

بررسی سؤالات امتحانی اعضای هیأت علمی دانشگاه علوم پزشکی جندی شاپور اهواز با تأکید بر ساختار سؤالات چند گزینه‌ای در سال ۱۳۸۶

عبدالحسین شکورنیا^{۱*}، آذر خسروی بروجنی^۲، علیرضا مظفری^۳، حسین الهام‌پور^۴

۱. فوق لیسانس ایمنی‌شناسی، مربی گروه ایمنولوژی، دانشکده پزشکی، دانشگاه علوم پزشکی جندی شاپور، اهواز، ایران
۲. دکترای میکروبی‌شناسی، دانشیار گروه میکروبی‌شناسی دانشکده پزشکی، دانشگاه علوم پزشکی جندی شاپور، اهواز، ایران
۳. متخصص داخلی، دانشیار گروه داخلی دانشکده پزشکی، دانشگاه علوم پزشکی جندی شاپور، اهواز، ایران
۴. فوق لیسانس علوم تربیتی، مربی گروه علوم تربیتی دانشکده علوم تربیتی و روان‌شناسی، دانشگاه شهید چمران، اهواز، ایران

● دریافت مقاله: ۸۸/۴/۸ ● پذیرش مقاله: ۸۸/۸/۱۶

زمینه و هدف: ارزشیابی پیشرفت تحصیلی بخش مهمی از فرآیند آموزشی است و برای تحقق آن از آزمون استفاده می‌شود. استفاده از سؤالات چند گزینه‌ای به عنوان یک روش آزمون عینی، متداول‌ترین روش ارزشیابی دانشجویان می باشد. این مطالعه با هدف بررسی سؤالات امتحانی اعضای هیأت علمی دانشگاه علوم پزشکی جندی شاپور صورت گرفت.

روش کار: در این مطالعه توصیفی سؤالات امتحانی ۱۹۷ عضو هیأت علمی غیر بالینی دانشگاه مورد بررسی قرار گرفت. داده‌ها با استفاده از چک لیستی که حاوی اطلاعات فردی، ویژگی‌های سؤالات امتحانی و همچنین موارد تطبیق سؤالات چند گزینه‌ای با قواعد طراحی این‌گونه سؤالات بود، جمع‌آوری گردید.

یافته‌ها: نتایج نشان داد ۸/۱ درصد از اعضای هیأت علمی از سؤالات انشایی، ۵۵/۳۲ درصد از سؤالات عینی و ۳۶/۵۴ درصد به‌صورت ترکیبی از سؤالات انشایی و عینی استفاده کرده بودند. از مجموع ۵۲۰۸ سؤال چند گزینه‌ای مورد بررسی، ۶۴/۹۶ درصد، فاقد هر گونه اشکال ساختاری بودند. بیشترین فراوانی اشکالات ساختاری در سؤالات شامل مشخص نبودن و یا عدم وضوح هدف در ساقه سؤال (۱۵/۹۴ درصد)، استفاده از افعال منفی در ساقه سؤال (۱۲/۲۲ درصد)، استفاده از عبارات‌های "همه موارد" یا "هیچ‌کدام" در گزینه‌ها (۱۰/۸۵ درصد) و بکار بردن جای خالی در اول یا وسط ساقه سؤالات (۱۰/۴۱ درصد) بود. حدود ۹۰ درصد سؤالات چند گزینه‌ای در حیطه دانش یعنی تاکسونومی I بود.

نتیجه‌گیری: با توجه به اهمیت ارائه سؤالات استاندارد و منطبق با قواعد طراحی سؤالات چند گزینه‌ای در برگزاری آزمون‌های مناسب، بازبینی مستمر سؤالات امتحانی و برگزاری دوره‌های آموزشی طراحی سؤالات امتحانی توصیه می‌شود.

کلید واژه‌ها: آزمون، عضو هیأت علمی، سؤال چند گزینه‌ای

* نویسنده مسؤول: گروه ایمنی‌شناسی، دانشکده پزشکی، دانشگاه علوم پزشکی جندی شاپور، اهواز، ایران

مقدمه

ارزشیابی پیشرفت تحصیلی دانشجویان یکی از اهداف عمده دانشگاه‌ها و مراکز آموزش عالی کشور است. زیرا بدین وسیله میزان یادگیری دانشجویان و در نهایت میزان دستیابی به اهداف آموزشی اندازه‌گیری می‌شود (۱). سوابق ارزیابی به ۲۲۰۰ سال قبل از میلاد بر می‌گردد. در آن زمان در چین از آزمون‌های کتبی برای گزینش افراد لایق برای احراز تصدی امور مملکتی استفاده می‌کردند. در ایران باستان در عهد شاپور ساسانی در دانشگاه جندی شاپور هم جلساتی برای آزمون دانشجویان پزشکی بر پا می‌کردند و اعطای گواهی‌نامه پزشکی موکول به موفقیت دانشجویان در این امتحانات بود (۲).

مهم‌ترین وسیله اندازه‌گیری پیشرفت تحصیلی، آزمون است. آزمون‌هایی که عملکرد تحصیلی دانشجویان را اندازه می‌گیرند، به انشایی (تشریحی) و عینی (تستی) تقسیم می‌شوند. در خصوص این تصمیم که از کدام یک از این آزمون‌ها برای اندازه‌گیری عملکرد تحصیلی دانشجویان استفاده شود بین اساتید، توافق نظر وجود ندارد. برخی عقیده دارند که آزمون‌های انشایی نشانه بهتری از پیشرفت واقعی دانشجویان فراهم می‌آورند و برخی دیگر به دلیل محدودیت‌هایی که آزمون‌های انشایی دارند استفاده از آزمون‌های عینی به‌ویژه چند گزینه‌ای را پیشنهاد کرده‌اند (۳).

