مجلـــه‌ی توســـعه‌ی آمـــوزش در علـوم پزشـــکی
دوره‌ی ۸، شماره ۱۸، تابستان ۹٤ ، صفحات ٤۳ تا ۵۳

مقایسه‌ی برخی شاخص های کیفیت سووالات چند گزینه‌ای آزمون‌های کتبی ارتقای دانشگاه علوم پزشکی تبریز و دانشنامه‌ی کشوری رشته های جراحی عمومی،کودکان،زنان ومامایی و داخلی (سال‌های ۸۹و۹۰)

دکتر منیژه سیاح ملی[۱] ، دکتر محمد برزگر[۲] ،دکتر نعمت بیلیان[۳]، دکتر سعید اصلان آبادی[٤]، دکتر منوچهر خوشباطن[۵]، دکتر عالیه قاسمزاده[٦]، دکتر محمد حسن کارگر ماهر[۷]

نویسنده مسوول: تبریز، مرکز تحقیقات آموزش پزشکی، دانشگاه علوم پزشکی تبریز mm_barzegar@yahoo.com

دریافت: ۹۳/٤/۱۷ پذیرش: ۹۳/۹/۲٤

چکیده

زمینه و هدف: *آزمون‌های چند گزینه ای رایج ترین نوع ارزیابی در آموزش علوم پزشکی بوده، اگر با دقت و مطابق با قواعد آزمون سازی تهیه شوند ، ابزار معتبری برای تضمین کیفیت آموزش خواهند بود. هدف این مطالعه مقایسه‌ی برخی شاخص های کیفیت سوالات چند گزینه ای امتحانات ارتقا دانشگاه علوم پزشکی تبریز و دانشنامه‌ی کشوری رشته های دستیاری جراحی عمومی ،کودکان ،زنان و مامایی و داخلی در سال‌های ۱۳۸۹ و ۱۳۹۰ با توجه خاص به مدیریت بهبود کیفیت و پایش آزمون های درون دانشگاهی بود.*

روش بررسی: *در این مطالعه‌ی توصیفی – مقطعی ، ۲٤۰۰ سوال چهارگزینه ای مربوط به آزمون ها کتبی ارتقا و دانشنامه‌ی کشوری رشته های مذکور از نظر تاکسونومی و رعایت اصول ساختاری توسط اساتید خبره در زمینه‌ی علوم بالینی و آموزش پزشکی ، مورد بررسی قرار گرفتند. اطلاعات جمع آوری شده به‌وسیله‌ی آمار توصیفی مورد تجزیه و تحلیل قرار گرفت.*

یافته‌ها: *کیفیت سوالات آزمون های درون دانشگاهی و دانشنامه‌ی کشوری رشته های مذکورمتفاوت و اختلاف بارزی داشتند. در مقایسه از نظر تاکسونومی و رعایت اصول ساختاری در آزمون‌های دانشنامه کشوری رشته‌ی جراحی عمومی و درآزمون‌های ارتقا رشته‌ی داخلی وضعیت بهتری داشتند..سوالات آزمون های ارتقای رشته‌های کودکان و جراحی عمومی در مقایسه با دانشنامه‌ی این رشته ها به طور قابل توجهی در سطوح پایین یادگیری (تاکسونومی۱) طراحی و دارای اشکالات ساختاری بیشتری بودند.*

نتیجه گیری: *با توجه به اختلاف قابل توجه در کیفیت سوالات آزمون های مورد بررسی ،پایش سیستماتیک این آزمون ها به منظور ارتقا کیفیت آن‌ها ضروری می‌با شد.*

واژگان کلیدی :سوالات چند گزینه ای، تاکسونومی، اشکال های ساختاری، آزمون ارتقا،آزمون دانشنامه

۱- متخصص زنان و مامائی، استاد دانشگاه علوم پزشکی تبریز
۲- فوق تخصص نورولوژی کودکان، کارشناس ارشد آموزش پزشکی، استاد دانشگاه علوم پزشکی تبریز
۳- فوق تخصص ریه کودکان، کارشناس ارشد آموزش پزشکی، استاد دانشگاه علوم پزشکی تبریز
٤- فوق تخصص جراحی کودکان، استاد مرکز تحقیقات سلامت کودکان دانشگاه علوم پزشکی تبریز
۵- متخصص گوارش، استاد مرکز تحقیقات سلامت کودکان دانشگاه علوم پزشکی تبریز
٦- متخصص زنان و مامایی، دانشیار دانشگاه علوم پزشکی تبریز
۷- متخصص کودکان، استادیار مرکز تحقیقات سلامت کودکان دانشگاه علوم پزشکی تبریز

[Downloaded from edujournal.zums.ac.ir on 2026-11-11]

[DOR: 20.1001.1.22519521.1394.8.18.3.7]

مقدمه

در هر برنامه‌ی آموزشی،ارزیابی دانشجویان یکی از مهم‌ترین بخش‌های آموزش را تشکیل می‌دهد. فرایند نظام دار گردآوری و تحلیل و تفسیر اطلاعات به منظور تعیین حصول هدف‌های آموزشی و میزان تحقق آن‌ها ارزیابی یا ارزشیابی گفته می‌شود(۱). فرایند ارزیابی باید تصویر مناسبی از پیشرفت تحصیلی هر یک از دانشجویان در مقاطع زمانی گوناگون و مشکلات و کاستی‌های موجود آموزش را مشخص نماید. در صورتی‌که نتایج ارزیابی رضایت بخش نباشد ممکن است نمایانگر فعالیت کم دانشجو یا معلول نارسایی در برنامه ریزی ، تدریس یا ارزیابی نامطلوب باشد(٤-١) .

