفاده مکرر محققین داخلی، روایی چک لیست مورد تأیید بود (۸، ۷، ۶ و ۱۱). همچنین دو مصحح که دوره کارشناسی ارشد آموزش پزشکی را گذرانده، در ابتدا و در طول مطالعه جلسات هماهنگی بین آنان برگزار شد و توافق نظر در مورد نحوه بررسی سوالات و تکمیل چک لیست به عمل آمد. با بررسی حدود ۵۷ درصد سوالات (۶۰۰ سوال) توسط مصححین، میزان همخوانی نتایج بین ارزیابی کننده ها (inter-rater) ۸۶/۵ درصد بدست آمد. موارد مبهم بین کارشناسان دوباره مطرح و توافق حاصل می‌گردید. این سوالات چهارگزینه‌ای مربوط به کلیه آزمون‌های کتبی دستیاری تخصصی پزشکی دانشگاه علوم پزشکی قزوین در ۷ رشته (اطفال، بیهوشی، جراحی، داخلی، روان پزشکی، زنان و زایمان و قلب) بود. هر گروه آموزشی در هر سال ۱۵۰ سوال طرح کرده بود. بدین طریق صد و پنجاه سؤال از هر رشته و جمعاً هزار و پنجاه سؤال آزمون ارتقا سال ۹۱ در مرحله دوم بررسی گردید. پس از تجزیه و تحلیل مجموعه سوالات و مشخص شدن نقاط قوت و ضعف آزمون مزبور، بررسی سوالات هر رشته بر اساس شاخص‌های کیفی در اختیار گروه‌ها قرار گرفت و بازخورد به گروه‌ها داده شد. در مرحله سوم کارگاه‌های آموزشی به تفکیک گروه‌های آموزشی جهت بررسی سوالات آن گروه، متفاوت از کارگاه‌های آموزشی که تاکنون داشته‌اند با عنوان "چگونگی طراحی سوالات چهارگزینه‌ای بهتر" قبل از برگزاری آزمون ارتقای سال ۹۲ تشکیل گردید. به این ترتیب که از ۷ رشته تخصصی به جز گروه بیهوشی و روان پزشکی برای ۵ گروه دیگر، ۵ کارگاه آموزشی اختصاصی آن گروه برگزار شد و در هر کارگاه سوالات آزمون ارتقاء سال ۹۱ مربوط به آن گروه آموزشی (تعداد صد و پنجاه سؤال) مورد نقد و بررسی قرار

آنالیز پس‌آزمون و تعیین شاخص دشواری و تمایز سوالات به بهبود کیفیت سوالات MCQ کمک کرده است (۴). Hammond و همکاران از دانشگاه هامپتون نشان دادند، درصد قابل توجهی از سوالات آزمون‌های چهارگزینه‌ای دارای اشکالات ساختاری است (۵). در ایران مطالعات مختلفی در علوم پزشکی به تجزیه و تحلیل آزمون‌های چندگزینه‌ای پرداخته‌اند. مطالعات حاکی از آن است که با انجام مداخلات آموزشی وضعیت طراحی سوالات، بهبودی قابل توجهی می‌یابد. به طور مثال، مطالعه معیاری و همکاران تفاوت آماری معنی داری را در ارتقای کیفیت طراحی سوالات از دو منظر تاکسونومی و ساختار پس از برگزاری کارگاه آموزشی جهت استادان جوان طراح سوالات در آزمون‌های ارتقای دستیاری دندان پزشکی نشان داد (۶). وحید شاهی و همکاران نیز نشان دادند که با انجام مداخله آموزشی از نوع ارائه بازخورد و متون تحصیلی و آموزشی به دبیران و اعضای هیئت‌های ممتحنه، می‌توان تأثیر مثبتی بر طراحی سوالات آزمون‌های ارتقای دستیاری گذاشت (۷).