به نظر می‌رسد که نحوه ارزشیابی اساتید از دانشجویان نیز همانند نحوه تدریس آنها با یکدیگر تفاوت داشته باشد. سیاسی در سال ۱۳۸۴ با بررسی شاخص‌های روان‌سنجی سؤالات امتحانی اعضای هیأت علمی دانشگاه شهید چمران نشان داد که اعضای هیأت علمی با توجه به نوع مطلبی که تدریس می‌کنند از سؤالات امتحانی یکسانی استفاده نمی‌کنند. برای مثال اعضای هیأت علمی دانشکده مهندسی بیشتر از سؤالات تشریحی و اعضای هیأت علمی دانشکده اقتصاد و علوم اجتماعی بیشتر از سؤالات عینی چند گزینه‌ای استفاده کرده بودند. در حالی که در دانشکده دام‌پزشکی ۵۰ درصد اساتید از

سؤالات تشریحی و ۵۰ درصد از سؤالات عینی چند گزینه‌ای برای ارزیابی دانشجویان استفاده کرده بودند (۳).

تحقیقات متعددی در خصوص آزمون‌های برگزار شده در دانشگاه‌ها و مؤسسات آموزش عالی داخل و خارج از کشور صورت گرفته است. نتایج این بررسی‌ها نشان داده است که در این آزمون‌ها از سؤالات مناسب کمتر استفاده می‌شود و در اجرا و ساخت آزمون‌ها مسائل و مشکلاتی وجود دارد و بسیاری از مدرسان بدون این‌که آموزش کافی در سنجش و اندازه‌گیری و شیوه ساخت آزمون‌ها کسب کنند، آماده تدریس می‌شوند (۷-۴). قدیمی در سال ۱۳۸۱ در مطالعه‌ای با بررسی میزان اطلاعات اعضای هیأت علمی گروه‌های آموزشی دانشکده‌های مختلف درباره ساخت سؤالات امتحانی گزارش نموده است که اعضای هیأت علمی گروه‌های آموزشی دانشکده علوم تربیتی و روان‌شناسی در مقایسه با سایر گروه‌های آموزشی، اطلاعات وسیع‌تری در ساخت و طراحی سؤالات امتحانی دارند و اطلاعات اعضای هیأت علمی گروه‌های آموزشی دانشکده علوم پایه در این خصوص از سایر دانشکده‌ها کمتر بوده است (۸). عرب در سال ۱۳۷۸ در بررسی سؤالات چند گزینه‌ای در دانشگاه علوم پزشکی همدان نشان داد که ۳۷ درصد سؤالات امتحانی تهیه شده دارای نقص ساختاری بودند (۹). شکونیا و همکاران نیز در یک مطالعه مشابه در دانشگاه علوم پزشکی جندی شاپور در سال ۱۳۸۴ نشان دادند که ۳۴/۸ درصد سؤالات چند گزینه‌ای تهیه شده توسط اساتید، دارای اشکالات ساختاری بودند (۱۰). همچنین نتایج یک بررسی که توسط تارانت (Tarrant) و همکاران در دانشکده پزشکی دانشگاه هنگ‌کنگ در سال ۲۰۰۶ میلادی انجام شد، نشان می‌دهد که حدود نیمی از سؤالات چند گزینه‌ای طراحی شده توسط اعضای هیأت علمی دارای اشکال ساختاری بوده و عمدتاً سطوح پایین شناختی را مورد سنجش قرار داده بودند (۱۱).

یافته‌های بیشتر این مطالعات نشان داده است که اساتید، مهارت کافی در ساخت آزمون‌های درسی ندارند؛ برای

اندازه‌گیری پیشرفت تحصیلی دانشجویان، اغلب مفاهیم ساده و سطحی درس را اندازه‌گیری می‌کنند و در اکثر مواقع سؤالات مبهم هستند (۷). کارشناسان معتقدند سؤالات امتحانی باید به گونه‌ای ساخته شوند که از نتایج آنها بتوان احتمال موفقیت فراگیران را با دقت پیش‌بینی کرد.

بنابراین با توجه به اهمیت ارزشیابی در نظام آموزش عالی و لزوم آگاهی و توانمندی اساتید دانشگاه‌ها در تهیه و طراحی آزمون‌های مناسب، بر آن شدیم تا با بررسی سؤالات امتحانی در آزمون‌های برگزار شده در دانشگاه علوم پزشکی جندی شاپور و مشخص نمودن کاستی‌های موجود، اطلاعات و رهنمودهای لازم جهت رفع اشکالات احتمالی و بهبود شرایط برگزاری آزمون‌های مناسب را فراهم آوریم. پرسش اساسی این است که سؤالات آزمون‌های موجود اساتید تا چه میزان با اصول طراحی و شاخص‌های آزمون‌سازی انطباق دارد؟ هدف عمده این تحقیق بررسی سؤالات امتحانی اعضای هیأت علمی دانشگاه جندی شاپور اهواز با تأکید بر ساختار سؤالات چند گزینه‌ای در سال ۱۳۸۶ از طریق تعیین میزان رعایت اصول آزمون‌سازی بوده است.

روش کار

در این مطالعه توصیفی-مقطعی، جامعه پژوهش کلیه اعضای هیأت علمی غیر بالینی دانشگاه علوم پزشکی جندی شاپور اهواز (۲۳۱ نفر) بودند که در سال تحصیلی ۸۶-۸۵ به امر تدریس و آموزش دانشجویان اشتغال داشتند.

جهت جمع‌آوری نمونه سؤالات امتحانی اعضای هیأت علمی، با ارسال نامه‌ای از طرف مرکز مطالعات و توسعه آموزش پزشکی دانشگاه به اداره آموزش دانشکده‌های مختلف اعلام گردید که در زمان برگزاری امتحانات پایان ترم به اعضای هیأت علمی اطلاع داده شود که قرار است سؤالات امتحانی پایان ترم توسط کارشناسان مرکز مطالعات مورد بررسی قرار گیرد، و لذا از هر عضو هیأت علمی که با انجام این برنامه موافقت دارد یک نمونه سؤال امتحانی دریافت و

