

ارزشیابی به وسیله آزمون‌های چندگزینه‌ای

محسن ملاهادی^{*} *BSc*

^{*} گروه اطفال، دانشکده پرستاری، دانشگاه علوم پزشکی بقیه‌ا... (۶۰)، تهران، ایران

چکیده

مقدمه. ارزشیابی، فرآیند نظاممند گردآوری، تحلیل و تفسیر داده‌ها به منظور ارزش‌گذاری و قضاوت است. استیضاح و معلمان باید از میان روش‌های مختلف ارزشیابی، متناسب با اهداف آموزشی خود، بهترین روش را برگزینند. سئوالات چندگزینه‌ای، از رایج‌ترین آزمون‌ها در علوم پزشکی هستند. این آزمون‌ها به علت عینیت کامل، از حداکثر پایابی ممکن برخوردارند، اما روایی آنها هنوز مورد توجه صاحب‌نظران بحث ارزشیابی قرار دارد. با این که آزمون‌های چندگزینه‌ای به صورت گستره بهخصوص در پرستاری و پزشکی مورد استفاده قرار می‌گیرند، اما مطالعات در این زمینه حاکی از ایرادهای اساسی در طراحی این نوع سئوالات است.

نتیجه‌گیری. با توجه به آنکه آزمون‌های چندگزینه‌ای از نظر پایابی و عینی بودن، که از چالش‌انگیزترین جنبه‌های آزمون است، از کفایت بالایی برخوردارند، با تلاش در زمینه ارتقای جنبه‌های روایی محتوا و ساختار، به یکی از مطلوب‌ترین شیوه‌های ارزشیابی دانش فرآگیران تبدیل می‌شوند. بنابراین، تجدیدنظر و مطالعه بیشتر استیضاح در رابطه با قواعد طراحی سئوالات چندگزینه‌ای امری الزامی به نظر می‌رسد.

کلیدواژه‌ها: ارزشیابی، سئوالات چندگزینه‌ای، پایابی، روایی

Evaluating by multichoice tests

Mollahadi M.* *BSc*

*Department of Pediatrics, Faculty of Nursing, Baqiyatallah University of Medical Sciences, Tehran, Iran

Abstract

Introduction. Evaluation is a systematically gathering, analysis and interpretation of data process to valuation and judge. Teachers should select the best method between the different methods of evaluation, according to their educational objectives. Exams and multiple-choice questions are one of the most common tests in medical sciences. These tests due to complete identity have the maximum reliability, but their validity is still considered in evaluation by experts. Although the multiple-choice tests are widely used especially in the nursing and medicine, however, studies in this area are indicated fundamental defects in designing this type of questions.

Conclusion. Considering that multiple-choice questions have high quality of reliability and objectivity, which are the most controversial aspects of testing, with efforts in promoting the main aspects of content validity and structure, turn into one of the most desirable methods to assess students' knowledge. Therefore, revision and studying more about the designing rules of multiple-choice questions by teachers is necessary.

Keywords: Evaluation, Multiple Choice Question, Reliability, Validity

مقدمه

در مورد سوالات‌های چندگزینه‌ای تحقیقات مختلفی انجام شده و نتایج مختلفی هم بدست آمده است که در این مقاله خواهد آمد، اما در این میان مطلب مهم این است که، معلمان این آمادگی را داشته باشند که بهسادگی و به بهترین نحو ممکن، آزمون‌های چندگزینه‌ای را که برای اندازه‌گیری و ارزشیابی شرایط آموزشی مناسب است، انتخاب و اجرا نمایند و قادر به تعبیر و تفسیر نتایج حاصل از آن باشند [۷]. با توجه به آنچه گفته شد، هدف از نگارش مقاله حاضر، بررسی چگونگی طراحی آزمون‌های چندگزینه‌ای و اشتباهات رایج در طراحی این نوع سوالات و محاسن و معایب آنها بود.

