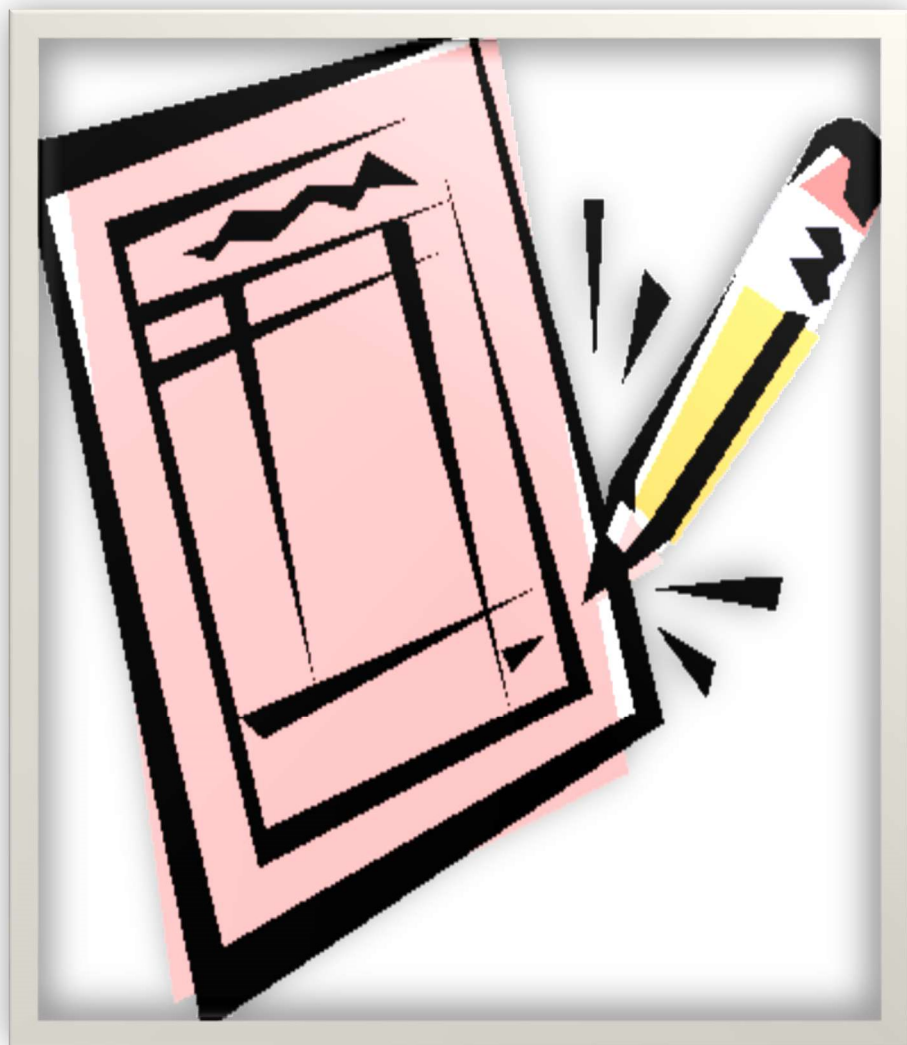


۱۳۸۹



استاندارد سازی در ارزشیابی دانشجو

درس ارزشیابی دانشجو (۲)

جلسه هفدهم

فهرست مطالب

اهداف کلی	۳
۱ اصول و انواع استاندارد	۳
۱,۱ چرا به استاندارد نیازمندیم؟	۳
۱,۲ روش درصد ثابت	۵
۱,۳ تفاوت استانداردهای هنجار مدار با استانداردهای معیار مدار	۵
۱,۴ تفاوت در استاندارد جبرانی و استاندارد مستقل در هر حیطة	۶
۲ روشهای استانداردسازی	۸
۲,۱ مدل های آزمون محور و آزمودنی محور	۸
۳ مدل‌های آزمودنی محور	۱۱
۳,۱ روش گروههای تقابلی (گروه مرزی)	۱۱
۳,۲ روش آنگوف	۱۴
۳,۳ آنگوف تعدیل شده	۱۵
۳,۴ ارزشیابی فرایند استانداردسازی و نتیجه گیری	۱۶
خلاصه	۱۸
منابع	۱۹

اهداف کلی

در تدوین این درس اهداف کلی زیر مورد نظر است.

انتظار می رود شما بعد از خواندن مطالب این نوشته به اهداف زیر نائل آید.

آشنایی با:

۱. استانداردهای هنجار مدار با استانداردهای معیار مدار
۲. استانداردهای جبرانی با استانداردهای غیرجبرانی
۳. روشهای استاندارد سازی (مدلهای آزمون محور و مدل‌های آزمودنی محور)

۱ اصول و انواع استاندارد

➤ مفاهیم کلیدی

۱. طراحان آزمون نیاز به یک ابزار آموزشی دارند که در نمره دهی داوطلبان باصلاحیت را از بدون صلاحیت جدا سازد.
۲. در یک آزمون نقطه برش^۱ نمره ای است که دانشجوی قوی را از دانشجوی ضعیف جدا می سازد.


➤ اهداف یادگیری

شما پس از مطالعه مطالب این بخش قادر خواهید بود:

۱. دلایل نیاز به استاندارد را شرح دهید
۲. تفاوت استانداردهای هنجار مدار با استانداردهای معیار مدار را با ذکر مثال توضیح دهید.
۳. استانداردهای جبرانی را با استانداردهای غیرجبرانی را با هم مقایسه نمایید.

۱،۱ چرا به استاندارد نیازمندیم؟

تعیین استانداردها برای سنجش عملکرد یک حیطة نسبتا جدید مطالعاتی است. هدف از این درس آشنا شدن با چارچوب، اهداف، مفاهیم کلیدی و ملاحظات عملکردی رویکردهای استانداردسازی آزمون ها و توانا ساختن افراد جهت بکار بردن مناسبترین رویکرد استاندارد سازی در آزمونها می باشد.