به‌طور محرمانه به دفتر مدیریت مطالعات و توسعه آموزش پزشکی ارسال نمایند. در پایان، ۱۹۷ نفر داوطلبانه سؤالات امتحانی پایان ترم خود را در اختیار اداره آموزش قرار دادند. ابتدا مجریان طرح با بررسی منابع موجود (۱۴-۱۲) و مطابق با نظرات متخصصین سنجش و اندازه‌گیری، اصول تهیه و طراحی سؤالات چند گزینه‌ای را در قالب چک‌لیستی مشتمل بر ۱۵ بند تدوین نمودند. در تهیه این مقیاس از نظرات ۳ نفر از اساتید متخصص در زمینه سنجش و اندازه‌گیری استفاده شد. طبقه‌بندی سطوح شناختی (تاکسونومی) سؤالات امتحانی معمولاً بر اساس تقسیم‌بندی بلوم (Bloom) انجام می‌شود. در طبقه‌بندی بلوم اطلاعات شناختی در شش حوزه دانش، درک و فهم، بکارگیری، تجزیه و تحلیل، ترکیب و استنتاج و ارزشیابی تقسیم‌بندی شده است که فقط چهار حوزه اول از این طبقه‌بندی می‌توانند با استفاده از سؤالات چند گزینه‌ای مورد ارزیابی قرار گیرند. برای ساده‌تر کردن دسته‌بندی سؤالات چند گزینه‌ای در سطوح شناختی و نتیجه‌گیری بهتر، با ایجاد تغییراتی در طبقه‌بندی بلوم، سطوح شناختی سؤالات به سه دسته شامل تاکسونومی I (شناسایی و به‌یاد آوردن اطلاعات)، تاکسونومی II (تفسیر و کاربرد اطلاعات) و تاکسونومی III (قدرت تجزیه و تحلیل و توان حل مسئله و تصمیم‌گیری) تقسیم گردید.

تک تک سؤالات چند گزینه‌ای از نظر اشکالات ساختاری و سطوح شناختی توسط چهار نفر از کارشناسان مورد بررسی قرار گرفتند. این کارشناسان قبلاً در دوره‌های آموزشی طرح سؤالات امتحانی شرکت نموده و در حال حاضر به‌عنوان مدرس کارگاه‌های ارزشیابی دانشجو و طرح سؤال در دانشگاه در آموزش طراحی و تهیه سؤالات امتحانی به اعضای هیأت علمی مشارکت دارند. در ابتدا تعداد ۳۰۰ سؤال چند گزینه‌ای توسط هر چهار کارشناس به‌طور جداگانه و مستقلاً مورد ارزیابی قرار گرفتند. سپس به منظور هماهنگ نمودن دیدگاه کارشناسان و ایجاد درک مشترک از مفاهیم مورد مطالعه، در یک نشست مشترک هر ۳۰۰ سؤال از نظر اشکالات ساختاری

است. بدین منظور انواع اشکالات ساختاری و تاکسونومی در کل سؤالات چند گزینه‌ای به صورت مجزا تعیین و فراوانی آنها به طور کلی و صرف نظر از طراح سؤال به صورت درصد فراوانی و میانگین گزارش گردید.

داده‌ها توسط نرم افزار SPSS.V.15 و با استفاده از شاخص‌های مرکزی و پراکندگی و ضریب همبستگی پیرسون تجزیه و تحلیل گردید و $P \leq 0.05$ به عنوان سطح معنی داری در نظر گرفته شد.

یافته‌ها

نتایج نشان داده بیشتر اعضای هیأت علمی مورد مطالعه مرد (۶۱/۴ درصد) و میانگین (انحراف معیار $\pm$) سابقه تدریس آنها ۱۲/۱۳ (۸/۶ $\pm$) بود. ۸/۱ درصد از اعضای هیأت علمی فقط از سؤالات انشایی، ۵۵/۳۲ درصد از سؤالات عینی و ۳۶/۵۴ درصد به صورت ترکیبی از سؤالات انشایی و عینی استفاده کرده بودند.

از مجموع ۶۳۶۵ سؤال مورد بررسی ۹/۸ درصد انشایی و ۹۰/۲ درصد عینی بودند. از ۵۲۰۸ سؤال عینی، ۹۰/۷ درصد چند گزینه‌ای، ۶/۴ درصد صحیح/غلط، ۱/۹ درصد کامل کردنی و ۰/۹۴ درصد جور کردنی بودند.

از شیوه‌های متفاوتی برای ارائه سؤالات امتحانی پایان ترم استفاده شده بود، ۹۱/۴ درصد افراد سؤالات خود را به صورت تایپ شده و ۸/۶ درصد به صورت دست‌نویس در اختیار دانشجویان قرار داده بودند. بیشتر آزمون‌ها فاقد برگه راهنما بودند. (۶۰/۴ درصد).

سؤالات امتحانی چند گزینه‌ای ۵۲/۸ درصد از اعضای هیأت علمی مناسب و مطابق با اصول طراحی سؤالات چند گزینه‌ای تهیه شده و فاقد هر گونه اشکال ساختاری بود. در سؤالات ۲۲/۳۴ درصد از افراد، حداکثر دو اشکال ساختاری و در سؤالات ۲۴/۸۷ درصد اعضای هیأت علمی، بیش از دو اشکال ساختاری وجود داشت.

و سطوح شناختی مرور شده و مورد بحث قرار گرفت. در ادامه کار بقیه سؤالات به طور جداگانه و مستقلاً توسط کارشناسان مورد ارزیابی قرار گرفته و نتایج در جداول مخصوص ثبت می‌شد. هر کارشناس موارد مبهم و بغرنج از سؤالات مورد بررسی را مشخص می‌نمود. و در یک جلسه مشورتی موارد مبهم و بغرنج نیز مطرح و تیم محققین روی اشکالات ساختاری و سطوح شناختی سؤالات توافق می‌نمودند.

در نهایت، اطلاعات مربوط به سؤالات هر عضو هیأت علمی در چک‌لیست بررسی سؤالات امتحانی که حاوی اطلاعات فردی و ویژگی‌های سؤالات امتحانی وی بود ثبت و گردآوری می‌شد. جهت تعیین اشکالات آزمون هر عضو هیأت علمی، تک تک سؤالات امتحانی او با اصول پانزده گانه مورد توافق مقایسه شده و در صورت مغایرت، آن سؤال نامناسب تشخیص داده شده و در لیست سؤالات دارای اشکال ساختاری قرار می‌گرفت. در نهایت نوع اشکالات ساختاری و درصد سؤالات دارای اشکال برای هر عضو هیأت علمی تعیین می‌شد. جهت تعیین تاکسونومی سؤالات نیز دقیقاً به همین نحو عمل می‌شد. البته در مورد تعیین تاکسونومی سؤالات برخی از گروه‌ها که تیم کارشناسی با مشکل مواجه می‌شد و توافق کامل بین کارشناسان وجود نداشت با دعوت از اساتید متخصصین آن رشته و در حضور آنها آن دسته سؤالات مورد بررسی قرار گرفته و در آن مورد تصمیم‌گیری می‌شد.