آزمون‌های چندگزینه‌ای

یکی از روش‌های آزمون دانش، سوالات چندگزینه‌ای است که اولین بار در سال ۱۹۱۴ توسط فردریک در دانشگاه کانتراس ابداع گردید [۱]. اکنون سوالات چندگزینه‌ای از شایع‌ترین نوع آزمون‌ها در دنیا، بهویژه در ایران هستند. در این نوع سوالات، مجموعه‌های از جواب‌ها که برای سوال مطرح شده و بالقوه درست به‌نظر می‌رسند، به‌شکل ستاریو در اختیار فراغیر قرار می‌گیرد و فراغیر از میان آنها جواب درست را انتخاب می‌کند. این نوع سوالات چندگزینه‌ای علاوه بر سنجش دانش، توانایی حل مسئله را نیز می‌سنجند [۸]. دو شرط اصلی سوال چندگزینه‌ای، (الف) پرسش درباره موضوع مهم؛ و (ب) طراحی خوب است بهنحوی که در آن از خطاهایی که موجب می‌شود آزمون دهنده بدون داشتن اطلاعات علمی موردنظر و صرفاً به‌واسطه آشنایی با فنون تست‌زنی، بتواند پاسخ دهد، پرهیز شده باشد. آزمونی می‌تواند این خصوصیات را داشته باشد که از روایی و پایایی لازم برخوردار باشد [۳]. آزمون‌های چندگزینه‌ای به‌علت عینیت کامل، از پایایی بالا (حداکثر پایایی ممکن) برخوردار هستند، اما روایی آنها از مسائل مورد توجه در مطالعات و پژوهش‌های آموزشی است، البته این مسئله را نیز می‌توان با رعایت اصول طراحی ساختار سوالات چندگزینه‌ای و اصول طراحی سوالات چندگزینه‌ای، براساس سطوح طبقه‌بندی شناختی، کاهش داد [۱]. لذا لازم است سوالات چندگزینه‌ای براساس یکی از طبقه‌بندی‌های اهداف شناختی طراحی شوند [۹]؛ مثلاً طبقه‌بندی بلوم، بلوم در سال ۱۹۵۹، طبقه‌بندی شامل سطوح دریافت و ادراک، کاربرد، تجزیه و تحلیل و ارزشیابی بود. بهدلیل آن اساتید و ارزیابان این طبقه‌بندی را پذیرفته و در نهایت آن را در ۳ طبقه جای دادند: ۱) سطح دانش، شامل یادآوری مطلب و اطلاعات خاص است؛ ۲) سطح ادراک و کاربرد، شامل فهم و توانایی تشریح اطلاعات قبلی و استفاده از اطلاعات جدید، قوانین، روش‌ها، مفاهیم و تئوری‌ها است؛ و ۳) سطح حل مسئله، شامل انتقال و تبدیل دانش و مهارت‌ها به موقعیت‌های جدید است [۱۰].

گاهی بعضی اهداف شناختی به‌وسیله سوالات چندگزینه‌ای قابل اندازه‌گیری نیست مثلاً در پرستاری از طبقه‌بندی بلوم فقط ۴ سطح

ارزیابی یا ارزشیابی، عبارت از فرآیند نظامند گردآوری، تحلیل و تفسیر داده‌ها به‌منظور ارزش‌گذاری و قضایت است. اگر این فرآیند در مورد افراد، گروه‌ها و سازمان‌ها انجام گردد، عموماً به عنوان ارزیابی و اگر در مورد برنامه انجام شود، به آن ارزشیابی اطلاق می‌گردد. آزمون‌ها از نوع رایج ارزیابی هستند که در سیستم آموزشی صورت می‌گیرند [۱]. ارزشیابی آموزشی شامل جمع‌آوری و تجزیه و تحلیل اطلاعات برنامه آموزشی برای بررسی، تصمیم‌گیری و مشخص کردن موقفيت یا عدم موقفيت در اهداف مشخص شده برای آن برنامه آموزشی است [۲]. هدف از امتحان، می‌تواند نشان دادن اهمیت مباحثت به فراغیران، ترغیب فراغیران به مطالعه، شناسایی موارد ضعیف در یادگیری و اصلاح آنها، تصمیم‌گیری درباره ارتقای فراغیر و شناسایی موارد ضعف دوره یا برنامه آموزشی باشد.

هرچه نقش آزمون در تعیین سرنوشت فرد آزمون دهنده کمتر باشد، بیشتر می‌توان هدف از آن آزمون را هدایت یادگیری دانشجویان قرارداد و بر عکس هرچه نقش آزمون در تعیین سرنوشت فرد بیشتر باشد، تعیین ویژگی‌های روان‌سنگی اهمیت بیشتری پیدا می‌کند [۳]. در امتحان باید، ۱- اهداف آموزشی مورد آزمون قرار بگیرند؛ ۲- به مباحثت مهم بیشتر توجه شود؛ و ۳- مدت زمانی که برای امتحان گرفتن هر مطلب اختصاص داده می‌شود، با اهمیت نسبی آن موضوع تناسب داشته باشد [۳].