گزارش دبیرخانه شورای آموزش پزشکی و تخصصی وزارت بهداشت از تحلیل نتایج و ساختار آزمون کتبی دانشنامه تخصصی در سال‌های ۸۸-۱۳۸۶ بیانگر لزوم تداوم و گسترش برنامه ارتقای کیفیت آزمون‌ها جهت بهبود کیفیت ساختاری و ضریب تمایز سوالات بوده است (۸). به منظور طراحی و اجرای آزمون‌های استاندارد دستیاری در دانشگاه علوم پزشکی قزوین چندین کارگاه آموزشی طراحی سؤال جهت اعضای هیئت‌علمی بدون تفکیک گروه‌های آموزشی برگزار شده بود. نتایج تحقیقات مرکز مطالعات و توسعه آموزش دانشگاه علوم پزشکی قزوین، وجود نقص در طراحی سوالات چندگزینه‌ای آزمون‌های ارتقای دستیاری در این دانشگاه را نشان داد (۹). لذا برنامه توانمندسازی اساتید متفاوت با دوره های آموزشی گذشته و بر اساس تفکیک گروه های آموزشی صورت پذیرفت. این مطالعه با هدف کلی بررسی تاثیر دوره آموزشی به تفکیک گروه های آموزشی در رعایت اصول ساختاری سوالات آزمون ارتقاء دستیاری تخصصی پزشکی دانشگاه علوم پزشکی قزوین قبل از برگزاری دوره آموزشی (سال ۱۳۹۱) و بعد از برگزاری دوره آموزشی (سال ۱۳۹۲) انجام گرفت.

## روش‌ها

این پژوهش مطالعه کارآزمایی نیمه تجربی quasi-experimental از نوع قبل و بعد مداخله است. برنامه توانمندسازی اساتید در سه مرحله صورت پذیرفت. در

مورد متغیرهای کیفی از آزمون مجذور کای استفاده شد. سطح معنادار آزمون ۰/۰۵ در نظر گرفته شد.

### یافته‌ها

۱۰۵۰ سوال مربوط به سال ۹۱ و ۱۰۵۰ سوال مربوط به سال ۹۲ در ۷ رشته تخصصی اطفال، جراحی، داخلی، قلب، روانپزشکی، زنان، بیهوشی در دانشگاه علوم پزشکی قزوین مورد تجزیه و تحلیل قرار گرفت. یافته‌ها نشان داد در مجموع دو سال برگزاری آزمون ارتقاء بهترین گروه‌های آموزشی در رعایت اصول ساختاری سوالات، گروه‌های زنان و داخلی به ترتیب با ۸۶/۷ درصد و ۷۲/۳ درصد و ضعیف‌ترین گروه‌های قلب و بیهوشی به ترتیب با ۴۵/۷ درصد و ۳۷ درصد بودند. در بررسی میزان رعایت اصول ساختاری سوالات بر اساس چک‌لیست میلن به تفکیک گزینه‌ها نتایج مطابق جدول شماره یک به دست آمد. اشکالات ساختاری در ۹ آیتم (آیتم های ۱، ۵، ۶، ۷، ۱۶، ۱۲، ۱۳ و ۱۸ بر اساس چک لیست) به طرز معنی‌داری کاهش یافته بود (P < ۰/۰۵).

گرفت. در کارگاه‌های مزبور در مورد آزمون‌های چندگزینه‌ای، ویژگی‌ها و شرایط یک آزمون چندگزینه‌ای استاندارد بحث گردید همچنین به بحث اصول میلن پرداخته شد و با مشارکت طراحان سؤال به تعیین نقاط ضعف و قوت سؤالات آن گروه پرداخته شد. پس از برگزاری آزمون ارتقای سال ۹۲، سؤالات آزمون و در مجموع ۲۱۰۰ سوال چهارگزینه‌ای یک پاسخ درست (single best answer) مربوط به آزمون کتبی ارتقا دستیاران تخصصی مورد تجزیه و تحلیل قرار گرفت. لازم به ذکر است آزمون کتبی رشته تخصصی عفونی که در سال ۱۳۹۱ وجود نداشت از دور مطالعه خارج شد. اطلاعات مربوط به اشکالات ساختاری سوالات به تفکیک گروه‌های آموزشی وارد برنامه Excel گردید. سپس نتایج حاصله وارد نرم افزار آماری SPSS.Ver.19 گردید. ابتدا اطلاعات توصیفی (جداول توزیع فراوانی، شاخص‌های آمار توصیفی و نمودار) محاسبه تا برآورد کلی از داده‌ها در مرحله بعد حاصل گردد. مقایسه بین داده های کمی با استفاده از آزمون t زوجی و ANOVA و در

