مقایسه نمرات آزمون تشریحی و چندگزینه‌ای درس بیماری‌های کلیه و بررسی نقطه نظرات دانشجویان پزشکی در خصوص تأثیرآزمون‌ها بر مطالعه

شهرزاد شهیدی، وجیهه وفامهر*، محمد دادگسترنیا

چکیده

مقدمه: برای بهره مندی از منافع آزمون‌های تشریحی، باید مطمئن شد که می‌توان براساس نمرات این آزمون‌ها، قضاوت پایایی درباره دانشجویان داشت. این مطالعه با هدف بررسی همبستگی نمرات بخش‌های تشریحی و چندگزینه‌ای آزمون پایانی درس پاتوفیزیولوژی کلیه و نقطه نظرات دانشجویان درباره تأثیر نوع آزمون بر مطالعه آنها انجام شد.

روش‌ها: این مطالعه‌ی توصیفی همبستگی، از پاییز ۹۴ تا زمستان ۹۵ در دانشکده پزشکی اصفهان انجام گرفت. ۴۵۵ دانشجو در ۳ دوره، طی ارزشیابی پایانی درس پاتوفیزیولوژی کلیه، در آزمونی با ۳۶ سؤال چندگزینه‌ای و۱۰ سؤال تشریحی شرکت کردند. آزمون چندگزینه‌ای با نرم افزار OPSCAN تصحیح شد. همبستگی نمرات آزمون چندگزینه‌ای و تشریحی از طریق Bivariate Correlation و مقایسه میانگین نمرات تشریحی و چهارگزینه‌ای با تی زوج صورت پذیرفت. نقطه نظرات دانشجویان نسبت به این نوع ارزشیابی و تأثیر آن بر نوع مطالعه آنها توسط پرسشنامه با مقیاس لیکرت، بررسی شد.

نتایج: ارتباط قوی و مثبت بین نمرات بخش چندگزینه‌ای و تشریحی ۳ دوره وجود داشت؛ گروه اول با تعداد ۱۵۱ نفر ($p<۰/۰۰۰۵$, $r=۰/۶۳۱$)، گروه دوم با تعداد ۱۳۶ نفر ($p<۰/۰۰۰۵$, $r=۰/۷۱۰$) و گروه سوم با تعداد ۱۶۸ نفر ($p<۰/۰۰۰۵$, $r=۰/۷۱۶$). بیش از ۵۰٪ دانشجویان معتقد بودند زمانی که صرف مطالعه‌ی این درس نموده‌اند بیش‌تر از زمان صرف شده در دروس دیگری که آزمون آنها صرفاً به صورت چندگزینه‌ای برگزار شده، بوده است.

نتیجه‌گیری: با توجه به همبستگی بالا و مثبت نمرات بخش تشریحی با بخش چندگزینه‌ای آزمون، می‌توان در کنار آزمون چندگزینه‌ای از آزمون تشریحی نیز برای ارزیابی میزان یادگیری دانشجو در درس پاتوفیزیولوژی کلیه و همچنین افزایش انگیزه‌ی دانشجویان برای مطالعه عمیق‌تر استفاده نمود.

واژه‌های کلیدی: آزمون تشریحی، آزمون چند گزینه ای، ارزشیابی دانشجو، دوره مقدمات پزشکی بالینی

مجله ایرانی آموزش در علوم پزشکی/ ۱۳۹۶؛ ۱۷(۵۹): ۵۵۲ تا ۵۶۰

مقدمه

نوع سنجش دانشجویان تاثیر به سزایی در فرایند آموزش آنها دارد؛ به گونه‌ای که نحوه تدریس، نحوه یادگیری و نحوه ارزشیابی دانشجویان را تحت تأثیر قرار می‌دهد(۱). تأثیر سنجش بر آموزش و یادگیری به این

* **نویسنده مسؤول:** دکتر وجیهه وفامهر، پزشک عمومی و دانشجوی PhD آموزش پزشکی، واحد EDO دانشکده پزشکی، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران.
vafamehr.v@gmail.com
دکتر شهرزاد شهیدی(استاد)، گروه داخلی(نفرولوژی)، مرکز تحقیقات بیماری های کلیه، دانشکده پزشکی، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران.
(shahidi_shahrzad@yahoo.com) ؛ دکتر محمد دادگسترنیا، پزشک عمومی و کارشناس ارشد آموزش پزشکی، مرکز تحقیقات آموزش علوم پزشکی، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران. (mdadgostarnia@gmail.com)

تاریخ دریافت مقاله: ۹۵/۷/۱۰، تاریخ اصلاحیه: ۹۵/۱۲/۱۷، تاریخ پذیرش: ۹۶/۵/۱۴

دلیل ایجاد می‌شود که آنچه مورد سنجش قرار می‌گیرد تبدیل به چیزی می‌شود که برای آن ارزش قائل می‌شویم و طبیعی است که همان هم آموزش داده می‌شود(۲).