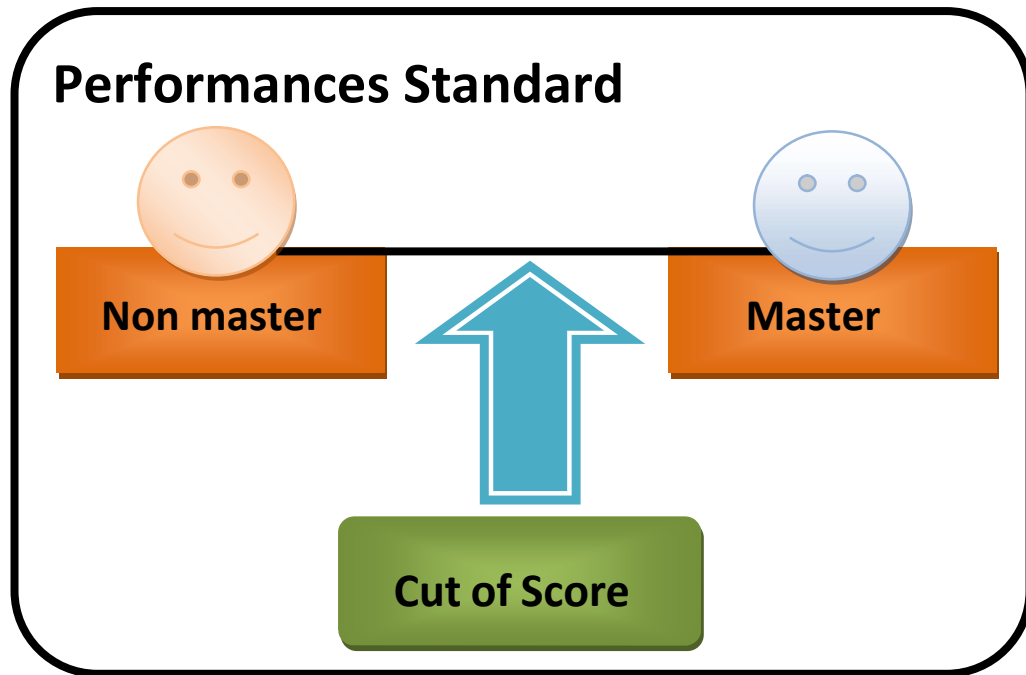


شناسه جستار
عنوان جستار: استاندارد سازی ارزشیابی دانشجو
نویسنده: لیلا بذرافکن
مری مرکز مطالعات و توسعه آموزش علوم پزشکی
آخرین تاریخ به روز رسانی: ۱۳۸۹/۹/۷
طراح آموزشی: دکتر مانوش مهربانی
کارشناس آموزشی معاونت آموزشی دانشگاه علوم پزشکی
ویرایش: مریم فخرزاد / زهرا صفاری
کارشناسان مرکز مطالعات و توسعه آموزش علوم پزشکی
با همکاری:
مرکز مطالعات و توسعه آموزش پزشکی
قطب علمی آموزش الکترونیکی
دانشگاه علوم پزشکی شیراز

¹ - Cut of point

❖ چرا ما بایستی روشهای استانداردسازی را بکار ببریم؟

مؤسسات دانشگاهی برای اعتباربخشی، و صدور گواهی در جستجوی رویکردهای نوآور جهت سنجش کفایت حرفه ای می باشند. تمرکز بر این اقدام نوآورانه یک نیاز جهت تعیین استانداردهای عملکرد است که داوطلبان با صلاحیت را از بدون صلاحیت جدا سازد . طراحان آزمون نیاز به یک ابزار آموزش دارند که در نمره دهی داوطلبان باصلاحیت را از بدون صلاحیت جدا سازد .



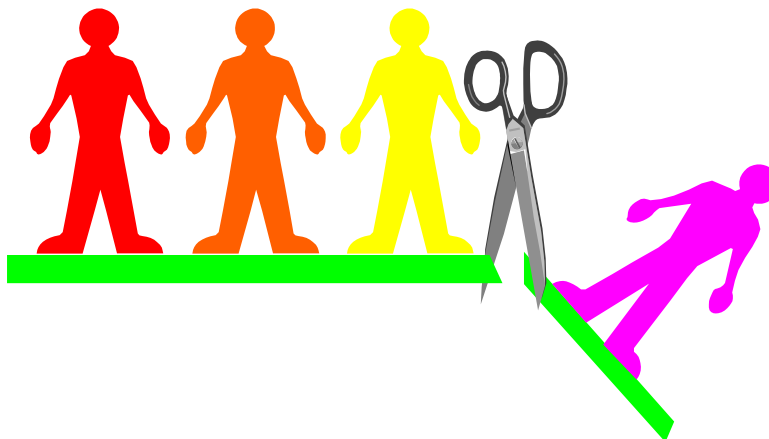
مشکل ما در اینجا این است که در استاندارد های سنتی ممکن است فردی با وجود شایستگی لازم مردود شود و فرد دیگری بدون صلاحیت لازم از آزمون ما نمره قبولی بگیرد

So what is the problem ?

