

آگاهی و عملکرد اعضای هیأت علمی در سنجش روایی و پایایی سؤال های امتحانی

شهناز نجار* پروین عابدی*

Knowledge and practice of faculty members in assessing the validity and reliability of examinations

Sh.Najar P.Abedi

*Abstract

Background: Evaluation is one of the important aspects of teaching activity; and makes it possible to detect the positive and negative aspects of teaching. So evaluation is very important for improvement of quality of higher education in university.

Objective: Determining the knowledge and practice of faculty members about validity and reliability of exams.

Methods: In this study; 100 faculty members in Ahvaz medical university were selected randomly. Data were gathered via a questionnaire which included 22 questions about validity and reliability. 70 questionnaires returned after 2 months. We used descriptive statistical method for analyzing data.

Findings: 80% of subjects knew about content validity; 42.8% about surface validity; 25% about construct validity and 41.5% about possession validity.

41.5% of subjects had knowledge about test-retest; 35.7% about parallel; 42.9% about halves split and 24.2% about koudier-Richardson reliability index. In assessment of practice 45.7% of subjects were using content validity; and 20% test-retest; for validity and reliability of exams. But about other techniques of determining validity and reliability, subjects seemed to fail using them correctly.

Conclusion: Knowledge of faculty members about validity and reliability of exams is low; and using of various kinds of validity and reliability in exams is very low too.

Key words: Measurement, Exam Question, Teaching Assessment

*چکیده

زمینه: ارزشیابی یکی از بخش های مهم فرآیند آموزشی است و این امکان را فراهم می سازد تا بر اساس نتایج آن، نقاط قوت و ضعف را مشخص نمود. لذا نقش خطیری در ارتقاء و تضمین کیفیت آموزش عالی ایفا می کند.

هدف: مطالعه با هدف ارزیابی آگاهی و عملکرد اعضای هیأت علمی در مورد چگونگی سنجش اعتبار و پایایی سؤال های امتحانی انجام شد.

مواد و روش ها: این مطالعه در سال ۱۳۸۰ بر روی ۱۰۰ نفر از اعضای هیأت علمی دانشگاه علوم پزشکی اهواز انجام شد که به صورت تصادفی انتخاب شدند، پرسش نامه ای شامل ۲۲ سؤال میان افراد توزیع و پس از ۲ ماه، ۷۰ پرسش نامه جمع آوری شد. اطلاعات به دست آمده با نرم افزار SPSS تجزیه و تحلیل شد.

یافته ها: ۸۰٪ افراد از روش اعتبار محتوی، ۴۱/۵٪ از روش اعتبار ملاکی، ۴۲/۸٪ از اعتبار صوری و ۲۵٪ از اعتبار سازه جهت تعیین اعتبار سؤال های امتحانی آگاهی داشتند. درخصوص آگاهی از روش های تعیین پایایی ۵۴/۱٪ از پایایی باز آزمایی، ۳۵/۷٪ از پایایی موازی یا همتا ۴۲/۹٪ از پایایی به روش دونیمه کردن و ۲۴/۲٪ از پایایی به روش کودر ریچاردسون آگاهی داشتند. در بررسی عملکرد افراد در خصوص تعیین اعتبار ۴۵/۷٪ اظهار داشتند که از طریق محتوی، ۱۰٪ ملاکی، ۵/۷۲٪ صوری و ۱/۵٪ از طریق محتوی و صوری اعتبار سؤال ها را می سنجند و در خصوص تعیین پایایی ۲۰٪ اظهار کردند که از طریق باز آزمایی، ۱۲/۸۵٪ موازی یا همتا، ۷/۱۵٪ دونیمه کردن و ۲/۸۵٪ کودر ریچاردسون پایایی سؤال های امتحانی را تعیین می کنند.

نتیجه گیری: با توجه به یافته ها، درصد کمی از افراد مورد مطالعه از شیوه های انجام پایایی و روایی سؤال های امتحانی آگاهی داشتند و عملکرد آنها نیز در این مورد ضعیف بود. لذا باید با برگزاری کارگاه های اختصاصی در این زمینه آگاهی اعضا را افزایش داد.

کلید واژه ها: سنجش آموزشی، پرسش های امتحانی، تدریس

*** مقدمه :**

نسبت به روایی آن نیز افزوده خواهد شد.^(۵) بنابراین یک آزمون مناسب که وسیله ارزیابی کار معلم و فراگیر محسوب می شود باید از پایایی و روایی قابل قبولی نیز برخوردار باشد. با توجه به اهمیت ارزیابی در امر آموزش این مقاله با هدف تعیین چگونگی سنجش اعتبار و پایایی سؤال های امتحانی توسط اعضای هیأت علمی انجام شد.

