مجله راهبردهای آموزش
دوره ۲، شماره ٤، زمستان ۱۳۸۸
صفحات: ۱۷۷ تا ۱۸۲

مروری

ارزشیابی به‌وسیله آزمون‌های چندگزینه‌ای

محسن ملاهادی * *BSc*

* گروه اطفال، دانشکده پرستاری، دانشگاه علوم پزشکی بقیه‌ا...(عج)، تهران، ایران

چکیده

مقدمه. ارزشیابی، فرآیند نظام‌مندِ گردآوری، تحلیل و تفسیر داده‌ها به‌منظور ارزش‌گذاری و قضاوت است. اساتید 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 به‌علت عینیت کامل، از حداکثر پایایی ممکن برخوردارند، اما روایی آنها هنوز مورد توجه صاحب‌نظران بحث ارزشیابی قرار دارد. با این که آزمون‌های چندگزینه‌ای به‌صورت گسترده 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 پایایی 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 فراگیران تبدیل می‌شوند. بنابراین، تجدیدنظر و مطالعه بیشتر اساتید در رابطه با قواعد طراحی سئوالات چندگزینه‌ای امری الزامی به‌نظر می‌رسد.

کلیدواژه‌ها: ارزشیابی، سئوالات چندگزینه‌ای، پایایی، 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

[Downloaded from edcbmj.ir on 2026-11-25]

نویسنده مسئول: محسن ملاهادی. تمام درخواست‌ها باید به نشانی mmollahadi@gmail.com فرستاده شوند.

دریافت مقاله: ۱۳۸۸/۱۰/۱۶ پذیرش مقاله: ۱۳۸۸/۱۱/۱۸

مقدمه

ارزیابی یا ارزشیابی، عبارت از فرآیند نظامندِ گردآوری، تحلیل و تفسیر داده‌ها به‌منظور ارزش‌گذاری و قضاوت است. اگر این فرآیند در مورد افراد، گروه‌ها و سازمان‌ها انجام گردد، معمولاً به‌عنوان ارزیابی و اگر در مورد برنامه انجام شود، به آن ارزشیابی اطلاق می‌گردد. آزمون‌ها از انواع 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 مشخص‌کردن موفقیت یا عدم موفقیت در اهداف مشخص‌شده برای آن برنامه آموزشی است [۲]. هدف از امتحان، می‌تواند نشان دادن اهمیت مباحث به فراگیران، ترغیب فراگیران به مطالعه، شناسایی موارد ضعیف در یادگیری و اصلاح آنها، تصمیم‌گیری درباره ارتقای فراگیر و شناسایی موارد ضعف دوره یا برنامه آموزشی باشد.

هرچه نقش آزمون در تعیین سرنوشت فرد آزمون‌دهنده کمتر باشد، بیشتر می‌توان هدف از آن آزمون را هدایت یادگیری دانشجویان قرارداد و برعکس هرچه نقش آزمون در تعیین سرنوشت فرد بیشتر باشد، تعیین ویژگی‌های روان‌سنجی اهمیت بیشتری پیدا می‌کند [۳]. در امتحان باید، ۱- اهداف آموزشی مورد آزمون قرار بگیرند؛ ۲- به مباحث 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؛ به‌عنوان مثال در امتحان تشریحی، داوطلب می‌تواند قالب جملات و بیان مطالب را خودش انتخاب کند و آزاد است هر اندازه که می‌خواهد در پاسخ سئوال، 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 روشن می‌کند [۴].

در میان روش‌های متعدد اندازه‌گیری و ارزشیابی، یکی از روش‌های بسیار پرکاربرد، سئوالات چندگزینه‌ای (MCQ) است. سئوالات و آزمون‌های چندگزینه‌ای عموماً برای ارزشیابی سطح دانش دانشجویان به‌کار می‌رود [۵]، اما در صورت طراحی صحیح و قوی، این سئوالات قادر به سنجش و ارزشیابی قدرت تجزیه و تحلیل و حل‌مسأله 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، سئوالات جورکردنی گسترده جای آنها را گرفته‌اند، زیرا این‌گونه سئوالات عادلانه‌ترین نوع سئوال‌ها هستند [۳]. عادلانه بودن سئوال به ۱)کفایت از نظر شاخص‌های روان‌سنجی، ۲) دقت و کوشش در ساختن آنها، ۳) روایی منتج‌شده (منظور اثر حاصله از نمرات آزمون بر آزمون‌دهنده، اعضای هیات‌علمی، بیماران، جامعه و بر یاددهی و یادگیری است)؛ و ۴) تعیین استاندارد به روش مناسب بستگی دارد [۳].

