

## ارتباط شاخص های کمی و کیفی آزمون های چند گزینه ای ارتقای دستیاری تخصصی پزشکی دانشگاه علوم پزشکی کرمانشاه در سال ۱۳۹۱

### چکیده

زمینه: آزمون های چند گزینه ای متداول ترین ابزارهای ارزشیابی در علوم پزشکی هستند. این مطالعه با هدف بررسی ارتباط شاخص های کمی و کیفی آزمون های چند گزینه ای ارتقای دستیاری تخصصی پزشکی دانشگاه علوم پزشکی کرمانشاه در سال ۱۳۹۱ انجام شد.

**روش ها:** برای انجام این مطالعه توصیفی-همبستگی، ۹۵۰ سوال از آزمون های ارتقاء دستیاری تخصصی پزشکی شش گروه آموزشی که به صورت تصادفی انتخاب شده بودند، مورد آنالیز کمی و کیفی قرار گرفتند. داده های کمی شامل ضریب دشواری و ضریب تمیز و داده های کیفی شامل درصد تاسونومی I و II و III سوالات بودند. آنالیز داده ها با نرم افزار SPSS نسخه ۱۸ و آزمون های تی و کای دو و ضریب همبستگی پیرسون انجام شد.

**یافته ها:** میانگین ضریب دشواری کل سوالات  $(\pm 0/24)$  و میانگین ضریب تمیز کل سوالات  $(\pm 0/23)$  بود. ضریب همبستگی اسپیرمن نشان داد که ارتباط معکوسی بین ضریب تمیز و ضریب دشواری سوالات وجود داشت به این معنی که هر چه سوالات ساده تر بودند ضریب تمیز سوال کمتری شد ( $P=0/001$ ). ارتباط آماری معنی داری بین ضریب دشواری سوالات و سطوح تاسونومی و بین ضریب تمیز سوالات و سطوح تاسونومی نیز وجود نداشت.

**نتیجه گیری:** نتایج این مطالعه نشان داد که ضریب تمیز و ضریب دشواری با هم ارتباط معکوسی دارند و طراحی سوالات ساده و بسیار سخت باعث کاهش ضریب تمیز می شود. پیشنهاد می شود تا طراحان آزمون های ارتقاء دستیاری با انتخاب سوالات دارای ضریب های دشواری و تمیز مناسب کیفیت این امتحانات جامع را افزایش دهند.

**کلید واژه ها:** سوالات چند گزینه ای، ضریب دشواری، ضریب تمیز

رضا پورمیرزا کلهری<sup>۱</sup>، منصور رضایی<sup>۲</sup>،

امیر رضا شجاعی مقدم<sup>۳</sup>، ویدا سپاهی<sup>۴\*</sup>،

لیدا معمار افتخاری<sup>۴</sup>

۱. گروه فوریت های پزشکی، دانشکده پیراپزشکی، دانشگاه علوم پزشکی کرمانشاه، کرمانشاه، ایران.

۲. گروه آمار حیاتی، دانشکده بهداشت، دانشگاه علوم پزشکی کرمانشاه، کرمانشاه، ایران.

۳. ستاد مرکزی جمعیت هلال احمر جمهوری اسلامی ایران، تهران، ایران.

۴. معاونت آموزشی دانشگاه علوم پزشکی کرمانشاه، ایران.

\* **عهده دار مکاتبات:** کرمانشاه، معاونت آموزشی دانشگاه علوم پزشکی.

Email: vidasep2002@yahoo.com

### مقدمه:

یکی از مهارت های پس از تدریس مدرسین ارزشیابی یادگیری است<sup>۱</sup>. ارزشیابی یادگیری با آزمون انجام می شود<sup>۲</sup>. در علوم پزشکی دو نوع سنجش تراکمی (تصمیم گیری در مورد آینده دانشجویان) و تکوینی (برای راهنمایی دانشجویان) وجود دارد. به دلیل نقش مهم آزمون در سنجش یادگیری ضروری است که هر آزمون دارای سه شاخص روایی، پایایی و عملی بودن باشد<sup>۳</sup>. آزمون ها در دو دسته کلی آزمون های معلم ساخته و آزمون های استاندارد طبقه بندی می شوند<sup>۴</sup>. دسته بندی دیگر آزمون ها از لحاظ عینیت در تصحیح است. این خصوصیت آزمون را به دو نوع عینی و شفاهی تقسیم می نماید. منظور از آزمون عینی

آزمونی است که در صورت تصحیح مکرر و توسط افراد مختلف دارای نتیجه ای یکسان باشد. از آزمون های عینی می توان آزمون های صحیح - غلط و چند گزینه ای را نام برد<sup>۵</sup>. شایع ترین سوالات عینی مورد استفاده سوالات چند گزینه ای هستند. هر سوال دارای بخش ریشه و گزینه های پیشنهادی است که یکی از آنها پاسخ صحیح است و فراگیر باید پاسخ صحیح را از بین گزینه های پیشنهادی/انحرافی انتخاب کند<sup>۴</sup>. سوالات چند گزینه ای در حیطه شناختی می توانند دانش، فهم، ادراک، کاربرد و اطلاعات را ارزشیابی کنند<sup>۶</sup>. به طور معمول ضروری است که بعد از اجرای هر آزمون تجزیه و تحلیل سوالات از لحاظ کمی و کیفی انجام شود.