*** مواد و روش ها :**

این مطالعه مقطعی در سال ۱۳۸۰ بر روی ۱۰۰ نفر از اعضای هیأت علمی دانشگاه علوم پزشکی اهواز انجام شد که به صورت تصادفی از میان ۴۰۰ عضو هیأت علمی انتخاب شدند. سپس پرسش نامه تهیه شده که حاوی ۲۲ سؤال بود میان این افراد توزیع شد. بخش اول پرسش نامه مربوط به اطلاعات دموگرافیک، بخش دوم در مورد سنجش آگاهی از روایی و پایایی سؤال های امتحانی و بخش سوم مربوط به عملکرد اعضای هیأت علمی در تعیین روایی و پایایی سؤال های امتحانی بود. پس از دو ماه تعداد ۷۰ پرسش نامه جمع آوری گردید. داده های به دست آمده پس از رمزگذاری با استفاده از نرم افزار SPSS استخراج و با آمار توصیفی تجزیه و تحلیل شدند.

*** یافته ها :**

۳۹ نفر از افراد مورد مطالعه مذکر و بقیه مؤنث بودند. اکثر افراد (۵۷/۱ درصد) رتبه علمی مربی داشتند و میانگین سابقه تدریس افراد ۱۱/۳ سال بود. متوسط واحد تدریس شده در طول نیمسال ۹ واحد بود. ۵۷ نفر از افراد مورد مطالعه در کارگاه سنجش و ارزیابی شرکت کرده بودند. در مورد روش های تعیین اعتبار، ۸۰ درصد افراد از اعتبار محتوا آگاهی داشتند (جدول شماره ۱).

در مورد آگاهی از روش های تعیین پایایی، ۴۲/۹ درصد از پایایی به روش دو نیمه کردن و تنها ۲۴/۲ درصد با روش کودریچاردسون آشنایی داشتند (جدول شماره ۲).

ارزشیابی بخش جدایی ناپذیر فرآیند آموزش به شمار می رود و چنانچه بر اساس اصول و موازین علمی و به طور مستمر انجام گیرد، می توان آن را مهم ترین رکن آموزش به شمار آورد.^(۱) ارزشیابی در آموزش، فرآیندی نظام دار برای جمع آوری، تحلیل و تفسیر اطلاعات است که نشان می دهد اهداف مورد نظر تحقق یافته اند یا در حال تحقق هستند و به چه میزانی تحقق یافته اند. از طرف دیگر ارزشیابی پیشرفت تحصیلی عبارت است از فرآیندی منظم برای تعیین و تشخیص میزان پیشرفت فراگیران در رسیدن به اهداف آموزشی که کیفیت روش تدریس معلم و میزان مهارت وی را در این زمینه مورد ارزیابی قرار می دهد.^(۲)

بدیهی است اندازه گیری صحیح، پایه و اساس ارزشیابی درست را تشکیل می دهد. مهم ترین وسیله اندازه گیری پیشرفت تحصیلی، آزمون است و هر آزمون باید از ویژگی هایی مانند عینیت، سهولت اجرا، روایی و پایایی برخوردار باشد. دقت در ساخت آزمون به افزایش اعتبار ارزشیابی منجر می شود. روایی یا اعتبار به هدفی که آزمون برای تحقق بخشیدن به آن درست شده است، اشاره می کند. آزمونی دارای روایی یا اعتبار است که برای اندازه گیری آنچه که مورد نظر است مناسب باشد. انواع روایی عبارت اند از: روایی محتوایی، ملاکی، صوری و سازه که روایی ملاکی بر دو نوع است: روایی پیش بینی و همزمان.

پایایی یا اعتماد یک وسیله اندازه گیری به دقت آن اشاره می کند. آزمون در صورتی دارای پایایی است که اگر آن را در فاصله زمانی کوتاه چندین بار به گروه واحدی از افراد بدهیم نتایج حاصل به هم نزدیک باشند.^(۳) برای محاسبه ضریب قابلیت اعتماد آزمون ها می توان از روش های بازآزمایی، موازی، دو نیمه کردن و کودریچاردسون استفاده نمود.^(۴)

به طور کلی اعتبار و پایایی از همدیگر جدا نیستند و هر چه به پایایی یک آزمون افزوده شود به همان

جدول ۱- فراوانی اعضای هیات علمی بر حسب آگاهی از انواع روش های تعیین اعتبار سؤال های امتحانی