	Competent	Incompetent
Pass		
Fail		

شکل شماره یک : مفهوم شایستگی و عدم شایستگی

در یک آزمون نقطه برش^۲ نمره ای است که دانشجوی قوی را از دانشجوی ضعیف جدا می سازد. اگر این مرز، قضاوتی فردی باشد در اغلب مواقع نشان دهنده ی واقعیت نیست. بدین منظور روشهای مختلفی ابداع شده است که حتی الامکان سهم قضاوت (به عبارت دیگر پیش داوری و خطاهای انسانی) را کم می کند. برای این کار روشهای زیر مطرح می شود:



۱,۲ روش درصد ثابت^۳

روشی که اغلب در موقعیت های آموزشی مورد استفاده است، داوری تخصصی استاد می باشد. استاد با محتوایی که باید در آن تسلط صورت گیرد و سطح عملکردی که برای موفقیت های بعدی در رابطه با هدفهای عینی ضروری است آشنا می باشد. مثلاً می داند که آیا ۷۰ درصد پاسخ صحیح نشانگر این است که فراگیر در رابطه با موارد دیگر آموزشی دچار مشکل شده یا آموزش جبرانی باید داشته باشد، یا اینکه ۷۰ درصد پاسخ صحیح برای پیشروی فراگیر و شروع کار با هدفهای جدید کافی است.

منتقدان این رویکرد، ذهنی بودن آن را تذکر داده و می گویند که اگر از معلم توضیح خواسته شود او ممکن است دلایل محدودی را برای تصمیمش در زمینه استانداردها بتواند ذکر کند. هنگامی که استاندارد برای همه هدفهای عینی مثلاً ۸۰ درصد به عنوان حداقل قابل قبول تعیین می شود، چنین تصمیم گیری هایی نشانه پاسخی بدون اندیشه و قراردادی است و نه یک تصمیم دقیق و ملاحظه شده که حداقل عملکرد قابل قبول را در هر یک از هدف های مختلف تعیین می کند. بنابراین لازم است تا انواع استاندارد را بررسی نماییم.

۱,۳ تفاوت استانداردهای هنجار مدار با استانداردهای معیار مدار

در رویکرد هنجار مدار استاندارد بر اساس عملکرد یک نمونه بزرگ خارجی (گروه هنجاری) است ، که برابر با داوطلبان شرکت کننده در آزمون است ، تعیین می شود. رویکرد هنجارمدار در سطح مدرسه یا موسسه آموزشی یک استاندارد نسبی^۴ در مورد عملکردهای متوسط داوطلبین تعیین می کند ، که با تعیین انحراف معیار از میانگین مشخص می گردد. این استانداردها ممکن است از سالی به سال دیگر بدلیل تغییر در توانایی گروه هنجاری تغییر کند.

^۲ - Cut of point

^۳ - percentage method

^۴ - Relative standard

➤ رویکرد معیار مدار

رویکرد معیار مدار استاندارد را به محتوای سطح کفایت که مورد انتظار است، ارتباط می دهد. این استاندارد بعنوان یک نقطه مطلق در نظر گرفته می شود که در حقیقت شامل دانش و مهارتهایی که یک دانشجو باید کسب نماید، تابتواند یک دوره را با موفقیت بگذراند، اطلاق می شود. معیار در این موقعیت عبارت از رفتارهای مطلوب و مورد انتظار است.

تفسیر ملاک- مرجع بیش از آن که تفسیری نسبی باشد تفسیری مطلق است که به مجموعه تعریف شده‌ای از رفتارهای فراگیر و یا همان طور که غالباً انجام می شود، به سطوحی معین از عملکرد فراگیر ارتباط دارد.

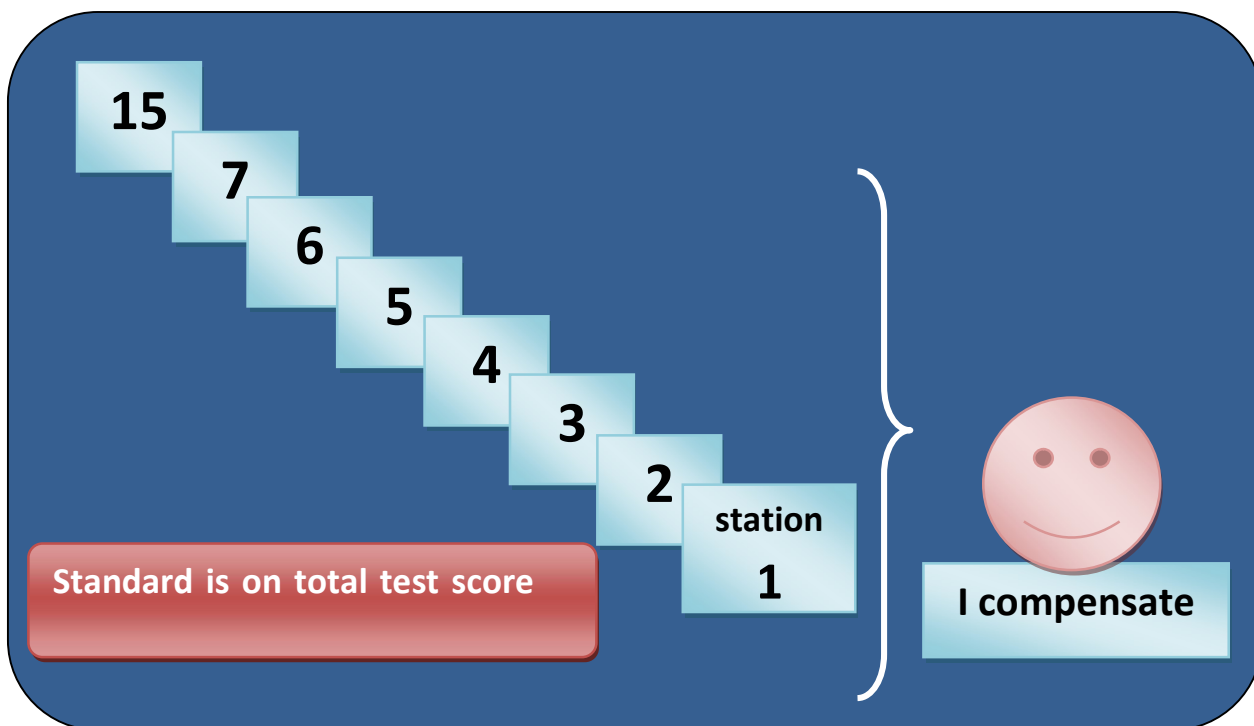
استاندارد ملاک- مرجع مستلزم این است که رفتارهای فراگیر قبل از تهیه آزمون مشخص شود. همانطور که میزان شکست یا موفقیت در آزمون ممکن است بدلیل تغییرات در توانایی گروه از واحد اجرایی به واحد دیگر متفاوت باشد، به عبارتی ممکن است تمام افراد استاندارد لازم را کسب کنند، یا هیچکدام از افراد به حد استاندارد دست نیابند.