دانشجویان برای روش‌های مختلف سنجش از رویکردهای متفاوت مطالعه بهره می‌گیرند(۳). استفاده از آزمون‌های عینی چندگزینه‌ای در ارزشیابی دانشجویان باعث می‌شود که آنها هنگام مطالعه و یادگیری مطالب درس، در جهت پیداکردن و به یاد سپردن نکات جزیی و پراکنده درس کوشش کنند و به مسائل کلی و ساختاری آن توجه زیادی نکنند. در مقابل، دانشجویانی که خود را برای شرکت در یک امتحان تشریحی آماده می‌کنند، به اطلاعات پراکنده درس چندان اهمیتی نمی دهند و بیش‌تر سعی می‌کنند تا مسائل کلی و استنباطی را فرا بگیرند(٤). پژوهشگران معتقدند که در هنگام آماده شدن برای امتحان، سؤال‌های تشریحی روش‌های برتر مطالعه را برمی انگیزند، اما آزمون‌های چندگزینه‌ای باعث یادگیری سطحی جزییات می‌شوند. زیرا تأکید این آزمون‌ها بر بازشناسی است و لازم نیست که یادگیرندگان مطالب را سازمان‌دهی کنند یا درباره آنها استدلال نمایند(۲تا۵). نظام سنجش می‌تواند فرایند آموزش و یادگیری را، هم به‌صورت مثبت وهم به صورت منفی متأثر سازد. پژوهشگرانی که به بررسی تأثیر آزمونهای ملی و سراسری پرداخته‌اند، از اصطلاحات واش بک (Wash back) مثبت و منفی برای نشان دادن این اثرات استفاده کرده‌اند(۵).

تأثیر سنجش بر جریان آموزشی به پدید آمدن مفهومی به نام "آموزش جهت داده شده به وسیله اندازه‌گیری (measurement driven instruction) انجامیده است؛ چنانچه روش‌های آموزشی معلمان به وسیله سنجش‌های سطحی هدایت شود، یادگیرندگان نیز به صورت سطحی تربیت می‌شوند(٤). منتقدین به‌آزمون‌های سنتی، معتقدند که این آزمون‌ها به ویژه آزمون‌های چندگزینه‌ای، مهارت‌هایی را می‌سنجند که در دنیای واقعی مانند ندارند(٦). توماس و بین(Thomas & Bain)، تراب و مک‌روی (Traub & MacRury) و شکورنیا و همکاران در پژوهش‌های خود نشان داده‌اند که انتظار دانشجویان از نوع آزمون (چندگزینه‌ای یا تشریحی) نحوه مطالعه آنها را تحت تأثیر قرار می‌دهد وآنها به روش‌های کاملاً متفاوتی برای این دو نوع امتحان آماده می‌شوند. یافته جالب دیگر این است که وقتی دانشجویان برای یک امتحان از نوع تشریحی آماده می‌شوند، در صورت تغییر فرمت آزمون به چندگزینه‌ای می‌توانند از پس آن برآیند، اما اگر برای یک امتحان از نوع چندگزینه‌ای آماده شوند، در صورت تغییر فرمت آزمون به تشریحی، نمی‌توانند به آسانی از عهده آن برآیند(۳و۷تا۹). اما در این میان در چند مطالعه از جمله مطالعه اسکولر و پروسر (Scoullar & Prosser)، رابطه ای بین بکارگیری آزمون چندگزینه ای و استفاده از رویکرد سطحی مطالعه وجود ندارد(۱۰و۱۱).

با وجود تمام محاسنی که در مطالعات مختلف برای آزمون‌های تشریحی بیان شده است، اما در عمل کمتر از این آزمون‌ها استفاده می‌شود. یکی از دلایل عمده مخالفت برخی از صاحب‌نظران با کاربرد این آزمون‌ها ناشی از بی‌ثباتی نمره‌گذاری این آزمون‌ها است. آزمون‌های تشریحی در گروه آزمون‌های ذهنی یا Subjective طبقه‌بندی می‌گردد. این بدان معناست که بخش اعظم اندازه‌گیری توانایی مورد نظر به شخصی که پاسخ‌های سؤال را می‌خواند و به آنها نمره می‌دهد وابسته است. اگرچه به کاربردن قواعد تصحیح آزمون‌های تشریحی که در متون علمی به آنها اشاره شده است تا حدودی از میزان ذهنی بودن نمره‌دهی می‌کاهد اما بازهم به کار بردن این نوع سؤال منجر به نارضایتی شدید دانشجویان می‌گردد(۱۲).