جدول ۱: مقایسه اشکالات ساختاری در سوالات چندگزینه‌ای بر اساس چک لیست میلن در سال‌های ۹۱ و ۹۲

ردیف	مواردی که باید در طراحی سوالات رعایت شود	تعداد سوالاتی که رعایت نشده		p
		سال ۹۱	سال ۹۲	
۱	سوال روشن و واضح است و از کلمات قابل فهم آزمون‌شوندگان استفاده شده است. (عدم استفاده از کلمات و علائم اختصاری)	۳۷	۹	۰/۰۰۰
۲	بیش از یک مساله یا مطلب در یک سوال قرار ندارد.	۱	۲	۰/۵۰۰
۳	مطالب اصلی سوال به طور کامل در تنه سوال نوشته شده است.	۸۱	۸۱	۰/۵۳۳
۴	جای خالی (نقطه چین) در انتهای تنه سوال قرار داده شده است.	-	-	-
۵	سوال حاوی اطلاعات اضافه نیست.	۵۵	۱۶	۰/۰۰۰
۶	کلمات به کار رفته در تنه سوال و گزینه‌ها از نظر املایی و انشایی درست است.	۱۱۵	۸۳	۰/۰۱۰
۷	کلمات منفی در تنه برجسته جلوه داده شده یا زیر آن خط کشیده شده است.	۳۹	۲۴	۰/۰۳۶
۸	تنه سوال و گزینه‌ها هر دو منفی (منفی مضاعف) نیست.	۱۲	۸	۰/۲۵۰
۹	گزینه همه موارد فوق یا هیچ‌کدام از موارد فوق بکار برده نشده است.	۳	۸	۰/۱۱۳
۱۰	در گزینه‌ها قیود هدایت کننده نیست (ممکن است، معمولاً، هرگز، همیشه و ...)	۳۵	۴۰	۰/۳۱۹
۱۱	گزینه‌ها کوتاه بوده و گزینه درست از سایر گزینه‌ها کوتاه تر یا طولانی تر نیست.	۱۹	۱۳	۰/۱۸۷
۱۲	گزینه‌های گیج کننده (مثال: الف وب درست ولی ج غلط) به کار نرفته است.	۱۵	-	۰/۰۰۰
۱۳	گزینه‌های سوال از نظر دستوری و جمله بندی مکمل متن سوال است.	۲۷	-	۰/۰۰۰
۱۴	چهار گزینه هر سوال به صورت عمودی زیر هم نوشته شده است.	۱۳	۱۷	۰/۲۹۱
۱۵	گزینه‌ها از لحاظ طول، واژگان، ساختار جمله، محتوی و پیچیدگی موضوع با یکدیگر متجانس هستند.	۵۰	۱۸	۰/۰۰۰
۱۶	از تکرار مطالب در گزینه‌ها خودداری شده است.	۱۳۰	۹۲	۰/۰۰۴
۱۷	در سوال، واحد اندازه‌گیری ذکر شده است.	۲۴	۶۰	۰/۰۰۰
۱۸	در گزینه‌هایی که عدد آورده شده، ترتیب عددی رعایت شده است.	۱۶	۵	۰/۰۱۳
۱۹	تنه سوال مثبت است.	۱۷۸	۱۷۶	۰/۴۷۷

۴۷/۳ درصد سوالات رشته‌های مختلف در آزمون ارتقاء سال ۹۱، فاقد اشکال ساختاری بودند، در سال ۹۲ این میزان ۵۲/۷ درصد به دست آمد که اختلاف آماری معنی‌داری را قبل و بعد از مداخله آموزش نشان داد ( $P=0/001$ ).