۱،۴ تفاوت در استاندارد جبرانی^۵ و استاندارد مستقل^۶ در هر حیطه

در استاندارد جبرانی در محاسبه نمرات و امتیاز دانشجویان نمره کلی تست لحاظ می گردد. بنابراین داوطلبین در عملکرد نسبتاً بد در قسمتهایی از امتحان می توانند با عملکرد خوب در قسمتهای دیگر جبران کنند. در استاندارد مستقل و مجزا استانداردها برای قسمتهای مجزای امتحان لحاظ می گردند و داوطلبین عملکرد نسبتاً بد در یک قسمت از امتحان را نمی توانند جبران نمایند. پایایی اجزاء امتحان یک مشکل در استانداردهای مرتبط می باشد. و مهمترین مشکل این است که برای مثال در یک OSCE یک ایستگاه ممکن است پایایی لازم را نداشته باشد.

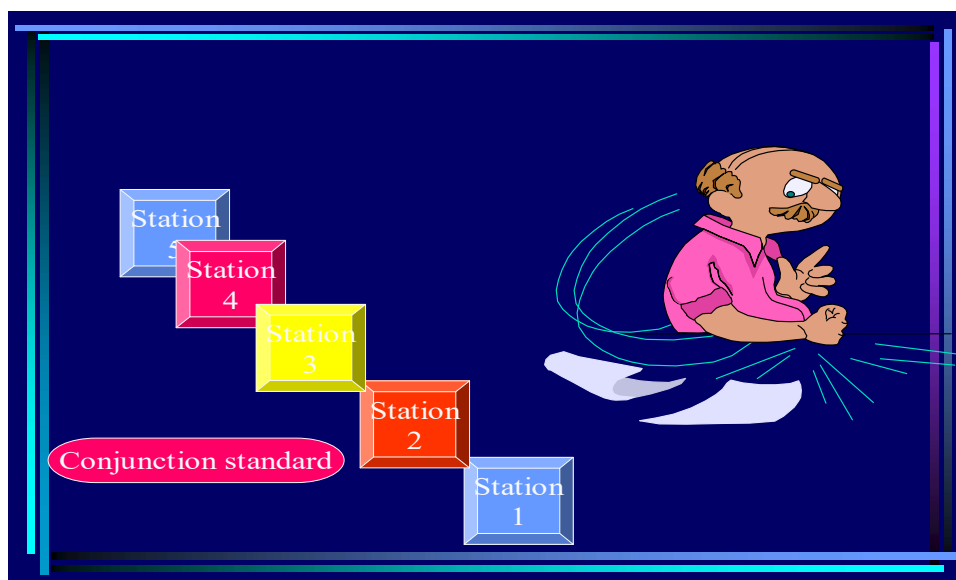
⁵ Compensatory standard

⁶ Conjunctive standard



شکل شماره دو : مفهوم استاندارد جبرانی

استاندارد جبرانی در گروه ایستگاه ها یا سوالاتی که کفایت های مشابه را سنجش می کنند ممکن است قابل قبول باشند . امتیاز مهم استاندارد مجزا، این است که بازخورد تشخیصی در مورد نقاط قوت و ضعف به داوطلبین می دهد . از آنجاییکه هر قسمت مهارت بطور مجزا در نظر گرفته می شود ضریب همبستگی بالاتر میان اجزای آزمون توافق بیشتر جهت استاندارد جبرانی را نشان می دهد . یک تست با همبستگی بالاتر اجزایش ممکن است یک سازه رایج یا بعدی از عملکرد را نشان دهد،



شکل شماره دو: مفهوم استاندارد مجزا

➤ آزمون

۱- در کدام یک از رویکردهای استاندارد مطلق مورد نظر است؟

- الف) هنجارمدار
ب) معیارمدار
ج) جبرانی
د) پیوسته

گزینه "ب" صحیح است.

۲- اگر در یک آزمون آسکی با ۱۰ ایستگاه در دو ایستگاه مردود گردد، از کدام نوع استاندارد استفاده شده است؟

- الف) جبرانی
ب) هنجاری
ج) میدانی
د) مستقل

گزینه "د" صحیح است.

۲ روشهای استانداردسازی

➤ مفاهیم کلیدی

۱. روش های استاندارد سازی به دو گروه آزمون محور و آزمودنی محور تقسیم می شود.

➤ اهداف یادگیری

شما پس از مطالعه مطالب این بخش قادر خواهید بود :

۱. مدل‌های آزمون محور را لیست نمایید.
۲. مدل‌های آزمودنی محور را لیست نمایید.
۳. کاربرد هریک از مدل های فوق را شرح دهید.