برای بهره‌مندی از منافع آزمون‌های تشریحی در دروس بالینی، باید اساتید و دانشجویان مطمئن باشند که می‌توان بر اساس نمرات این آزمون‌ها قضاوت پایایی درباره

دانشجویان داشت. در متون علمی راه‌های مختلفی برای افزایش روایی و پایایی این نوع آزمون پیشنهاد شده است. از جمله این که سؤالات تشریحی در کنار سؤالات عینی‌تر مانند سؤالات چهارگزینه‌ای استفاده شود. همچنین قواعد و اصولی برای تصحیح سؤالات تشریحی ارائه شده است. اما سؤال اینجاست که با رعایت این اصول تا چه میزان می‌توانیم بر پایایی نمرات آزمون تشریحی اطمینان کنیم. از جمله راه‌هایی که می‌توان برای بررسی میزان اطمینان به نمرات آزمون تشریحی استفاده کرد، مقایسه نمرات دانشجویان در بخش‌های چهار گزینه‌ای و تشریحی یک آزمون است. بنابراین، مطالعه حاضر با هدف بررسی میزان همبستگی نمرات بخش تشریحی با بخش چندگزینه‌ای آزمون پایانی درس پاتوفیزیولوژی کلیه و بررسی نقطه نظرات دانشجویان درباره تأثیر نوع آزمون بر کیفیت مطالعه آنها انجام گردید.

روش‌ها

مطالعه حاضر یک مطالعه توصیفی از نوع همبستگی و پیمایشی بود که از پاییز ۹٤ تا زمستان ۹۵ در ۳ دوره از درس پاتوفیزیولوژی کلیه مقطع مقدمات پزشکی بالینی(ICM) دوره پزشکی عمومی دانشگاه علوم پزشکی اصفهان انجام گرفت. در این بررسی ٤۵۵ دانشجو در ۳ دوره به روش سرشماری انتخاب شدند و برای ارزشیابی پایانی درس پاتوفیزیولوژی کلیه، در آزمونی کتبی مشتمل بر ۳٦ سؤال چندگزینه‌ای و ۱۰ سؤال تشریحی شرکت کردند. بدین صورت که بر خلاف ترم‌های پیش از آن، در ابتدای بسته آموزشی با ارائه طرح درسی شامل اطلاعات کافی در مورد منبع درس، نحوه‌ی تدریس اساتید درس و نحوه ارزشیابی پایانی به دانشجویان داده شد. همچنین فایل الکترونیک آخرین چاپ کتاب منبع (Cecil Essential 2016) نیز از طریق شبکه تلگرام در اختیار دانشجویان و اساتید درس قرار گرفت. از دانشجویان خواسته شد که در طول دوره علاوه بر جزوات کلاسی، کتاب را نیز مطالعه کنند. در طول دوره دانشجویان از طریق شبکه تلگرام می‌توانستند با مسؤول درس ارتباط داشته باشند و راهنمایی‌های لازم را دریافت کنند. اساتید دوره نیز درباره تدریس بر اساس کتاب و طرح سؤال مبتنی بر کتاب توجیه شدند. در دوره‌های قبل برای ارزشیابی پایانی درس کلیه از آزمون چندگزینه‌ای استفاده می‌شد و برای هر جلسه درس براساس Blue print درس ٦-٤ سؤال چندگزینه‌ای طراحی می‌گردید. اما در این دوره از اساتید خواسته شد که از هر یک از جلسات ۱۲ گانه درس، ۳ سؤال چندگزینه‌ای و ۱ سؤال تشریحی مطابق با آخرین چاپ کتاب منبع طرح نمایند. ۵۰ درصد نمره پایانی به آزمون چند گزینه‌ای (۱۰ از ۲۰ نمره) و ۵۰ درصد به آزمون تشریحی اختصاص داده شد. در پایان دوره آزمون پاتوفیزیولوژی کلیه مشتمل بر ۳٦ سؤال چندگزینه‌ای و ۱۰ سؤال تشریحی برگزار گردید.

پس از آزمون، نمرات آزمون چندگزینه‌ای از طریق OPSCAN تصحیح و محاسبه شد. برای حل مشکل بی‌ثباتی در نمره دهی در زمان تصحیح اوراق آزمون تشریحی، قواعدی که در متون علمی آمده، در نظر گرفته شد (۱۲و۱۳).

