



کارگروه تخصصی مدیریت جامع کیفیت آموزش علوم پزشکی
با همکاری مرکز مطالعات و توسعه آموزش
دانشگاه علوم پزشکی جندی شاپور اهواز

تهیه و تنظیم

دکتر فاطمه احمدی (مدیر مرکز مطالعات و توسعه آموزش)

دکتر نسرين خواجه علي (دبیر کارگروه دانشگاهی مدیریت جامع کیفیت آموزش)

سارا حیدری (دانشجوی دکتری تخصصی آموزش پزشکی)

مهر ماه ۱۴۰۱

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ

جامع کیفیت آموزش

کارگروه

"تحلیل آزمون"

مراحل تحلیل آزمون

بررسی جدول دو بعدی آزمون

- آیا نمونه گیری مناسبی از محتوا انجام شده است؟
- آیا نمونه گیری مناسبی از اهداف انجام شده است؟
- آیا اهداف یا تاکسونومی های مختلف پوشش داده شده اند؟
- آیا ابزار و روش مناسب انتخاب شده است؟

واریسی تک تک سؤالات و تعیین میزان به کارگیری اصول صحیح طراحی سؤالات و عدم اشکالات علمی سؤالات

- استفاده از چک لیست میلمن
- استفاده از چک لیست سؤالات تشریحی و کوتاه پاسخ

بررسی نحوه اجرای آزمون

- انتخاب زمان مناسب
- انتخاب مکان مناسب
- رعایت اصول پیشگیری از تقلب
- ...

بررسی نمره دهی آزمون

بلوپرینت یا جدول دو بعدی آزمون

برای یک دوره آموزشی به صورت بالقوه تعداد سؤالات زیادی می توان در نظر گرفت اما هنگام برگزاری آزمون از نظر اجرایی با محدودیت تعداد سؤال مواجه هستیم و باید از بین سؤالات، تعداد مشخصی را انتخاب کنیم. مسئله مهم، انتخاب سؤالات به گونه ای است که به خوبی معرف دوره باشند.

بلوپرینت (Blueprint) یا جدول مشخصات آزمون به ما کمک می کند تا قبل از اینکه وارد فرایند طرح سؤال شویم، استراتژی خود را برای انتخاب نمونه خوبی از سؤالات مشخص کنیم و در واقع مسیر ارزیابی را روشن کنیم. این کار مخصوصاً در دوره هایی که چندین مدرس وجود دارد، واقعا ضروری است اما برای آزمون های دیگر نیز شدیداً توصیه می شود.

اولین اقدامی که برای داشتن نمونه معرف و خوب به نظر می رسد این است که سؤالات از «تمام» مباحث ارائه شده در طول دوره طرح شود. در جدول زیر نمونه ای از بلوپرینت آزمون کورس قلب را برای دانشجویان پزشکی مشاهده می کنید که شامل ۱۲ جلسه بوده و کل آزمون حاوی ۶۰ سؤال است. این مثال ساده ترین شکل بلوپرینت است که تعداد سؤالات هر مبحث از تعداد جلسات آن مشخص می شود.

جدول ۱. بلوپرینت بر اساس تعداد جلسه

عنوان جلسه	تعداد جلسه	تعداد سؤال
فشار خون	۲	۱۰
آریتمی	۲	۱۰
نارسایی	۱	۵
ایسکمی	۳	۱۵
دریچه ای	۲	۱۰
سایر	۲	۱۰
کل	۱۲	۶۰

بلوپرینت همچنین به ما کمک می کند ابزار مناسب برای آزمون را انتخاب کنیم. به عنوان مثال، در صورتی که هدف آزمون، سنجش بُعد دانشی فراگیران است، از انواع آزمون های کتبی استفاده خواهد شد و چنانچه این انتظار می رود که دانشجو علاوه بر کسب دانش، بتواند اصول فراگرفته شده را در عمل نیز به کار گیرد، در این صورت استفاده از انواع آزمون های مبتنی بر محیط کار را می توان توصیه کرد. جدول ۲ مثالی از بلوپرینت و ابزارهای مناسب برای هر مبحث را نشان می دهد.

به صورت کلی بلوپرینت می تواند برای یک ابزار مشخص تدوین شود یا برای دوره ای که شامل ابزارهای متنوع است. مزیت استفاده از بلوپرینت این است که انتخاب ابزارهای مختلف را به صورت معنادار و در قالب یک چارچوب مفهومی برای ما مشخص می کند. پس از تهیه بلوپرینت آزمون، هر عضو هیات علمی برای هر فعالیت مشخص شده، سؤالات آزمون را طراحی می کند.