جمع		بدون پاسخ		اشتباه		صحیح		نوع اعتبار
درصد	تعداد	درصد	تعداد	درصد	تعداد	درصد	تعداد	
۱۰۰	۷۰	۴/۳	۳	۱۵/۷	۱۱	۸۰	۵۶	محتوا
۱۰۰	۷۰	۷/۱	۵	۵۱/۴	۳۶	۴۱/۵	۲۹	ملاکی
۱۰۰	۷۰	۱۴/۳	۱۰	۶۰	۴۲	۲۵/۷	۱۸	سازه
۱۰۰	۷۰	۱۷/۲	۱۲	۴۰	۲۸	۴۲/۸	۳۰	صوری

جدول ۲- فراوانی اعضای هیات علمی بر حسب آگاهی از انواع روش های تعیین پایایی سؤال های امتحانی

جمع		بدون پاسخ		اشتباه		صحیح		نوع پایایی
درصد	تعداد	درصد	تعداد	درصد	تعداد	درصد	تعداد	
۱۰۰	۷۰	۲۰	۱۴	۳۸/۵	۲۷	۴۱/۵	۲۹	بازآزمایی
۱۰۰	۷۰	۲۷/۱۵	۱۹	۳۷/۱۵	۲۶	۳۵/۷	۲۵	موازی یا همتا
۱۰۰	۷۰	۲۴/۲	۱۷	۳۲/۹	۲۳	۴۲/۹	۳۰	دو نیمه کردن
۱۰۰	۷۰	۵۱/۶	۳۶	۲۴/۲	۱۷	۲۴/۲	۱۷	کودرریچاردسون

۲۰ درصد افراد از روش بازآزمایی در مورد چگونگی عملکرد اعضای هیات علمی در تعیین پایایی سؤال ها استفاده می کردند (جدول شماره ۴).

در مورد چگونگی عملکرد، اکثر افراد (۴۵/۷ درصد) از روش اعتبار محتوا برای تعیین اعتبار سؤال های امتحانی استفاده می کردند (جدول شماره ۳).

جدول ۴- فراوانی چگونگی عملکرد اعضای هیات علمی در تعیین پایایی سؤال های امتحانی

درصد	تعداد	فراوانی نوع پایایی
۱۲/۸۵	۹	موازی یا همتا
۲۰	۱۴	بازآزمایی
۷/۱۵	۵	دو نیمه کردن
۲/۸۵	۲	کودرریچاردسون
۱۷/۱۵	۱۲	هیچ کدام
۳۲/۸۵	۲۳	بدون پاسخ
۷/۱۵	۵	نامفهوم
۱۰۰	۷۰	جمع

جدول ۳- فراوانی چگونگی عملکرد اعضای هیات علمی در تعیین اعتبار سؤال های امتحانی

درصد	تعداد	فراوانی نوع اعتبار
۴۵/۷	۳۲	محتوا
۱۰	۷	ملاکی
۵/۷۲	۴	سازه
۱/۵	۱	صوری
۱/۵	۱	صوری و محتوا
۴/۲	۳	هیچ کدام
۲۴/۲	۱۷	بدون پاسخ
۷	۵	نامفهوم
۱۰۰	۷۰	جمع

در مورد میزان تأثیر تعیین روایی و پایایی سؤال های امتحانی بر فرآیند ارزشیابی، ۴۲/۸ درصد افراد پاسخ دادند که این تأثیر زیاد است. تقی پور معتقد است در بررسی یک وسیله ارزشیابی، اکتفا به عینی بودن آن کافی نیست، بلکه باید دید که آیا نمونه های رفتاری مورد ارزشیابی در وسیله مذکور به حد کافی منظور شده اند یا نه.^(۹)

بررسی عملکرد افراد در مورد سنجش اعتبار سؤال های امتحانی نشان داد ۴۵/۷ آنها از طریق روایی محتوا و بقیه از سایر روش ها استفاده می کردند. بنابراین علی رغم آن که ۸۰ درصد افراد از اعتبار محتوای آگاهی داشتند، فقط نیمی از آنها از روش مذکور استفاده می نمودند.

در بررسی عملکرد افراد در خصوص تعیین پایایی سؤال های امتحانی، ۲۰ درصد از روش بازآزمایی و بقیه از سایر روش ها استفاده می کردند. مقایسه میزان آگاهی افراد از تعیین پایایی و عملکرد آنان نشان داد تنها نیمی از افرادی که از پایایی بازآزمایی آگاهی داشتند، از این روش استفاده می کردند.