تصحیح بخش تشریحی بدین صورت انجام شد که ۲ سؤال توسط طراح مربوطه برای همه دانشجویان تصحیح شد و بقیه سؤالات توسط مسؤول درس و با همکاری اساتید و دستیاران گروه کلیه و براساس پاسخ‌های ارائه شده توسط طراحان سؤال‌ها، تصحیح گردید. نمره‌بندی سؤالات تشریحی توسط مسؤول درس و با توجه به تعداد موارد مربوط به پاسخ هر سؤال انجام گرفت. در نهایت همبستگی بین نمرات بخش چندگزینه‌ای و تشریحی آزمون کلیه از طریق همبستگی دو متغیره Bivariate Correlation مورد بررسی قرار گرفت. البته آنالیزهای مقدماتی به منظور اطمینان از normality، linearity

انجام گرفت. برای تحلیل همبستگی بین نمرات آزمون چندگزینه‌ای و تشریحی، از گایدلاین (Cohen1988) به صورت زیر استفاده شد:

r=.10 to.29 or r= -.10 to -.29 Small
r=.30 to.49 or r= -.30 to -.49 Medium
r=.50 to 1.0 or r= -.50 to -1.0 large

مقایسه میانگین نمرات آزمون تشریحی با آزمون چهارگزینه‌ای از طریق آزمون تی جفت Paired Samples T Test صورت پذیرفت.

همچنین از طریق یک پرسشنامه محقق ساخته ٦ سؤالی با مقیاس لیکرت (از کاملاً مخالف تا کاملاً موافق با امتیاز ۱ تا ٥)، رضایت اولین گروه دانشجویان شرکت‌کننده در این مطالعه نسبت به این نوع ارزشیابی (چندگزینه‌ای در کنار تشریحی)، تأثیر آن بر نوع مطالعه آنها، زمان صرف شده برای مطالعه و رضایت از میزان یادگیری درس جمع‌آوری شد. روایی پرسشنامه از طریق کارشناسان و اساتید و پایایی آن از طریق همبستگی درونی با آلفای کرونباخ ۰/۸۵ در استفاده از پرسشنامه برای ۳۰ نفر از گروه هدف تعیین شد. پرسشنامه در پایان آزمون چندگزینه‌ای-تشریحی مابین کلیه ۱۵۱ نفر شرکت‌کننده در آزمون اول توزیع شد. نتایج پرسشنامه رضایت‌سنجی از طریق آمار توصیفی، میانگین و فراوانی تحلیل گردید. فراوانی موافقان و کاملاً موافقان بالای ٥٠ درصد به عنوان رضایت از آیتم ذکر شده در نظر گرفته شد.

نتایج

شرکت کنندگان در بخش آزمون چندگزینه ای- تشریحی مطالعه در ۳ ترم متوالی به ترتیب۱۵۱ نفر [۷۷ نفر(٥١٪)زن، ٧٤ نفر(٤٩٪) مرد]، ١٣٦نفر [٥٩ نفر(٤٣٪) زن، ٧٧ نفر(٥٧٪) مرد] و١٦٨ نفر [٩٤ نفر(٥٦٪) زن، ٧٤ نفر(٤٤٪) مرد] دانشجوی ترم اول دوره مقدمات پزشکی بالینی بودند. با توجه به این که این مداخله در آزمون پایان دوره آنها صورت گرفت، بنابراین ریزشی در تعداد نمونه در بخش آزمون چندگزینه ای- تشریحی مطالعه نداشتیم. مقایسه میانگین نمرات آزمون چندگزینه‌ای و آزمون تشریحی به طور جداگانه در ۳ گروه نشان داد که میانگین نمرات دانشجویان در بخش چندگزینه‌ای آزمون در گروه اول به طور معناداری از بخش تشریحی آزمون بالاتر بوده است. اما در گروه دوم و سوم میانگین نمره تشریحی بالاتر بود و این تفاوت در گروه دوم معنادار بود و در گروه سوم معنادار نبود(جدول۱).

جدول ۱: میانگین نمرات بخش‌های چندگزینه‌ای و تشریحی آزمون کلیه در سه نیمسال

	تعداد کل	تعداد موارد رد شده در آزمون(درصد)	میانگین نمره از ۱۰		sig	t
			بخش چند گزینه‌ای	بخش تشریحی		
گروه ۱	۱۵۱	۲۸(٪۱۹/٤٤)	٦/٤٤±۱/٤۳	٥/۱٥±۱/۷٤	۰/۰۰۰۱	۱۱/۳۷٤
گروه ۲	۱۳٦	۱٥(٪۱۱/٦۲	٦/۳٥±۱/۳۳	٦/۹۳±۱/۸۸	۰/۰۰۰۱	-٥/۱۰۲
گروه ۳	۱٦۸	۳٦(٪۲۱/٤۲	٥/۸۸±۱/۲۱	٦/۰٤±۱/٦٦	۰/۷۲	-۱/۸۰۸