وجود گزارش مبنی بر عدم ارتباط بین ضریب های دشواری و سطوح شناختی سوالات و وجود ارتباط بین ضریب تمیز و سطوح شناختی سوالات دوره مقدماتی پزشکی<sup>۱۱</sup> می توان انتظار داشت که که مدرسین علوم پزشکی با بالاتر بردن سطوح تاکسونومی سعی در طراحی سوالاتی با قدرت تمیز بیشتری هستند و این می تواند که نشانه استقبال بیشتر مدرسین از سوالات چند گزینه ای به جای سوالات شفاهی در شناسایی مهارت های شناختی دانشجویان باشد<sup>۱۳</sup>. از دیدگاهی دیگر با بررسی گزارشاتی که از عدم خطی بودن ارتباط ضریب دشواری و تمیز در سوالات آموزش پزشکی وجود دارد<sup>۱۴</sup> می توان انتظار داشت که مدرسین علوم پزشکی با هدف بالاتر بردن قدرت تمیز سوالات آنها را مشکل تر طراحی می کنند. با وجود این مطالعات ضروری به نظر می رسد که در مورد ارتباط شاخص های کمی و کیفی آزمون های چند گزینه ای بخصوص در مقاطع دستیاری تخصصی پزشکی بررسی های بیشتری انجام داد لذا این مطالعه با هدف بررسی ارتباط شاخص های کمی و کیفی آزمون های چند گزینه ای ارتقای دستیاری تخصصی پزشکی دانشگاه علوم پزشکی کرمانشاه در سال ۱۳۹۱ انجام شد.

### مواد و روش ها :

برای انجام این مطالعه توصیفی - همبستگی از مجموع کل آزمون های ارتقای دستیاری تخصصی رشته های پزشکی، که توسط اساتید هیات علمی دانشگاه علوم پزشکی کرمانشاه در سال ۱۳۹۱ طراحی شده بودند، شش آزمون به طور تصادفی انتخاب شدند و حدود ۹۵۰ سوال مورد تجزیه و تحلیل کمی و کیفی قرار گرفتند. ابزار جمع آوری داده های گزارشات نرم افزار کشوری مستقر در دانشکده پزشکی دانشگاه علوم پزشکی کرمانشاه بود که تحت برنامه ویندوز XP آزمون های وابسته به هنجار و ملاک را با استفاده از نظریه کلاسیک آزمون ها CTT تحت تحلیل کمی قرار می داد. بر اساس مطالعه وفا مهر و همکارانش، ضریب دشواری بین ۰/۷-۰/۳ مناسب، بیشتر از ۰/۷ آسان و کمتر از ۰/۳ سخت و ضریب تمیز بالاتر از ۰/۳ مناسب، صفر تا ۰/۳ کم و زیر صفر ضعیف طبقه بندی شدند<sup>۱۱</sup>. برای آنالیز کیفی آزمون و تعیین سطوح تاکسونومی سوالات از نظرات چهار نفر کارشناس

بعد کیفی آنالیز سوالات منطبق بر اهداف آموزشی بنیامین بلوم و شامل: تاکسونومی I (سطوح دانشی)، تاکسونومی II (یاد آوری) و تاکسونومی III (کاربرد مفاهیم) و تعیین درصد سوالات بدون اشکال ساختاری است. در بعد کمی اعتبار کل آزمون و ضریب های دشواری و تمیز آزمون سنجش می شود. در سالیان اخیر این کار با استفاده از نرم افزار هایی که با استفاده از نظریه کلاسیک آزمون ها Classical Test Theory (CTT) برنامه ریزی شده اند تحلیل کمی آزمون ها انجام می شود.

در آزمون کتبی بوردهای تخصصی پزشکی ایالات متحده از آزمون های چند گزینه ای استفاده می شود<sup>۱۵،۱۶</sup>. در ایران نیز آزمون های چند گزینه ای متداول ترین آزمون های کتبی در آموزش علوم پزشکی هستند. باید به این اصل توجه داشت که هدف از به کارگیری تکنیک های نوین آموزشی، افزایش سطح علمی و توانایی عملی دانشجویان است. این هدف در رشته های علوم پزشکی از اهمیت بیشتری برخوردار است. مطالعات مختلفی گزارش کرده اند که کیفیت آزمون بر کیفیت فرآیند آموزش موثر است<sup>۱۷،۱۸</sup>. یکی از روش های نوین ارزشیابی مدرسین بررسی کیفیت آزمون های طراحی شده توسط مدرسین است. در ایران این بررسی در مراکز توسعه و آموزش علوم پزشکی انجام می شود و در طول سالیان اخیر مورد توجه بسیاری از مدرسین قرار گرفته است به طوری که علاوه بر تحلیل آزمون های ارتقاء دستیاری، آزمون های معلم ساخته سایر مدرسین نیز مورد تحلیل کمی و کیفی قرار می گیرند<sup>۱۱</sup>.