داده‌های حاصل از بررسی سؤالات امتحانی اعضای هیأت علمی که در چک‌لیست‌های مربوطه ثبت شده بود وارد رایانه شده و در دو بخش مورد ارزیابی قرار می‌گرفت. در بخش اول ویژگی‌های کلی سؤالات امتحانی از قبیل نوع سؤالات، فراوانی اشکالات ساختاری و رعایت یا عدم رعایت شرایط ارائه سؤالات امتحانی فقط به صورت درصد و فراوانی اعضای هیأت علمی گزارش گردیده است و در بخش دوم، ساختار سؤالات چند گزینه‌ای و میزان رعایت اصول آزمون‌سازی در این گونه سؤالات به عنوان متغیر اصلی مورد ارزیابی قرار گرفته

۹۱/۹ درصد از اعضای هیأت علمی جهت ارزیابی دانشجویان صرفاً از سؤالات چند گزینه‌ای و یا به صورت ترکیبی از سؤالات چند گزینه‌ای و سایر انواع سؤالات استفاده کرده بودند. تعداد سؤالات چند گزینه‌ای استفاده شده توسط این اساتید در مجموع ۵۲۰۸ سؤال بود، ۶۴/۹۶ درصد این سؤالات فاقد اشکال ساختاری و ۳۵/۰۴ درصد آنها دارای اشکالات ساختاری در ساقه و گزینه‌ها، بودند. بیشترین فراوانی اشکالات ساختاری در سؤالات مربوط به مشخص نبودن و یا عدم وضوح هدف در ساقه سؤال (۱۵/۹۴ درصد)، استفاده از افعال منفی در ساقه سؤال (۱۲/۲۲ درصد)، استفاده از عبارت‌های "همه موارد" یا "هیچ کدام" در گزینه‌ها (۱۰/۸۵ درصد) و به کار بردن جای خالی در اول یا وسط ساقه سؤالات (۱۰/۴۱ درصد) بود (جدول ۱).

از نظر تاکسونومی به ترتیب ۸۹/۴، ۸/۴ و ۲/۲ درصد در سطح تاکسونومی I، II و III طراحی شده بود. جهت بررسی رابطه بین تاکسونومی سؤالات با اشکالات ساختاری به تفکیک در سه سطح تاکسونومی I، II و III از ضریب همبستگی پیرسون استفاده شد. ضریب همبستگی به جز در دو مورد (موارد ۱۲ و ۱۳ در جدول ۲) در سؤالات با تاکسونومی I، بالا و معنی‌دار بود. با توجه به مثبت بودن ضرایب همبستگی می‌توان نتیجه گرفت که سؤالاتی که در تاکسونومی I قرار گرفته‌اند اشکالات ساختاری بیشتری داشته‌اند.

جدول ۱: میانگین، انحراف معیار و فراوانی اشکالات ساختاری سؤالات چند گزینه‌ای

اشکالات ساختاری	تعداد (درصد)	انحراف معیار $\pm$ میانگین
۱- مشخص نبودن و یا عدم وضوح هدف در ساقه	۲۹۱ (۱۵/۹۴)	$۳/۹۰ \pm ۱/۶۴$
۲- استفاده از افعال منفی در ساقه سؤال	۲۲۳ (۱۲/۲۲)	$۲/۲۰ \pm ۱/۲۶$
۳- استفاده از عبارت همه موارد یا هیچ کدام	۱۹۸ (۱۰/۸۵)	$۲/۰۶ \pm ۱/۱۲$
۴- بکار بردن جای خالی در اول یا وسط ساقه	۱۹۰ (۱۰/۴۱)	$۲/۹۹ \pm ۱/۰۷$
۵- وجود غلط املایی در ساقه یا گزینه‌ها	۱۶۹ (۹/۲۶)	$۱/۸۴ \pm ۰/۹۵$
۶- وجود اشکالات انشایی در ساقه یا گزینه‌ها	۱۶۹ (۹/۲۶)	$۲/۱۵ \pm ۰/۹۵$
۷- تکرار یک عبارت در همه گزینه‌ها	۱۲۸ (۷/۰۱)	$۱/۵۳ \pm ۰/۷۲$
۸- عدم رعایت طول گزینه‌ها	۱۱۹ (۶/۵۲)	$۱/۷۰ \pm ۰/۶۷$
۹- بکار بردن بیش از یک هدف در ساقه	۹۷ (۵/۳۱)	$۱/۳۵ \pm ۰/۵۵$
۱۰- استفاده از عبارت الف و ب یا ب و ج	۹۲ (۵/۰۴)	$۱/۱۲ \pm ۰/۵۲$
۱۱- عدم رعایت نکات دستوری در ساقه و گزینه‌ها	۶۱ (۳/۳۴)	$۲/۳۰ \pm ۰/۳۴$
۱۲- استفاده از علائم اختصاری در ساقه	۳۶ (۱/۹۷)	$۰/۷۵ \pm ۰/۲۰$
۱۳- استفاده از واژه‌های مشابه در ساقه و گزینه‌ها	۲۱ (۱/۱۵)	$۰/۷۱ \pm ۰/۱۲$
۱۴- وجود توضیحات اضافی و غیر ضروری در ساقه	۱۹ (۱/۰۴)	$۰/۴۳ \pm ۰/۱۱$
۱۵- استفاده از قیود مطلق مثل فقط هرگز همیشه	۱۲ (۰/۶۵)	$۰/۲۹ \pm ۰/۰۷$
کل	۱۸۲۵ (۱۰۰)	

جدول ۲: همبستگی بین درصد اشکالات ساختاری با درصد سطوح شناختی (تاکسونومی) سؤالات چند گزینه‌ای