آنالیز همبستگی میان نمرات بخش چندگزینه‌ای و تشریحی آزمون نشان داد که یک ارتباط قوی و مثبت میان نمرات بخش چندگزینه‌ای و بخش تشریحی آزمون‌های ۳ دوره بیماری‌های کلیه وجود داشته است: گروه اول با تعداد ۱۵۱ نفر (r=۰/٦۳۱ , p<۰/۰۰۰٥)، گروه دوم با تعداد ١٣٦ نفر (r=۰/۷۱۰ , p<۰/۰۰۰٥) و گروه سوم با تعداد ١٦٨ نفر (r=۰/۷۱٦ , p<۰/۰۰۰٥) همان‌گونه که مشاهده می‌شود، اگرچه درصد دانشجویان رد شده در گروه سوم بیش‌تر است، اما بیش‌ترین همبستگی مربوط به این گروه بوده و ۷۱ ٪ همبستگی در

نمونه‌ها، قابل تعمیم به جامعه است(جدول ۱).
در بخش دیگر مطالعه، تکمیل پرسشنامه‌ی اخذ نقطه نظرات دانشجویان درباره نحوه ارائه درس کلیه و نوع آزمون کلیه بود که توسط ۱۲۲ دانشجو [٦٧ نفر(٥٥٪) زن و ٥٥ نفر(٤٥٪) مرد] از ۱۵۱ نفر شرکت‌کننده در اولین دوره مداخله، تکمیل شد و بنابراین ۲۹ نفر(۱۹/۲۰ درصد) ریزش داشت که احتمالاً ناشی از خستگی پس از امتحان بود.
طبق نتایج فقط در دو آیتم زمان صرف شده برای مطالعه و لزوم مطالعه جزوه، فراوانی موافقان و کاملاً موافقان بالاتر از ۵۰ درصد بود (جدول ۲).

جدول۲: توزیع فراوانی نسبی و مطلق ومیانگین نمره نظرات دانشجویان درباره نحوه ارائه و شیوه آزمون درس کلیه

آیتم‌ها	کاملاً موافق (نمره ٥)	موافق (نمره ٤)	بدون نظر (نمره۳)	مخالف (نمره۲)	کاملاً مخالف (نمره۱)	میانگین انحراف معیار
اطلاعات ارائه شده در ابتدای بسته آموزشی (درباره قوانین کلاس و آزمون) راهنمای خوبی برای من بود	٦(٤/٨٪)	٣٨(٣٠/٦٪)	٤١(٣٣/١٪)	٢٣(١٨/٥٪)	١٣(١٠/٥٪)	٣±١/٠٦
نسبت به دروس دیگر، این درس را عمیق‌تر مطالعه نمودم	١٤(١١/٣٪)	٣١(٢٥٪)	٣١(٢٥٪)	٣٢(٢٥/٨٪)	١٣(١٠/٥٪)	٣±١/١٩
زمانی که صرف مطالعه‌ی این درس نمودم بیش‌تر از مدت زمان صرف شده در دروس دیگر بود	٢٦(٢١٪)	٣٩(٣١/٥٪)	٢٣(١٨/٥٪)	٢٨(٢٢/٦٪)	٥(٤٪)	٣/٤٣±١/١٨
فکر می‌کنم برای پاسخ به سؤالات لازم بود کتاب منبع رامطالعه کنم	٩(٧/٣٪)	٣٦(٢٩٪)	٣٩(٣١/٥٪)	٢٨(٢٢/٦٪)	٩(٧/٣٪)	٣/٠٦±١/٠٦
فکر می‌کنم برای پاسخ به سؤالات لازم بود جزوه را مطالعه کنم	٢٤(١٩/٤٪)	٦٤(٥١/٦٪)	٢١(١٦/٩٪)	١١(٨/٩٪)	١(٠/١٪)	٣/٨١±٠/٨٨
پس از پایان آزمون این درس از میزان یادگیری ام بیش‌تر از دروس دیگر راضی بودم	٧(٥/٦٪)	١٧(١٣/٧٪)	٤١(٣٣/١٪)	٣٢(٢٥/٨٪)	٢٤(١٩/٤٪)	٢/٥٩±١/١٢