۲,۱ مدل های آزمون محور و آزمودنی محور

- داوری تخصصی جیگر^۷
- روش ندلسکی^۸
- رویکرد ایبل^۹
- روش گروههای تقابلی^{۱۰}
- روش آنگوف^{۱۱}

⁷ Jaeger
⁸ Nedelsky
⁹ Ebel
¹⁰ Contrasting groups
¹¹ Angoff method

در مدل‌های آزمون محور داوران از طریق مرور اجزا آزمون و فراهم آوردن قضاوت بعنوان تنها سطح مکفی جهت عملکرد این اجزا استانداردها را تعیین می نمایند. از جمله این روش ها میتوان به سیستم قضاوت جیگر نام برد . در این روش بر اهمیت تشخیص سطح کفایت جهت تعیین پیامدهای آزمونها را دارد در حقیقت تاکید جیگر بر امتحان دهندگان موفق است تا امتحان دهندگان لب مرزی با حداقل شایستگی.

❖ روش ندلسکی

این روش را ندلسکی (۱۹۵۴) برای استفاده در آزمونهای چند گزینه ای ساخته ، و برای تعیین یک استاندارد مناسب از حداقل عملکرد قابل قبول، اعمال زیر باید انجام شود:

۱. هیأتی از کارشناسان واجد صلاحیت در زمینه محتوا در نظر گرفته می شود.
 ۲. از هر کارشناس خواسته می شود که هر سؤال آزمون را بررسی کرده و هر پاسخ چند گزینه ای را که فکر می کند آزمایش شونده می تواند در حداقل شایستگی آن را حذف کند، حذف می کند.
 ۳. پس از حذف گزینه های بالا احتمال حدس زدن پاسخ صحیح را محاسبه می کنند. این کار یعنی تقسیم ۱ بر تعداد گزینه های باقی مانده (در هر سؤال) .
 ۴. مجموع این احتمالات در طول مجموعه سؤالات، معادل نمره مورد انتظار آزمایش شونده ای با حداقل شایستگی است که بوسیله آن کارشناس، تعیین شده است.
- سپس معدل نمرات مورد انتظار همه کارشناسان، استاندارد حداقل قابل قبول عملکرد خواهد بود.

به عنوان مثال، فرض کنید که یکی از کارشناسان آزمون چهار گزینه ای ۲۰ سؤالی را زمینه یابی کرده است. احتمال حدس زدن پاسخ در هر سؤال ۱ به ۴ یا $0/25$ خواهد بود. اگر این کارشناس احساس کند که یک دانش آموز با حداقل شایستگی خواهد توانست یکی از گزینه های غلط یک را حذف کند، در آن صورت احتمال حدس پاسخ صحیح سؤال شماره ۱ معادل $0/33$ خواهد بود. اگر در سؤال شماره ۲، دو گزینه اشتباه بتواند بوسیله یک دانشجو با حداقل شایستگی حذف شود، احتمال حدسی زدن پاسخ $0/5$ می شود. اگر همه گزینه های سؤال شماره ۳ مناسب تشخیص داده شود احتمال درست زدن پاسخ همان $0/25$ باقی خواهد بود. کارشناس مورد نظر همه سؤالات را به همین طریق درجه بندی کرده و احتمال حدس صحیح را در هر سؤال بوسیله دانش آموزی با حداقل شایستگی تعیین می نماید. کارشناس هر سؤال را به همین طریق درجه بندی کرده و احتمالی را که یک دانشجو با حداقل شایستگی می تواند سؤال را حدس بزند تعیین و سپس احتمالات مربوط به ۲۰ سؤال را با یکدیگر جمع می کند.

سایر کارشناسان نیز این کار را انجام داده و نمراتشان را به همین ترتیب تعیین می کنند. معدل نمرات همه کارشناسان به عنوان استاندارد حداقل عملکرد قابل قبول مورد استفاده قرار می گیرد. این روش و روش های دیگر در تعیین استانداردها به هیأتی از کارشناسان واجد صلاحیت نیاز دارد. این کار الزاماً بدین معنی نیست که این افراد کارشناسان شناخته شده مملکت باشند. آنها می توانند همکارانی باشند که با زمینه محتوایی و مشخصات و سطوح عملکرد معمولی دانش آموزان آشنایی دارند. معلمان با تجربه به خوبی می توانند این نقش را ایفا نمایند.

The Nedelsky(1957) approach

Items	Response options	MPL
1	A B C D E	0.5= 1/2
2	A B C D E	1/1 =1
3	A B C D E	1/3=0.33
4	A B C D E	1/5=0.2
5	A B C D E	=0.25 1/4
Total	5	2.28

شکل شماره چهار : روش ندلسکی در استاندارد سازی

❖ رویکرد ایبل ۱۲

داوران رویکرد ایبل را جهت طبقه بندی اجزا آزمون بکار می برند ، بدین صورت که در آن ابتدا تعدادی از طبقه بندی ها بر طبق درجه سختی و سطوح اهمیت سوال در آزمون تعیین می شود . بعد از طبقه بندی اجزا ، داوران در مورد نسبتی از سوالات که در هر طبقه بطور فرضی یک داوطلب با حداقل شایستگی میتواند پاسخ درست دهد را مشخص می کنند . سپس تعداد سوالات در هر تقسیم بندی را در نسبتی از سوالات که در هر طبقه بطور فرضی یک داوطلب با حداقل شایستگی میتواند پاسخ درست دهد ، ضرب میکنند و استاندارد حداقل به دست می آید .

Ebels method

	easy	moderate	hard
essential	1,4,7,8,13 95%	11,15,22 85%	21 80%
important	2,6,9 90%	10,14,20 75%	16,25 60%
acceptable	5 80%	12,18 55%	19,23 35%
doubtful	3 50%	- -%	17,24 20%

شکل شماره چهار : روش ایبل در استاندارد سازی

➤ آزمون

۱- بهترین روش استانداردسازی برای آزمون های چهار گزینه ای کدام است؟

- الف (انگوف
ب (گروه مقابل
ج (ندلسکی
د (ابل

گزینه "ج" صحیح است.