روش آزمون می‌بایست رابطه نزدیکی با محتوای آزمون داشته باشد؛ به عنوان مثال در امتحان تشریحی، داولطلب می‌تواند قالب جملات و بیان مطالب را خودش انتخاب کند و آزاد است هر اندازه که می‌خواهد در پاسخ سوال، مطلب بنویسد. این روش، کمک می‌کند که دانشجو تفکرات و درک مطلب خود را از موضوع انتقال دهد. امتحانات بالینی و عملی تفاوت آنچه را که دانشجو می‌تواند انجام دهد و آنچه را که می‌گوید روش می‌کند [۴].

در میان روش‌های متعدد اندازه‌گیری و ارزشیابی، یکی از روش‌های بسیار پرکاربرد، سوالات چندگزینه‌ای (MCQ) است. سوالات و آزمون‌های چندگزینه‌ای عموماً برای ارزشیابی سطح دانش دانشجویان به کار می‌رود [۵]، اما در صورت طراحی صحیح و قوی، این سوالات قادر به سنجش و ارزشیابی قدرت تجزیه و تحلیل و حل مسئله نیز هستند [۶]. در عین حال، بسیاری از افراد آن را به‌اندازه کافی عادلانه نمی‌دانند و این موضوع یکی از دلایلی است که در حال حاضر در دنیا در بسیاری از امتحانات سرنوشت‌ساز، سوالات جوگردانی گسترده جای آنها را گرفته‌اند، زیرا این‌گونه سوالات عادلانه‌ترین نوع سوال‌ها هستند [۳]. عادلانه بودن سوال به (۱) کفایت از نظر شاخص‌های روان‌سنگی، (۲) دقت و کوشنش در ساختن آنها، (۳) روایی منتاج شده (منظور اثر حاصله از نمرات آزمون بر آزمون دهنده، اعضای هیأت‌علمی، بیماران، جامعه و بر یاددهی و یادگیری است)؛ و (۴) تعیین استاندارد به روش مناسب بستگی دارد [۳].

طراحی متن سوالات چندگزینه‌ای

سئوالی خوب است که اطلاعات آن تا آنجا که ممکن است در درون متن سؤال جای داده شود و گزینه‌های آن تا آنجا که ممکن است کوتاه باشند. متن سؤال باید حاوی تمام اطلاعات لازم برای پاسخ دادن باشد [۳]. در مطالعه ریچارد و همکاران مشخص شد که متن سؤال با سناریویی مختصر اما کامل، فرآیند یادگیری و آموخته‌های دانشجویان را بهتر منعکس می‌کند [۱۳]. متن سؤال می‌تواند نقشه، دیاگرام، نمودار و حتی عکس باشد. اما این موارد حتماً باید با توضیح کامل سؤال به عنوان راهنمای همراه باشد [۱۴]. در طرح سوالات چندگزینه‌ای در رشته‌های پزشکی و پرستاری، بهتر است که بیشتر سوالات به صورت معرفی کوتاه بالینی بیمار طراحی شوند. مطالب ارایه شده در معرفی بالینی بیمار باید از ترتیب خاصی تعیت کند. ولی به طور کلی رعایت موارد زیر در نوشتن متن سؤال مهم است:

- ۱- مطلب مهمی را در متن سؤال قرار دهید.
- ۲- متن به صورتی طراحی شود که بتوان بدون نگاه کردن به گزینه‌ها به آن پاسخ داد.

۳- تمام اطلاعات مرتبط در گزینه‌ها، در متن سؤال ذکر شده باشد و اطلاعات اضافه در گزینه‌ها وجود نداشته باشد [۳]. یک راه کسب اطمینان از این امر آن است که روی گزینه‌ها را پوشانده و متن سؤال را به تنها بخوانید تا متوجه شوید که مسأله مورد نظر به صورت کامل بیان شده است یا نه [۱۵].

۴- متن سؤال منحرف‌کننده و یا خیلی پیچیده نباشد.

۵- در متن سؤال عبارت منفی مانند "به جز" بکار نبرید [۳].

۶- بهتر است متن سؤال به صورت جمله سوالی طراحی شود و کمتر از عبارات و جملات نیمه‌تمام و ناکامل استفاده گردد [۱۶].

۷- میزان سختی محتوى و متن سؤال متناسب با سطح فرآگیر باشد [۱۷].

در مطالعه تارانت و همکاران ۴۷/۳٪ سوالات مرتبط با دروس پرستاری در قسمت آیتم دچار اشکال ساختاری بودند [۱۱]. نکته مهم دیگر در طراحی سوالات های چندگزینه‌ای ارایه راهنمایی مناسب در ابتدای سوالات است. توضیحاتی از قبیل تعداد سوالات، محاسبه نمره منفی، تعداد صفحات و توضیحات دیگری از قبیل اینکه ممکن است، بعضی از سوالات به ظاهر ۲ یا ۳ جواب صحیح داشته باشند، اما یکی از آنها صحیح‌تر است، آن را انتخاب کنید، در این قسمت آورده می‌شوند [۱۶].