در مورد سئوال‌های 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 سئوالات و محاسن و معایب آنها بود.

آزمون‌های چندگزینه‌ای

یکی از روش‌های آزمون دانش، سئوالات چندگزینه‌ای است که اولین بار در سال ۱۹۱۴ توسط *فردریک* در دانشگاه کانزاس ابداع گردید [۱]. اکنون سئوالات چندگزینه‌ای از شایع‌ترین نوع آزمون‌ها در دنیا، به‌ویژه در ایران هستند. در این نوع سئوالات، مجموعه‌ای از جواب‌ها که برای سئوال مطرح شده و بالقوه درست به‌نظر می‌رسند، به‌شکل سناریو در اختیار فراگیر قرار می‌گیرد و فراگیر از میان آنها جواب درست را انتخاب می‌کند. این نوع سئوالات چندگزینه‌ای علاوه بر سنجش دانش، توانایی حل‌مسأله را نیز می‌سنجند [۸]. دو شرط اصلی سئوال چندگزینه‌ای، الف) پرسش درباره موضوع مهم؛ و ب) طراحی خوب است به‌نحوی‌که در آن از خطاهایی که موجب می‌شود آزمون‌دهنده 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 مسایل مورد توجه در مطالعات و پژوهش‌های آموزشی است، البته این مسأله را نیز می‌توان با رعایت اصول طراحی ساختار سئوالات چندگزینه‌ای و اصول طراحی سئوالات چندگزینه‌ای، براساس سطوح طبقه‌بندی شناختی، کاهش داد [۱]. لذا لازم است سئوالات چندگزینه‌ای براساس یکی از طبقه‌بندی‌های اهداف شناختی طراحی شوند [۹]؛ مثلاً طبقه‌بندی *بلوم*. *بلوم* در سال ۱۹۵۹، طبقه‌بندی سطوح یادگیری شناختی را ارایه کرد. این طبقه‌بندی شامل سطوح دانش، دریافت و ادراک، کاربرد، تجزیه و تحلیل و ارزشیابی بود. به‌دنبال آن اساتید و ارزیابان این طبقه‌بندی را پذیرفته و در نهایت آن را در ۳ طبقه جای دادند: ۱) سطح دانش، شامل یادآوری مطالب 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 حل‌مسأله، شامل انتقال و تبدیل دانش و مهارت‌ها به موقعیت‌های جدید است [۱۰].

گاهی بعضی اهداف شناختی به‌وسیله سئوالات چندگزینه‌ای قابل اندازه‌گیری نیست مثلا در پرستاری از طبقه‌بندی *بلوم* فقط ۴ سطح

[Downloaded from edcbmj.ir on 2026-11-25]

اول آن یعنی، دانش، دریافت و درک، کاربرد و تجزیه تحلیل با سئوالات 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 ۸۶-۸۵ انجام دادند، از مجموع ۱۴۷۸ سئوال مربوط به ۲۵ آزمون، ۷۷/۴٪ سئوالات در سطح دانش و بقیه در سطح ادراک و کاربرد و سطح حل‌مسأله طراحی شده بودند [۱].

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

کاربرد سئوالات چندگزینه‌ای در پرستاری و پزشکی

سئوالات چندگزینه‌ای شکل رایج ارزشیابی در دانشجویان پرستاری و دیگر رشته‌های مرتبط با پزشکی است. اینگونه سئوالات به اساتید این اجازه را می‌دهد که همزمان اهداف آموزشی متنوع و گوناگونی را در دانشجویان مختلف و با زمان کمتری ارزشیابی کنند [۱۱]. به‌طور کلی آزمون‌های چندگزینه‌ای سطح دانش و تاحدودی توانایی حل‌مسأله 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 بر خلاف OSCE برای تعیین و پیش‌بینی عملکرد و صلاحیت بالینی دانشجویان مناسب نیستند [۱۲].

سئوالات 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 فراگرفته و آنها را تجزیه و تحلیل کنند. بنابراین سئوالات چندگزینه‌ای باید به‌گونه‌ای طراحی شوند که سطوح دانش و مهارت این دانشجویان را بسنجد [۹].

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

طراحی متن سئوالات چندگزینه‌ای

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

۱- مطلب مهمی را در متن سئوال قرار دهید .

۲- متن به‌صورتی طراحی شود که بتوان بدون نگاه کردن به گزینه‌ها به آن پاسخ داد.

۳- تمام اطلاعات مرتبط در گزینه‌ها، در متن سئوال 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 سئوال را به تنهایی بخوانید تا متوجه شوید که مسأله مورد نظر به‌صورت کامل بیان شده است یا نه [۱۵].