روش های ارزیابی متعدد بوده، بر اساس هدف ارزیابی نوع آن تعیین می‌گردد. آزمون‌های دانشنامه و ارتقای رشته‌های بالینی در دوره‌های دستیاری آموزش پزشکی سالیانه انجام می‌شود نوعی از ارزیابی تصدیقی می‌باشد که برای رتبه بندی و تصمیم گیری درباره‌ی ارتقای آنان به سال بالاتر و یا دریافت یک درجه‌ی علمی به کار می‌رود تا جامعه را در مقابل شروع به فعالیت افراد غیر کارآمد حفاظت کند. در ایران این آزمون‌ها شامل دو بخش کتبی و شفاهی می‌باشند. در حال حاضر برای آزمون کتبی از سوالات چند گزینه‌ای استفاده می‌شود. .آزمون های چند گزینه‌ای می‌توانند بخش وسیعی از دانش فراگیران را در مدت کوتاهی مورد ارزیابی قرار دهند، این نوع آزمون از نظر یکنواختی سوال ها و امکان حدس کورکورانه کم نسبت به آزمون های درست – نادرست و از لحاظ سهولت تصحیح پاسخ نامه ها بهترین نوع آزمون عینی است. سوال‌های چند گزینه‌ای از یک سوال یا مساله به عنوان پایه یا ساقه‌ی سوال و تعداد چهار یا پنج پاسخ به عنوان گزینه‌ها که فقط یکی از آن‌ها صحیح است، درست می‌شوند، پاسخ‌های غیر صحیح به‌عنوان جواب های احتمالی سوال و تحت عنوان گزینه های انحرافی در کنار گزینه‌ی صحیح ارایه می‌شوند.سوال های چند گزینه ای اگر با دقت و رعایت اصول آزمون‌سازی تهیه شوند، به‌خوبی می‌توانند بین دانشجویان با توانایی بالا و پایین را تفکیک نمایند(٦-١). تهیه‌ی هر سوال چند گزینه ای خوب علاوه بر داشتن دانش و مهارت و تجربه نیازمند صرف وقت و زمان کافی است (۷). عینیت و در نتیجه پایایی بالا و سهولت نسبی اجرا به‌ویژه در مقیاس‌های بزرگ از مزایای عمده این آزمون‌ها می‌باشد، معهذا نگرانی در مورد برخی شاخص‌های کیفیت سوالات چند گزینه‌ای شامل سطح دانش به کار رفته(تاکسونومی) و رعایت اصول ساختاری وجود دارد(۸و٦).

برای طراحی صحیح سوال چند گزینه ای در چند سال گذشته کارگاه‌های آموزشی تقریبا برای تمام اعضای هیات علمی دانشگاه علوم پزشکی تبریز برگزار شده است . از سال ۱۳۸۵ براساس برنامه‌ی دبیرخانه شورای آموزش پزشکی و تخصصی به منظور افزایش میزان مشارکت اعضای هیات علمی گروه های آموزشی دانشگاه های علوم پزشکی کشور، و در راستای سیر به سمت استقلال دانشگاه‌ها ، تعمیم واگذاری اختیارات آموزشی ، طراحی و برگزاری امتحانات ارتقای دستیاری رشته های بالینی به دانشگاه های علوم پزشکی تیپ I واگذار شده است،بنابراین ضرورت دارد که این فرآیند (طراحی سوالات و برگزاری آزمون های ارتقا توسط دانشگاه‌ها)مورد بررسی و ارزشیابی قرار گیرد. اگرچه مطالعات متعددی در مورد تجزیه و تحلیل سوالات چند گزینه ای در کشور انجام شده استف معهذا هدف اصلی آن‌ها ارایه‌ی بازخورد به طراحان سوال بوده است.(۱۱-۹). هدف این مطالعه مقایسه‌ی برخی شاخص‌های کیفیت (چگونگی توزیع تاکسونومی و میزان رعایت اصول ساختاری استاندارد) سوالات چند گزینه‌ای امتحانات ارتقای رشته‌های کودکان ، داخلی ، جراحی عمومی و زنان و مامایی جراحی عمومی (سال‌های۱۳۹۰و۱۳۸۹) برگزار شده در دانشگاه علوم پزشکی تبریز با سوالات امتحانات دانشنامه این رشته ها در سال‌های

[Downloaded from edujournal.zums.ac.ir on 2026-11-11]

[DOR: 20.1001.1.22519521.1394.8.18.3.7]

مذکور با توجه خاص به مدیریت بهبود کیفیت و پایش آزمون های درون دانشگاهی بود.

روش بررسی

روش مطالعه این تحقیق یک مطالعه‌ی توصیفی –مقطعی بود. سوالات آزمون‌های کتبی ارتقای دستیاری رشته های کودکان، داخلی ، جراحی عمومی و زنان و مامایی وجراحی عمومی دانشگاه علوم پزشکی تبریز ودانشنامه‌ی تخصصی همان رشته‌ها در سال‌های۱۳۹۰ و۱۳۸۹ از نظرچگونگی توزیع تاکسونومی و میزان رعایت اصول ساختاری استاندارد مورد بررسی قرارگرفتند. کل سوالات هر آزمون (۱۵۰ سوال)در هر سال مورد بررسی قرار گرفت. تک تک سوالات از نظر تاکسونومی در سطح III، II،I فرم تعدیل یافته‌ی بلوم، به‌طور مجزا توسط سه استاد خبره در رشته‌ی مربوطه و آموزش پزشکی مورد ارزیابی قرار گرفت. برای یکسان کردن برداشت اساتید در مورد تاکسونومی سوالات ،راهنمایی برای هر کدام از اساتید ارسال شد تا براساس آن تقسیم بندی سوالات ازنظر تاکسونومی به تعریف : تاکسونومی I - توانایی یادآوری حقایق (محفوظات) ،تاکسونومی II - توانایی تفسیر اطلاعات وتاکسونومی III - توانایی حل یک مساله جدید انجام گیرد(۱۲و۲).

در صورت نظرات متفاوت سه استاد ارزیابی کننده برای تعیین تاکسونومی یک سوال، مجددا توسط استاد چهارم (متخصص همان رشته و دوره دیده طراحی سوال) همان سوال تعیین تاکسونومی گردید و نظر وی مورد ملاک قرار گرفت.

بررسی رعایت اصول ساختاری میلمن توسط مجری طرح بر اساس فهرست وارسی میلمن (تعدیل شده)به شرح زیر مورد بررسی قرار گرفت.