طراحی گزینه‌های صحیح در سوالات چندگزینه‌ای

گزینه صحیح باید کاملاً صحیح باشد. اگر گزینه صحیح سؤال، بهترین گزینه صحیح است، این مورد حتماً باید در متن سؤال ذکر شود. اشتباهات رایج در طراحی گزینه صحیح شامل موارد زیر است:

۱- جواب صحیح معمولاً طولانی‌تر از گزینه انحرافی است.

اول آن یعنی، دانش، دریافت و درک، کاربرد و تجزیه تحلیل با سوالات چندگزینه‌ای، قابل اندازه‌گیری است [۹]. در مطالعه‌ای که حق‌شناس و همکاران بر روی آزمون‌های چندگزینه‌ای دانشکده پزشکی دانشگاه علوم پزشکی مازندران در نیمسال اول سال تحصیلی ۱۴۷۸-۸۵ انجام دادند، از مجموع ۱۴۷۸ سؤال مربوط به ۲۵ آزمون، ۴/۷۷٪ سوالات در سطح دانش و بقیه در سطح ادراک و کاربرد و سطح حل مسأله طراحی شده بودند [۱].

کیفیت آزمون‌های چندگزینه‌ای از نظر توزیع، براساس اهداف سطوح شناختی و رعایت اصول ساختاری سوالات چندگزینه‌ای بین آزمون‌های مقاطع مختلف متفاوت بوده و در برخی موارد نیازمند اتخاذ تدابیر اصلاحی است [۱]. ازین رو سوالات چندگزینه‌ای در سنجش دانش افراد باید با منابعی که از آنها طرح می‌شوند، همخوانی داشته باشند و موجب شوند دانشجویان اطلاعات زیادی را که فرآگرفته‌اند تجزیه و تحلیل کنند. بنابراین سوالات چندگزینه‌ای باید به گونه‌ای طراحی شوند که تمام سطوح مهارت‌های شناختی افراد را بسنجد.

کاربرد سوالات چندگزینه‌ای در پرستاری و پزشکی

سوالات چندگزینه‌ای شکل رایج ارزشیابی در دانشجویان پرستاری و دیگر رشته‌های مرتبط با پزشکی است. اینگونه سوالات به اساتید این اجازه را می‌دهد که همزمان اهداف آموزشی متنوع و گوناگونی را در دانشجویان مختلف و با زمان کمتری ارزشیابی کنند [۱]. به طور کلی آزمون‌های چندگزینه‌ای سطح دانش و تاحدوی توابی حل مسأله را می‌سنجند.

در پرستاری و پزشکی از این آزمون‌ها برای ارزشیابی پایانی و ارایه گواهی و مدرک بیشتر استفاده می‌گردد، اما در حوزه صلاحیت بالینی، آزمون‌های چندگزینه‌ای، آزمون‌های روایی نیستند، زیرا این آزمون‌ها نحوه عملکرد فرد را نشان نمی‌دهند [۸]. در مطالعه‌ای که با هدف مقایسه آزمون‌های تشریحی، چندگزینه‌ای و آزمون عینی ساختاردار بالینی (OSCE) در دانشجویان پزشکی برای تعیین و پیش‌بینی عملکرد و صلاحیت بالینی آنها انجام شد، مشخص شد که آزمون‌های چندگزینه‌ای و تشریحی برخلاف OSCE برای تعیین و پیش‌بینی عملکرد و صلاحیت بالینی دانشجویان مناسب نیستند [۱۲].

سوالات چندگزینه‌ای باید در سنجش دانش افراد در رشته‌های پزشکی و پرستاری با منابعی که از آنها طرح می‌شوند همخوانی داشته باشند. در رشته‌های تخصصی همانند پرستاری و پزشکی، دانشجویان باید اطلاعات زیادی را فرآگرفته و آنها را تجزیه و تحلیل کنند. بنابراین سوالات چندگزینه‌ای باید به گونه‌ای طراحی شوند که سطوح دانش و مهارت این دانشجویان را بسنجد [۹].

سوالات چندگزینه‌ای از آزمون‌های عینی هستند. عینی بودن این آزمون‌ها به این دلیل است که نظر و عقاید افراد دیگر و مصححان در آن تاثیری ندارند.