۲- معیارهای مورد استفاده داوران در روش ابل عبارتند از:

- الف (سختی و میزان اهمیت
ب (میزان سادگی و میزان مقبولیت
ج (میزان سختی و معیارمدار بودن
د (میزان سادگی و هنجارمدار بودن

گزینه "الف" صحیح است.

۳ مدل‌های آزمودنی محور

➤ مفاهیم کلیدی

۱. روش گروههای تقابلی در بسیاری از موارد استاندارد سازی کاربرد دارد.
۲. روش انگوف در استاندارد سازی آزمون های عملکردی کاربرد بیشتری دارد.

➤ اهداف یادگیری

- شما پس از مطالعه مطالب این بخش قادر خواهید بود :
۱. چگونگی روش گروههای تقابلی (مرزی) را در تعیین استاندارد توضیح دهید.
 ۲. تفاوت روش آنگوف را با روش آنگوف تعدیل شده توضیح دهید.
 ۳. نقش برقراری استاندارد در توانمندی اعضای هیات علمی را شرح دهید.
 ۴. خصوصیات داوران شرکت کننده در پانل را لیست نمایید.

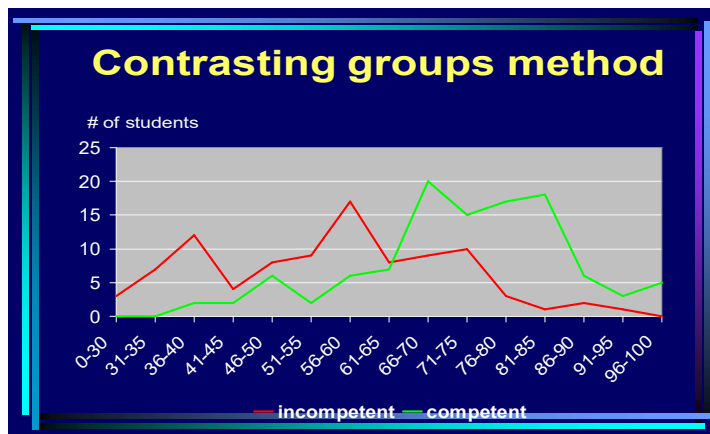
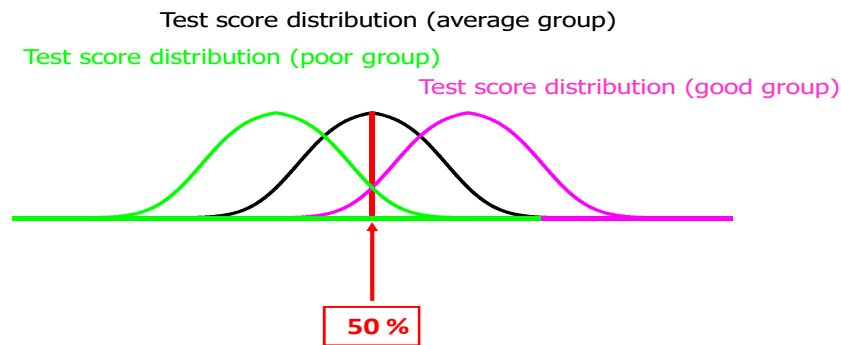
۳,۱ روش گروههای تقابلی (گروه مرزی)

روش گروههای تقابلی (گروه مرزی) از جمله روشهای آزمودنی محور است . ندلسکی (۱۹۵۴) همچنین یک روش برقراری استاندارد را پیشنهاد کرد که بر اساس عملکرد واقعی گروه هایی قرار دارد که از نظر تفاوت در متغیر مورد اندازه گیری شناخته شده، می باشند. یعنی در این روش داوران گروه مرزی واقعی (نه فرضی) را تشخیص می دهند.

بطور خلاصه این روش عبارت است از:

۱. هیأتی از کارشناسان واجد صلاحیت که با نوع شایستگی داوطلبانی، که قرار است مورد آزمون قرار گیرند آشنایی دارند در نظر گرفته می شوند.

۲. از کارشناسان خواسته می شود دو گروه از داوطلبان را مورد شناسایی قرار دهند. گروه اول آنهایی که به وضوح در محتوای مورد نظر تسلط دارند و گروه دوم آنهایی که فاقد تسلط هستند.
۳. آزمون بر روی هر دو گروه اجرا می شود.
۴. توزیعهای فراوانی نمرات آزمون هر دو گروه بر روی یک نمودار به صورت منحنی پلی گن رسم می شود.
۵. جایی که دو منحنی با یکدیگر تقاطع می کنند نمره ای است که نشان دهنده حداقل عملکرد قابل قبول می باشد.



شکل شماره پنج : روش گروههای تقابلی (گروه مرزی)

روش هافستی^{۱۳} یک رویکرد استاندارد سازی است که روشهای استانداردسازی مطلق و نسبی را با هم تلفیق می کند.

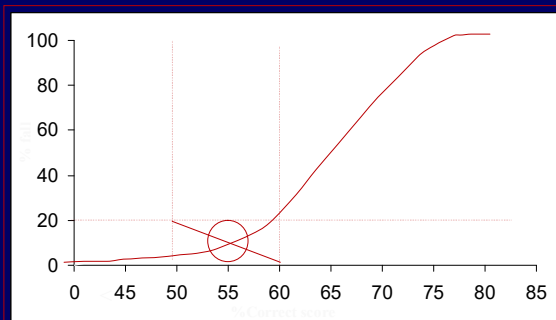
¹³ Hofstee

Relative /absolute compromise standards:
the Hofstee method

- Lowest acceptable percentage of failing examinee (**minimum failure rate**)
- Highest acceptable percentage of failing examinee (**maximum failure rate**)
- Lowest score which allows a candidate to pass (**minimum passing point**) ;
- Highest score required for a candidate to pass (**maximum passing point**)

داوران بعد از بررسی مواد آزمون چهار سوال مطرح میکنند.