بیشترین تعداد اشکال ساختاری در سال‌های ۹۱ و ۹۲ مربوط به طراحی سوالات منفی (آیتم ۱۹) بود که در سال ۹۱، ۱۷۸ سوال و در سال ۹۲، ۱۷۶ سوال منفی طراحی شده بود. میانگین تعداد خطاها در سوالاتی با تنه منفی به طور متوسط  $0/01 \pm 0/72$  و در سوالات با تنه سوال مثبت

$0/04 \pm 0/47$  بدست آمد که اختلاف معنی‌داری داشت ( $P < 0/001$ ).

اشکال ساختاری هر سوال در سال ۹۱ با میانگین و انحراف معیار  $0/02 \pm 0/58$  و در سال ۹۲،  $0/02 \pm 0/44$  بدست آمد که با استفاده از آزمون تی زوجی اختلاف معنی‌داری قبل و بعد از آموزش دیده شد ( $P < 0/001$ ). بیشترین میزان بهبود در اشکال ساختاری مربوط به گروه اطفال با کاهش میزان اشکال ساختاری  $0/08 \pm 0/47$  و ضعیف‌ترین گروه در میزان بهبود اشکال ساختاری، در گروه جراحی با افزایش میزان اشکال ساختاری  $0/07 \pm 0/06$  بود.

جدول ۲: فراوانی مطلق و نسبی رعایت اصول میلمن در سوالات آزمون ارتقای دستیاری به تفکیک گروه های آموزشی در سال های ۹۱ و ۹۲

گروه آموزشی	رعایت اصول میلمن فراوانی مطلق (نسبی)						
	مجموع	سال		۹۱	۹۲	p	
		۹۱	۹۲				
اطفال	۷۳	۴۸/۷	۱۱۳	۷۵/۳	۱۸۶	۶۲	۰/۰۰۰*
بیهوشی	۴۹	۳۲/۷	۶۲	۴۱/۳	۱۱۱	۳۷	۰/۱۲
داخلی	۱۰۰	۶۶/۷	۱۱۷	۷۸	۲۱۷	۷۲/۳	۰/۰۲۸*
جراحی	۹۰	۶۰	۷۶	۵۰/۷	۱۶۶	۵۵/۳	۰/۱۰۴
قلب	۵۹	۳۹/۳	۷۸	۵۲	۱۳۷	۴۵/۷	۰/۰۲۸*
روان	۹۸	۶۵/۳	۱۰۲	۶۸	۲۰۰	۶۶/۷	۰/۶۲۴
زنان	۱۳۰	۸۶/۷	۱۳۰	۸۶/۷	۲۶۰	۸۶/۷	۱/۰۰۰

ارتباط معنی‌دار آماری در افزایش رعایت اصول میلمن در گروه‌های آموزشی اطفال، داخلی، قلب در سال های ۹۱ و ۹۲ وجود داشت ( $P < 0/001$ ).

در گروه‌های آموزشی بیهوشی و روان‌پزشکی سال ۹۲ در مقایسه با سال قبل میزان رعایت اصول میلمن در سوالات افزایش و گروه جراحی کاهش داشته ولی در هیچ‌یک از گروه‌ها اختلاف آماری معنی‌داری وجود ندارد. همچنین نتایج نشان داد گروه زنان در مقایسه با سایر گروه‌ها از بالاترین میزان رعایت اصول میلمن برخوردار بوده اما تغییری قبل و بعد از مداخله آموزشی مشاهده نشد.