أنواع خطاهای تکنیکی در طراحی سوالات چندگزینه‌ای

در مطالعه‌ای که مارتین و همکاران در رابطه با سوالات چندگزینه‌ای سرنوشت‌ساز انجام دادند مشخص شد هنگامی که دانشجویان بعد از جواب دادن به این سوال‌ها برای بار اول جواب خود را تغییر می‌دهند، این تغییر جواب در بار اول معمولاً باعث انتخاب جواب‌های صحیح می‌گردد. اما تغییر جواب‌های انتخابی در دفعات بعدی معمولاً باعث انتخاب جواب‌های غلط می‌گردد. بنابراین نتایج مطالعه حاضر، این عقیده را که در سوالات چندگزینه‌ای اولین انتخاب همیشه درست است، را رد کرد [۱۶]. اما خطاهای تکنیکی در تستزنی متفاوت هستند، که آنها را به دو گروه تقسیم نموده‌اند:

(الف) خطاهای مربوط به مهارت یا خرد تستزنی: در سال ۱۹۶۵، میلمان و همکاران طی مقاله‌ای، خرد تستزنی را اینگونه تعریف کردند: توانایی آزمون‌دهنده در استفاده از خصوصیات و فرمتهای آزمون یا موقعیت برگزاری آزمون در جهت کسب نمره بالا در آزمون. بدیهی است هرچه قدر آزمون‌دهنده بهتر بتواند با استفاده از خرد تستزنی، پاسخ درست سوال را پیدا کند، روایی آن آزمون کمتر خواهد بود، زیرا این استراتژی‌های تستزنی ارتباطی به داشت فرد درباره محتوی مورد آزمون ندارد. بنابراین اگر در استراتژی‌های تستزنی ذکر می‌شود: "اگر پاسخ درست را نمی‌دانید، طولانی‌ترین گزینه را انتخاب کنید"، به این علت است که معمولاً احتمال درست بودن آن بیشتر است. ما باید گزینه‌های سوال را طوری طراحی کنیم که از نظر اندازه طول مساوی داشته باشند تا این نکته نتواند از سوی آزمون‌دهنگان زیرک (ولی کم‌سواد در موضوع مورد نظر)، مورد سوء استفاده قرار بگیرد [۳]. برخی از خطاهای مربوط به فراهم نمودن امکان خرد تستزنی برای آزمون‌دهنگان عبارت است از:

۱- وجود راهنمای دستور زبانی در سوال
۲- وجود راهنمای منطقی در سوال

۳- به کار بردن واژه‌های مطلق (مانند همیشه، هرگز و ...)
۴- به کار بردن واژه‌های مبهم (مانند غالباً، بهندرت و ...)

۵- استراتژی همگرایی

فرض اصلی در استفاده از این اصل (استراتژی همگرایی) این است که گزینه درست معمولاً گزینه‌ای است که بیشترین اشتراک را با سایر گزینه‌ها دارد. به عنوان مثال، اگر جواب سوال یک عدد است، معمولاً جواب درست عددی است که در وسط قرار دارد و نه خیلی خارج از گستره مورد نظر. همچنین اگر پاسخ بهصورت عدد نیست، معمولاً پاسخ درست گزینه‌ای است که بیشتر از بقیه تکرار شده است [۳].

(ب) خطاهای مربوط به دشواری نامتناسب آزمون: اینگونه خطاهای شامل موارد زیر است:

۱- گزینه‌ها طولانی، پیچیده و یا دو قسمتی باشند.

۲- اطلاعات عددی بهصورت یکدست و هماهنگ داده نشده باشند.

۳- زبان مورد استفاده در گزینه‌ها حالت موازی و متناسب نداشته باشند.

۲- معمولاً برای گزینه صحیح عین کلمات کتاب آورده می‌شود.
۳- کلمات راهنمایی‌کننده مانند، همیشه و هرگز در گزینه صحیح بیشتر از گزینه‌های انحرافی به کار می‌روند [۱۷].

گزینه صحیح باید بهصورت تصادفی میان گزینه‌های انحرافی قرار گیرد. در مطالعه مستزر و همکاران، ۲۹۱۳ سوال چندگزینه‌ای پرستاری که از بانک سوالات انتخاب شده بود از ۳ جنبه داشتن راهنمای سوالات، سطح سوالات براساس طبقه‌بندی بلوم و نحوه تصادفی قرار گرفتن جواب صحیح میان ۴ گزینه، بررسی شد. نتایج نشان داد که، ۲۲۳۲ مورد در قسمت آیتم سوال بهصورت کلی ایراد داشتند. ۴۷/۳٪ سوالات در سطح دانش، ۴۶/۵٪ در سطح ادراک، ۲۱/۸٪ در سطح کاربرد و فقط ۶/۵٪ در سطح تجزیه و تحلیل بودند. گزینه صحیح نیز بهصورت عادی میان ۴ گزینه قرار گرفته بود ۲۴/۶٪ گزینه ۱، ۲۹/۶٪ گزینه ۲، ۲۵/۸٪ گزینه ۳ و ۲۲/۸٪ گزینه ۴ [۹].