۱. حداقل در صد قابل قبول از مردودی در داوطلبان چقدر است؟
۲. حداکثر در صد قابل قبول از مردودی در داوطلبان چقدر است؟
۳. حداقل نمره ی مورد نیاز که داوطلب با آن قبول می شود ، چقدر است؟
۴. حداکثر نمره ی مورد نیاز که داوطلب با آن قبول می شود ، چقدر است؟



شکل شماره شش : روش هافستی در استاندارد سازی

همانطور که در شکل ملاحظه می کنید حداقل در صد قابل قبول از مردودی در داوطلبان صفر و حداکثر در صد قابل قبول از مردودی در داوطلبان ۲۰٪ می باشد، حداقل نمره ی مورد نیاز که داوطلب با آن قبول می شود ۵۰ است در حالیکه حداکثر نمره ی مورد نیاز که داوطلب با آن قبول می شود ۶۰ می باشد. با رسم منحنی نمرات آزمون با قطر مستطیل نقطه برش به دست می آید که برابر با ۵۵ است.

۳,۲ روش آنگوف ۱۴

در این مدل یک رویکرد آزمون محور را بکار می برد که در آن یک گروه از داوران متخصص در مورد اینکه داوطلبین در خصوص هر یک از اجزا آزمون در امتحان چگونه عمل می نمایند تخمین هایی را بکار می برند.

آنگوف (۱۹۷۱) روش تقریباً متفاوتی را توصیف کرده است.

مراحل این فرآیند به قرار زیر است:

۱. هیاتی از کارشناسان واجد صلاحیت در زمینه محتوایی در نظر گرفته می شود.
۲. از هر کارشناس خواسته می شود که تک تک سؤالات آزمون را بررسی کرده و درصد افراد یک گروه با حداقل شایستگی را درست جواب خواهند داد تخمین بزنند.
۳. درصدهای همه سؤالات را برای تعیین یک نمره حداقل قابل قبول برای آن کارشناس با یکدیگر جمع می کنند.
۴. سپس متوسط نمرات همه کارشناسان به عنوان استاندارد حداقل عملکرد قابل قبول تعیین می شود.

Items	Proportion of answer an item correctly
1	0.95
2	0.80
3	0.90
4	0.60
5	0.75
6	0.40
7	0.50
8	0.25
9	0.25
10	0.40
Expected total score= 5/80	Total= 5.80

شکل شماره هفت : روش آنگوف در استاندارد سازی

۳,۳ آنگوف تعدیل شده^{۱۵}

رویکرد استانداردسازی آنگوف یک رویکرد قضاوتی است که در آن یک گروهی از داوران متخصص در مورد ای موضوع که داوطلبین لب مرز چگونه در پاسخ گویی به سوالات امتحان عمل می نمایند تخمین می زنند . در رویکرد آنگوف تعدیل شده تعیین استاندارد آزمون با تعیین مشخصات گروههای مرزی توسط آزمون گیرندگان و مقایسه آن با عملکرد واقعی توسط اعضای پانل مشخص می شود.

مراحل اجرایی رویکرد آنگوف تعدیل شده چهار مرحله بکار می رود.

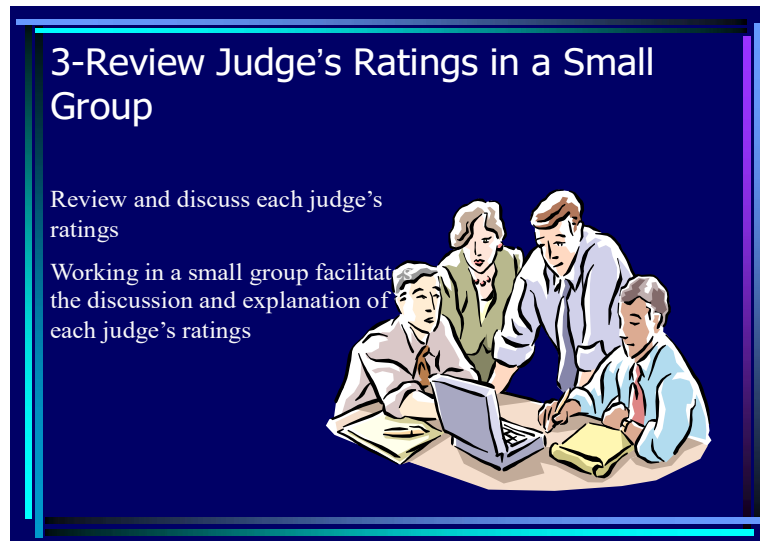
➤ مرحله اول: معرفی کلی

تهیه کننده یک شمای کلی از امتحان OSCE شامل هدف آزمون، طول هر ایستگاه، اجزای آزمون، روش های نمره دهی و دیگر اطلاعات مفید را به اعضای پانل ارایه می کند. یک آزمون OSCE کوچک بعنوان این فرایند معرفی می تواند بکار رود.

➤ مرحله دوم: معرفی ایستگاه حیطه عملی

طراحان آزمون ایستگاه عملی را به اعضای پانل ارایه می کنند. که شامل:

- ✓ توصیف کامل ایستگاهها شامل چک لیست های معاینات بدنی و اخذ تاریخچه
- ✓ نوارهای ویدیویی یک اجرای ضعیف و قوی برای ایستگاههای فیلد عملی



➤ مرحله سوم: تعیین خصوصیات داوطلبین لب مرزی توسط امتحان گیرنده ها

خصوصیات داوطلبین لب مرزی از امتحان گیرنده ها سوال می شود .. جهت بیان انتظاراتشان برای عملکرد گروه مرزی فرضی، بعد از بحث و رسیدن به یک وفاق عمومی برای هر جزء مهارت یک نمره مناسب برای داوطلب مرزی مشخص می گردد.