فهرست وارسی میلمن جهت بررسی رعایت اصول ساختاری در سوالات چند گزینه ای

موضوع	بلی	خیر
۱–هدف در شناسنامه سوال درج شده است و این سوال یک هدف آموزشی را اندازه می گیرد.		
۲–مطالب اصلی سوال به‌طور کامل در تنه‌ی سوال نوشته شده است.		
۳–در متن سوال از اشاره های دستوری نامربوط وتشابه ظاهری هدایت کننده وغیره استفاده نشده است.		
٤–گزینه های انحرافی را طوری بنویسید که توجه آزمون شوندگان بی اطلاع از موضوع سوال را به خود جلب کنند.		
۵–در سوال منفی کلمات منفی برجسته جلوه داده شده و زیر آن خط کشیده شده است.		
٦–اگر متن سوال منفی باشد گزینه ها هم منفی نیستند (منفی مضاعف)		
۷–گزینه همه موارد فوق یا هیچکدام از موارد فوق به‌کاربرده نشده است.		
۸–محتوای گزینه‌های انحرافی از لحاظ طول ، واژگان ، ساختار جمله، با گزینه‌ی درست بوده و به موضوع واحدی مربوط است.		
۹–تمام گزینه های هر سوال از نظر دستوری و جمله بندی به نحو درست مکمل متن سوال هستند .		
۱۰–سوال بدون مشاهده گزینه ها نیز قابل پاسخگویی است.		
۱۱–دو گزینه‌ی متضاد که یکی از آن‌ها درست باشد به‌کار برده نشده است.		
۱۲–از تکرار مطالب در گزینه ها خودداری شده است.		

[Downloaded from edujournal.zums.ac.ir on 2026-11-11]

[DOR: 20.1001.1.22519521.1394.8.18.3.7]

در صورت وجود بیش از یک اشکال ساختاری در مورد هر سوال فقط یک اشکال ساختاری لحاظ و مورد تجزیه و تحلیل قرار گرفت.

اطلاعات جمع آوری شده به‌وسیله نرم افزار SPSS 18 مورد تجزیه و تحلیل قرار گرفت ، از آزمون Chi Square برای بررسی معنی دار بودن تفاوت متغیرهای مورد بررسی بین دو آزمون استفاده گردید. p کمتر از ۰/۰۵ معنی دار تلقی شد.

یافته‌ها

در مجموع ۲٤۰۰ سوال چهارگزینه ای، یک پاسخ صحیح مربوط به چهار رشته‌ی تخصصی دستیاری رشته های کودکان ، داخلی ، جراحی عمومی و زنان و مامایی بررسی شدند.

در جدول ۱ توزیع سطح تاکسونومی سوالات آزمون های کتبی ارتقای دانشگاه علوم پزشکی تبریز و دانشنامه‌ی کشوری رشته‌های دستیاری تخصصی داخلی، جراحی عمومی، کودکان و زنان ومامایی مجموع سال‌های ۸۹ و ۹۰ مقایسه شده است.

همان‌طور که در جدول ۱ مشاهده می‌شود، اختلاف قابل توجهی در توزیع تاکسونومی سوالات در آزمون‌های ارتقا و دانشنامه‌ی رشته‌های مورد بررسی وجود دارد. در مقایسه آزمون‌های ارتقای رشته‌ی داخلی با ۷۲/۷ درصد سوالات در تاکسونومی II-III وضعیت بهتری داشته است(P <۰/۰۰۱)و در مقایسه‌ی آزمون‌های دانشنامه‌ی رشته جراحی عمومی با ۷۹ درصد سوالات در تاکسونومی II-III وضعیت مطلوبی داشته است (P<۰/۰۰۱)

جدول ۱: توزیع تاکسونومی سوالات آزمون های کتبی ارتقای دانشگاه علوم پزشکی تبریز و دانشنامه‌ی کشوری رشته های دستیاری تخصصی داخلی،جراحی عمومی،کودکان و زنان ومامایی مجموع سال‌های ۸۹ و ۹۰

نوع آزمون کتبی	ارتقا		دانشنامه		P.Value
	تاکسونومی I	تاکسونومی II-III	تاکسونومی I	تاکسونومی II-III	
	تعداد(درصد)	تعداد(درصد)	تعداد(درصد)	تعداد(درصد)	
جراحی عمومی	(٤٨/٣)١٤٥	(٥١/٧)١٥٥	(٢١)٦٣	(٧٩)٢٣٧	P=۰/۰۰۰
زنان و مامایی	(٣٥/٧)١٠٧	(٦٤/٣)١٩٣	(٤١/٣)١٢٤	(٥٨/٧)١٧٦	P=۰/١٥٨
کودکان	(٥٤/٧)١٦٤	(٤٥/٣)١٣٦	(٤٣)١٢٩	(٥٧)١٧١	P=۰/۰٠٤
داخلی	(٢٧/٣)٨٢	(٧٢/٧)٢١٨	(٣٠)٩٠	(٧٠)٢١٠ ٧٠	P=۰/٤٧٠

بررسی روند آزمون‌ها نشان دهنده‌ی افزایش درصد تاکسونومی II - III در هر دو نوع آزمون (ارتقا و دانشنامه) در سال ۹۰ در مقایسه با سال ۸۹ (به‌جز آزمون ارتقای ۹۰ رشته‌ی زنان و مامایی) بود. که نشان‌دهنده‌ی روند ارتقای کیفیت طراحی سووال در سطوح عمیق‌تر یادگیری در حیطه‌ی شناختی می‌باشد (جدول۲).

در جدول ۳ وضعیت رعایت اصول ساختاری میلمن در سووالات چهارگزینه ای آزمون های کتبی ارتقا و دانشنامه

[Downloaded from edujournal.zums.ac.ir on 2026-11-11]

[DOR: 20.1001.1.22519521.1394.8.18.3.7]

سال‌های ۸۹ و۹۰رشته های مذکور مورد مقایسه قرار گرفته است.