مقالات متعددی در داخل کشور به تجزیه و تحلیل آزمون های چند گزینه ای پرداخته اند و گزارشات متفاوتی از شاخص های کمی و کیفی آزمون ها وجود دارد<sup>۱۶-۱۲</sup> که به نامناسب بودن برخی شاخص های کمی و کیفی آزمون ها اشاره کرده اند، حتی سوالات آزمون های ارتقاء دستیاری تخصصی پزشکی نیز از این روند مستثنی نبوده اند<sup>۱۷،۱۸</sup>. البته باید در نظر داشت که این گزارشات فقط اختصاص به ایران ندارد و در مقاطع مختلف پزشکی سایر کشورها نیز گزارشات مشابهی وجود دارد<sup>۱۹-۲۲</sup>. اما نکته قابل توجه در این تجزیه و تحلیل ها توجه کمتر به بررسی ارتباط شاخص های کمی و کیفی با یکدیگر است. با

شماره دو فاصله اطمینان ضریب دشواری را در گروه های مورد مطالعه نشان می دهد.

در بررسی مجموع کل سوالات از لحاظ ضریب تمیز و میانگین ضریب تمیز کل سوالات (۲۳/۰ ± ۲۶/۰) بود. بهترین ضریب تمیز در سوالات گروه داخلی وجود داشت. ضریب همبستگی پیرسون نشان داد که بین میزان ضریب تمیز سوالات و گروه آموزشی ارتباط وجود دارد ( $P < ۰/۰۰۱$ ). نمودار شماره سه فاصله اطمینان ضریب تمیز را در گروه های مورد مطالعه نشان می دهد.

ضریب همبستگی اسپیرمن - رو نشان داد که ارتباط معکوسی بین ضریب تمیز و ضریب دشواری سوالات وجود دارد به این معنی که هر چه سوالات ساده تر (دارای ضریب دشواری بیشتر) بودند ضریب تمیز سوال کمتری شد ( $P = ۰/۰۰۱$ ). ارتباط آماری معنی داری بین ضریب دشواری سوالات و سطوح تاکسونومی سوالات وجود نداشت. این عدم ارتباط آماری معنی دار بین ضریب تمیز و سطوح تاکسونومی سوالات نیز وجود داشت ( $P < ۰/۰۰۱$ ).

#### بحث:

در این مطالعه دو شاخص کمی مهم ضریب دشواری و ضریب تمیز در سوالات گروه های آموزشی مورد مطالعه مورد بررسی قرار گرفتند. در مجموع کل سوالات از لحاظ ضریب دشواری ۹/۷٪ سوالات سخت و ۵۹/۵٪ مناسب و ۳۰/۸٪ آسان بودند. در مقایسه با مطالعه قبلی در مورد شاخص های کمی آزمون های ارتقاء دستیاری دانشگاه علوم پزشکی کرمانشاه در سال ۱۳۸۸ نتایج این مطالعه نشان دهنده مناسبت بیشتری در درجه دشواری سوالات در مطالعه حاضر است. در مطالعه وفامهر و همکارانش از نظر ضریب دشواری سوالات آسان ۴۲/۲٪، سوالات مناسب ۲۸/۸٪ و ۲۹٪ سوالات دشوار بودند<sup>۱۱</sup>.