بحث

هدف از انجام این مطالعه تعیین همبستگی نمرات بخش‌های تشریحی و چندگزینه‌ای آزمون پایانی درس پاتوفیزیولوژی کلیه و نقطه نظرات دانشجویان درباره تأثیر نوع آزمون بر مطالعه آنها بود. مهم‌ترین یافته‌ی این مطالعه، همبستگی بالا و معنادار نمرات عملکرد دانشجویان در بخش چندگزینه‌ای و تشریحی آزمون بیماری‌های کلیه طی ۳ دوره بود.
مهم‌ترین یافته‌ی این مطالعه، همبستگی بالا و معنادار نمرات عملکرد دانشجویان در بخش چندگزینه‌ای و تشریحی آزمون بیماری‌های کلیه طی ۳ دوره بود. مجیب (Mujeeb) (هند) و پپل (Pepple) (هند) و اویبولا (Oyebola) (هند) نیز نتایج مشابهی را از عملکرد دانشجویان در دو نوع آزمون چندگزینه‌ای و تشریحی کوتاه پاسخ و بلند پاسخ به دست آورده‌اند(١٤ تا١٦).
همان‌گونه که در نتایج نشان داده شده است، در مقایسه میانگین نمرات آزمون چندگزینه‌ای و آزمون تشریحی میانگین نمرات دانشجویان در بخش چندگزینه‌ای آزمون فقط در گروه اول به طور معناداری از بخش تشریحی آزمون بالاتر بود. در گروه دوم و سوم میانگین نمره

تشریحی دانشجویان بالاتر از میانگین نمره بخش چندگزینه‌ای آزمون بود که البته این اختلاف در گروه دوم معنادار بود و در گروه سوم معنادار نبود. این نتیجه بر خلاف انتظار بود چرا که همیشه انتظار داریم دانشجویان در آزمون‌های چندگزینه‌ای نمره بالاتری دریافت نمایند. این مسأله شاید به این دلیل باشد که در آزمون تشریحی، پاسخ‌های مختصر و ناکامل نیز تا حدی نمره می‌گیرد. یعنی اطلاعات اندک دانشجو راجع به موضوع نیز ارزش دارد. اما در آزمون چندگزینه‌ای اطلاعات نسبی بی ارزش هستند و آنچه اهمیت دارد، اطلاع مطلق از موضوع است که باعث هدایت فرد به گزینه صحیح می‌شود. پس شاید آزمون تشریحی از این لحاظ مزیتی بر آزمون چندگزینه‌ای داشته باشد(۱۲). در مطالعه مشابهی که در دانشجویان پرستاری دانشگاه شهرکرد انجام گرفته، نمرات بخش چهارگزینه‌ای به طور معناداری بالاتر از بخش تشریحی آزمون مادر و کودک بود(۱۷). در مطالعه پپل (Pepple) نیز میانگین نمره آزمون چندگزینه‌ای در هر دو گروه نمرات بالا و نمرات پایین، از میانگین نمرات بخش تشریحی بالاتر بوده است(۱۵). مکاتاش (Moqattash) و همکاران نیز نتایج مشابهی به دست آورده و نتیجه‌گیری کرده‌اند که اگر از آزمون‌های چهارگزینه‌ای به تنهایی استفاده شود، برخی از فراگیران بدون توانایی شناخت، تجزیه و تحلیل، تفکر و کاربرد دانش به مقاطع بالاتر می‌روند. بنابراین، براستفاده از روش‌های مختلف ارزیابی تأکید کرده‌اند(۱۸).

از طرف دیگر در گروه دوم و سوم مطالعه ما، نمره تشریحی بالاتر بوده است. به نظر می‌رسد این نتایج متفاوت به این دلیل است که در گروه دوم و سوم، دانشجویان از تجارب گروه اول استفاده کرده‌اند و مطالعه را طوری برنامه‌ریزی کرده‌اند که عملکرد آنها در آزمون تشریحی با موفقیت بیش‌تری همراه بوده است. این مسأله تاییدکننده این نظریه است که اگر مدرسین رویکرد ارزشیابی خود را تغییر دهند، عملکرد فراگیران نیز تغییر خواهد کرد(۱۹).

در بخش نظرات دانشجویان می‌بینیم که بیش از نیمی دانشجویان معتقد بودند که زمانی که صرف مطالعه‌ی این درس نموده‌اند بیش‌تر از مدت زمان صرف شده در دروس دیگری بوده است که آزمون آنها صرفاً به صورت چندگزینه‌ای برگزار شده است. همچنین دو سوم دانشجویان معتقد بودند که برای پاسخ به سؤالات لازم بوده است که جزوه را هم مطالعه کنند و بیش از یک سوم آنها با لزوم مطالعه کتاب موافق بودند. این نتیجه نشان می‌دهد، با وجودی که سؤالات از کتاب طراحی شده بودند اما دانشجویان به نقش حضور در کلاس و مطالعه جزوه کلاسی نیز واقف بوده اند.