همان‌طور که در جدول ۳ مشاهده می‌شود، بیشترین درصد سوالات بدون اشکال ساختاری در آزمون‌های ارتقا مربوط به رشته‌ی داخلی بود(P<۰/۰۰۱). رعایت اصول ساختاری در آزمون‌های دانشنامه‌ی کشوری رشته‌های مختلف در این بررسی تقریبا مشابه هم بود. روند رعایت اصول ساختاری آزمون‌ها در سال‌های ۸۹ و ۹۰ در هر نوع آزمون در جدول ٤ نشان داده شده است. بررسی روند آزمون‌های مختلف (به غیر از آزمون ارتقای زنان و مامایی) پیشرفت در صد سوالات بدون اشکال ساختاری در سال ۹۰ نسبت به سال ۸۹ را نشان داد.

بیشترین فراوانی از نظر اشکالات ساختاری ۳۹/۵ درصد در مجموع سووالات ارتقا ۸۹ و ۳۰ درصد در مجموع سوالات دانشنامه ۸۹ ، ۲۸ درصد در مجموع سووالات ارتقای ۹۰ و ۱۹/۲ درصد در مجموع سووالات دانشنامه‌ی ۹۰ مربوط به قرار نداشتن بخش اعظم اطلاعات در ساقه‌ی سوال و امکان پاسخ دادن به سوالات بدون مراجعه به گزینه ها بودند.

جدول ۲: مقایسه‌ی توزیع تاکسونومی آزمون های کتبی ارتقا دانشگاه علوم پزشکی تبریز و دانشنامه‌ی کشوری رشته های دستیاری تخصصی داخلی،جراحی عمومی،کودکان و زنان ومامایی در سال‌های ۸۹ و۹۰

نوع آزمون کتبی	سال ۸۹		سال ۹۰		P.Value
	تاکسونومی I	تاکسونومی II-III	تاکسونومی I	تاکسونومی II-III	
	تعداد(درصد)	تعداد(درصد)	تعداد(درصد)	تعداد(درصد)	
جراحی عمومی(ارتقا)	(٥٤/٧)٨٢	(٤٥/٣)٦٨	(٤٢)٦٣	(٥٨)٨٧	P=۰/۰۲۸
جراحی عمومی (دانشنامه)	(٢٦/٧)٤٠	(٧٣/٣)١١٠	(١٥/٣)٢٣	(٨٤/٧)١٢٧	P=۰/۰١٦
زنان و مامایی (ارتقاء)	(٣٤)٥١	(٦٦)٩٩	(٣٧/٣)٥٦	(٦٢/٧)٩٤	P=۰/٧٢٩
زنان ومامایی(دانشنامه)	(٤١/٣)٦٢	(٥٨/٧)٨٨	(٣٥/٣)٥٣	(٦٤/٧)٩٧	P=۰/٥٤٢
کودکان (ارتقا)	(٦٣/٣)٩٥	(٣٦/٧)٥٥	(٤٦)٦٩	(٥٤)٨١	P=۰/۰۲۸
کودکان(دانشنامه)	(٤٤/٧)٦٧	(٥٥/٣)٨٣	(٣١/٣)٤٧	(٦٩/٣)١٠٣	P=۰/۰١٦
داخلی(ارتقا)	(٣٤)٥١	(٦٦)٩٩	(٢٠/٧)٣١	(٧٩/٣)١١٩	P=۰/۰٢٩
داخلی (دانشنامه)	(٣٢/٧)٤٩	(٦٧/٣)١٠١	(٢٧/٣)٤١	(٧٢/٧)١٠٩	P=۰/٣١٣

[Downloaded from edujournal.zums.ac.ir on 2026-11-11]

[DOR: 20.1001.1.22519521.1394.8.18.3.7]

جدول۳ : مقایسه‌ی وضعیت رعایت اصول ساختاری میلمن در سووالات چهارگزینه ای آزمون های کتبی ارتقای دانشگاه علوم پزشکی تبریز و دانشنامه‌ی کشوری رشته‌های دستیاری تخصصی داخلی،جراحی عمومی،کودکان و زنان ومامایی مجموع سال‌های ۸۹ و ۹۰

آزمون کتبی	ارتقا		دانشنامه		P.Vaule
	بدون اشکال ساختاری	دارای اشکال ساختاری	بدون اشکال ساختاری	دارای اشکال ساختاری	
	تعداد (درصد)	تعداد (درصد)	تعداد (درصد)	تعداد (درصد)	
جراحی عمومی	(۵۵)۱۶۵	(۴۵)۱۳۵	(۷۰/۲)۲۱۰	(۲۹/۸)۸۹	P=۰/۰۰۱
زنان و مامایی	(۶۶)۱۹۸	(۳۴)۱۰۲	(۷۰/۷)۲۱۲	(۲۹/۳)۸۸	P=۰/۲۳۰
کودکان	(۶۲/۷)۱۸۸	(۳۷/۳)۱۱۲	(۷۳/۶)۲۲۱	(۲۶/۴)۷۹	P=۰/۰۰۶
داخلی	(۷۲/۷)۲۱۸	(۲۷/۳)۸۲	(۷۳/۳)۲۲۰	(۲۶/۷)۸۰	P=۰/۸۵۴

جدول ٤: مقایسه‌ی وضعیت رعایت اصول ساختاری میلمن آزمون های کتبی ارتقا و گواهی نامه‌ی دانشگاه علوم پزشکی تبریز و دانشنامه‌ی کشوری رشته های دستیاری تخصصی داخلی،جراحی عمومی،کودکان و زنان ومامایی در سال‌های ۸۹ و۹۰