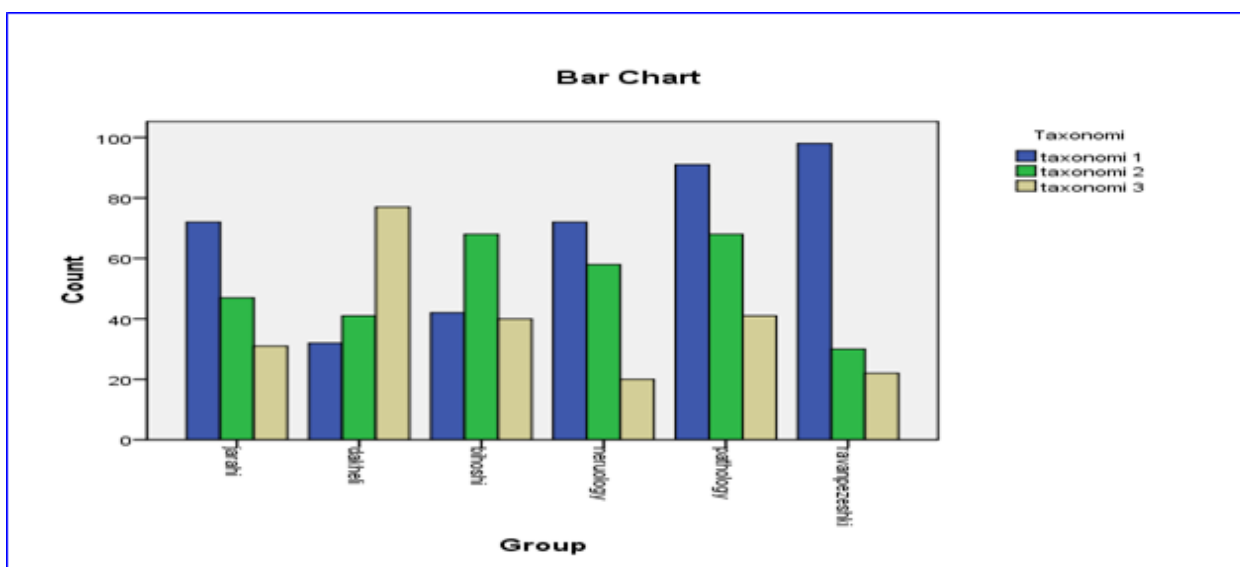
در این مطالعه بین دو شاخص ضریب دشواری و ضریب تمیز ارتباط معکوسی وجود داشت که از نظر آماری معنی دار بود به این مفهوم که با افزایش ضریب دشواری (آسان تر شدن سوال) ضریب تمیز سوال کمتر می شود.

مرکز مطالعات و توسعه آموزش علوم پزشکی دانشگاه کرمانشاه استفاده شد و مبنای تاکسونومی هر سوال توافق غالب آنها در مورد سطح حیطه شناختی سوالات بود. این گزارشات منطبق بر اهداف آموزشی بنجامین بلوم و شامل: تاکسونومی I (سطح دانش بکار رفته)، تاکسونومی II (یاد آوری) و تاکسونومی III (کاربرد مفاهیم) است. تجزیه و تحلیل داده ها با نرم افزار SPSS نسخه ۱۸ انجام شد. برای تحلیل سوالات برحسب ویژگی های کمی و کیفی جداول دو بعدی به صورت تعداد و درصد و برای مقایسه با استاندارد از آزمون t یک نمونه ای و مجذور کای استفاده شد. برای بررسی همبستگی شاخص های کمی و کیفی از ضریب همبستگی پیرسون استفاده شد. در این مطالعه به منظور حفظ شئون معلمی و محرمانه ماندن اطلاعات آنها به اسامی طراحان سوالات اشاره نشد و اجازه انجام پژوهش از ریاست مرکز توسعه و آموزش پزشکی اخذ شد.

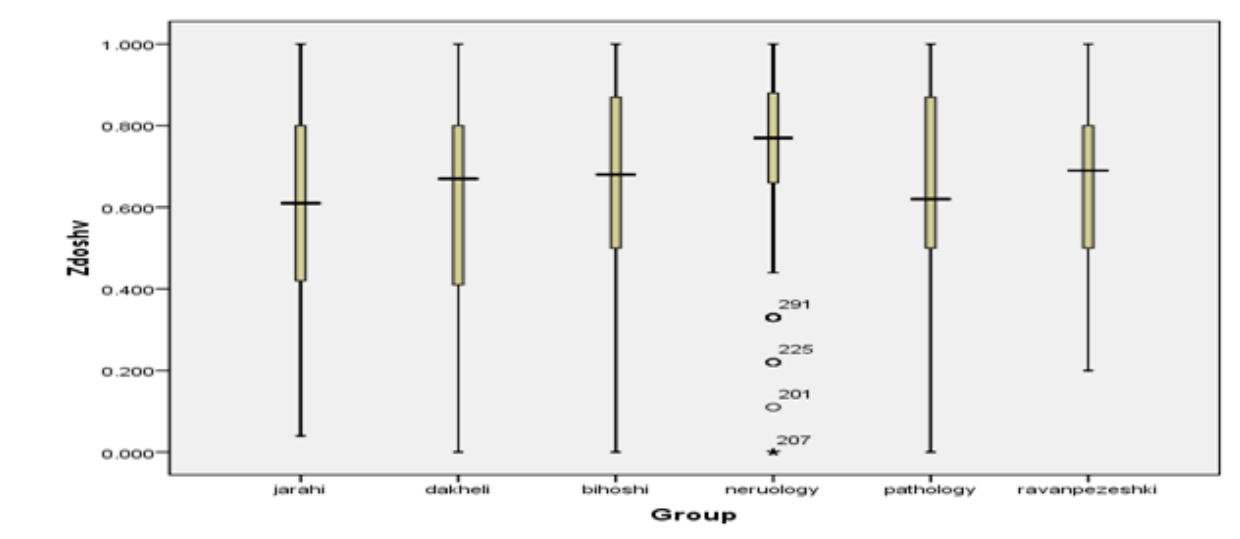
#### یافته ها:

در این مطالعه سوالات آزمون ارتقاء دستیاری شش گروه آموزشی که به طور تصادفی از بین گروه های موجود انتخاب شده بودند مورد بررسی قرار گرفتند و در مجموع ۹۵۰ سوال مورد بررسی کمی و کیفی قرار گرفتند و حجم نهایی نمونه را تشکیل دادند. بر اساس توافق نظرات کارشناسی غالب کارشناسان مرکز مطالعات و توسعه آموزش علوم پزشکی دانشگاه کرمانشاه از مجموع کل سوالات ۴۲/۸٪ سوالات در تاکسونومی I، ۳۲/۸٪ در تاکسونومی II و ۲۴/۴٪ در تاکسونومی III طراحی شده بودند. ضریب همبستگی پیرسون نشان داد که بین میزان سطوح تاکسونومی سوالات و گروه آموزشی ارتباط وجود دارد ( $P = ۰/۰۰۱$ ). نمودار شماره یک سطوح تاکسونومی را در گروه های مورد مطالعه نشان می دهد.

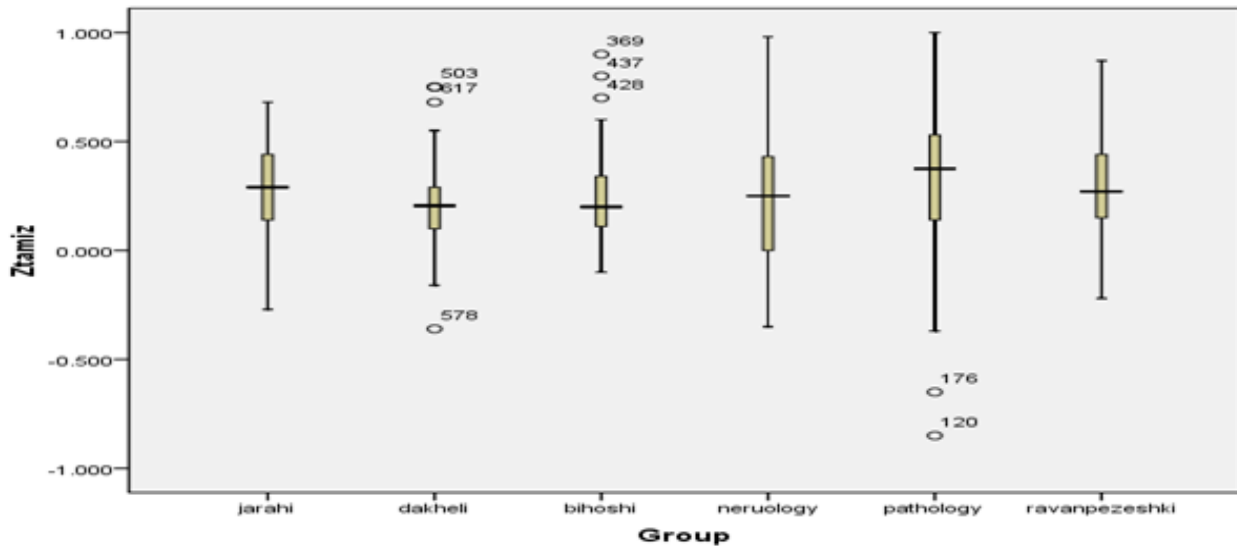
در بررسی مجموع کل سوالات از لحاظ ضریب دشواری ۹/۷٪ سوالات سخت و ۵۹/۵٪ مناسب و ۳۰/۸٪ آسان بودند و میانگین ضریب دشواری کل سوالات (۲۴/۰ ± ۶۴/۰) بود. سخت ترین سوالات در گروه آسیب شناسی طراحی شده بود. ضریب همبستگی پیرسون نشان داد که بین میزان ضریب دشواری سوالات و گروه آموزشی ارتباط وجود دارد ( $P = ۰/۰۰۱$ ). نمودار



نمودار ۱. سطوح تاکسونومی را در گروه های مورد مطالعه



نمودار ۲. فاصله اطمینان ضریب دشواری را در گروه های مورد مطالعه



نمودار ۳. فاصله اطمینان ضریب تمیز را در گروه های مورد مطالعه

در این مطالعه مجموع کل سوالات از لحاظ ضریب تمیز ۳۸/۲٪ ضعیف، ۳۴/۲٪ متوسط و ۲۷/۶٪ مناسب بودند. نکته که باید به آن توجه داشت این است که در تفسیر داده های حاصل از تجزیه و تحلیل آزمون ها باید احتیاط کرد زیرا ضریب تمیز و یا تفاوت عملکرد دو گروه قوی و ضعیف از دانشجویان همیشه بیان کننده کیفیت سوالات نمی باشد و نباید این سوالات از بانک سوالات حذف شوند بلکه باید به دقت بررسی شوند و به نقش آنها در اندازه گیری دستیابی به اهداف آموزشی توجه داشت. در حقیقت بودن سوالاتی که دارای درجه دشواری آسان یا بسیار سخت هستند که دارای درجه تمیز پایینی می باشند و در آزمون ها وجود آنها برای دستیابی به نمونه کاملی از محتوا و اهداف آموزشی ضروری است.<sup>۵</sup> نکته ی قابل توجهی که در مطالعه Hingorjo و همکارانش که در سال ۲۰۱۲ انجام گرفته، این است که یک آزمون چند گزینه ای ایده آل باید دارای درجه دشواری متوسط، ضریب تمیز بالا و گزینه های انحرافی عملکردی functioning distractors کمتر از ۳ باشد.<sup>۷</sup> ضروری به نظر می رسد که در مطالعات آینده درپایش مناسب یک آزمون در کنار پایایی آزمون، ضریب دشواری و ضریب