### بحث و نتیجه‌گیری

در زمینه شاخص کیفی رعایت اصول طراحی سوالات، مطالعه حاضر افزایش معنی‌دار رعایت اصول طراحی در ساختار سوالات به عنوان یک شاخص کیفی در آزمون ارتقای دستیاری را در سال ۹۲ به نسبت سال ۹۱ نشان داد که مؤید تأثیر برنامه ارتقای توانمندسازی اساتید بر این شاخص بود. مطالعه حاضر با تحقیقات خوشرنگ و همکاران در دانشکده پزشکی گیلان طی سال‌های ۸۹ تا ۹۰ که تأثیر کارگاه

آموزشی طراحی سوالات چندگزینه‌ای بر روی ۳۶۰۰ سوال آزمون ارتقا دستیاری را بررسی کرد مطابقت دارد تحقیقات آنان نشان داد از نظر مراعات اصول میلمن بین کیفیت سوالات آزمون سال‌های ۸۹ و ۹۰ تفاوت معنی‌دار وجود داشت (۱۱). همچنین مطالعات معیاری که در آن تأثیر ارائه بازخورد بر روی کیفیت سوالات آزمون ارتقاء دستیاری دانشکده دندان‌پزشکی در سال ۸۷ و سوالات آزمون ارتقاء دستیاری پزشکی در سال ۸۸ در دانشگاه علوم پزشکی همدان را بررسی نموده بود نیز موید این نکته هست (۶، ۱۲). وحیدشاهی و همکاران مطالعه‌ای بر روی برخی شاخص‌های کیفیت پرسش‌های چند گزینه‌ای آزمون کتبی دانشنامه تخصصی در سال‌های ۸۷-۱۳۸۶ انجام دادند که تحقیقات آنان نیز نشان داد برنامه ارتقای کیفیت آزمون دانشنامه باعث بهبود معنی‌دار رعایت اصول ساختاری در سوالات شده است (۷). در پژوهشی که با هدف تأثیر اجرای برنامه ارتقای دانش اعضای هیات علمی طراح این سوالات بر نتایج آنالیز این آزمون‌ها انجام شد نتایج نشان دهنده افزایش دانش طراحان سؤال با استفاده از پمفلت‌های آموزشی و ارائه بازخورد در مورد نتایج آزمون‌های

دوره های گذشته بود (۱۳). نتایج تحقیقات ذکر شده با مطالعه حاضر همخوانی دارد.

نتایج مطالعه حاضر حاکی از آن بود که گروه زنان با این که بالاترین درصد رعایت اصول میلمن را در بین گروه‌های آموزشی در هر دو سال داشته (۸۶/۷ درصد) ولی قبل و بعد از دوره آموزشی تغییری در رعایت اصول ساختاری سؤالات نداشته است. این امر می‌تواند مؤید این نکته باشد که در گروه‌هایی وضعیت طراحی سؤالات به حد مطلوب نزدیک بود می‌بایست به منظور افزایش کیفیت طراحی سؤالات از شیوه‌های آموزشی تخصصی و دقیق‌تر استفاده نمود.

در مطالعه حاضر گروه بیهوشی در هر دو سال پایین‌ترین میزان رعایت اصول طراحی سؤالات را داشته است. ۶۳ درصد سؤالات این گروه دارای اشکال ساختاری بود. یکی از دلایل آن می‌تواند طراحی سؤالات محدود به یک دوره آزمون دستیاری باشد. در صورتی که دیگر گروه‌ها در طول سال سؤالات چندگزینه‌ای بیشتری طرح می‌کنند و تمامی سؤالات پایان‌ترم آن‌ها توسط دفتر توسعه آموزش دانشکده، مورد بازنگری قرار می‌گیرد و به گروه‌ها بازخورد داده می‌شود که این مسئله می‌تواند در بهبود سؤالات گروه‌های دیگر نسبت به گروه بیهوشی مؤثر باشد، به عبارت دیگر عدم بازخورد سؤالات در حین سال تحصیلی و عدم برگزاری دوره آموزشی برای گروه بیهوشی، می‌تواند از جمله عواملی باشد که ضعیف بودن نتایج به دست آمده در گروه مذکور را باعث شده است.

