

## تصحیح ماشینی آزمون‌های پیشرفت تحصیلی چندگزینه‌ای و آنالیز اوراق امتحانی: چند نکته

رضا بهنام‌فر\*

\* دانشجوی دکتری مدیریت آموزشی، کارشناس مرکز مطالعات و توسعه آموزش علوم، دانشگاه علوم پزشکی شهید صدوقی یزد

تاریخ پذیرش: ۱۳۹۲/۲/۲۰

تاریخ دریافت: ۱۳۹۱/۱۲/۱۰

دشواری و تمیز صرفاً برای آزمون‌های وابسته به هنجار مناسب است و در آزمون‌های وابسته به ملاک این ضریب‌ها مورد استفاده‌ای ندارند؛ با این حال متخصصان روانسجی معتقدند که محاسبه ضرایب دشواری و تمیز برای این گونه سؤال‌ها نیز مفید است (۳). سؤال اینجاست که آیا تنها با دسترسی به این شاخص‌ها می‌توان در مورد "کیفیت" یک آزمون "قضاوت" نمود و به تبع این قضاوت به ارزشیابی عملکرد آموزشی اساتید در بحث آزمون و آزمون‌سازی پرداخت؟ به نظر می‌رسد که برای قضاوت عادلانه‌تر و صحیح‌تر در این زمینه می‌بایست به عوامل دیگری نیز به طور هم‌زمان توجه داشت.

آزمون‌های پیشرفت تحصیلی جزو آزمون‌های ملاکی محسوب می‌شوند. آزمون‌های ملاکی حوزه محدودی از هدف‌هایی را که اختصاصاً مربوط به مواد تدریس است مورد اندازه‌گیری قرار می‌دهند. تمرکز این آزمون‌ها بر خود فرد است و با توصیف آنچه فرد می‌تواند و چیزهایی که نمی‌تواند انجام دهد به ارزیابی عملکرد او می‌پردازند و به این ترتیب در تصمیم‌گیری راجع به چگونگی تدریس مفید واقع می‌شوند.

روش‌های مختلفی برای ارزشیابی وجود دارند که از آن جمله می‌توان به مشاهده مستقیم، آزمون شفاهی یا کتبی، آزمون‌های چندگزینه‌ای و آزمون عملی دوره بالینی اشاره کرد (۱). آزمون چندگزینه‌ای (تستی) متداول‌ترین آزمون عینی است که هم در سنجش توانایی‌های مختلف و هم در سنجش یادگیری در زمینه‌های گوناگون تحصیلی مورد استفاده واقع می‌شود و بر اساس نتایج حاصله از آن، در مورد ارتقا دانشجویان، تشخیص مشکلات یادگیرندگان و رفع نواقص یادگیری و آموزشی، راهنمایی و مشاوره دانشجویان و برنامه‌ریزی درسی اتخاذ تصمیم می‌شود (۲). متداول‌ترین آزمون عینی مورد استفاده در علوم پزشکی، آزمون‌های چندگزینه‌ای است (۱).

با تصحیح آزمون‌های پیشرفت تحصیلی چندگزینه‌ای به کمک دستگاه‌های تصحیح و آنالیز اوراق امتحانی، علاوه بر نمره دانشجو در آن درس، می‌توان به اطلاعاتی در مورد پایایی آزمون، ضریب دشواری و ضریب تمیز سؤال‌ها نیز با سرعت و سهولت دست یافت. با وجودی که استفاده از ضریب‌های

\* (نویسنده مسئول)؛ تلفن: ۰۳۵۱-۸۲۶۵۵۵۸، آدرس الکترونیکی: reza82br@yahoo.com

قرار می‌گیرند که توانایی‌های تحصیلی مشابه یا نسبتاً مشابه با یکدیگر دارند، انتظار داشتن یک توزیع نمرات با کجی منفی نسبت به توزیع نرمال منطقی‌تر به نظر می‌رسد.

زمانی که صحبت از دروس تخصصی باشد، به دلیل اهمیت ویژه‌ای که این دروس برای دانشجویان دارند، و همچنین تأکید بیشتر نظام آموزشی بر این دروس، و تأکیدی که در طول یک ترم بر اهداف آموزشی آن‌ها می‌گردد؛ تلاش بیشتری از جانب دانشجو در ارتباط با این دروس مشاهده می‌گردد. این امر می‌تواند در تعامل با آنچه که در ارتباط با موارد قبلی ذکر شد، در نرمال نبودن توزیع نمرات در یک کلاس درس مؤثر واقع شود.

