



! اللہ
الرحمن

اندازه گیری، سنجش و ارزشیابی آموزشی

منبع

دکتر علی اکبر سیف

مقدمه :

هدف از ارزیابی پیشرفت تحصیلی چیست؟

۱- ارزیابی از روش تدریس معلم و موفقیت آموزشی وی

۲- ارزیابی محتوای آموزشی کتب و منابع آموزشی

۳- بررسی سطح یادگیری و میزان موفقیت دانش آموزان در کسب

مهارت های تحصیلی

فصل اول

تعاریف و اصطلاحات

تعریف اندازه گیری:

« اندازه گیری عبارت است از فرایندی که تعیین می کند یک شخص یا یک شی تا چه میزان از یک ویژگی برخوردار است.»

«گی ۱۹۹۱»

« اندازه گیری از کاربرست ابزارهای پیشرفته و پیچیده الکترونیکی تا امتحانات مداد، کاغذی کتبی، مقیاس های درجه بندی و فهرست های واریسی را شامل می شود.»

«پین ۲۰۰۳»

« شیوه نسبت دادن یک عدد (که معمولاً به آن نمره گفته می شود) به یک صفت یا ویژگی معین یک شخص به گونه ای که آن عدد درجه ای از آن صفت یا ویژگی را که آن شخص از آن برخوردار است را نشان دهد.»

«نیکتو ۲۰۰۱»

تعریف آزمون:

مقدمه:

اندازه گیری یک فرایند است، این فرایند نیاز به یک وسیله دارد که به آن آزمون می گویند.

آزمون وسیله ای است برای اندازه گیری صفات فیزیکی (جسمی) و روانی، چون متخصصین اندازه گیری به ویژگی های روانی مورد نظر برای اندازه گیری دسترسی ندارند ناچار به اندازه گیری رفتارهایی که معرف ویژگی روانی است می پردازند.

رفتار چون متنوع و فراوان است نمی توان به اندازه گیری همه آن ها دست یافت، بنابراین نمونه ای از رفتار مورد بررسی قرار می گیرد.

تعریف:

« آزمون وسیله یا روشی است نظام دار (یا نظام مند) برای اندازه گیری نمونه ای از رفتار است» .

«گرانلاند ولین ۱۹۹۰»

نیکتو (۲۰۰۱) در تعریف اصطلاح آزمون گفته است:

« آزمون به عنوان یک ابزار با یک شیوه نظام دار مشاهده و توصیف ویژگی های دانش آموزان با یک مقیاس عددی یا یک طرح طبقه بندی را به کار می بندد تعریف می شود».

تعریف سنجش:

« سنجش یک اصطلاح کلی است و به صورت فرایندی تعریف می شود که برای گردآوری اطلاعات مورد نیاز تصمیم گیری درباره دانش آموزان، برنامه های درسی و سیاست های آموزش مورد استفاده قرار می گیرد.

«اگن و گاوچاک ۲۰۰۱»

ونگلینز (۱۹۹۳) در تعریف سنجش گفته است:

« سنجش به یک تحلیل جامع و چند وجهی از عملکرد گفته می شود». او همچنین در مقایسه سنجش و آزمون از قول کرانباخ نقل کرده است: « سنجش شامل استفاده از فنون مختلف است و تاکید زیادی بر مشاهده عملکرد دارد...»

سنجش نوعی تحلیل بالینی و پیش بینی عملکرد است در حالی که آزمون یک وسیله اندازه گیری است.

تعریف ارزشیابی:

ارزشیابی به معنی تعیین ارزش و داوری کردن است.

« ارزشیابی به یک فرآیند نظام دار برای جمع آوری، تحلیل و تفسیر اطلاعات گفته می شود به این منظور که تعیین کند آیا هدف های مورد نظر تحقق یافته اند یا در حال تحقق یافتن هستند و به چه میزان؟ »

« گی ۱۹۹۱ »

یکی از ویژگی های مهم ارزشیابی تعیین کیفیت است.

تعریف کیفیت: (در ارزشیابی پیشرفت تحصیلی)

« کیفیت عبارتست از دانش، مهارت ها و توانایی هایی که از دانش آموزان پس از آموزش انتظار می رود ».

« نیکتوا ۲۰۰۱ »

مثال: تعیین کیفیت افراد در یک مسابقه مقاله نویسی یا مرور نوشته های دانش آموز و مرور نمره های وی

علاوه بر ارزشیابی پیشرفت درسی دانش آموزان، ارزشیابی با موضوع های دیگری سرو کار دارد.

مثال: عملکرد معلمان و مدیران، روش های آموزشی، برنامه های درسی دوره های آموزشی، مواد آموزشی، پروژه های آموزشی و سازمان های آموزشی.