بر اساس مفهوم این یافته با آسان طراحی کردن سوالات قدرت سوال در تفکیک دانشجویان زرنگ و درس خوان از دانشجویانی که به طور شانسی به سوال پاسخ می دهند کمتر می شود. این یافته ای است که در بخشی از یافته های مطالعه Sim و همکارانش نیز به آن اشاره شده است.<sup>۲۴</sup> مطالعاتی که در سال ۲۰۱۳ انجام گرفته است به ارتباط ضعیف ضریب دشواری و ضریب تمیز ( $r = 0/19$ ) و قرار داشتن بهترین ضریب تمیز در محدوده ۰/۸-۰/۶ با سوالات دارای درجه دشواری متوسط<sup>۲۵</sup> در سوالات رشته فیزیولوژی اشاره می کنند. نکته خاصی که باید به آن توجه داشت این است که گزارش این ارتباط ضعیف / عدم ارتباط بین ضریب دشواری و ضریب تمیز را نه تنها در یک ترم بلکه در طول آزمون های سال ۲۰۱۲-۲۰۰۸ در سوالات داروشناسی دانشکده راجکوت هندوستان می توان مشاهده کرد.<sup>۲۶</sup> این تفاوت در نتایج مطالعات مختلف می تواند منعکس کننده این اصل باشد که مدرسین طراح سوالات چند گزینه ای علاوه بر درجه دشواری سوال باید به قدرت تمیز سوالات نیز توجه کنند چون نمی توان به صرف سخت طراحی کردن سوالات آزمون قدرت تمیز آن را افزایش داد.

تمییز به گزینه های انحرافی عملکردی / غیر عملکردی نیز توجه خاص داشت.

در این مطالعه ارتباط معنی داری بین سطوح تاکسونومی سوالات و ضریب تمییز و ضریب دشواری مشاهده نشد. ولی در مطالعه وفامهر و همکارانش سوالاتی که در سطح تاکسونومی بالاتری قرار داشتند دارای ضریب تمییز بهتری بودند ولی سطح تاکسونومی سوالات با درجه دشواری ارتباطی نداشت<sup>۱۱</sup>. Kibble و همکارش در سال ۲۰۱۱ گزارش کرده اند که سطوح تاکسونومی سوالات نه تنها با ضریب تمییز بلکه با نمرات دانشجوی نیز ارتباط ندارند<sup>۲۸</sup>. اما Kim و همکارانش در سال ۲۰۱۲ گزارش نموده اند که سوالات دارای سطوح تاکسونومی بالاتر در حیطه شناختی بلوم Bloom دارای ضریب تمییز بالاتر و درجه دشواری سخت تری هستند و می توانند جایگزین سوالات شفاهی در ارزشیابی مهارت های تفکر انتقادی در جمعیت های بزرگ دانشجویی شوند<sup>۲۹</sup>. در این مطالعه شاخص های کمی ضریب تمییز و ضریب دشواری و سطوح تاکسونومی سوالات در گروه های مختلف آموزشی با هم دارای تفاوت آماری معنی داری بودند که مشابه نتایج مطالعاتی بود که به بررسی سوالات

آزمون های ارتقاء دستیاری دانشگاه گیلان<sup>۳۰</sup> و کرمانشاه<sup>۳۱</sup> می پرداخت.

### نتیجه گیری:

نتایج این مطالعه نشان داد که شاخص های ضریب تمییز و ضریب دشواری با هم ارتباط دارند و طراحی سوالات ساده و بسیار سخت باعث کاهش ضریب تمییز می شوند. در این مطالعه سطوح تاکسونومی سوالات نیز با شاخص های ضریب تمییز و ضریب دشواری ارتباطی نداشت توصیه می شود تا طراحان محترم سوال با مراجعه به بانک های سوالات خود نسبت به انتخاب سوالات دارای ضریب دشواری متوسط برای آزمون های آینده اقدام کنند و در طراحی آزمون ها با فرض افزایش دشواری سوالات سایر شاخص های آزمون را دستخوش تغییرات ناخواسته نکنند. پیشنهاد می شود در مطالعات آینده ارتباط شاخص های ضریب تمییز و ضریب دشواری سوالات با اعتبار آزمون ها و ساختار سوالات مورد بررسی قرار گیرد. این گزارش منتج از نتایج طرح تحقیقاتی مصوب در مرکز مطالعات و توسعه آموزش علوم پزشکی دانشگاه علوم پزشکی کرمانشاه در تاریخ ۱۳۹۲/۳/۱ و شماره ۹۲۱۳۱ است.