۴- متن سئوال منحرف‌کننده و یا خیلی پیچیده نباشد.

۵- در متن سئوال عبارت منفی مانند "به‌جز" بکار نبرید [۳].

۶- بهتر است متن سئوال به‌صورت جمله سئوالی طراحی شود و کمتر از عبارات و جملات نیمه‌تمام و ناکامل استفاده گردد [۱۶].

۷- میزان سختی محتوی و متن سئوال متناسب با سطح فراگیر باشد [۱۷].

در مطالعه *تارانت* و همکاران ۴۷/۳٪ سئوالات مرتبط با دروس پرستاری در قسمت آیتم دچار اشکال ساختاری بودند [۱۱]. نکته مهم دیگر در طراحی سئوال‌های چندگزینه‌ای ارایه راهنمایی مناسب در ابتدای سئوال‌ها است. توضیحاتی از قبیل تعداد سئوالات، محاسبه نمره منفی، تعداد صفحات و توضیحات دیگری از قبیل اینکه ممکن است، بعضی از سئوال‌ها 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 سئوالات چندگزینه‌ای

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

۱- جواب صحیح معمولاً طولانی‌تر از گزینه انحرافی است.

[Downloaded from edcbmj.ir on 2026-11-25]

۲- معمولاً برای گزینه صحیح عین کلمات کتاب آورده می‌شود.
۳- کلمات راهنمایی‌کننده مانند، همیشه و هرگز در گزینه صحیح بیشتر از گزینه‌های انحرافی به‌کار می‌روند [۱۷].
گزینه صحیح باید به‌صورت تصادفی میان گزینه‌های انحرافی قرار گیرد. در مطالعه *مسترز* و همکاران، ۲۹۱۳ سئوال چندگزینه‌ای پرستاری که از بانک سئوالات انتخاب شده بود از ۳ جنبه داشتن راهنمای سئوالات، سطح سئوالات براساس طبقه‌بندی بلوم و نحوه تصادفی قرار گرفتن جواب صحیح میان ۴ گزینه، بررسی شد. نتایج نشان داد که، ۲۲۳۲ مورد در قسمت آیتم سئوال به‌صورت کلی ایراد داشتند. ۴۷/۳٪ سئوالات در سطح دانش، ۲۴/۸٪ 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 سئوالات چندگزینه‌ای

به‌طور کلی گزینه‌های یک سئوال چندگزینه‌ای اعم از گزینه‌های انحرافی و گزینه صحیح می‌تواند به‌صورت جمله کامل، جمله ناکامل که با متن سئوال 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 سئوال 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 محتوا 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 سئوال هم‌خوان باشد [۳].
۶- در آن از غلط‌های معمول دانشجویان استفاده شود.
نکته مهم در طرح گزینه انحرافی این است که این گزینه‌ها باید پاسخ دهندگان بی‌اطلاع از موضوع سئوال را منحرف کند نه افراد مسلط بر موضوع سئوال [۱۵]. به‌طور کلی اشتباهات رایج در طرح گزینه‌های انحرافی شامل موارد زیر است:
۱- گزینه‌های انحرافی و گزینه صحیح از نظر محتوی و مفهوم یکنواخت و متجانس نیستند (به‌عنوان مثال جواب صحیح روش درمانی است، در حالیکه گزینه‌های انحرافی تست تشخیصی هستند).
۲- گزینه‌های انحرافی از نظر دستور زبانی با متن سئوال متجانس نیستند.
۳- گزینه‌های انحرافی و گزینه صحیح از نظر طول واژگان متناسب نیستند [۱۷].