یک ارزشیاب، ابتدا هدف های کلی برنامه درسی را مشخص می کند و سپس میزان تحقق آن ها را بررسی می کند.

ورتن و سندرز (۱۹۸۷) دو تن از صاحب نظران حوزه ارزشیابی آموزش گفته اند:

« در آموزش و پرورش، ارزشیابی به یک فعالیت رسمی گفته می شود که برای تعیین کیفیت، اثربخشی یا ارزش یک برنامه، فرآورده پروژه، فرآیند، هدف یا برنامه درسی به اجرا در می آید».

مقایسه اندازه گیری، آزمون، سنجش و ارزشیابی با یکدیگر:

ارزشیابی برای داوری درباره کیفیت موضوع است و ارزش اطلاعات به دست آمده را بررسی می کند. که مستلزم تحلیل و تفسیر اطلاعات جمع آوری شده است.

اطلاعات مورد نیاز ارزشیابی آموزشی که کمک اندازه گیری (آزمون) و سنجش جمع آوری می شود.

البته سنجش یک فرآیند مبتنی بر مشاهده عملکرد و شیرینی عملکرد است در حالی که آزمون یک وسیله اندازه گیری است.

مقیاس های اندازه گیری

تعریف و انواع مقیاس های اندازه گیری:

بنا به تعریف:

«مقیاس مجموعه ای عدد (یا نمادهای دیگر) است که ویژگی های آن ها بر ویژگی های تجربی اشیایی که اعداد به آن ها نسبت داده می شوند منطبق است».

«کوهن و سورد یک ۲۰۰۲»

انواع مقیاس های اندازه گیری در چهار سطح دسته بندی شده است:

اسمی، ترتیبی، فاصله ای، نسبتی

«استیونس ۱۹۵۱»

مقیاس اسمی

۲- طبقه بندی

اعداد و ارقام برای طبقه بندی اشیاء و افراد استفاده می شود بدون این که مفهوم ریاضی داشته باشند.

۱- اسم گذاری

در این مقیاس از ارقام و اعداد برای اسم گذاری و تشخیص اشیاء و افراد استفاده می شود. بدون این که مفهوم ریاضی داشته باشند

مثال:

شماره بازی کنان فوتبال

تعیین جنسیت با شماره

تعیین نژاد، ملیت یا مذهب

مقیاس اسمی

عملیات مجاز آماری:

شمارش فراوانی و تعیین نما

عملیات مجاز ریاضی:

انجام هیچ یک از چهار عمل اصلی جمع، تفریق، ضرب و تقسیم در این مقیاس میسر نیست.

مقیاس ترتیبی

۱- دسته ای از افراد و اشیاء با توجه به یک صفت از بزرگ به کوچک (یا بالعکس) مرتب شده اند.

۲- معلوم نیست هر یک از اشیاء چقدر از آن صفت را دارند.

۳- و نیز مشخص نیست این اشیاء یا افراد از لحاظ صفت مورد اندازه گیری چقدر با هم فاصله دارند.

«فاتالی ۱۹۷۸»

مثال: مرتب کردن دانش آموزان یک کلاس به ترتیب قد و شماره گذاری آنها

- پرسشنامه استرس هولمزکه بر اساس تعداد تغییرات زندگی فرد در ۶ ماه گذشته مشخص میشود

- نگرش سنجها

مقیاس ترتیبی

عملیات مجاز آماری:

شمارش فراوانی، تعیین نما، محاسبه میانه، محاسبه درصدها
و محاسبه ضریب همبستگی رتبه ای اسپیرمن

عملیات مجاز ریاضی:

هیچ یک از چهار عمل اصلی جمع، تفریق، ضرب و تقسیم

مقیاس فاصله ای

۱- در آن رتبه اشیاء (یا افراد) با توجه به یک صفت مشخص است.

۲- معلوم است که اشیاء (یا افراد) از لحاظ صفت مورد اندازه گیری چقدر از هم فاصله دارند

۳- اما هیچ اطلاعی از مقدار مطلق صفت مورد نظر برای هیچ یک از افراد در دست نیست

« ناتالی ۱۹۷۸ »

مثال:

مرتب کردن قد دانش آموزان یک کلاس بر اساس قد کوچکترین انها(کوچکترین نمره صفر و بقیه به نسبت اندازه های بیشتر مشخص میشوند)

-اندازه های دماسنج

مقیاس فاصله ای

عملیات مجاز آماری:

محاسبه نما، میانه، میانگین، انحراف معیار، ضریب همبستگی
رتبه ای اسپیرمن، ضریب همبستگی گشتاوری پیرسون

عملیات مجاز ریاضی:

جمع و تفریق مجاز است، ضرب و تقسیم مجاز نیست.