جدول ۲. بلوپرینت و تعیین ابزار ارزیابی مناسب

سیستم	مشکل یا بیماری	اپیدمیولوژی	شرح حال	معاینه	تشخیص	پروسیجر	درمان	پروگنوز	ارتباطی
قلب	درد قفسه سینه		OSCE		KF				
	نارسایی قلبی						MCQ		
تنفس	آمبولی						MCQ		
	آسم		OSCE		OSCE				OSCE
گوارش	خونریزی فوقانی		OSCE			DOPS	KF		
	التهابی روده	MCQ						MCQ	
عصبی	سکته مغزی				OSCE			MCQ	
	صرع		MCQ			KF		MCQ	

مرکز مدیریت جامع کیفیت آموزش

آنالیز کیفی سوالات آزمون

جدول ۳. آنالیز کیفی سوالات چند گزینه ای

ردیف	فهرست واریسی	بله	خیر
		+۱	-۱
۱	آیا سؤال به روشنی موضوع مورد نظر را بیان می کند؟		
۲	موضوع اصلی سؤال را در تنه سوال نوشته اید؟		
۳	مطالب تکراری گزینه ها را در تنه سؤال قرار داده اید؟		
۴	سؤال یک موضوع مهم را اندازه می گیرد؟		
۵	گزینه های سؤال با هم تجانس دارند؟		
۶	در سؤال فقط یک گزینه درست یا گزینه که از همه بهتر است وجود دارد؟		
۷	در سؤالی که برای سنجش فرآیندهای پیچیده ذهنی نوشته اید موقعیت های جدید را مطرح کرده اید؟		
۸	گزینه های انحرافی آزمون شوندگان بی اطلاع از موضوع را به خود جلب می کند؟		
۹	گزینه های سؤال از نظر دستوری و جمله بندی با تنه سؤال مطابقت دارد؟		
۱۰	از کاربرد اشاره های نامربوط خودداری کرده اید؟		
۱۱	در سؤال منفی زیر عبارات منفی خط کشیده اید؟		
۱۲	از کاربرد منفی مضاعف در نوشتن سؤال ها خودداری کرده اید؟		
۱۳	از کاربرد عباراتی چون: همه آن ها، همه آن چه در بالا گفته شد، خودداری کرده اید؟		
۱۴	سؤال مستقل از سؤال های دیگر است؟		
۱۵	آیا از طرح سؤال به شکل گمراه کننده (که شکل و جواب مختلفی دارند) اجتناب کرده اید؟		
۱۶	از کاربرد دو گزینه متضاد که یکی از آن ها درست است خودداری کرده اید؟		
۱۷	از کاربرد دو گزینه هم معنی خودداری کرده اید؟		
۱۸	محل گزینه درست را در میان گزینه های انحرافی به طور تصادفی انتخاب کرده اید؟		
۱۹	آیا طول گزینه درست با سایر گزینه ها متفاوت است؟		
۲۰	سؤال در حد درک و فهم آزمون شوندگان هست؟		
۲۱	گزینه، هیچ یک از موارد فوق، را مورد استفاده قرار نداده اید؟		
۲۲	جای خالی را در انتخاب تنه سؤال قرار داده اید؟		
	جمع امتیازات		
	درصد امتیازات		

جدول ۴. آنالیز کیفی سوالات تشریحی

ردیف	فهرست واریسی	بله	خیر
		+۱	-۱
۱	آیا سؤال به روشنی موضوع مورد نظر را بیان می کند؟		
۲	موضوع اصلی سؤال را در تنه سؤال نوشته اید؟		
۳	بارم در نظر گرفته برای سؤال مناسب است؟		
۴	زمان آزمون برای پاسخ به این سؤال مناسب است؟		
۵	منبع طرح سؤال مشخص است؟		
۶	پاسخ به این سؤال هدف اختصاصی مشخصی را ارزیابی می کند؟		
۷	پاسخ به این سؤال حیطه یادگیری مشخصی را ارزیابی می کند؟		
۸	ضرورت یادگیری موضوع سؤال مشخص است؟		
۹	سؤال مستقل از سؤال های دیگر است؟		
۱۰	آیا از طرح سؤال به شکل گمراه کننده (که شکل و جواب مختلفی دارند) اجتناب کرده اید؟		
۱۱	سؤال در حد درک و فهم آزمون شوندگان هست؟		
جمع امتیازات			
درصد امتیازات			

جامع کیفیت آموزش

تحلیل کمی آزمون

- محاسبه ضریب دشواری
- محاسبه ضریب تمیز
- تحلیل گزینه های انحرافی

ضریب دشواری (Difficulty index)

- ضریب دشواری که به صورت P نشان داده می شود و به آن P value نیز گفته می شود، برای یک آیتم (به عنوان مثال، یک سوال در آزمون چند گزینه ای یا یک ایستگاه در OSCE) تعریف می شود.
- ضریب دشواری، درصد آزمون شوندگانی است که به سوال مورد نظر جواب درست داده اند.
- می تواند مقداری بین صفر تا یک داشته باشد. به این ترتیب، هر چقدر تعداد دانشجویانی که به آیتم مورد نظر پاسخ صحیح داده باشند، بیشتر باشد، منطقاً آن آیتم آسان تر است در حالی که بنا به تعریف ضریب دشواری آن بزرگتر است.