عناوین	تاکسونومی I	تاکسونومی II	تاکسونومی III
۱- مشخص نبودن و یا عدم وضوح هدف در ساقه	* ۰/۱۶	-۰/۰۶	-۰/۰۵
۲- استفاده از افعال منفی در ساقه سؤال	* ۰/۱۶	۰/۰۴	۰/۰۵
۳- استفاده از عبارت همه موارد یا هیچکدام	** ۰/۱۸	-۰/۰۱	-۰/۰۴
۴- بکار بردن جای خالی در اول یا وسط ساقه	** ۰/۲۳	۰/۱	-۰/۰۳
۵- وجود غلط املائی در ساقه یا گزینه‌ها	** ۰/۳۴	۰/۰۲	۰/۰۱
۶- وجود اشکالات دستوری در ساقه یا گزینه‌ها	** ۰/۲۶	-۰/۰۹	-۰/۰۳
۷- تکرار یک عبارت در همه گزینه‌ها	** ۰/۳۲	۰/۱	-۰/۱۳ *
۸- عدم رعایت طول گزینه‌ها	** ۰/۲۹	۰/۱۲	-۰/۱۱
۹- بکار بردن بیش از یک هدف در ساقه	** ۰/۱۹	-۰/۰۷	-۰/۰۸
۱۰- استفاده از عبارت الف و ب یا ب و ج	* ۰/۱۶	-۰/۰۰۹	-۰/۰۰۵
۱۱- عدم رعایت نکات گرامری در ساقه و گزینه‌ها	** ۰/۲۲	-۰/۰۵	-۰/۰۳
۱۲- استفاده از علائم اختصاری در ساقه	۰/۰۸	-۰/۰۱	-۰/۰۰۱
۱۳- استفاده از واژه‌های مشابه در ساقه و گزینه‌ها	۰/۰۲	-۰/۰۵	-۰/۰۴
۱۴- وجود توضیحات اضافی و غیر ضروری در ساقه	** ۰/۲۰	* -۰/۱۵	-۰/۰۳
۱۵- استفاده از قیود مطلق مثل فقط هرگز همیشه	** ۰/۲۳	۰/۰۱	-۰/۰۴

* و ** به ترتیب در سطح ۰/۰۱ و ۰/۰۰۱ معنی‌دار است.

بحث و نتیجه‌گیری

بررسی نوع سؤالات امتحانی در این مطالعه نشان داد که اعضای هیأت علمی با توجه به موضوعی که تدریس می‌کنند از آزمون‌های یکسانی استفاده نمی‌کنند. اما در هر حال بیش از دو سوم آنها از سؤالات عینی و عمدتاً چند گزینه‌ای استفاده کرده بودند. آزمون‌های چند گزینه‌ای که به آزمون‌های تستی شهرت دارند تنها نوع آزمون عینی هستند که اخیراً به دلیل محاسبه ویژگی‌های روان‌سنجی آن مورد توجه متخصصان آزمون‌سازی قرار گرفته‌اند. این نوع آزمون‌ها اگر با دقت و مهارت کافی ساخته شوند برای اندازه‌گیری تمام اهداف درس در تمام سطوح فکری و در تمام رشته‌های تحصیلی مناسب هستند. اکثر آزمون‌های استاندارد که در سطح بین‌المللی از جمله تافل که برای گزینش دانشجو در سطوح مختلف تحصیلی انجام می‌پذیرد، از نوع چند گزینه‌ای می‌باشند (۱۶-۱۵).

یافته‌های پژوهش بیانگر آن است که عملکرد بیش از یک سوم اعضای هیأت علمی دانشگاه در ساخت آزمون‌های چند گزینه‌ای خالی از اشکال نبوده است که با نتایج مطالعات مشابه در همدان، اهواز و یزد مطابقت دارد (۹،۱۰،۱۷). نتایج پژوهشی در دانشگاه علوم پزشکی ایران نیز میزان رعایت قواعد ساختاری سؤالات آزمون جامع علوم پایه پزشکی را در دوره‌های اول و دوم به ترتیب ۴۹/۱۹ و ۳۹/۵۲ درصد گزارش نموده است (۱۸). نتایج یک مطالعه مشابه در سال ۲۰۰۵ در دانشگاه هنگ‌کنگ نیز نشان داده است که ۴۶/۲ درصد سؤالات چند گزینه‌ای دارای اشکالات ساختاری بوده‌اند (۱۱).

شواهد موجود نشان می‌دهد عدم آشنایی کامل اعضای هیأت علمی با اصول آزمون‌سازی از علل اصلی آمار بالای اشکالات ساختاری در سؤالات امتحانی می‌باشد. یافته‌های یک مطالعه که با هدف بررسی توانایی‌های اعضای هیأت علمی در زمینه ارزشیابی آموزشی در دانشگاه همدان انجام شده نشان

می‌دهد که ۴۸/۸ درصد اعضای هیأت علمی در طراحی و تهیه سؤالات امتحانی دارای آگاهی کافی نمی‌باشند (۱۹). نتایج مطالعه دیگری در دانشگاه علوم پزشکی مشهد نیز نشان داد که روش‌های ارزشیابی موجود نه تنها در سنجش مهارت‌های بالینی، بلکه در سنجش دانش و محفوظات نیز از کفایت لازم برخوردار نبوده‌اند (۲۰). نتایج فوق لزوم تدارک دوره‌های آموزشی مناسب در زمینه طراحی سؤالات امتحانی برای اعضای هیأت علمی را ضروری می‌سازد.

بر اساس یافته‌های به‌دست آمده، بیشترین مشکل سؤالات مورد بررسی، ابهام یا عدم وضوح در تبیین هدف و موضوع سؤال در ساقه سؤالات چند گزینه‌ای بود. در ۱۵/۹۴ درصد از سؤالات یا هدف سؤال مشخص نشده و یا هدف به‌طور واضح و شفاف تبیین نگردیده بود. نتایج مطالعات مشابه در ایران و سایر کشورها نیز بر این دلالت دارد که عدم تبیین کامل هدف سؤال در ساقه، مهم‌ترین نارسایی در تهیه سؤالات چند گزینه‌ای بوده است. این اشکال در مطالعات مشابه نیز با کسب بالاترین درصد به‌عنوان بارزترین علت نقص ساختاری سؤالات چند گزینه‌ای گزارش شده است (۹، ۱۱، ۱۸).