نوع آزمون کتبی	سال ۸۹		سال ۹۰		P.Value
	بدون اشکال ساختاری	دارای اشکال ساختاری	بدون اشکال ساختاری	دارای اشکال ساختاری	
	تعداد(درصد)	تعداد(درصد)	تعداد(درصد)	تعداد(درصد)	
جراحی عمومی(ارتقا)	(۴۶)۶۹	(۵۴)۸۱	(۶۴)۹۶	(۳۶)۵۴	P=۰/۰۰۲
جراحی عمومی (دانشنامه)	(۶۵/۳)۹۸	(۳۴/۷)۵۲	(۷۵/۲)۱۱۳	(۲۴/۸)۳۷	P=۰/۰۶۳
زنان و مامایی (ارتقا)	(۶۸/۷)۱۰۳	(۳۱/۳)۴۷	(۶۳/۳)۹۵	(۳۶/۷)۵۵	P=۰/۳۲۹
زنان ومامایی(دانشنامه)	(۷۰)۱۰۵	(۳۰)۴۵	(۷۰/۷)۱۰۶	(۲۹/۳)۴۴	P=۰/۵۴۲
کودکان (ارتقا)	(۵۴/۷)۸۲	(۴۵/۳)۶۸	(۷۰/۷)۱۰۶	(۲۹/۳)۴۴	P=۰/۰۰۲
کودکان(دانشنامه)	(۶۴)۹۱	(۳۶)۵۹	(۸۲/۷)۱۲۴	(۱۷/۳)۲۶	P=۰/۰۶۳
داخلی(ارتقا)	(۷۰)۱۰۵	(۳۰)۴۵	(۷۵/۳)۱۱۳	(۲۴/۷)۳۷	P=۰/۱۹۲
داخلی (دانشنامه)	(۷۶/۶)۱۱۵	(۲۳/۳)۳۵	(۷۰)۱۰۹	(۳۰)۴۱	P=۰/۳۱۳

[Downloaded from edujournal.zums.ac.ir on 2026-11-11]

[DOR: 20.1001.1.22519521.1394.8.18.3.7]

بحث

تغییر برنامه‌ی آموزشی و یا روش‌های آموزش بدون تغییر نحوه‌ی ارزیابی به نتیجه‌ی مطلوب نخواهد رسید . از طرف دیگر تغییر نظام ارزیابی حتی بدون تغییر برنامه‌ی آموزشی اثر بسیار عمیق‌تری نسبت به تغییر برنامه‌ی آموزشی بدون تغییر نظام ارزیابی ، بر ماهیت یادگیری باقی می‌گذارد . (۷ و ٤-۱) در تجزیه وتحلیل سووال هر دو جنبه‌ی کیفی و کمی سووال مورد توجه قرار می‌گیرد. از نظر کیفی شکل سوال ، روش تهیه ، دقت در ساختن متن سووال و گزینه‌های درست و نادرست وتاکسونومی آن ارزیابی می‌شود.از لحاظ کمی تجزیه و تحلیل درجه‌ی دشواری و قدرت تمیز هر سووال وتحلیل گزینه های انحرافی بررسی می‌شوند(١٤-١٣و٧،٤-١).

براساس نتایج این مطالعه توزیع تاکسونومی سوالات در آزمون‌های ارتقا و دانشنامه‌ی رشته‌های مورد بررسی متغیر بود. اگرچه استاندارد ثابتی برای توزیع درصد سووالات در سطوح مختلف تاکسونومی وجود ندارد، ولی توصیه می‌شود، این رقم کمتر از ۵۰ درصد باشد. در مطالعه‌ی محققی و همکاران در ارزیابی تاکسونومی سووالات آزمون کتبی ۲۵ رشته‌ی تخصصی‌بالینی‌در آزمون کتبی‌سال ۸٦ ، ۱۸/۹ ± ۳۸/۷ درصد سووالات در تاکسونومی II - IIIطراحی شده بود که این رقم برای سال ۸۷ ، ۱۹/۳± ٤۵ و برای سال ۸۸ ، ۱۵/۵۱ ± ۵٦ درصد بوده است که نشانگر روند رو به رشد طراحی سووالات در تاکسونومی بالاتر بوده است(۱۵). در بررسی ۲٤۰۰ سوال مربوط به آزمون های دستیاری دانشگاه علوم پزشکی جندی شاپور در سال ۱۳۸٦ فقط ۲۸ درصد سوالات در تاکسونومی III-IIطراحی شده بودند.(۹) متاسفانه این رقم برای امتحانات درون دانشگاهی غیر پزشکی کمتر نیز می‌باشد. در بررسی ثنا گو وهمکاران از ۵۲۳ سوال مربوط به رشته‌ی پرستاری کمتر از ۵ درصد آن‌ها در تاکسونومی III-II قرار داشتند(۱۰).به نظر می‌رسد این مشکل یعنی غلبه سوالات مبتنی بر حافظه (تاکسونومی I) در آزمون‌ها ، معضلی ریشه دار در سیستم آموزش پزشکی بوده و دامنه‌ی درگیری آن از آزمون‌های پیشرفت تحصیلی درون دانشگاهی تا آزمون‌های ملی را فراگرفته است.برخی مطالعات نشان داده اند، وجود چنین مشکلی می‌تواند روایی آزمون را با مشکل مواجه ساخته و موجب سوق دادن دانشجویان به دانش سطحی و محفوظات می‌گردد. در صورت عدم دقت درطراحی سوالات، استفاده بیش از حد سوالات چند گزینه ای برای سطوح پایین یادگیری روی می‌دهد (۷-۱). اختلاف معنی داری در مقایسه‌ی توزیع سطح تاکسونومی سوالات مجموع آزمون های کتبی ارتقا و گواهی نامه‌ی دانشگاه علوم پزشکی تبریز و دانشنامه‌ی کشوری دستیاری رشته‌های دستیاری جراحی عمومی و کودکان وجود داشت. با توجه به جدول ۳ پیشرفت قابل توجهی در زمینه‌ی طراحی سووالات در سطوح عمیق یادگیری (تاکسونومی II و III) در آزمون ارتقای سال ۹۰ دانشگاه علوم پزشکی تبریز در مقایسه با سال ۸۹ این دو رشته اتفاق افتاده است که بیانگر روند افزایش مهارت اعضای هیات علمی در طراحی سووال می‌باشد.