طراحی گزینه‌های انحرافی در سوالات چندگزینه‌ای

به‌طور کلی گزینه‌های یک سوال چندگزینه‌ای اعم از گزینه‌های انحرافی و گزینه صحیح می‌تواند بهصورت جمله کامل، جمله ناکامل که با متن سوال تکمیل می‌شود و یا بهصورت یک کلمه‌ای، نوشته شود [۱۸]. بدیهی است که طراحی گزینه‌های انحرافی خوب، بسیار اهمیت دارد. یک گزینه انحرافی خوب باید دارای ویژگی‌های زیر باشد:

۱- توجیه‌پذیر باشد و آزمون‌شونده نامطلع و کم‌دانش را بهسوی خود جلب کند. اگر یک گزینه انحرافی، آزمون شوندگان قوی را بیش از آزمون شوندگان ضعیف به خود جلب کند، نتیجه می‌گیریم که یا کلید سوال نادرست بوده و یا گزینه انحرافی خوب طرح نشده است.

۲- گزینه انحرافی باید نادرست باشد، یا در قیاس با گزینه پاسخ، کمتر درست باشد.

۳- از نظر محبتوا با گزینه پاسخ مشابه باشد.

۴- از نظر ساختار و طول جمله، مشابه گزینه درست باشد.

۵- از نظر دستور زبانی و منطقی با متن سوال هم خوان باشد [۳].

۶- در آن از غلط‌های معمول دانشجویان استفاده شود.

نکته مهم در طرح گزینه انحرافی این است که این گزینه‌ها باید پاسخ دهنگان بی‌اطلاع از موضوع سوال را منحرف کند نه افراد مسلط بر موضوع سوال [۱۵]. به‌طور کلی اشتباهات رایج در طرح گزینه‌های انحرافی شامل موارد زیر است:

۱- گزینه‌های انحرافی و گزینه صحیح از نظر محتوى و مفهوم یکنواخت و متجانس نیستند (به عنوان مثال جواب صحیح روش درمانی است، در حالیکه گزینه‌های انحرافی تست تشخیصی هستند).

۲- گزینه‌های انحرافی از نظر دستور زبانی با متن سوال متجانس نیستند.

۳- گزینه‌های انحرافی و گزینه صحیح از نظر طول واژگان متناسب نیستند [۱۷].

نتیجه‌گیری

به‌نظر می‌رسد روش ارزیابی از نظر علمی تاثیر زیادی در نحوه مطالعه دانشجویان دارد و اطلاع آنان از شکل رایج سوالات امتحانی، سوال‌خوانی را جانشین تعمق و تفکر در هنگام یادگیری نموده است. در سال ۱۹۹۸ هاموند و همکاران گزارش نمودند که نتایج مطالعات بر روی دانشجویان شرکت‌کننده در کلاس مروری قبل از امتحان، نشان داد که بخش مهمی از آمادگی فرد برای امتحان، ارزیابی او از آن روش است [۲۱]. روش‌های ارزشیابی نقاط ضعف و قوت متفاوت و خاص خود را دارند [۲۲]. آزمون‌های عینی نیز که به صورت گسترشده در تمام سطوح آموزش پژوهشی استفاده می‌شوند از این قاعده مستثنی نیستند. سوالات چندگزینه‌ای جز آزمون‌های عینی هستند، که به خاطر ویژگی‌های روان‌سنجدی مثبت و تحقیقات و ارایه نظرات تکمیلی که در رابطه با اصلاح این آزمون‌ها بیان شده است، کاربرد زیادی دارند. اما در عین حال ایراداتی نیز به این نوع سوالات‌ها وارد است. از جمله می‌توان به عدم روایی بالای این آزمون‌ها اشاره کرد [۲۳]. همچنین مشکل اساسی دیگری که همراه سوالات چندگزینه‌ای است، این است که در صورتیکه آزمون‌شونده جواب را اشتباه انتخاب کند، می‌توان اطمینان داشت که وی جواب را نمی‌دانسته است. اما اگر جواب صحیح را انتخاب کند، نمی‌توان اطمینان داشت که وی جواب صحیح را می‌دانسته، زیرا که ممکن است، آنرا حدس زده باشد [۲۴]. البته توجه به این نکته نیز ضروری است که با توجه به محدودیت‌های آزمون‌های چندگزینه‌ای، همانطور که در بسیاری از دانشگاه‌های متبر جهان از جمله شده است، باید حرکت استایید به استفاده از سایر همکاران نشان داده شده است، ولی بهترین روش‌های استاندارد سنجش دانش در کنار سوالات چندگزینه‌ای سوق تا ارزیابی جامع‌تر و دقیق‌تری انجام گردد [۱]. تاکنون اعتبار روش‌های مختلف برای ارزیابی دانش فرآگیران، مستقل از یکدیگر بررسی و تعیین شده است. هر چند که نحوه ارزیابی در پژوهشی و پرستاری عمدهاً برای طرح سوالات چندگزینه‌ای است، ولی بهترین شرایط استفاده از این روش استفاده از آن همراه با دیگر روش‌های ارزشیابی است. با توجه به ضریب تفاضل بالا میان سوالات چندگزینه‌ای و صحیح/نادرست و اعتبار مناسب سوالات کوتاه پاسخ، بهتر است که از این سه روش به صورت مکمل یکدیگر استفاده شود تا بدین وسیله بتوان معایب سوالات چندگزینه‌ای را کاهش داد و یادگیری موثر دانشجویان را تقویت کرد [۲۵].