علی‌رغم توصیه کارشناسان مبنی بر عدم استفاده از افعال منفی در ساقه سؤالات چند گزینه‌ای (۱۴-۱۲) در ۱۲/۲۲ درصد از سؤالات مورد بررسی در این مطالعه از افعال منفی استفاده شده بود. استفاده از افعال منفی در نتایج مطالعات انجام شده در دانشگاه‌های علوم پزشکی همدان، اهواز و یزد نیز نسبتاً بالا گزارش گردیده است (۲۲، ۲۱، ۹). این اشکال در مطالعه انجام گرفته در دانشگاه هنگ‌کنگ (۶/۹ درصد) به‌عنوان دومین نقص ساختاری گزارش گردیده است. استفاده از افعال منفی در ساخت سؤالات چند گزینه‌ای، ممکن است به‌خاطر سادگی طراحی این گونه سؤالات باشد. اگرچه نتایج تحقیقات نشان داده است که وجود افعال منفی در ساقه سؤالات باعث گیجی و سردرگمی دانشجو می‌شود، با این وجود به‌دلیل سهولت تهیه، این نوع سؤالات به کرات مورد استفاده قرار می‌گیرند.

نتایج دیگر این مطالعه نشان داد که عبارات "همه موارد" یا "هیچ‌کدام" در ۱۰/۸۵ درصد از سؤالات مورد استفاده قرار گرفته است. این مورد در دو مطالعه در دانشگاه‌های علوم پزشکی همدان و اهواز به‌ترتیب ۱۸/۳ و ۱۵/۷ درصد گزارش گردیده است (۹، ۱۰). اصولاً استفاده از این واژه‌ها در گزینه‌ها حاکی از عدم توانایی طراح سؤال در تهیه گزینه‌های انحرافی مناسب است و یا ممکن است به‌دلیل استفاده از مضامینی باشد که فقط سه گزینه مرتبط برای آن مضامین قابل طرح است. بنابراین طراح برای فرار از کاوش بیشتر و یا کوشش برای تغییر مضمون سؤال، به عبارات‌های "همه موارد" یا "هیچ‌کدام" متوسل می‌شود. بکارگیری این واژه‌ها در گزینه‌ها باعث افت کیفیت سؤال شده و طبعاً این قبیل سؤالات نمی‌توانند به درستی آموخته‌های دانشجویان را مورد سنجش قرار دهند.

یکی از موارد دیگری که به‌عنوان اشکال ساختاری در ۱۰/۴۱ درصد از سؤالات وجود داشت، بکار بردن جای خالی در اول یا وسط ساقه سؤال می‌باشد. مطابق با نظر متخصصان سنجش و اندازه‌گیری، ساقه هر سؤال چند گزینه‌ای باید یک هدف آموزشی را به‌طور واضح و با یک جمله کامل مورد پرسش قرار دهد و اگر از جمله کامل استفاده نمی‌شود بایستی در انتهای عبارت فضای خالی آورده شود. بنابراین استفاده از فضای خالی در ساقه سؤالات چند گزینه‌ای توصیه نمی‌شود و به‌عنوان نقص تلقی شده و به‌خاطر این‌که باعث صرف انرژی زیادتری از دانشجو برای فهم محتوای سؤال و پاسخ‌دهی به آن می‌شود، از کیفیت سؤال می‌کاهد.

۹/۲۶ درصد غلط املائی و همین تعداد اشکال دستوری در ساقه و یا گزینه‌های سؤالات چند گزینه‌ای مورد بررسی نشان‌دهنده عدم دقت و توجه کافی در ارائه سؤالات امتحانی می‌باشد. عدم بازیابی و تصحیح سؤالات ماشین شده، احتمالاً یکی از علل بروز این اشکال می‌باشد. این مطلب را سپاسی (۱۳۸۴) در مطالعه خود در بررسی سؤالات امتحانی دانشگاه شهید چمران مورد تأکید قرار داده است (۲). نتایج یک مطالعه

اگر چه تاکنون نسبت خاصی برای سطوح مختلف شناختی سؤالات چند گزینه‌ای یک آزمون، پیشنهاد نشده است، اما انتظار می‌رود در یک آزمون، سؤالات امتحانی بتوانند جنبه‌های مختلف توانایی دانشجویان را مورد سنجش قرار داده و قابلیت‌های کسب شده‌ای که برای آن تربیت شده‌اند را به‌خوبی نمایان سازد (۳۰).

ضریب همبستگی مثبت بین سطوح شناختی (تاکسونومی) سؤالات چند گزینه‌ای با اشکالات ساختاری موجود در این سؤالات نشان می‌دهد که اشکالات ساختاری در سؤالاتی که در سطوح پایین شناختی (تاکسونومی I) طراحی شده‌اند بیشتر است. یعنی سؤالاتی که در سطوح بالای شناختی (تاکسونومی II و III) طراحی شده‌اند اشکالات ساختاری کمتری داشته‌اند. این یافته با نتایج برخی از مطالعات همخوانی دارد. تارانت نشان داده است که اشکالات ساختاری در سؤالات مطرح شده در سطوح بالای حیطه شناختی کمتر است. او گزارش نموده است که اشکالات ساختاری در سؤالات با سطوح شناختی پایین، ۴۸/۹ درصد در حالی که در سؤالات با سطوح شناختی بالا تنها ۱۸/۶ درصد بوده است (۱۱). این وضعیت احتمالاً به این دلیل است که افرادی که سؤالات چند گزینه‌ای با تاکسونومی بالا طراحی می‌کنند از آگاهی و مهارت بیشتری برخوردار می‌باشند.

بیش از ۶۰ درصد اعضای هیأت علمی برگه راهنمای آزمون در اختیار دانشجویان قرار نداده بودند. این یافته با نتایج مطالعه سیاسی در دانشگاه شهید چمران اهواز که رقم ۶۷ درصد را گزارش نموده، کاملاً همخوانی دارد (۲).

نتایج این مطالعه همچون مطالعات مشابه دیگر (۱۶-۱۹) نشان می‌دهد سؤالات چند گزینه‌ای عمدتاً در سطوح شناختی پایین طراحی شده و اشکالات ساختاری قابل ملاحظه‌ای دارند. با این وجود شواهد موجود حاکی از آن است که آموزش اعضای هیأت علمی و دادن بازخورد مناسب به آنان باعث ارتقاء کیفیت سؤالات چند گزینه‌ای و بهبود شرایط ارزشیابی دانشجویان شده است (۲۲، ۲۳، ۳۱).