تفسیر ضریب دشواری آزمون:

- ضریب دشواری: از صفر تا ۱
 - ضریب دشواری سوال بزرگتر: سوال آسان تر
 - ضریب دشواری کوچکتر: سوال سخت تر
 - حد مطلوب در آزمون: $0.7 - 0.3$
- 

ضریب تمیز (Discrimination Index)

- به صورت D نمایش داده می شود.
- قدرت سوال در تمایز گذاری یا تشخیص بین گروه قوی و گروه ضعیف آزمون شوندگان
- مشخص می کند که سوال تا چه اندازه می تواند گروه قوی را از گروه ضعیف جدا سازد.

تفسیر ضریب تمیز

- مقادیر این ضریب بین -1 تا $+1$ می باشد.
- حد مطلوب: نزدیک یک، هر چه بالاتر بهتر
- هر قدر ضریب تمیز سؤال بزرگتر و به صفر نزدیکتر باشد، به این معنا است که قوه تمیز آن سوال بیشتر است و سؤال دانشجویان قوی و ضعیف را بهتر از یکدیگر جدا می کند.
- هر قدر این ضریب کوچکتر باشد، قوه تمیز آن کمتر است.
- ضریب تمیز صفر: سوال به هیچ وجه نتوانسته بین گروه قوی و ضعیف تمایز قائل شود.

- ضریب تمیز منفی: ضریب تمیز می تواند مقادیر منفی نیز به خود بگیرد که این حالت به این معنا است که دانشجویان گروه پایین، بهتر به سؤال جواب داده اند. در آزمون ها این وضعیت مطلوب نیست و باید به دنبال علت آن بود.

یکی از مواردی که در سؤال با ضریب منفی به ذهن می رسد، این احتمال است که کلید سؤال اشتباه اعلام شده است. احتمال دیگر، وجود نکته انحرافی یا خطای ساختاری در سؤال است که موجب شده دانشجویان قوی برداشت اشتباهی داشته باشند و غلط پاسخ دهند. همچنین می تواند نشانه این موضوع باشد که دانشجویان طی دوره مبحث مورد نظر را به صورت اشتباه یاد گرفته اند.

مثال:

اگر ضریب تمیز سوالی ۰/۹۰ باشد، آن سوال آزمون شوندگان قوی و ضعیف را خیلی خوب از هم جدا خواهد کرد.
اگر ضریب ۰/۱۰ باشد، آن سوال از عهده جداسازی دانشجویان قوی و ضعیف به خوبی برنخواهد آمد.

تحلیل گزینه های انحرافی

هدف از قرار دادن گزینه انحرافی در سؤال ها منحرف کردن آزمون شوندگانی است که جواب درست سؤال را نمی دانند. در صورتی یک سؤال به خوبی عمل می کند که افراد ضعیف بیشتر از افراد گروه قوی گزینه های انحرافی آن سؤال را انتخاب نمایند.

هر گزینه انحرافی باید دست کم یک نفر از افراد ضعیف را به خود جلب کند.
اگر هر دو گروه قوی و ضعیف را به خود جلب کند تعداد افراد ضعیف باید بیشتر از افراد قوی باشد.

یک نمونه سؤال دشوار

سؤال: در گرفتن شرح حال یک کودک همه موارد زیر شبیه گرفتن شرح حال بزرگسالان است به استثناء:

الف- نمو

ب- تغذیه

ج- واکسیناسیون

د- معاینه گوش

DIFF-Index	DIFF-Index	RPbis	DISC-Index
۰/۲۳	دشوار	۰/۰۳	کم

یک نمونه سؤال آسان

سؤال: کلیه موارد زیر در ابتدای معاینه بیمار مورد نظر است بجز:

الف- اندازه گیری درجه حرارت

ب- اندازه گیری فشار خون

ج- اندازه گیری قند خون

د- اندازه گیری تعداد تنفس

DIFF-Index	DIFF-Index	RPbis	DISC-Index
۰/۹۹	آسان	۰/۱۱	کم

محاسبه ضریب دشواری و ضریب تمیز در آزمون چند گزینه ای

برای محاسبه ضریب دشواری سؤالات در آزمون چند گزینه ای، باید تعداد کل افرادی که به هر سؤال پاسخ صحیح داده اند، مشخص کنیم و بر تعداد کل دانشجویان تقسیم کنیم:

$$\text{ضریب دشواری سؤال} = \frac{\text{دانشجویانی که به سوال پاسخ صحیح داده اند}}{\text{تعداد کل دانشجویان}}$$