گروه‌های قلب و روان‌پزشکی در مقایسه با گروه‌های دیگر، تجربه کمتری در طراحی سؤال آزمون ارتقاء داشته و جزء ضعیف‌ترین گروه‌های آموزشی در زمینه طراحی ساختار سؤالات بودند. اما گروه قلب در مقایسه با گروه روان‌پزشکی که دوره آموزشی را طی نکردند از لحاظ آماری پیشرفت معنی‌دار در استانداردهای سؤالات این گروه داشتند.

نتایج مطالعه حاضر نشان داد که در ۵۰ درصد آیت‌ها بعد از کارگاه‌های آموزشی اشکالات ساختاری به طرز معنی‌داری کاهش یافته است. با این که مطالعه حاضر پیشرفت معنی‌دار رعایت اصول طراحی در ساختار سؤالات به‌عنوان یک شاخص کیفی در آزمون ارتقای دستیاری را در دو سال تأیید می‌نماید، هنوز نیمی از سؤالات آزمون دارای اشکال ساختاری است. مطالعات داخل و خارج کشور در زمینه رعایت اصول ساختاری سؤالات نیز نشان می‌دهد که ضعف در طراحی ساختار سؤالات وجود دارد. در مطالعه‌ای که Tarrant بر روی ۲۷۷۰ سؤال از دانشگاه هنگ‌کنگ انجام داد تقریباً نیمی از سؤالات (۴۶/۲)

درصد) دارای اشکالات ساختاری بودند (۱۴). مطالعات دیگر در داخل کشور نیز نشان داد که هنوز در طراحی سؤالات، اشکالات ساختاری در بسیاری از سؤالات آزمون‌های چندگزینه‌ای مشاهده می‌گردد (۱۵-۱۷).

بر اساس چک‌لیست میلمن، تعداد اشکالات گزینه «در سؤال، واحد اندازه‌گیری ذکر شده است» در سال ۹۲ نسبت به سال ۹۱ افزایش یافته است که با بررسی‌های دقیق‌تر مشخص گردید افزایش تعداد این آیت‌ها در سال ۹۲ مربوط به گروه بیهوشی بوده که کارگاه آموزشی نداشته است.

یکی از مواردی که جزء اشکالات ساختاری سؤالات است طراحی سؤالات منفی است. تحقیق حاضر نشان داد ارتباط معنی‌داری بین تعداد خطاها در سؤالاتی که منفی طراحی می‌شوند نسبت به سؤالاتی که تنه سوال مثبت است وجود دارد که با مطالعه Haladyna و همکاران مطابقت دارد. مطالعه مذکور نشان داد سؤالاتی که کلمات منفی دارند معمولاً سریع‌تر و آسان‌تر طراحی می‌شوند در نتیجه احتمال بروز خطا در طراحی سؤالات با تنه منفی بیشتر است (۱۸). لذا پیشنهاد می‌شود طراحان سوال در طراحی سؤالات به این مسئله توجه بیشتری نمایند و از سؤالاتی که تنه سوال مثبت است استفاده نمایند.

نتایج مطالعه حاضر بیانگر آن بود که برگزاری کارگاه‌های آموزشی در رعایت طراحی سؤالات آزمون مؤثر بوده است. نتایج مداخله آموزشی که در دانشگاه علوم پزشکی گرگان در سال‌های ۹۱ و ۹۲ انجام شده بود نیز نشان داد سؤالات آزمون دستیاری پیشرفت معنی‌داری در رعایت اصول ساختاری میلمن داشته است که با نتیجه ما همخوانی دارد (۱۶).