منابع

- حق‌شناس محمدرضا، وحیدشاهی کوروش، محمودی میترا، شهرآزاد لیلا، پروین‌نژاد نیکو، عمامی اعظم، بررسی آزمون‌های چندگزینه‌ای دانشکده پژوهشی دانشگاه علوم پژوهشی مازندران در نیمسال اول سال تحصیلی ۱۳۸۵-۸۶. گام‌های توسعه در آموزش پژوهشی. ۱۱۹:۲۶-۲۵.

- گزینه‌ها به صورت منطقی پشت سرهم نیاشند.
- عبارت "هیچ کدام از موارد فوق" و "همه موارد بالا" به عنوان یک گزینه به کار رفته باشد [۳، ۱۳].
- متن سوال گمراحتکننده یا به گونه‌ای غیرضروری پیچیده باشد.
- پاسخ به یک سوال در گرو پاسخ دادن به سوال دیگری باشد.
- دستورالعمل‌هایی برای طراحی یک سوال چندگزینه‌ای مناسب:

 - بر روی مطالب مهم تمرکز نمایید.
 - متن سوال را بدون ابهام طرح کنید.
 - گزینه‌های انحرافی را طوری طرح کنید که توجیه‌پذیر باشند (آشکارا غلط نباشند).
 - از بکاربردن عبارت‌های غیردقیق مانند همراه است، مهم است و ... خودداری کنید.
 - گزینه‌ها را یکدست و از یک جنس و ماهیت بنویسید.
 - از بکاربردن واژه‌های منفی در متن سوال بپرهیزید.
 - بهتر است نمره منفی تدھید، زیرا نمره منفی بین دانشجویان براساس رفتار خطرپذیری آنها تفاوت قابل می‌شود [۳].
 - از بکاربردن اشاره‌های دستوری نامربوط خودداری کنید (اشاراتی از قبیل زمان افعال، جمع یا مفرد بودن آنها و ...).
 - از بکاربردن عباراتی مانند اغلب، مکررا و گاهی استفاده نکنید [۱۷].

محاسن و محدودیت‌های آزمون‌های چندگزینه‌ای

این آزمون‌ها قادر به اندازه‌گیری غالب بازده‌های یادگیری، از دانش گرفته تا فهمیدن، قضاوت‌کردن و حل مسئله هستند. به‌طور کلی محاسن آزمون‌های چندگزینه‌ای شامل موارد زیر است:

- این آزمون‌ها از سایر آزمون‌های عینی انعطاف‌پذیرترند.
- با استفاده از این آزمون‌ها در یک زمان محدود می‌توان اهداف آموزشی زیادی را سنجید.

- تصحیح کردن این آزمون‌ها کاملاً عینی است.
- در صورت طرح گزینه‌های انحرافی با توجه به اشتباهات دانشجویان، این سوالات منبع مناسبی برای پی‌بردن به مشکلات آموزشی آنها هستند [۱۵].

- آزمون‌های چندگزینه‌ای از چند جهت نیز محدودیت‌هایی دارند:
- ساختن این آزمون‌ها به صورت مناسب بسیار دشوار است.
 - معلمان اغلب سوالاتی را در حد سنجش اطلاعات جزئی و کم‌همیت می‌نویسند.
 - زمانی که آزمون نمره منفی داشته باشد، معمولاً دانش‌آموزانی که خطر می‌کنند از سایر دانش‌آموزان نمره بهتری می‌گیرند.
 - دانش‌آموزان قوی بابت تیزی‌بینی و تجزیه و تحلیل خود در سوالات مشکل تشویق نمی‌شوند [۱۵].
 - بهوسیله این آزمون‌ها مهارت و صلاحیت بالینی فرآگیران مشخص نمی‌شود [۲۰].