در درازمدت نوع فعالیت‌هایی که دانشجویان برای یادگیری انجام می‌دهند و آنچه که یاد می‌گیرند ، به وسیله‌ی نوع امتحاناتی که باید بگذرانند تعیین می‌شود(۲). بنابراین اگر معلمین ارزیابی را مورد استفاده قرار دهند که فقط احتیاج به حافظه داشته باشد ، دانشجویان تشویق به حفظ کردن مطالب می‌شوند (۲). برعکس هرگاه امتحان به نحوی تنظیم شده باشد که برای جواب دادن به آن لازم باشد دانشجو اصولی را رعایت کند ، اطلاعاتی را تفسیر کند و یا مساله‌ای را حل کند ، خودبخود موجب می شود تا دانشجو به فراگیری چنین نکاتی گرایش پیدا کند تا بتواند به خوبی از عهده‌ی امتحانات برآید. در صد کم سوالات با تاکسونومیII-III ممکن است منجر به مطالعه‌ی سطحی دستیاران و کاهش میزان موفقیت آن‌ها در آزمون دانشنامه شود.

[Downloaded from edujournal.zums.ac.ir on 2026-11-11]

[DOR: 20.1001.1.22519521.1394.8.18.3.7]

طبق نتایج این بررسی سوالات بدون اشکال ساختاری در آزمون ها ارتقا ۵۵-۷۲/۷ درصد و در آزمون‌های دانشنامه ۷۰/۲-۷۲/۷ درصد بود. نتایج مطالعه‌ی محققی و همکاران در بررسی آزمون کتبی دانشنامه‌ی ۲۵ رشته‌ی تخصصی در سال‌های ۱۳۸٦ تا ۱۳۸۸ از نظر رعایت اصول ساختاری میلمن در سووالات چند گزینه‌ای نشان داد که ۱۵/۱± ۵۷/۵ درصد سووالات آزمون‌های کتبی دانشنامه سال ۸٦ و ، ۱۵/۵± ٦۳/۸ آزمون‌های کتبی دانشنامه سال ۸۷ درصد سووالات و ، ۱۸/۹± ٦۰/٦ آزمون‌های کتبی دانشنامه سال ۸۸ از نظر ساختاری بدون اشکال بودند. (۱۵). مطالعه‌ی شکورنیا و همکاران و برخی مطالعات دیگر میزان‌های متفاوتی از اشکالات ساختاری آزمون‌های چند گزینه‌ای را گزارش داده‌اند(۱۱-۹). مطالعه‌ای در آمریکا ، اشکالات ساختاری سوال های چند گزینه ای را ٤٦ درصد گزارش نموده است (۱۸)و همچنین دو مطالعه‌ی مشابه در سال های ۲۰۰٦ و ۲۰۰۸ در کشور هنگ کنگ نشان داده است که به ترتیب ٤٦/۲ و ٤۷/۳ درصد از سوال های چند گزینه‌ای به‌کار گرفته شده در ارزشیابی دانشجویان ، اشکال ساختاری داشته و بیشتر در سطوح پایین شناختی تهیه شده اند(۲۰و۱۹) . به طور کلی رعایت اصول ساختاری در آزمون های ارتقا در مقایسه با تهیه‌ی سوال با تاکسونومی بالاتر قابل قبول‌تر بود که احتمالا مربوط به کاربرد عملی‌تر چک لیست میلمن می‌باشد. با وجود وضعیت قابل قبول نسبی ،هنوز میزان قابل توجهی از سوالات ارتقا حاوی اشکالات ساختاری هستند ، در مورد رایج‌ترین اشکالات ساختاری بر اساس اصول میلمن این مطالعه نشان داد آیتم های نداشتن بخش اعظم اطلاعات در ساقه‌ی سوال و عدم امکان پاسخ دادن به سوالات بدون مراجعه به گزینه‌ها شایع‌ترین موارد اشکال بوده است که مشابه مطالعات دیگر (۱۵،۱۱-۹) و احتمالا مربوط به عادات استقرار یافته‌ی قدیمی در طراحی سوال می‌باشد.آزمون‌های ارتقا و دانشنامه‌ی رشته های بالینی یکی از مهم‌ترین ارزیابی‌ها در رشته‌ی دستیاری تخصصی می‌باشند. به جهت اهمیت آزمون‌های دانشگاهی در تصمیم گیری برای ارتقا یا رد دانشجویان ، ضرورت قطعی دارد که این آزمون ها مورد تجزیه و تحلیل وپایش قرار گیرند .مسلما کار تیم آزمون با طراحی سوالات پایان نمی‌یابد و ضرورت دارد بعد از اجرای آزمون و تعیین نمرات آن ، تجزیه و تحلیل سوالات از هر دو جنبه‌ی کیفی و کمی صورت گیرد. در حال حاضر این تجزیه و تحلیل به طور جامع و سیستماتیک صورت نگرفته و بازخورد لازم به طراحان سوال داده نمی‌شود. در بررسی محققی و همکاران درصد سوالات با تاکسونومی II-III در آزمون‌های کتبی دانشنامه ۲۵ رشته‌ی تخصصی سال ۱۳۸۸ از ۲۷ درصد تا ۹۲ متغیر بوده، همچنین در صد سوالات بدون اشکال ساختاری در همان آزمون‌ها از ۳٦ تا ۹۰ درصد متغیر بوده است (۱۵). در مطالعه‌ی ما(جدول۱) مشابه بررسی فوق توزیع تاکسونومی سوالات در ازمون‌های ارتقای درون دانشگاهی و دانشنامه‌ی کشوری رشته‌های کودکان ، داخلی ، جراحی عمومی و زنان و مامایی جراحی عمومی متفاوت و اختلاف بارزی داشتند. در آزمون‌های ارتقا ۷۲/۸ درصد سوالات آزمون رشته‌ی داخلی در مقایسه با ٤۵/۳ درصد آزمون ارتقای رشته کودکان و ۵۱/۷ درصد آزمون ارتقای رشته جراحی عمومی در تاکسونمی II-III طراحی شده بودند. در آزمون‌های دانشنامه کشوری ۷۹ درصد سوالات آزمون رشته‌ی جراحی عمومی در مقایسه با ۵۷ درصد سوالات آزمون دانشنامه‌ی رشته‌ی کودکان و ۵۸/۷ درصد آزمون دانشنامه‌ی رشته‌ی زنان ومامایی در تاکسونومی II-III طراحی شده بودند . با توجه به اینکه مدت زمان پاسخ دادن به سوالات و حد نصاب ارتقا و قبولی در رشته‌های مورد بررسی یکسان می‌باشد و با نظر به اینکه پاسخ دادن به سوالات با تاکسسونومی بالاتر زمان بیشتری لازم دارد ، به نظر می‌رسد که عدالت لازم در آزمون ها رعایت نشده است .به عبارت ساده‌تر امتحانات بعضی از رشته‌ها آسان‌تر می‌باشد. با توجه به نقش مهم نتایج این امتحانات در تعیین کفایت