در دانشگاه علوم پزشکی اصفهان نشان داد که بازخورد مناسب به طراحان سؤالات امتحانی منجر به کاهش ۷۰ درصدی اشتباهات املائی و دستوری گردیده است (۲۳).

عدم رعایت طول گزینه‌ها (۶/۵۲ درصد)، عدم رعایت نکات دستوری در ساقه و گزینه‌ها (۳/۳۴ درصد) و استفاده از الفاظ مشابه در ساقه و گزینه‌ها (۱/۱۵ درصد) از اشکالاتی است که در این مطالعه با فراوانی کمتری مشاهده شده و در مطالعات مشابه نیز از درصد بالایی برخوردار نبوده است (۲۴). جوزفویکز (Jozefowicz) علت وجود اشکالات زیاد در سؤالات امتحانی اساتید را عدم صرف وقت کافی در این ارتباط می‌داند. او می‌گوید اساتید وقت زیادی را برای آماده کردن مطالب درسی و رفتن به کلاس اختصاص می‌دهند اما برای طرح سؤالات امتحانی خود وقت زیادی صرف نمی‌کنند و اطلاعات آنها از اصول ارزشیابی و تکنیک‌های طرح سؤالات امتحانی، ناکافی است (۲۵).

علی‌رغم اهمیت طرح سؤال از سطوح مختلف حیطه شناختی، واقعیت موجود حاکی است که اکثر سؤالات امتحانی سطوح پایین حیطه شناختی را ارزشیابی می‌کنند (۴). یافته‌های این پژوهش نشان می‌دهد که حدود ۹۰ درصد سؤالات امتحانی، سطوح پایین طبقه‌بندی بلوم (تاکسونومی I) را ارزشیابی می‌کنند و تنها ۱۰/۶ درصد سؤالات در سطوح بالاتر یعنی تفسیر و حل مسئله (تاکسونومی II و III) می‌باشند. این یافته با نتایج مطالعات زیادی مطابقت دارد. نتایج مطالعات انجام شده در اصفهان و همدان و مازندران نشان داد که بیشتر سؤالات در سطح تاکسونومی I بوده‌اند (۹، ۲۶، ۲۷)، در حالی که نتایج یک مطالعه در دانشکده پرستاری و مامایی گرگان میزان سؤالات طراحی شده در سطح پایین (تاکسونومی I) را ۵۷/۷ درصد گزارش نموده است (۲۸). همچنین نتایج دو مطالعه در دانشگاه‌های هنگ‌کنگ و استرالیا نشان داد که به‌ترتیب ۹۱/۱ درصد و ۹۱/۲ درصد سؤالات در سطوح پایین حیطه شناختی (تاکسونومی I) بوده است (۲۹، ۱۱).

در اصلاح سؤالات امتحانی و بهبود شرایط موجود مؤثر باشد.

بررسی سؤالات امتحانی اعضای هیأت علمی در پایان ترم و

دادن بازخورد مناسب به آنان، از راهکارهایی است که می‌تواند

References

1. Guilbert JJ. Educational handbook for health personnel. 6th ed. Geneva; World Health Organization: 1987.
2. Sepasi H, Attari YA. The study of psychometric characteristics of Shahid Chamran University Faculty members final test scores. Journal of Education and Psychology (Special Issue an Education) 2006; 3(4): 1-20. [In Persian]
3. Sepasi H. A comparison between multiple-choice and True-False test item in measuring Shahid Chamran University undergraduate students' academic achievement in measurement and evaluation course with the control of students' absences. Journal of Education and Psychology 1998; 5(1&2): 31-56. [In Persian]
4. Moatamedi A, Seif A. Survey on evaluation methods of faculties of evaluation method course in Training Teacher Center and related factors. Master's dissertations. Allameh Tabatabaei University 1994. [In Persian]
5. Kaveh Tabatabaei MS, Bahreini Tousi MH, Modabber Azizi MJ, Ebrahimzadeh S, Bahreini Tousi V, Bahreini Tousi K. Analyses survey on multiple choice question in medical school in Mashhad University of medical sciences in 2002. Medical Journal of Mashhad University of Medical Sciences 2003; 45 (76): 89-95. [In Persian]
6. Holsgrave G, Elzubeir M. Imprecise terms in UK medical multiple-choice questions: What examiners think they mean. Med Educ 1998; 32 (4): 343-50.
7. Sepasi H. A study of cognitive domain and analysis of psychometric characteristics of test terms in final examinations of Arabic, calculus and biology courses of the third grade high school girls in different Socio-economic status in Khuzestan, Iran. Journal of Education and Psychology 2006; 13(4): 57-78. [In Persian]
8. Ghadimi Moghaddam MM. Survey on special knowledge Faculties of Tehran Azad University (9 Area). Master theses. Azad University of Tehran; 2004. [In Persian]
9. Arab M. Survey on defect of multiple choice questions in medical school of Hamadan Medical University in 2000. 4th National Congress of Medical Education. Tehran: University of Medical Sciences; 2001: 51. [In Persian]
10. Shakurnia A, Khosravi A, Shariati A, Zareei A. Survey on multiple choice questions of academic members in Ahvaz Jundishapur University of Medical Sciences in 2005. Abstract of the 8th National Congress of Medical Education. Kerman University of Medical Sciences; 2007: 44. [In Persian]
11. Tarrant M, Knierim A, Hayes SK, Ware J. The frequency of item writing flaws in multiple-choice questions used in high stakes nursing assessments. Nur Educ Today 2006; 26 (8): 662-71.
12. Seif A. Assessment methods and educational evaluation. Tehran; Nashre Douran: 1999. [In Persian]
13. Rezaei H, Shoughia Shafagh Aria F, Doulati Y, Mohagheghi MA, Malekzadeh R. Evaluation techniques in medical education. Tehran: Ammeh Publication; 1999. [In Persian]
14. Zolfaghari B, Adibi N, Darakhshanfar S, Tansaz M, Karbasi A, Niroumand P. Academic achievement tests in medical sciences. Isfahan: Medical University Educational Development center; 2001. [In Persian]
15. Mehrens AW, Lehman JI. Measurement and evaluation in education and psychology. New York: Holt, Rinehart and Winston Inc; 1984.
16. Kiyamanesh A. Multiple choice questions and its effect on student achievements', content and teaching methods. Taalim va Tarbiat 1995; 8 (3): 57-77. [In Persian]
17. Vaezi AA, Morovvati MA, Heidari MR. Obeying pertinent rules for making multiple choice questions in final exams designed by