انواع خطاهای تکنیکی در طراحی سئوالات چندگزینه‌ای

در مطالعه‌ای که *مارتین* و همکاران در رابطه با سئوالات چندگزینه‌ای سرنوشت‌ساز انجام دادند مشخص شد هنگامی که دانشجویان بعد از جواب دادن به این سئوال‌ها برای بار اول جواب خود را تغییر می‌دهند، این تغییر جواب در بار اول معمولا باعث انتخاب جواب‌های صحیح می‌گردد. اما تغییر جواب‌های انتخابی در دفعات بعدی معمولا باعث انتخاب جواب‌های غلط می‌گردد. بنابراین نتایج مطالعه حاضر، این عقیده را که در سئوالات 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 سئوال را پیدا کند، روایی آن آزمون کمتر خواهد بود، زیرا این استراتژی‌های تست‌زنی ارتباطی به دانش 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 معمولا احتمال درست بودن آن بیشتر است. ما باید گزینه‌های سئوال را طوری طراحی کنیم که از نظر اندازه طول مساوی داشته باشند تا این نکته نتواند از سوی آزمون‌دهندگان زیرک (ولی کم‌سواد در موضوع مورد نظر)، مورد سوء استفاده قرار بگیرد [۳]. برخی از خطاهای مربوط به فراهم نمودن امکان خرد تست‌زنی برای آزمون‌دهندگان عبارت است از:
۱- وجود راهنمای دستور زبانی در سئوال
۲- وجود راهنمای منطقی در سئوال
۳- به‌کار بردن واژه‌های مطلق (مانند همیشه، هرگز و ...)
۴- به‌کار بردن واژه‌های مبهم (مانند غالبا، به‌ندرت و ...)
۵-استراتژی همگرایی
فرض اصلی در استفاده از این اصل (استراتژی همگرایی) این است که گزینه درست معمولا گزینه‌ای است که بیشترین اشتراک را با سایر گزینه‌ها دارد. به‌عنوان مثال، اگر جواب سئوال یک عدد است، معمولا جواب درست عددی است که در وسط قرار دارد و نه خیلی خارج از گستره مورد نظر. همچنین اگر پاسخ به‌صورت عدد نیست، معمولا پاسخ درست گزینه‌ای است که بیشتر از بقیه تکرار شده است [۳].
ب) خطاهای مربوط به دشواری نامتناسب آزمون: اینگونه خطاها شامل موارد زیر است:
۱- گزینه‌ها طولانی، پیچیده و یا دوقسمتی باشند.
۲- اطلاعات عددی به‌صورت یکدست و هماهنگ داده نشده باشند.
۳- زبان مورد استفاده در گزینه‌ها حالت موازی و متناسب نداشته باشند.

[Downloaded from edcbmj.ir on 2026-11-25]

۴- گزینه‌ها به‌صورت منطقی پشت سرهم نباشند.
۵- عبارت "هیچ کدام از موارد فوق" و "همه موارد بالا" به‌عنوان یک گزینه به‌کار رفته باشد [۳، ۱۳].
۶- متن سئوال گمراه‌کننده یا به گونه‌ای غیرضروری پیچیده باشد.
۷- پاسخ به یک سئوال در گرو پاسخ دادن به سئوال دیگری باشد.
دستورالعمل‌هایی برای طراحی یک سئوال چندگزینه‌ای مناسب:
۱- بر روی مطالب مهم تمرکز نمایید.
۲- متن سئوال را بدون ابهام طرح کنید.
۳- گزینه‌های انحرافی را طوری طرح کنید که توجیه‌پذیر باشند (آشکارا غلط نباشند).
۴- از بکاربردن عبارت‌های غیردقیق مانند همراه است، مهم است و ... خودداری کنید.
۵- گزینه‌ها را یکدست و از یک‌جنس و ماهیت بنویسید.
۶- از بکاربردن واژه‌های منفی در متن سئوال بپرهیزید.
۷- بهتر است نمره منفی ندهید، زیرا نمره منفی بین دانشجویان براساس رفتار خطرپذیری آنها تفاوت قایل می‌شود [۳].
۸- از بکاربردن اشاره‌های دستوری نامربوط خودداری کنید (اشاراتی از قبیل زمان افعال، جمع یا مفرد بودن آنها و ...) [۱۹].
۹- از بکاربردن عباراتی مانند اغلب، مکررا و گاهی استفاده نکنید [۱۷].

محاسن ومحدودیت‌های آزمون‌های چندگزینه‌ای

این آزمون‌ها قادر به اندازه‌گیری غالب بازده‌های یادگیری، از دانش گرفته تا فهمیدن، قضاوت‌کردن و حل‌مسأله هستند. به‌طور کلی محاسن آزمون‌های چندگزینه‌ای شامل موارد زیر است:
۱- این آزمون‌ها از سایر آزمون‌های عینی انعطاف‌پذیرترند.
۲- با استفاده از این آزمون‌ها در یک زمان محدود می‌توان اهداف آموزشی زیادی را سنجید.
۳- تصحیح کردن این آزمون‌ها کاملا عینی است.
۴- در صورت طرح گزینه‌های انحرافی باتوجه به اشتباهات دانشجویان، این سئوالات منبع مناسبی برای پی‌بردن به مشکلات آموزشی آنها هستند[۱۵].
آزمون‌های چندگزینه‌ای از چند جهت نیز محدودیت‌هایی دارند:
۱- ساختن این آزمون‌ها به‌صورت مناسب بسیار دشوار است.
۲- معلمان اغلب سئوال‌های چندگزینه‌ای را درحد سنجش اطلاعات جزیی و کم‌اهمیت می‌نویسند.
۳- زمانی که آزمون نمره منفی داشته باشد، معمولا دانش‌آموزانی که خطر می‌کنند از سایر دانش‌آموزان نمره بهتری می‌گیرند.
۴- دانش‌آموزان قوی بابت تیزبینی و تجزیه و تحلیل خود در سئوالات مشکل تشویق نمی‌شوند [۱۵].
۵- به‌وسیله این آزمون‌ها مهارت و صلاحیت بالینی فراگیران مشخص نمی‌شود [۲۰].