[Downloaded from edujournal.zums.ac.ir on 2026-11-11]

[DOR: 20.1001.1.22519521.1394.8.18.3.7]

توانمندی و صلاحیت آن‌ها برای فارغ التحصیلی یا ارتقا به سال بالاتر ، در صورتی که سوالات بیشتر در سطح‌های پایین شناختی طراحی شوند، روایی آزمون مورد سوال خواهد شد. پیشنهاد می‌شود که دبیرخانه‌ی شورای آموزش پزشکی و تخصصی باتوجه به اهمیت امتحانات دانشنامه با مدیریت و پایش این آزمون‌ها زمینه‌ی برطرف شدن کاستی های موجود را فراهم نماید.

رویکرد واگذاری امتحانات ارتقا به دانشگاه‌های علوم پزشکی در راستای روند استقلال دانشگاه رویکردی ارزشمند می‌باشد . جهت پیشگیری از افت کیفیت آزمون‌های درون داتشگاهی پایش آزمون‌ها توسط خود دانشگاه‌ها موجب برطرف شدن کاستی های موجود خواهد شد. با توجه به اینکه تقریبا تمام طراحان سوالات ارتقای دانشگاه علوم پزشکی تبریز کارگاه های آموزشی طراحی سوال را گذرانده‌اند، متاسفانه کیفیت سوالات رشته‌های مختلف متفاوت بود.به نظر می‌رسد اتخاذ تدابیر اصلاحی دیگری جهت ارتقای کیفیت سوالات لازم است.پیشنهاد می‌شود مرکز مطالعات و توسعه‌ی آموزش پزشکی دانشگاه علوم پزشکی تبریز با همکاری گروه آموزش پزشکی این دانشگاه نتایج تجزیه و تحلیل سوالات هر رشته را در گروه مربوطه ارایه و با الگو قرار دادن سوالات با کیفیت خوب از همان گروه در ارتقای مهارت طراحی سوال اعضا هیات علمی اقام نماید، به شرطی این رویکرد موثر خواهد شد که با تهیه‌ی شناسنامه طراحی سوال برای اعضای هیات علمی (طراح سوال) وتاثیر آن در فرایندهای ارتقای اعضای هیات علمی، مهارت طراحی سوال با کیفیت خوب در امتیازات ارتقای آن‌ها لحاظ شود.

محدودیت های مطالعه : جهت تجزیه و تحلیل کمی نمرات فراگیران و پاسخ نامه‌های آن‌ها ضروری می‌باشد ، که متاسفانه در این بررسی این اطلاعات در دسترس نبود و تجزیه وتحلیل کمی آزمون‌ها مقدور نشد. به علت دردسترس نبودن شناسنامه سوالات دانشنامه که معمولادر آن هدف سوال قید می‌شود بررسی بودجه بندی سوالات از نظر اهداف دوره آموزشی مقدور نشد. مورد " آیا هدف در شناسنامه سوال درج شده است " به‌علت در دسترس نبودن شناسنامه سوالات و مورد " گزینه‌های انحرافی طوری نوشته شود که توجه آزمون شوندگان بدون اطلاع از موضوع را به خود جلب کنند " به علت در دسترس نبودن پاسخ نامه‌های آزمون شوندگان و نمرات آن‌ها ارزیابی نشدند.

نتجه گیری

کیفیت سوالات آزمون‌های درون دانشگاهی و دانشنامه‌های رشته‌های مذکورمتفاوت و اختلاف بارزی داشتند. در مقایسه از نظر تاکسونومی و رعایت اصول ساختاری در آزمون‌های دانشنامه‌ی کشوری رشته‌ی جراحی عمومی و درآزمون های ارتقا رشته‌ی داخلی وضعیت بهتری داشتند. سوالات آزمون‌های ارتقای رشته‌های کودکان و جراحی عمومی در مقایسه با دانشنامه این رشته‌ها به طور قابل توجهی در سطوح پایین یادگیری (تاکسونومیI) طراحی و دارای اشکالات ساختاری بیشتری بودند. مدیریت بهبود کیفیت آزمون‌های دانشنامه و ارتقا جهت پیشگیری از افت کیفیت آزمون‌ها و ارایه‌ی بازخورد به طراحان سوال و برنامه ریزی برای توانمند سازی طراحان سوال در زمینه‌ی اصول آزمون سازی ضروری می‌با شد.

تقدیر و تشکر

و این مقاله نتیجه‌ی طرح تحقیقاتی مصوب مرکز تحقیقات آموزش پزشکی دانشگاه علوم پزشکی تبریز با کد ۱۰۰۳۸ به تاریخ ۹۱/۳/۱۱ است. از جناب آقای دکتر آیدین تبریزی که در تهیه‌ی سوالات آزمون ها کمک نمودند و از سرکار خانم ذکیه عبادی که زحمت وارد کردن داده ها در نرم افزار آماری را تقبل نمودند، تشکر و سپاس‌گزاری می‌شود.

[Downloaded from edujournal.zums.ac.ir on 2026-11-11]

[DOR: 20.1001.1.22519521.1394.8.18.3.7]

References

1- Seif A. Educatinal measurement, assessment and evaluation. 5th ed Tehran: Doran; 2008 ,30-41

2- Guilbert JJ. Education handbook for health personnel. 7thed.Geneva : Word Health Organization; 1998,237-54.