### References:

- Shabani H. [Instructional Skills, Methods and Techniques of teaching] 15 thed. Ghom.SAMT. 2007: 367. [Persian]
- Saif AA. Educational measurements, assessment and evaluation. 3th ed. Tehran: Doran Publication; 2010:32,205. [Persian]
- Newble DI, Cannon RA. A handbook for medical teachers. 4th ed. Boston: Kluwer academic publications; 2001.
- Pourmirza kalhori R, Rezaei M, Karami Matin B, Roshanpour F. A survey of quality and quantity indexes of multiple choice question (MCQ) exams of medical residents at Kermanshah University of Medical Sciences: 2008-2012. J Med Edu Dev 2014; 8(4): 64-75.
- Pourmirza Kalhori R, Roushanpour F, Rezaei M, Esmaeili K. Analysis of multiple choice tests designed by faculty members of Kermanshah

University of Medical Sciences. Edu Res Med Sci 2013; 2(2):18-19.

6. Rezaei H, Shogi F, Dolati Y, Mohagheghi M, Malak zadeh R. [Measurement and Evaluation in Medicine]. 1 sted Tehran. Ameh. 1998:75. [Persian]

7. McCoubrie P. Improving the fairness of multiple choice questions: a literature review. Med teach 2004; 26(8): 709-12.

8. Newell FW, Leopold IH, Adler FH. The multiple choice question test of the American Academy of ophthalmology. The American ophthal. Society Publication 1970; 68: 163-170.

9. Larsen DP, Butler AC, Ruediger HL. Test-enhanced learning in medical education. Med Edu 2008; 42(10): 959-66.

10. Reid WA, Duvall E, Evans P. Relationship between assessment results and approaches to learning and studying in year two medical students. Med Edu 2007; 41(8): 754-62.

11. Vafamehr V, Dadgostarnia M. Reviewing the results of qualitative and quantitative analysis of MCQs in Introduction to clinical medicine course. *IJME* 2010; 10(5):1146-1152. [Persian]
12. Grak yaraghi M, Toghyanifar N, Esmaili A, Avizhg M, Mazaheri M. [Evaluation of MGQ on clinical Trianing in Esfahan Medical University 2005 – 2006]. *Strides in Development of Medical Education*. 8 th National Congress of Education – March 2007: 73. [Persian]
13. Javadi M, Abbaszade A, Borhani F, Abdoli R. [Analytic Evaluation of MCQ in End Termister Exams in Nursing School of Baam] *Strides in Development of Medical Education*. 8 th National Congress of Medical Education – March 2007: 217. [Persian]
14. Drakhsan F, Jahanihashemi H. [Evaluation of End Term Exams of Ghazvin Medical University in 2007-2008] *Special Supplement*. 10 th National Congress of Medical Education – May 2009: 308-309. [Persian].
15. Latifi SM, Shakornia A. [Comporision of Experts behaviors and Wives about MCQ Disigend in Ahvaz Medical University]. *Strides in Development of Medical Education*. 8 th National Congress of Medical Education – March 2007: 159-160. [Persian].
16. Haghshenas M, Vahid Shahi K, mahmudi M, Shahbazenejod L, Parvinne jad 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 Eduction* 2008 ; 5(2) :120-127. [Persian]
17. PourmirzaKalhori R, Razazian N, Naderipour A, et al. Educational purpose and medical students'achievements in clinical training in an university hospital in Kermanshah. *J Kermanshah Univ Med Sci Behhood* 2010; 13(4):344-52. [Persian]
18. Mohagheghi MA, Vahid shahi K, Shakeri S, Sabori 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]. (Cited Sep, 2011).available at:[Http://cgme.behdasht.gov.ir](http://cgme.behdasht.gov.ir).
19. Mitra NK, Nagaraja HS, Ponnudurai G, Judson JP. The Levels of Difficulty and Discrimination Indices in Type a Multiple Choice Questions Of Pre-clinical Semester 1, Multidisciplinary Summative Tests. *IeJSME* 2009; 3 (1): 2-7.
20. Quadrelli S, Davoudi M, Galíndez F, Colt HG. Reliability of a 25-item low-stakes multiple-choice assessment of bronchoscopic knowledge. *Chest* 2009; 135(2):256-8.
21. Paulino AC, Kurtz E. American College of Radiology in training examination for residents in radiation oncology. *Int J Radiat oncol Biol phys* 2008; 70(3): 666-670..
22. 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.
23. Khan MU, Aljarallah BM. Evaluation of Modified Essay Questions (MEQ) and Multiple Choice Questions (MCQ) as a tool for assessing the Cognitive Skills of Undergraduate Medical Students. *Int J Health Sci (Qassim)* 2011; 5(1):39-43.
24. Sim SM, Rasiah RI. Relationship between item difficulty and discrimination indices in true/false-type multiple choice questions of a Para-clinical multidisciplinary paper. *Ann Acad Med Singapore* 2006; 35(2):67-71.
25. Pande SS, Pande SR, Parate VR, Nikam AP, Agrekar SH. Correlation between difficulty & discrimination indices of MCQs in formative exam in Physiology . *South-East Asian Journal of Medical Education* 2013; 7(1):45-50.
26. Karelia BN, Pillai A, Vegada BN. The levels of difficulty and discrimination indices and relationship between them in four-response type multiple choice questions of pharmacology summative tests of Year II M.B.B.S students. *IeJSME* 2013 7(2): 41-46.
27. Hingorjo MR, Jaleel F. Analysis of One-Best MCQs: the Difficulty Index, Discrimination Index and Distractor Efficiency. *J Pak Med Assoc* 2012; 62(2):142-147.
28. Kibble JD, Johnson T. Are faculty predictions or item taxonomies useful for estimating the outcome of multiple-choice examinations? *Adv Physiol Educ* 2011; 35(4):396-401.
29. Kim MK, Patel RA, Uchizon JA, Beck L. Incorporation of Bloom's Taxonomy into Multiple-Choice Examination Questions for a Pharmacotherapeutics Course. *Am J Pharm Edu* 2012; 76(6): 114.
30. Khoshrang H, Taheri M, Asadi A, Heidarzadeh A. [ Quality of Residents' Promotion Exams before & after Educational Intervention in 2010-2011] .*Iranian J Med Edu* 2013; 13(7): 559-560. [Persian]
31. Pourmirza Kalhori R, Darabi F, Roshanpour F, Rezaei M. Analytic assessment of multiple choice questions (MCQ) of faculties' members' scholarship in Kermanshah University of Medical Sciences, 2010-2011. *J Clin Res Paramed Sci* 2013; 2(1):43-51. [Persian]