استفاده از آزمون‌های تشریحی برای ارزشیابی دانشجویان در مقاطع پس از علوم پایه، مسأله رایجی در دانشگاه‌های علوم پزشکی نیست. چرا که در کشور ما به علت تعدد وظایف، اساتید بالینی فرصت‌های کمی را به طراحی ارزشیابی اختصاص می‌دهند. با توجه به نتایج فوق می‌توان گفت که اگر آزمون‌ها فقط برای صدور مجوز رفتن به مرحله بعدی برای دانشجویان اجرا می‌گردد، استفاده از الگوهای متفاوت آزمون توجیهی ندارد و با توجه به همبستگی نمرات در دو الگوی به کار رفته در این پژوهش، می‌توانیم از همان روش قدیمی سؤالات چندگزینه‌ای به تنهایی استفاده کنیم. پپل (Pepple) نیز از نتایج مطالعه خود، چنین استنباطی کرده است(۱۵). اما اگر هدف از آزمون تعیین میزان پیشرفت تحصیلی دانشجو و دستیابی به سطوح متفاوت یادگیری باشد، بهتر است از روش‌هایی استفاده شود که هم برآوردی از سطوح بالاتر یادگیری به استاد بدهد و هم به عنوان عاملی برای افزایش انگیزه‌ی دانشجویان برای مطالعه عمیق‌تر عمل نماید. به نظر می‌رسد که شیوه برگزاری آزمون در سیستم علوم پزشکی کشورمان نه تنها منجر به صعود افراد توانمند به مراتب بالاتر نمی‌شود بلکه به عنوان مانعی بر سر راه آموزش و

مطالعه صحیح و از همه مهم‌تر هدر رفت فرصت‌های یادگیری دانشجویان عمل می‌کند.

انجام مطالعات تکمیلی و با قدرت علمی بالا در جهت بررسی بهترین شیوه‌های ارزشیابی در مقاطع مختلف دوره پزشکی ضروری است.

نقطه‌ی قوت این پژوهش این بود که برای اولین بار در دانشگاه علوم پزشکی اصفهان و شاید در ایران، از آزمون تشریحی در کنار آزمون چهار گزینه‌ای در یک گروه بالینی استفاده گردید؛ که با توجه به مقاومتی که در برابر استفاده از آزمون‌های تشریحی وجود دارد، می‌تواند به عنوان گام آغازینی برای تسری استفاده از این نوع آزمون و سایر انواع آزمون‌ها در قالب پژوهش‌های آموزشی در سایر گروه‌های بالینی باشد.

نتیجه‌گیری

با توجه به این که نمرات حاصل از آزمون تشریحی همبستگی بالا و مثبتی با نمرات آزمون چندگزینه‌ای داشت، می‌توان برای عمقی‌تر کردن مطالعه دانشجویان، در کنار آزمون چندگزینه‌ای از آزمون تشریحی نیز برای ارزیابی میزان یادگیری دانشجو در درس پاتوفیزیولوژی کلیه مقطع ICM (Introduction to Clinical Medicine) استفاده نمود. به طور کلی بهتر است در آزمون‌هایی که هدف تعیین میزان پیشرفت تحصیلی و دستیابی به سطوح متفاوت یادگیری است، از آزمون‌های تشریحی که هم برآوردی از سطوح بالاتر یادگیری ارائه می‌دهد و هم به عنوان عاملی برای افزایش انگیزه‌ی دانشجویان برای مطالعه عمیق‌تر عمل می‌نماید، استفاده نمود.

قدردانی

پژوهشگران از همکاری اعضای هیأت‌علمی گروه فوق‌تخصصی نفرولوژی دانشکده پزشکی اصفهان که با وجود مشغله کاری، در اجرا و تصحیح بخش تشریحی آزمون‌ها همکاری بی‌دریغ داشتند، کمال تشکر و قدردانی را دارند. همچنین از همکاری عوامل اجرایی دوره مقدمات پزشکی بالینی و تمامی مشارکت کنندگان در این پژوهش، سپاس‌گزاری و قدردانی می‌گردد.

منابع

1. Cheng L. How does wash back influence teaching? Implications for Hong Kong. Language and Education. 1997; 11(1): 38- 5
2. James PW. The merits of measurement-driven instruction. Phi Delta Kappa. 1987; 68(9): 679-682.
3. Thomas PR, Bain JD. Contextual differences of learning approaches: the effects of assessment. Human Learning.1984; 3: 227-240.
4. Seif AA. [Measurement, assessment and training evaluation]. 3th ed. Tehran: Doran; 2011.[Persian]
5. Taylor L. Washback and impact. ELT Journal. 2005; 59(2): 154-155
6. Woolfolk-Hoy AE. Educational Psychology. 9th ed. London, England: Pearson; 2004.
7. Birenbaum M, Tatsuoka KK, Gutvirtz Y. Effects of Response Format on Diagnostic Assessment of Scholastic Achievement. Applied Psychological Measurement. 1990; 16(4): 353-363.
8. Shakurnia A, Alijani H, Najjar S, komeili H, Elhampour H. [The Effect of Two Assessment Methods on Exam Preparation and Study Strategies: Multiple Choice and Essay Questions]. Iranian Journal of Medical Education. 2013; 13(4) :306-318.[Persian]
9. Aslfatahi B, Karimi Y, Seif A, Pashasharifi H. [The Effect of Mindfulness-Based Techniques in Reduction of Students Test Anxiety]. Journal Of Instruction And Evaluation. 2009; 2(6): 7-33.[Persian]
10. Scoullar KM, Prosser M. Students' experiences in studying for multiple-choice question examinations.Studies in Higher Education. 1994; 19(3): 267-79.
11. Smith SN, Miller RJ. Learning Approaches: Examination type, discipline of study, and gender. Educational psychology. 2005; 25(1): 43-53.
12. Zolfaghari B, Adibi N, Derakhshanfar S,Tansaz M, Karbasi A, Niroomand P. [Academic achievement rests in medical sciences]. 1th ed. Isfahan: Isfahan university of medical science; 1379.[Persian]