در نظرسنجی افراد در مورد علت عدم انجام پایایی و روایی سؤال های امتحانی، ۴۳/۲ درصد علت را عدم آشنایی با نحوه تعیین روایی و پایایی سؤال های امتحانی و ۱۸/۶ درصد علت را عدم درخواست مسؤولین از اعضای هیأت علمی می دانستند. ۴۷/۱ درصد نمونه ها اظهار داشتند که برگزاری کارگاه سنجش و ارزشیابی و ۱۲/۸ درصد با تشکیل کمیته ارزشیابی دانشکده ها می توان اعضای هیأت علمی را به انجام اعتبار و پایایی سؤال های امتحانی ملزم نمود. البته علی رغم آن که اکثر اساتید در کارگاه سنجش و ارزشیابی شرکت نموده بودند، متأسفانه عملکرد آنها خوب نبود. در این جا سه عامل مهم به عنوان علت به نظر می رسد: نخست آن که کارگاه سنجش و ارزشیابی به اندازه مورد نیاز به مطلب مورد نظر اشاره نداشته است و تفهیم مطالب به طور کامل صورت نگرفته

در مورد علت عدم انجام اعتبار و پایایی سؤال های امتحانی، ۳۲/۲ درصد افراد ذکر کردند که با نحوه تعیین اعتبار و پایایی آشنایی ندارند و ۴۷/۵ درصد افراد ذکر کردند که تشکیل کارگاه سنجش و ارزشیابی می تواند در این خصوص به آنها کمک کند.

* بحث و نتیجه گیری :

یافته های این مطالعه نشان داد که افراد بیش تر با اعتبار محتوا (۸۰ درصد) و از همه کمتر با اعتبار سازه آشنایی داشتند. گروس معتقد است سنجش اعتبار مشکل است و یک راه عملی برای تمرکز بر اعتبار محتوا این است که آیا سؤال ها نمونه کاملی از دانش و مهارت های تدریس شده در طول نیمسال تحصیلی هستند یا نه؟ در واقع سؤال ها باید بر اساس اهمیت سخنرانی ها یا تدریس های معلم طرح شود.^(۶) بارل نیز اظهار می دارد که مهم ترین خصوصیت یک آزمون، اعتبار یا میزان تناسب، معنی دار بودن و فایده تفسیر سؤال های آن آزمون در کاربرد مورد نظر از نتایج آن است.^(۹)

در بررسی پاسخ های افراد به سؤال های سنجش آگاهی از انواع پایایی نشان داد که افراد بیش تر با پایایی به روش دو نیمه کردن و کمتر از همه با پایایی به روش کودریچاردسون آشنایی داشتند. بارل به نقل از گران لوند می نویسد « زمانی به پایایی خیلی قوی نیاز داریم که تصمیم گیری ما مهم، نهایی، غیر قابل برگشت، غیر قابل انعطاف، مربوط به افراد و دارای نتایج دراز مدت باشد».^(۷) ۶۴/۲۹ درصد افراد مطالعه معتقد بودند که تعداد سؤال ها بر پایایی امتحان تأثیر دارد. یتربی می نویسد هر اندازه طول آزمون بیش تر شود، نمره های آن از ثبات و پایایی بیش تری برخوردار است و در واقع امتحان بسیار کوتاه، پایایی خوبی ندارد.^(۸)

۴. سرمد زهره، بازرگان عباس، حجازی الهه. روش های تحقیق در علوم رفتاری. تهران، انتشارات آگاه، ۱۳۷۶، ۱۵

5. Tochim, William M,K. Reliability and validity. Medical Teacher J 2001 Jan; 55

6. Gross Davis Barbara. Tools for teaching. Medical Education J 1999 Oct; 78

7. Burrell B et al. Comprehensive assessment in special education. lippincott co 1990; 123

۸. یثربی رقیه. تغییرات افزایش طول آزمون بر پایایی آن. مجله علوم تربیتی و روانشناسی، شماره دوم، ۱۳۷۶، ۳۹-۵۰

۹. تقی پور ظهیری علی. مقدمه بر برنامه ریزی آموزشی و درسی. تهران، انتشارات آگاه، ۱۳۷۵، ۵۵

است؛ دوم این که لازم است کارگاه تکرار شود و سوم این که مسؤولین دانشگاه و دانشکده ها باید بیش از پیش به این موضوع بها دهند.

* مراجع :

۱. عابدی احمد. ارزشیابی سؤال های هماهنگ دروس از متوسطه تا نظری در امتحان های خرداد ۷۳ وزارت آموزش و پرورش. اداره آموزش و پرورش استان اصفهان، ۱۳۷۲، ۷۵

۲. شعبانی حسن. مهارت های آموزش و پرورش، روش ها و فنون تدریس. تهران، انتشارات سمت، ۱۳۷۴، ۱۲۳

۳. سیف علی اکبر. روش های اندازه گیری و ارزشیابی آموزشی. تهران، انتشارات دوران، ۱۳۷۵، ۱۱۶