با در نظر گرفتن مواردی که اشاره شد، به نظر می‌رسد برای "قضاوت" صحیح در مورد کیفیت یک آزمون، تنها توجه به نتایج حاصل از تحلیل آزمون در قالب شاخص‌های دشواری و تمیز، روش مناسبی نمی‌باشد. باید به طور هم‌زمان به نظر دانشجویان در ارتباط با کیفیت تدریس و سایر ویژگی‌های اثربخشی اساتید که منجر به دستیابی به اهداف آموزشی می‌شوند نیز توجه کرد. همچنین پیشنهاد دیگری که می‌توان در قالب یک روشی جدید ارائه داد تشکیل گروه‌هایی (پنل) از متخصصین "آن رشته" به منظور ارزشیابی کیفی سؤالات آزمون می‌باشد. چه بسا سؤالات یک آزمون به لحاظ کیفی (نه بر مبنای استفاده از شاخص دشواری) و نظر تخصصی در سطح دشواری مناسبی قرار داشته باشند اما بررسی کمی سؤالات به کمک دستگاه‌های تصحیح و آنالیز اوراق امتحانی نتایج متفاوت ارائه نماید. در اینجا می‌توان دستیابی به اهداف آموزشی را عامل این اتفاق قلمداد نمود و در واقع در این شرایط، استاد مربوطه شایسته تقدیر و تشکر می‌باشد حال آنکه با توجه صرف به تحلیل کمی، کیفیت آزمون ضعیف قلمداد می‌گردید و در ارزشیابی آموزشی استاد مربوطه تأثیر منفی به جای می‌گذاشت.

تشکیل تیمی از اساتید متخصص برای بررسی دشواری مناسب یک آزمون به لحاظ نظری و کیفی در کنار آنچه که به صورت کمی به دست می‌آید ملاک عادلانه‌تری برای قضاوت در اختیار گروه‌های آموزشی قرار می‌دهد.

اگرچه یک تعریف مورد قبول همگان از آزمون‌های ملاکی وجود ندارد، اما می‌توان گفت که این آزمون‌ها عموماً عملکرد یک فرد را با یک معیار مشخص مقایسه می‌کنند. به همین دلیل سؤالات یک آزمون ملاکی باید نمونه‌ای از دانش و رفتار در یک قلمرو معین باشند (۴). نتیجه آرمانی آزمون‌های وابسته به ملاک این است که در آن‌ها اکثر آزمون شونده‌گان اکثر سؤال‌ها را درست جواب دهند (۳). در واقع دسترسی به این اهداف آموزشی، در محدوده معین، غایت آموزش یک درس می‌باشد و به دست آوردن نمرات بالا در یک درس، در مجموع یک کلاس، می‌تواند نشانه‌ای از دستیابی به اهداف آموزشی باشد. هرچند روی دیگر آن را می‌توان در کیفیت پایین سؤالات جستجو کرد. در این زمینه می‌توان به کیفیت بالای تدریس یک استاد اشاره کرد که می‌تواند دسترسی به اهداف آموزشی را تسهیل نماید به نحوی که علی‌رغم طرح سؤال‌های با کیفیت بالا، سطوح بالای نمرات نیز قابل دسترس باشد.

در بحث آزمون‌ها می‌بایست به نکته‌ای ظریف توجه نمود: توزیع دانشجویان در یک کلاس.

آزمون هنجاری‌ای که دانشجویان برای قرار گرفتن در یک کد رشته/محل از آن عبور می‌کنند به عنوان یک غربال، افرادی با ویژگی‌های تحصیلی نسبتاً مشابه را در کنار یکدیگر قرار می‌دهد. به این ترتیب در نظر گرفتن یک توزیع با کشیدگی (غیرنرمال) که بیشتر دانشجویان را (در طول تحصیل) در یک طرف توزیع مشاهده نماییم منطقی‌تر از یک توزیع نرمال به نظر می‌رسد. به لحاظ نظری، رشته و مقطع تحصیلی را نیز می‌توان از عواملی در نظر گرفت که در توزیع نمرات در یک کلاس درس تأثیرگذار باشند. در ارتباط با رشته‌هایی که حضور در آن‌ها نیازمند قرار گرفتن متقاضی در طبقات بالای عملکردی آزمون هنجاری ورودی است، یا در ارتباط با دانشجویان تحصیلات تکمیلی به واسطه عبور موفق از آزمون‌های ملاکی و هنجاری بیشتر، با واقعیتی تحت عنوان "تجمع نخبگان" روبه‌رو هستیم. این رشته‌ها و مقاطع تحصیلی تکمیلی علی‌القاعده دربرگیرنده نخبگان جامعه دانشجویی خواهند بود و در شرایطی که در یک کلاس درس افرادی نخبه

### **References**

- 1- Toodehzaiem M. H, Daneshkazemi A. R, Aghili H, Haerian A, Mahdavi F. Evaluation of face-validity of multiple choice questions in special lessons of dentistry at Shahid Sadoughi University of Medical Sciences. *Journal of Medical Education and Development*. 2012; 7(2): 55-65.
- 2- karimi movasaghi H, Binaghi T, Abdolahi H. Survey of effect of changing answers on multiple choice test by students on their scores. *Future of Medical Education Journal*. 2009; 2(4): 3-5.
- 3- Saif A.A. *Educational Measurement, Assessment and Evaluation*. 3rd ed. Tehran: Dowran. 2005: 326.
- 4- LotfAbadi H. *Assessment and Measurement in Education and Psychology*. Tehran: Samt. 2008: 59.