13. Fathabadi J, Saif A. [Investigating the effects of type of assessment (Essay and Multiple-Choice) on students' approaches to studying and exam perception strategies between students with high and low achievement]. Journal of Education and Psychology. 2007; 3(4): 21-46.[Persian]
14. Mujeeb AM, Pardeshi ML, Ghongane BB. Comparative assessment of multiple choice questions versus short essay questions in pharmacology examinations. Indian J Med Sci. 2010; 64(3): 118-24.
15. Pepple DJ, Young LE, Carroll RG. A comparison of student performance in multiple-choice and long essay questions in the MBBS stage I physiology examination at the University of the West Indies (Mona Campus). Adv Physiol Educ. 2010; 34(2): 86-9.
16. Oyebola DDO, Adewoye OE, Iyaniwura JO, Alada ARA, Fasanmade AA, Raji Y. A comparative study of students' performance in preclinical physiology assessed by multiple choice and short essay questions. Afr J Med Med Sci. 2000; 29(3-4): 201-5.
17. Delaram M, Sharifi A. A Comparison of student's performances in multiple- choice and essay questions in mother and child Health examination. The Future of Medical Education Journal. 2014; 7(14): 31-37.
18. Moqattash S, Harris PF, Gumaa KA, Abu-Hijleh MF. Assessment of basic medical sciences in an integrated systems-based curriculum. Clin Anat. 1995; 8(2): 139-47.
19. Zeidner M. Essay versus Multiple-Choice Type Classroom Exams: The Student's Perspective. The Journal of Educational Research. 1987; 80(6): 352-358.

Comparison of the scores in essay part and MCQ part of renal pathophysiology exam and surveying students' views about the effect of such exams on their study

Shahrzad Shahidi[1], Vajiheh Vafamehr[2], Mohammad Dadgostarnia[3]

Abstract

Introduction: *In order to benefit from the advantages of essay exams, one must be sure to reliably judge students based on essay test scores. The aim of this study was to examine the correlation of scores in the essay and MCQ parts of renal pathophysiology final exam and students' views about the effect of the type of test on their study.*

Methods: *This descriptive correlational survey was performed in Isfahan School of Medicine from fall 2015 to winter 2017. Three groups of students (n=455) participated in the renal pathophysiology and took a test consisting of 36 multiple-choice and 10 essay questions. The MCQ part was graded in OPSCAN software. The correlation of MCQ and essay scores was investigated via bivariate correlation and paired t-test was used to compare the mean scores. Students' views about the type of test and its effect on their study were surveyed by means of a Likert-scale questionnaire.*

Results: *There was a strong positive correlation between the scores in essay and MCQ parts in the three groups (group 1: r=0.631 n=151 p<0.0005; group 2: r=0.710 n=136 p<0.0005; group3: r=0.716 n=168 p<0.0005). More than 50 percent of students stated that they had spent more time for studying this course than for those courses with MCQ-only mode of exam.*

Conclusion: *Considering the high and positive correlation between the scores in the MCQ and essay parts of the exam, essay tests can be used along with MCQ tests to assess students' level of learning in renal pathophysiology and to increase their motivation for deeper study.*

Keywords: Essay test, MCQ test, student evaluation, introduction to clinical medicine course

Addresses:

1. Professor, Internal Medicine Department (Nephrology), Isfahan Kidney Diseases Research, Medical School, Isfahan University of Medical Sciences, Isfahan, Iran. Email: shahidi_shahrzad@yahoo.com
2. (✉) MD & PhD Student In Medical Education, EDO Office, Medical School, Isfahan University Of Medical Sciences, Isfahan, Iran. Email: vafamehr.v@gmail.com
3. MD & MSc in Medical Education, Isfahan University of Medical Sciences, Isfahan, Iran. Email: mdadgostarnia@gmail.com