با توجه به نتایج این مطالعه، برگزاری برنامه‌های توانمندسازی اساتید به شرط آن که به صورتی هدفمند بوده و بر اساس اصول فنی مربوطه تمرکز بر تقویت نقاط قوت و رفع نقاط ضعف آزمون‌های قبلی همان طراحان مدنظر قرار گیرد، می‌تواند باعث کاهش تعداد اشکالات ساختاری سؤالات گردد و در ارتقای کیفیت طراحی آزمون‌های دستیاری مؤثر باشد. لذا پیشنهاد می‌شود طراحان سؤال پس از اجرای هر آزمون سؤال‌های آن را تحلیل کنند و با استفاده از نتایج حاصل به تجدیدنظر در آزمون و بهبود کیفیت آن اقدام نمایند.

پیشنهاد می‌گردد با توجه به این که بررسی تاکسونومی سؤالات آزمون نیز همانند ساختار سؤالات از اهمیت بالایی برخوردار است، موضوع تاکسونومی نیز در مطالعات بعدی مورد توجه قرار گیرد.

## References

- 1- David N. E-assessment by design: using multiple – choice tests to good effect. *J Further High Educ* 2007; 31 (1): 53–64.
- 2- Begum T. A Guideline on Developing Effective Multiple Choice Questions and Construction of Single Best Answer format. *J Bangladesh Coll Phys Surg* 2012; 30 (3): 159-66.
- 3- Masters JC, Hulsmeyer BS, Pike ME, Leichty K, Miller MT, Verst AL. Assessment of multiple-choice questions in selected test banks accompanying text books used in nursing education. *J Nurs Educ* 2001; 40 (1): 25-32.
- 4- Quinn FM. *The Principles and Practice of Nurse Education*. 4<sup>th</sup> ed. Cheltenham: Stanley Thorne (Publishers) Ltd; 2000.
- 5- Hammond EJ, McIndoe AK, Sansome AJ, Spargo PM. Multiple-choice examinations: adopting an evidence-based approach to exam technique. *Anaesthesia* 1998; 53(11): 1105-8.
- 6- Meyari A, Beiglarkhani M. [The effect of education on improvement of multiple choice questions' designing in annual residency exams of dental faculty]. *Iran J Med Educ* 2012; 12(1): 36-45 [Persian].
- 7-Vahidshahi K, Mohagheghi MA, Shakeri S, Sabouri M, Razavi SM, Mohammadi M, et al. [Comparison of some quality criteria of multiple choice questions in residency annual exam 2007 & 2008]. Available From: [http://cgme.behdasht.gov.ir/uploads/264\\_781\\_N4\\_Moghayese.html](http://cgme.behdasht.gov.ir/uploads/264_781_N4_Moghayese.html). [Persian]
- 8 -Mohagheghi MA, Vahidshahi K, Shakeri S, Sabouri M, Razavi SM, Mohammadi M, et al. [Evaluation of Some Quality index of Multiple choice Questions in the board exams in Medicine from 2005-2007]. Ministry of Health. [cited 2013 July 20]. Available from: [http://cgme.behdasht.gov.ir/uploads/264\\_975\\_Maghale\\_ShakhesKeifiyat.pdf](http://cgme.behdasht.gov.ir/uploads/264_975_Maghale_ShakhesKeifiyat.pdf). [Persian]
- 9- Ziaiha M, Aslani mehr M, Danesfehane Z, Ziaii A. [motale va tahlil azmoon erteghaye dastyari daneshghah oloom pezesheki Qazvin dar salhaye 88-90]. 13<sup>th</sup> national conference on medical education; 2012 April 30- May 3; Sari: mazandaran university of medical sciences; 2012. [Persian]
- 10- Allami A, Derakhshan F, Safdari F. [Design and evaluation of multiple choice questions (MCQ)]. Qazvin: Qazvin University of medical science; 2011. [Persian]
- 11- Khoshrang H, Taheri M, Asadi A, Heidarzadeh A. [Quality of Residents' Promotion Exams before & after Educational Intervention in 2010-2011]. *Iran J Med Educ* 2013; 13 (7): 551-60. [Persian]
- 12- Meyari M, Beiglarkhani M. [Improvement of Design of Multiple Choice Questions in Annual Residency Exams by Giving Feedback]. *Strides Dev Med Educ*. 2013; 10 (1): 109-118. [Persian]
13. Pour mirza kalhori R, Roshan pour F, Rezaei M, Naderipour A. [The effect of university board members' scholarship program in multiple choice questions designing on Results of medical residency promotion Exams analysis in 2009]. *Behbood* 2011; 15 (2): 112-18. [Persian]
- 14- 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.