## Correlation of quality and quantity index of multiple choice questions exams of residency promotion in Kermanshah University of medical sciences, 2013

Reza Pourmirza Kalhori<sup>1</sup>,  
Mansour Rezaie<sup>2</sup>, Amir Reza  
Shojee Moghadam<sup>3</sup>, Vida  
Sepahi<sup>4\*</sup>, Lida Memar  
Eftekhari<sup>4</sup>

1. Department of Emergency Medicine, Paramedicine School, Kermanshah University of Medical Sciences, Kermanshah, Iran.
2. Department of Biostatistics and Epidemiology, School of Health, Kermanshah University of Medical Sciences, Kermanshah, Iran.
3. Red Crescent society of Islamic republic of IRAN, Tehran, Iran.
4. Deputy of education in Kermanshah University Medical Sciences, Kermanshah, Iran.

**\*Corresponding Author:**  
Kermanshah, Deputy of  
education in Kermanshah  
University Medical Sciences.

**Email:** vidasep2002@yahoo.com

### Abstract

**Background:** Multiple choice questions are the most current tools to evaluate the learning in medical science. The purpose of this study is determining Correlation of quality and quantity index of multiple choice questions exams of residency promotion in Kermanshah University of medical sciences in 2013.

**Methods:** For done of this descriptive collative study 950 MCQs tests that designed by in six residency medical groups were analyzed. Quantity data was collected by discrimination index and difficulty level and the quality data by percent of Taxonomy I & II & III questions. Data was analyzed by SPSS-18 software and independent T test and Chi-square and person's test.

**Results:** In total tests Taxonomy I 42.8, Taxonomy II 32.8 and Taxonomy III 24.4 were been. Mean of difficulty level was 0.64 ( $\pm 0.24$ ) and Mean of discrimination index was 0.26 ( $\pm 0.23$ ). A significant statistically reversed correlation was observed with Pierson's test ( $P=0.001$ ) in difficulty level and discrimination index. This finding was seen in percent of questions without structural defects with mean of 75.23 % ( $\pm 13.23$  ( $P=0.041$ )). No significant statistically correlation was observed between difficulty level and Taxonomy levels and between discrimination index and Taxonomy levels.

**Conclusion:** The study showed that reversed correlation was observed in difficulty level and discrimination index and design of simple and very hard tests induced decreasing discrimination index. In future suggested to designer of exams in medicine with selection of appropriate difficulty level and discrimination index obtain better quality of residency promotion exams.

**Keywords:** Multiple choice questions, discrimination index, difficulty level

### How to cite this article

Pourmirza Kalhori R, Rezaie M, Shojee Moghadam AR, Sepahi V, Memar Eftekhari L. Correlation of quality and quantity index of multiple choice questions exams of residency promotion in Kermanshah University of medical sciences, 2013. J Clin Res Paramed Sci 2015; 4(1):71-78.