- faculty members. Abstract of the 9th National congress of Medical Education. Yazd: University of Medical Sciences; 2008: 250. [In Persian]
18. Khoshkam M, Shajari Zh. Trend in Processing of structural designing general science exam in the study in 1999 in Tehran. Research in Medical Sciences 1997; 3(Supple. 1): 160. [In Persian]
 19. Allah Verdi-pour H, Bashirian S. Knowledge and attitude of faculty members about evaluation of educational progress. Scientific Journal of Hamadan University of Medical Sciences 2000; 7 (4): 20-4. [In Persian]
 20. Kaveh Tabatabaei MS, Bahreini Tousi MH, Modabber Azizi MJ, Ebrahimzadeh S, Bahreini Tousi V, Bahreini Tousi K. Evaluation methods in clinical ward and medical student's views' in 2002. Medical Journal of Mashhad University of Medical sciences. 2003; 45 (76): 97-103. [In Persian]
 21. Latifi M, Shakurnia A. Comparing faculty members and experts' points of view about questions designed in Ahvaz University of Medical Sciences. Abstract of the 8th National Congress of Medical Education. Kerman University of Medical Sciences; 2007: 159. [In Persian]
 22. Mohammadi Zeidi A, Heydarnia A, Khalaj M. Comparing test preparation methods in the school of public health before and after test standardization workshop. Abstract of the 9th National Congress of Medical Education. Yazd University of Medical Sciences; 2008: 146. [In Persian]
 23. Gorg Yaraghi M, Nazem M, Ebrahimi A, Avizhgan M, Mazaheri M, Esmaeili A. The effect of an intervention on test preparation in surgery ward, 2007. Abstract of the 9th National Congress of Medical Education. Yazd University of Medical Sciences; 2008: 25. [In Persian]
 24. Dorripour F, Karimi S, Pakaein J. Evaluation of multiple choice questions in hygiene course of nursing students, 2006. Abstract of the 8th National Congress of Medical Education. Kerman University of Medical Sciences; 2007: 67. [In Persian]
 25. Jozefowicz RF, Koeppen BM, Case S, Galbraith R, Swanson D, Glew RH. The quality of in-house medical school examinations. Acad Med 2002; 77 (2): 156-61.
 26. Gorg Yaraghi M, Ebrahimi A, Ashirioun V, Azadbakht S, Avizhgan M, Esmaeili A, et al. An analysis of multiple choice questions for undergraduate medical students and designing necessary steps to correct students' evaluation process. Abstract of the 9th National Congress of Medical Education. Yazd University of Medical Sciences; 2008: 27. [In Persian]
 27. Vahidshahi K, Mahmoudi M, Shahbaz Nezhad L, Shakeri S. The quality of multiple choice questions in the first semester, 2006, School of Medicine, Mazandaran University of Medical Sciences. Abstract of the 9th National Congress on Medical Education. Yazd: Yazd University of Medical Science; 2008: 41. [In Persian]
 28. Khoddam H, Jouybari L, Sanagoo A. Nursing clinical instructors' skills in designing question. IJME 2005; 5 (2): 191-2. [In Persian]
 29. Sellappah S, Hussey T, Blackmore AM, McMurray A. The use of questioning strategies by clinical teacher. J Adv Nurs 1998; 28 (1): 142-8.
 30. 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.
 31. Fattahi Bafghi A, Karimi H. The effect of education on academic members' designing multiple choice questions regarding face validity and content validity. Abstract of the 7th National Congress of Medical Education. Tabriz University of Medical Sciences; 2005: 63. [In Persian]

An Evaluation of Exam Questions Designed by Faculty Members, Emphasizing on the Multiple Choice Question Structure, Ahvaz Jundishapur University of Medical Sciences, 2007

Abdolhussein Shakurnia^{1*}, Azar Khosravi Borojeni², Alireza Mozaffari³, Houssein Elhampour⁴

1. M.Sc. in Immunology, Instructor of Immunology Dept., School of Medicine, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran

2. Ph.D. in Microbiology, Associate Professor of Microbiology Dept., School of Medicine, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran

3. Specialist in Internal Medicine, Associate Professor of Internal Medicine Dept., School of Medicine, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran

4. M.Sc. in Educational Sciences, Instructor of Educational Sciences Dept., School of Educational Sciences and Psychology, Shahid Chamran University, Ahvaz, Iran

• Received: 29 Jun, 2009 • Accepted: 7 Nov, 2009

Background & Objective: Evaluation of educational achievement is an important part of an educational process; such an evaluation is possible via holding exams. Multiple Choice Question (MCQ), as an objective testing method, is the most common tool for evaluating students. Hence, this study was conducted to evaluate the exam questions designed by faculty members with emphasis on the MCQ structure in Ahvaz Jundishapur University of Medical Sciences.

Methods: In this descriptive study, questions designed by 197 non-clinician faculty members were analyzed. Data was collected using a checklist including demographic information, properties of exam questions, and items related to comparing MCQs with design rules of such questions.

Results: The results revealed that 8.1% of the questions developed by faculty members were essay, 55.32% and 36.54% were objective and a combination of essay and objective, respectively. Among a total of 5208 evaluated questions, 64.96% contained no structural flaw. The most common flaws included ambiguous or unclear question stems (15.94%), use of negative verb in the stem (12.22%), using "all" or "none" among option (10.85%), and use of blank space in the beginning or middle of the stem (10.41%). Almost 90% of the MCQs were at cognitive level (taxonomy I).

Conclusion: Due to the importance of developing standard questions corresponding to guidelines for MCQs in acceptable exams, regular revision of exam questions and further training courses for designing questions is recommended.

Key words: Exam, Faculty members, MCQ

*Correspondence: Immunology Dept., School of Medicine, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran

• Tel: 0611-3330074

• Fax: 0611-3332036

• Email: shakurnia@yahoo.com