مرحله چهارم: درجه بندی اعضای پانل

در این مرحله بعد از ملاحظه عملکرد واقعی اعضای پانل نمره می دهند ، و سپس با برگزاری چند راند و مقایسه نمرات افراد مختلف و انتخاب اعضای پانل نمره یا نمرات استاندارد تعیین می گردد.

انتخاب اعضای پانل در استانداردسازی از مهمترین مراحل است بطور خلاصه اعضای پانل بایستی

- در حیطه امتحان مربوطه متخصص باشند
 - باروشهای آزمونها آشنا باشند.
 - حلال مشکلات خوب باشند
 - با سطوح داوطلبین آشنا باشند.
 - علاقمند به آموزش باشند(اساتید)
- ✓ ارتقاء هیات علمی

روش های استانداردسازی می تواند بعنوان شکلی از ارتقا هیات علمی بکار رود. اطلاعات دست اول تجربه استاد از عملکرد داوطلبین بر روی یک وظیفه یا هدف مشخص آنها راقادر می سازد که انتظارات خودشان را در ارتباط با کفایت داوطلبین مقایسه نمایند. عملکرد داوطلبین ضعیف و عالی با ملاحظات اساتید می تواند مقایسه شود.

✓ کنترل کیفیت مطالب آزمون

فرایند مواجهه سازی استاد با محتویات آزمون سیاست نمره دهی و تعیین نمره عملکرد یک روش کنترل کیفیت می باشد. اعضای پانل در فرایند مرور مطالب آزمون سوالات نامناسب که غیر مرتبط و غیر شفاف می باشند راتشخیص می دهند. بنابراین ضمن آشنایی با انواع سوالات با مراحل ساخت یک آزمون استاندارد نیز آشنا می شوند.

۳,۴ ارزشیابی فرایند استانداردسازی و نتیجه گیری

خود فرایند استانداردسازی نیز بایستی ارزیابی گردد .. کارهای بیشتری هنوز برای ایجاد روشهای استانداردسازی موثر در آموزش مورد نیاز است . طول مدت باید به عنوان یک فاکتور مهم مورد توجه قرار گیرد و روشهایی جهت کاهش زمان این فرایند لازم است . ملاحظات بیشتری برای مدل‌های جبرانی کامل که در آن اجزاء آزمون بطور متوسط جهت تعیین استاندارد ها بکار می روند باید بکار رود . استانداردهای بدست آمده بایستی با اطلاعات در دسترس امتحان گیرنده ها جهت تضمین اعتبار آزمون مورد بررسی قرار گیرد. روش های موثر آموزش اعضای پانل جهت تشخیص مشخصات گروه مرزی ضروری است . هر چه روشهای استانداردسازی برای انواع آزمونها بیشتر بکار رود ، کیفیت بالاتری را در آزمون های کفایت حرفه ای تضمین خواهد نمود.

➤ آزمون

۱- اعضای پانل در استانداردسازی بالینی می بایست:

- الف) آشنا باشند.
- ب) علاقه مند به یادگیری باشند.
- ج) متخصص باشند.
- د) حلال مشکلات باشند.

گزینه "د" صحیح است.

۲- نشانه روش انگوف و ندلسکی در کدام مورد است؟

- الف (استفاده از نظر داوران
ب (استفاده از جدول
ج (استفاده از متوسط نمرات
د (داوطلب لب مرز
- گزینه "الف" صحیح است.

خلاصه

به طور کلی استاندارد برای سنجش عملکرد ، حیطة جدید مطالعاتی است و برای این منظور روش های مختلفی به کار می رود. این روش عبارتند از :

- روش درصد ثابت و شیوه ها استانداردسازی با استفاده از داوری تخصصی
- روش ندلسکی
- روش انگوف
- و گروههای متقابل در استاندارد سازی شهرت بیشتری دارند. برای استاندارد سازی عملکرد در علوم پزشکی روش انگوف مقبولیت بیشتر دارد و برای استاندارد سازی آزمون های چندگزینه ای روش ندلسکی مناسب تر است.

1. Standard setting in student assessment – An extended summary of AMEE Medical Education Guide No 18. *Medical Teacher* (2000) 22, 2
2. Berk, R.A.. A consumer's guide to setting performance standards on criterion-referenced tests. *Review of Educational Research*, (1986) 56, 137-172.
3. Cizek, G. J. *Setting Performance Standards: Concepts, Methods, and Perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates. (2001).
4. Jaeger, R.M. Certification of student competence. In R.L. Linn (Ed.), *Educational Measurement*. New York: American Council on Education and Macmillan Publishing Company. (1989).
5. Kane, M. Validating the performance standards associated with passing scores. *Review of Educational Research*, (1994). 64, 425-461.
6. Livingston, S.A. and Zeiky, M.J. *Passing scores: A manual for setting standards of performance on educational and occupational tests*. Princeton, NJ: Educational Testing Service. . (1982).
7. Norcini, J.J. and Guille, R.A. (2002). Combining tests and setting standards. In Norman, G., van der Vleuten, C., and Newble, D. (Eds.): *International Handbook of Research in Medical Education* (pp. 811-834). Dordrecht: Kluwer Press.
8. Norcini, J. J. (2003). Setting standards on educational tests. *Medical Education*, 37, 464-469.
9. Norcini, J. J. & Shea, J. A. (1997). The credibility and comparability of standards. *Applied Measurement in Education*, 10, 39-59.
10. Zeiky, M. J. (2001). So much has changed. How the setting of cutscores has evolved since the 1980s. In G.J.Cizek (Ed.), *Setting Performance Standards: Concepts, Methods, and Perspectives* (pp. 19-52). Mahwah, NJ: Lawrence Erlbaum Associates.