به منظور محاسبه ضریب تمیز در سؤالات چند گزینه ای، شرکت کنندگان باید بر اساس نمره ای که در کل آزمون کسب کرده اند، به دو دسته پایین (ضعیف) و بالا (قوی) تقسیم شوند. سپس باید مشخص شود که چه تعداد از هر دسته به آیت مورد نظر جواب درست داده اند. تفاضل این دو مقدار، ضریب تمیز را مشخص می کند:

$$\text{ضریب تمیز سؤال} = \frac{\text{انتخاب های درست گروه ضعیف} - \text{انتخاب های گروه قوی}}{\text{تعداد افراد یک گروه}}$$

آنچه در اینجا ذکر شد، برای آزمون هایی کاربرد دارد که جواب دو ارزشی دارند مانند سؤالات چند گزینه ای، به عبارت دیگر، دانشجویان در مواجهه با هر سؤال یا به آن پاسخ صحیح داده اند و تمام نمره را دریافت کرده اند یا پاسخ اشتباه داده اند و صفر شده اند. بنابراین همین روش برای تعیین ضریب دشواری و ضریب تمیز در یک آزمون با سؤالات درست-نادرست و جورکردنی گسترده نیز می تواند به کار رود.

مثال: در یک آزمون با شرکت ۱۲۰ دانشجو، توزیع فراوانی انتخاب هر یک از گزینه ها در سوال شماره ۵۰ که جواب صحیح آن گزینه ب است، به صورت زیر است. ضریب دشواری سوال را حساب کنید.

گزینه الف: ۳۰ نفر

گزینه ب: ۶۵ نفر

گزینه ج: ۱۵ نفر

گزینه د: ۱۰ نفر

$$\text{ضریب دشواری سؤال} = \frac{\text{دانشجویانی که به سوال پاسخ صحیح داده اند}}{\text{تعداد کل دانشجویان}} = \frac{65}{120} = 0/54$$

مثال: در آزمون دستیاری ۲۰۰۰ داوطلب شرکت کردند. بر اساس نمره کل، آنها را به دو دسته قوی و ضعیف تقسیم می کنیم. از دسته قوی، ۷۵۰ نفر و از دسته ضعیف ۵۰ نفر به سوال شماره ۲۰ پاسخ صحیح داده اند. ضریب تمیز این سوال را حساب کنید.

$$\text{ضریب تمیز سؤال} = \frac{\text{انتخاب های درست گروه ضعیف} - \text{انتخاب های گروه قوی}}{\text{تعداد افراد یک گروه}} = \frac{750 - 50}{1000} = 0/7$$

محاسبه ضریب دشواری و ضریب تمیز در آزمون تشریحی و عملکردی

در آزمون هایی که پاسخ به صورت صفر و یک نیست، قضیه کمی متفاوت است. آزمون تشریحی، سؤالات شفاهی، OSCE، Mini CEX و DOPS از این دست سؤالات هستند که نمره ای که می تواند به پاسخ آن ها تعلق بگیرد، به صورت یک محدوده نمره است.

• روش ویتنی و سیبرز

$$\text{ضریب دشواری} = \frac{\sum H + \sum L - (2N \text{ Score min})}{2N (\text{Score max} - \text{Score min})}$$

$$\text{ضریب تمیز} = \frac{\sum H - \sum L}{N (\text{Score max} - \text{Score min})}$$

$\sum H$: مجموع نمره های افراد یا برگه های گروه بالا

$\sum L$: مجموع نمره های افراد یا برگه های گروه پائین

N: تعداد افراد یا برگه های ۲۵ درصدی گروه مورد نظر

Score max: بالاترین نمره ممکن در سؤال

Score min: پائین ترین نمره ممکن در سؤال

مثال: ضریب دشواری و ضریب تمیز سؤال را با توجه به اطلاعات جدول زیر محاسبه کنید

گروه پایین		گروه بالا		نمره سوال
نمره	تعداد افراد	نمره	تعداد افراد	
۱۰	۲	۱۵	۳	۵
۱۲	۳	۲۴	۶	۴
۱۵	۵	۱۵	۵	۳
۱۲	۶	۴	۲	۲
۲	۲	۴	۴	۱
۰	۲	۰	۰	۰
۵۱	۲۰	۶۲	۲۰	مجموع

ضریب تمیز = ۰/۱۱

ضریب دشواری = ۰/۵۶

سوال تقریباً خوبی می باشد زیرا ضریب دشواری آن بین ۰/۷-۰/۳ می باشد و ضریب تمیز آن بزرگتر از صفر می باشد.