- 13- Rasaian N, Nakhaie S, Sadeghi Ghandehari N. Comparison of three exam thecniques in medical students: Multiple choice, true/false and short answer questions. *Hakim J.* 1381;5(4):273-80. [Persian]
- 14- Hays RB, Coventry P, Wilcock D, Hartley K. Short and long multiple-choice question stems in a primary care oriented undergraduate medical curriculum. *Educ Prim Care.* 2009;20(3):173-7.
- ۱۵- سیف علی‌اکبر. سنجش فرآیند و فرآورده یادگیری روش‌های جدید و قدیم. چاپ اول. تهران: نشر دوران؛ ۱۳۸۴.
- 16- Martin RF, Sibyl H, Veronika K, Answering multiple-choice questions in high-stakes medical examinations. *Med Educ.* 2005;39:890-4.
- 17- Tim W, Gary C. Developing multiple choice questions for the RCPSC certification examinations. The royal college of physicians and surgeons of Canada [Report]; 2004.
- 18- Patricia M, Nugent BA. Vitale, test success, test taking techniques for beginning nursing students. 5th ed. Garden City, New York: Davis company; 2008.
- ۱۹- سیف علی‌اکبر. اندازه‌گیری، سنجش و ارزشیابی آموزشی ویرایش پنجم. تهران: نشر دوران؛ ۱۳۸۷.
- 20- Padraig CD, Srinivas MS, Nadeem YK. Relationship between dental students' performance on standardized multiple-choice examinations and OSCEs. *J Dental Educ.* 2008;72(5):585-93.
- 21- Ronald ME, Malcolm C, David MI. Assessment in medical education. *N Engl J Med.* 2007;356:387-96.
- 22- Steven MD. Threats to the validity of locally developed multiple-choice tests in medical education: Construct-irrelevant variance and construct underrepresentation. *Adv Health Sci Educ.* 2002;(7):235-41.
- 23- Rajeeva LK. Multiple-choice tests, negative marks and an alternative. *Reson.* 2006;11(3):86-94.
- ۲- دادرس پیگانه محمد. ارزشیابی کاربردی برنامه‌های آموزشی در سازمان‌های تولیدی و خدماتی. چاپ اول. تهران: نشر مولف؛ ۱۳۸۶.
- ۳- ملکان‌راد الله، عین‌اللهی بهرام، حسینی سیدجیلی، ممتازمنش نادر. آموزش و ارزشیابی بالینی، آنچه هر استاد بالینی باید بداند. چاپ اول. تهران: انتشارات تحفه و بشری؛ ۱۳۸۵.
- ۴- جلیلی زهرا. چگونه دانشجویان را سنجش کنیم؟ نشریه علمی پک. ۱۳۸۶؛۱۲(۲):۱-۶.
- 5- Chirayu A, Steven M D, Uapong J, Rungnirand P. Sources of validity evidence for an internal medicine student evaluation system: An evaluative study of assessment methods. *Med Educ.* 2005;39:276-83.
- 6- Kathy D. Common formats of multiple choice questions testing analytical in clinical laboratory science. *Clin Lab Sci.* 2000;13(1):40-51.
- ۷- سیف‌نراقی مریم، نادری عزت‌اله. سنجش و اندازه‌گیری و بنیادهای تحلیلی و ابزارهای آن در علوم تربیتی و روان‌شناسی تهران: نشر ارسپاران؛ ۱۳۸۶.
- 8- Dacre J, Haq I. Assessing competencies in rheumatology. *Ann Rheum Dis.* 2005;(64):3-6.
- 9- Joan CM, Barbara SH, Mary EP, Kathy L, Margaret TM. Assessment of multiple choice questions in selected test banks Accompanying text books used in nursing education. *J Nurs Educ.* 2001;40(1):25-33.
- 10- Jannette C. Writing multiple-choice questions for continuing medical education activities and self-assessment modules. *Radiol Graphics.* 2006;26(2):543-51.
- 11- Marie T, James W. Impact of item-writing flaws in multiple-choice questions on student achievement in high-stakes nursing assessments. *Med Educ.* 2008;42(4):198-206.
- 12- Tim JW, Christopher MF. Comprehensive undergraduate medical assessments improve prediction of clinical performance. *Med Educ.* 2004;38:1111-6.