نتیجه‌گیری

به‌نظر می‌رسد روش ارزیابی از نظر علمی تاثیر زیادی در نحوه مطالعه دانشجویان دارد و اطلاع آنان از شکل رایج سئوالات امتحانی، سئوال‌خوانی 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 [۱۳]. روش‌های ارزشیابی نقاط ضعف و قوت متفاوت و خاص خود را دارند [۲۱]. آزمون‌های عینی نیز که به‌صورت گسترده در تمام سطوح آموزش پزشکی استفاده می‌شوند از این قاعده مستثنی نیستند. سئوالات چندگزینه‌ای جز آزمون‌های عینی هستند، که به خاطر ویژگی‌های روان‌سنجی مثبت و تحقیقات و ارایه نظرات تکمیلی که در رابطه با اصلاح این آزمون‌ها بیان شده است، کاربرد زیادی دارند. اما درعین حال ایراداتی نیز به این نوع سئوال‌ها وارد است. از جمله می‌توان به عدم روایی بالای این آزمون‌ها اشاره کرد [۲۲]. همچنین مشکل اساسی دیگری که همراه سئوالات 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 [۲۳]. البته توجه به این نکته نیز ضروری است که باتوجه به محدودیت‌های آزمون‌های چندگزینه‌ای، همانطور که در بسیاری از دانشگاه‌های معتبر جهان از جمله توسط *سیمکین* و همکاران در انگلستان و *بوش* و همکاران نشان داده شده است، باید حرکت اساتید به استفاده از سایر روش‌های استاندارد سنجش دانش در کنار سئوالات چندگزینه‌ای سوق یابد تا ارزیابی جامع‌تر و دقیق‌تری انجام گردد [۱]. تاکنون اعتبار روش‌های مختلف برای ارزیابی دانش فراگیران، مستقل از یکدیگر بررسی و تعیین شده است. هر چند که نحوه ارزیابی در پزشکی و پرستاری عمدتاً برپایه طرح سئوالات چندگزینه‌ای است، ولی بهترین شرایط استفاده از این روش استفاده از آن همراه با دیگر روش‌های ارزشیابی است. با توجه به ضریب توافق بالا میان سئوالات چندگزینه‌ای و صحیح/نادرست و اعتبار مناسب سئوالات کوتاه پاسخ، بهتر است که از این سه روش به‌صورت مکمل یکدیگر استفاده شود تا بدین وسیله بتوان معایب سئوالات چندگزینه‌ای را کاهش داد و یادگیری موثر دانشجویان را تقویت کرد [۱۳].

منابع

۱- حق‌شناس محمدرضا، وحیدشاهی کوروش، محمودی میترا، شهبازنژاد لیلا، پروین‌نژاد نیکو، عمادی اعظم. بررسی آزمون‌های چندگزینه‌ای دانشکده پزشکی دانشگاه علوم پزشکی مازندران در نیم‌سال اول سال تحصیلی ۸۶-۱۳۸۵. گام‌های توسعه در آموزش پزشکی. ۱۳۸۷؛۵(۲):۲۶-۱۱۹.

[Downloaded from edcbmj.ir on 2026-11-25]

۲- دادرس‌یگانه محمد. ارزشیابی کاربردی برنامه‌های آموزشی در سازمان‌های تولیدی و خدماتی. چاپ اول. تهران: نشر مولف؛ ۱۳۸۶.
۳- ملکان‌راد الهه، عین‌اللهی بهرام، حسینی سیدجلیل، ممتازمنش نادر. آموزش و ارزیابی 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؛(۱۳):۲-۱.
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 ارسباران؛ 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):198–206.
12- Tim JW, Christopher MF. Comprehensive undergraduate medical assessments improve prediction of clinical performance. Med Educ. 2004;(38):1111–6.
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.

[Downloaded from edcbmj.ir on 2026-11-25]