3- Azizi F. Medical education, mission, vision, and challenges first ed. Tehran: Shahid Beheshti University of Medical Sciences; 2003, 659-705

4- Zolfaghari B, Asadollahi GH. Academic achievement tests in medical sciences. Isfahan : Isfahan University of Medical Sciences; 2000 , 35-52

5- Burton SJ, Sudweeks RR, Merrill PF, Wood B. How to prepare better multiple choice tests items :Guidline for university faculty. Birgham Young University Testing Services

6- Haladyna TM, DowningSM,Rodriguez MC . A review of multiple-choice item-writing guidelines for classroom assessment .*Applied Measurement in Education*. 2002; 15(3): 309-34.

7- Roediger HL 3rd, Marsh EJ. The positive and negative consequences of multiple choice testing. *J Exp Psychol Learn Mem Cogn*.2005 31(5):1155-9

8- Farley JK. The multiple-choice test: developing the test blueprint. *Nurse Educator*. 1989; 14(5):3-5.

9- Shakurnia AH, Mozaffari AR, Khosravi Brougeni. Survey on structural of MCQs of residency exam in AJUMS . *Judishapour Scientific Medical Journal*. 2009; 8(4): 491-502

10- Sanagoo A, Jouybari L, Ghanbari Gorji M. Quantitative and qualitative analysis of achivement tests in Golestan University of Medical Sciences. *J Research in Medical Education.* 2010; 2(2): 24-32

11- Pourmirza Kalhori R, Roshanpour F, Rezaei M, Naderipour A. Academic members knowledge improvement effect on results of multiple choice questions in residency exams analysis. *J Kermanshah Med Sci*; 2010, 15(2): 112-118

12- Buckwalter KA, Schumacher R, Albright JP, Cooper RR .Use of an educational taxonomy for evaluation of cognitive performance .*J Medical Education*. 1981; 56: 115-21.

13- Collins J. Education Techniques for lifelong learning. Writing MCQs for continuous medical education activities and self assessment modules. *Radio Graphics*. 2006; 26(2): 543-551

14- Case SM, Swanson DB. Constructing written test questions for the basic and clinical sciences. Third edition, Philadelphia; 1998

15- Mohaggegi MA , Vahid Shahi K, SHakeri S, et al Comparision some aspect of quality of MCQs Board Examination 2007-2009. Available from:
http://www.*cgme*.behdasht.gov.ir/uploads/264_781_N4

16- Shaban M, Ramazani Badr F. Effect of test item analysis on summative exams quantity of test designing. *HAYAT*; 2007; 13(1);5-15.

17- Danish KF, Ahmad Khan R. Role of effective feed back in multiple choice questions designing for faculty development. *Journal of Rawalpindi*

[Downloaded from edujournal.zums.ac.ir on 2026-11-11]

[DOR: 20.1001.1.22519521.1394.8.18.3.7]

Medical College: 2010, 14 (2):98-100

18- Downing SM. The effects of violating standard item writing principles on tests and students: the consequences of using flawed test item on achievement examinations in medical education. *Adv in Health Sci Educe*. 2005; 10(2):133-43.

19- Tarrant M, Knierim A, Hayes SK, WareJ. The frequency of item writing flaws in multiple-choice questions used in high stakes nursing assessments. *Nurse Education today*. 2006; 26(8): 662-71.

20- Tarrant M, Ware J. Impact of item-writing flaws in multiple-choice questions on student achievement in high-stakes nursing assessment. *Medical Education*. 2008; 42(2) : 198-206

[Downloaded from edujournal.zums.ac.ir on 2026-11-11]

[DOR: 20.1001.1.22519521.1394.8.18.3.7]

Comparing Multiple-Choice Questions Quality Parameters of Pediatric, General Surgery, Internal Medicine and Genecology and Obstetrics Residency Tests in Preboard and Promotion Examinations of Tabriz University of Medical Sciences and National Board Examination in 2010 and 2011

Sayyahmelli M[1], Barzegar M[1], Bilan N[2], Aslanabadi S[2], khoshbaten M[1], Gasemzadeh A[1], Karegar Maher MH[2]

[1]Medical Education Research Center, Tabriz University of Medical Sciences, Tabriz, Iran
[2]Pediatric Health Research Center, Tabriz University of Medical Sciences, Tabriz, Iran

Corresponding Author: Barzegar M, Medical Education Research Center, Tabriz University of Medical Sciences, Tabriz, Iran
Email: mm_barzegar@yahoo.com
Received: 5 Apr 2014 ***Accepted:*** 17 Sep 2014

Background and objective: The multiple-choice questions (MCQs) are the most commonly used types of test in medical education. These tests could be a reliable measurement instrument to evaluate educational programs if developed carefully and according to standard test preparation criteria. This study aimed to compare some quality indicators of multiple-choice questions of pediatric, general surgery, internal medicine and genecology and obstetrics residency in pediatric residency promoting test program in 2010 and 2011 which was held at Tabriz University of Medical Science (TUMS) with a special consideration to improve test quality management and monitoring the university examinations.

Materials and Methods: In this cross-sectional study, 2400 MCQs of preboard and national board examinations were evaluated at TUMS regarding taxonomy levels (Bloom levels I, II and III) and structural principles (based on Millman check list). Data was analyzed by SPSS (version 18).

Results*:* The quality of MCQs in preboard examination of (TUMS) and national board examination was different and the difference was significant. National board tests were more reliable in general surgery and in preboard examination of (TUMS) in internal medicine. Pediatric and general surgery residency preboard MCQs of (TUMS) had significantly low level of learning comparing to national board MCQs and had more structural problems.

Conclusion*:* There was a significant difference in the quality of MCQs in preboard examination of (TUMS) and national board examination according to the disciplines. Improving test quality management and monitoring these exams are recommended.

Keywords: Multiple choice questions, Structural flaws, Taxonomy, Board examination, Preboard examination

[Downloaded from edujournal.zums.ac.ir on 2026-11-11]

[DOR: 20.1001.1.22519521.1394.8.18.3.7]