- 15- Shakurnia A, Khosravi Borojeni A, Mozaffari A, Elhampour H. [An Evaluation of Exam Questions Designed by Faculty Members, Emphasizing on the Multiple Choice Question Structure, Ahvaz Jundishapur University of Medical Sciences, 2007]. *Strides Dev Med Educ* 2009; 6(2): 129-38. [Persian]
- 16- Arab M. [Survey on defect of multiple choice question in medical school of Hamadan Medial University in 2000]. 4<sup>th</sup> National Congress of Medical Education; Tehran: Tehran University of Medical Sciences ;2000. [Persian]
- 17- Yosefi MR, Taziki MH, Rajaii S, Ghanbari MR. [moghayese keifiyat soalat azmoon ertegha dastyari ghabl va baad az modakhele dar daneshkade pezeshti Ghorghan 2012-2013]. 15<sup>th</sup> National Congress of Medical Education; 2014 April 29- May 1; Yazd: Yazd University of Medical Sciences;2014 [Persian]
- 18- Haladyna TM, Downing SM. Validity of taxonomy of multiple-choice item-writing rules. *Applied Measurement in Education* 1989; 2 (1): 51-78.

# Effect of Faculty Training Programs on Improving Quality of Residency Exams in 2013-2014

Derakhshan F<sup>1</sup>, Allami A<sup>2</sup>, Ahmadi S<sup>\*3</sup>

Received:2014/8/4

Accepted:2014/12/28

## Abstract

**Introduction:** Multiple choice questions are the most common and objective test in medical science education. The aim of this study is to determine effectiveness of faculty development program on quality of residency exams in medical school of Qazvin University of Medical Sciences (QUMS) in Iran.

**Methods:** This was a Quasi-experimental study (before & after intervention) .Faculty development program was conducted in three phases: First, a book titled "design and evaluation of multiple-choice questions "was given to all departments. All questions from residency exams of QUMS for 7 specialty programs in 2013 (there were 150 questions for each specialty, totaling 1050 questions) were structurally evaluated using a valid and reliable checklist based on Millman principles. In the second phase, feedback of survey result were given to all 7 groups. Finally, educational workshops, one for each department were organized before residency exams in 2014. Data related to questions designed by faculties who had participated in the interventional workshop were analyzed using statistical T-test and ANOVA.

**Results:** Overall, 2100 questions related to 7 departments including: pediatric, surgery, internal medicine, cardiology, psychiatry, obstetrics and gynecology, anesthesiology were analyzed. Findings showed that poorly structural questions in 2013 were  $0.58 \pm 0.02$ , and in 2014, were  $0.44 \pm 0.02$ . Using a t-test of independent groups showed significantly different before and after training ( $p < 0.001$ ).

**Conclusion:** Current study showed that faculty development programs are significantly effective in promoting and it needs to be developed.

**Key words:** Workshops , Examination Questions, Residency, Training Programs, Medical

**Corresponding Author:** Ahmadi S, Department of Medical Education, Shahid Beheshti University of Medical Sciences, Tehran, Iran. soleiman.ahmady@gmail.com

Derakhshan F, Department of Medical Education, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

Allami A , Department of Infectious Diseases, Qazvin University of Medical Sciences, Qazvin, Iran.