مقیاس نسبتی

۱- در آن رتبه اشخاص با توجه به یک صفت معلوم است.

۲- فاصله اشخاص معلوم است.

۳- علاوه بر این ها فاصله حداقل یکی از اشخاص از یک صفر منطقی معلوم است.

✓ اندازه گیری طول، وزن، حجم و سایر صفات فیزیکی افراد با مقیاس نسبتی عملی است. اما اندازه گیری متغیرهای روانی و پرورشی با این مقیاس به ندرت میسر است.

«ناتالی ۱۹۷۸»

مثال:

اعداد از صفر تا بینهایت

اندازه گیریهای طول، وزن، حجم

مقیاس نسبتی

عملیات مجاز آماری و ریاضی:

در این مقیاس همه عملیات آماری و ریاضی مجاز است.

تقسیم بندی مقیاسها از نظر پیوسته یا گسسته بودن

✓ ۱- **مقیاس گسسته:** در این نوع مقیاس ها متغیر تنها می تواند یکی از روش های معین را به خود اختصاص می دهد.

✓ ویژگی مهم این نوع مقیاس ها تساوی واحد های اندازه گیری است. مانند: ۱ و ۲ و ۳ (تعداد افراد که اعشار پذیر نیست)

✓ مثلا: اگر بخواهیم تعداد فرزندان خانوادهای شهری را حساب کنیم.

✓ یا صفحات مدرج

مقیاس پیوسته:

برخلاف مقیاس گسسته، در مقیاس پیوسته متغیر می تواند
در فاصله بین نقاط مقیاس هر تعداد ارزش را به خود
اختصاص دهد. مانند: $3/04$ ، $1/75$ ، $2/04$ ، و...

- حد واقعی اعداد:
- حد واقعی اعداد: حد واقعی ارزش عددی یک متغیر پیوسته برابر است با آن عدد به اضافه یا منهای نصف واحد اندازه گیری **(رونیون و هابر ۱۹۷۶)**

انواع آزمون های مورد استفاده

در

آموزش و پرورش

دسته بندی آزمون های مورد استفاده در آموزش و پرورش

۲- آزمون های عاطفی

۱- آزمون های توانایی

۱- آزمون های توانایی:

آزمون های توانایی آموخته ها، استعدادها و مهارت های افراد را می سنجد:

دسته ای از این آزمون ها توانایی های شناختی را می سنجد که برای فعالیت های ذهنی و فکری و تحصیلی کاربرد دارند.

دسته ای دیگر توانایی های روان - حرکتی را می سنجد که برای فعالیت های صنعتی، تربیت بدنی، نظامی و ... به کار می روند.

ویژگی مهم این آزمون ها این است که از آزمون شونده خواسته می شود حداکثر نمره ممکن را کسب کند.

انواع آزمون های توانایی

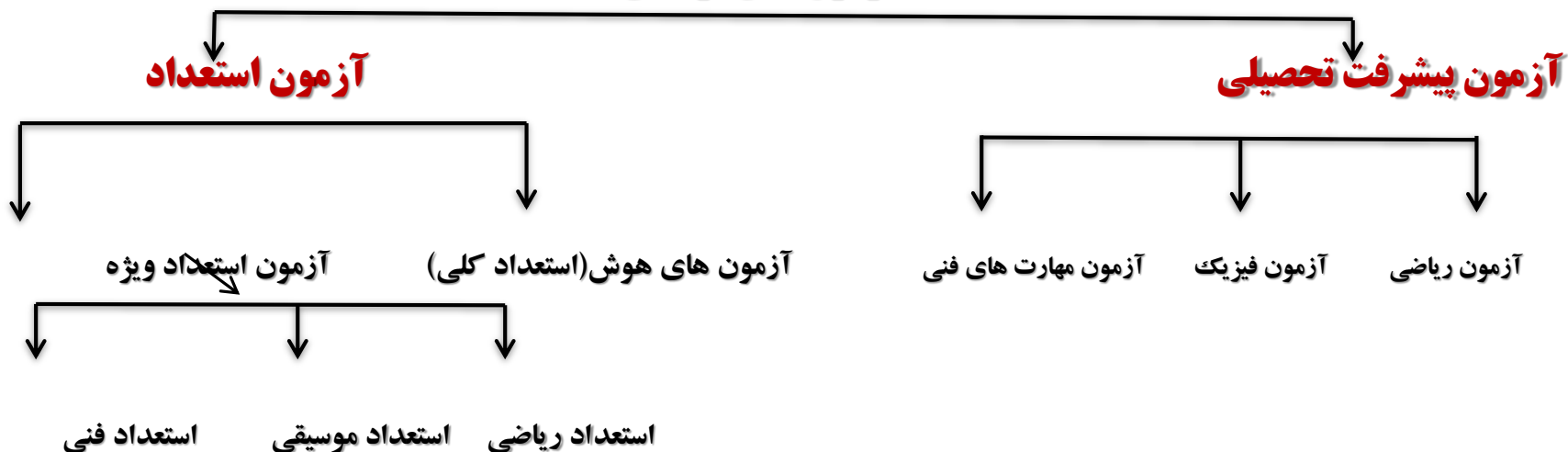
۱- آزمون های استعداد:

میزان توانایی فرد را برای کارهایی که باید انجام دهد می سنجد.

۲- آزمون های پیشرفت تحصیلی:

مهارت های کسب شده فرد را اندازه گیری می کنند.

آزمون های توانایی



ثورندایک و هیگن (۱۹۷۹) در کتاب معروف خود با عنوان اندازه گیری و ارزشیابی در روان شناسی و آموزش و پرورش در رابطه با آزمون های استعداد و پیشرفت تحصیلی به گونه زیر نظر داده اند:

« باید به خاطر داشت که همه آزمون های توانایی، صرف نظر از نامی که دارند توانایی های موجود فرد را اندازه می گیرند. ما نمی توانیم مستقیماً توانایی «ذاتی» یا « ارثی» کسی را اندازه گیری کنیم. آن چه را که ما می توانیم در فرد اندازه بگیریم توانایی او برای انجام دادن اعمال در زمان حال است. بنابراین تمایز بین آزمون های استعداد و پیشرفت تحصیلی به قصد و منظوری مربوط می شود که نتایج آزمون برای آن قصد و منظور مود استفاده قرار می گیرند، نه به ماهیت یا محتوای آزمون »

۲- آزمون های عاطفی

مقایسه	
آزمون های توانایی	آزمون های عاطفی
۱- توانایی های فکری و عملی افراد را اندازه گیری می کنند	۱- ویژگی های عاطفی و شخصیتی افراد را می سنجند.
۲- برای مقاصد پیش بینی موفقیت های تحصیلی و شغلی مود استفاده قرار می گیرند	۲- بیشتر به امور مشاوره در حل مسائل روانی و تربیتی و حرفه ای کمک می کنند.
معایب	
این است که به اندازه آزمون های شناختی عینی و دقیق نیستند	

✓ با وجود این به قول آیکن (۱۹۸۵) بسیاری از ابزارهای اندازه گیری عاطفی دارای ویژگی های خوب از جمله روایی و پایایی هستند و می توان از این ابزارها برای اندازه گیری شخصیت و سایر ویژگی های عاطفی استفاده کرد.

انواع آزمون های عاطفی



آزمون های شخصیت:

تعریف:

شخصیت مجموعه ای از رفتارهای فرد در موقعیت های اجتماعی است (ایبل ۱۹۷۹)

از شخصیت تعاف گوناگون شده و شامل ابعاد مختلفی می باشد از جمله: هوش، پیشرفت، قد و قامت، وضع سلامتی، کیفیت صدا، وضع ظاهر و ...

آزمون های شخصیت به طور عمده اندازه گیری متغیرهای

Motivation

انگیزش

Temperament

مزاج

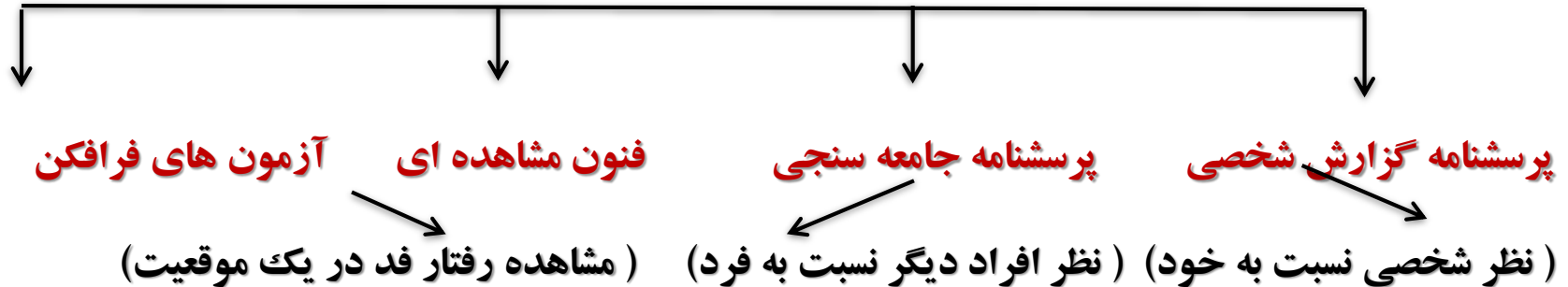
Adjustment

سازگاری

Character

منش

ابزارهای اندازه گیری شخصیت



۲ و ۳- آزمون های علاقه و نگرش:

شباهت ها:

- ۱- هم علاقه و هم نگرش به دوست داشتن ها و دوست نداشتن های فرد اشته می کند.
- ۲- هر دو به انتخاب فرد درباره فعالیت ها و نهادهای اجتماعی و یا گوه ها ارتباط دارند.
- ۳- هر دو شامل احساس های شخصی نسبت به امور هستند.

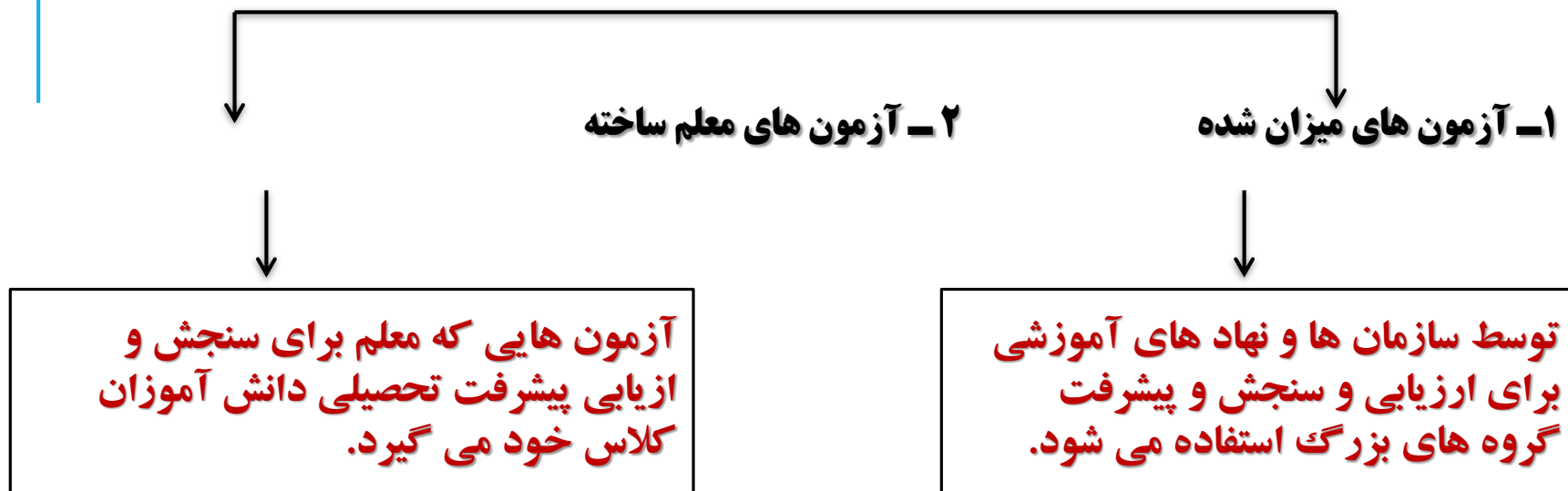
تفاوت علاقه و نگرش:

علاقه به احساس یک فرد نسبت به فعالیت اشاره می کند درحالی که **نگرش** حاکی از احساس فرد نسبت به شی یا نهاد اجتماعی یا یک گروه است (مهرنز و لهما ۱۹۸۴).

ابزار اندازه گیری:

پرسشنامه علائق و پرسشنامه نگرش

دسته بندی آزمون ها با توجه به چگونگی تهیه آن ها



استاندارد بودن یا میزان بودن آزمون (کران باخ ۱۹۹۰):

به این صوت است که شیوه های اجاء، شیوه های مشاهده تجهیزات و مواد و قواعد نمره گذاری تا آن اندازه تثبیت شده اند که تا حد امکان د شرایط و زمان های مختلف نتایج یکسان به دست آید.

سه تفاوت آزمون های استاندارد (میزان شده) و معلم ساخته:

- ۱- آزمون های میزان شده به زبان و تخصص بیشتری نیاز دارند چون برای گروه های زیاد و زبان های مختلف است، هزینه تخصص بیشتری می طلبد. محتوای درسی و مهارت نمونه گیری می شو، آزمون آزمایشی اجرا می شود، سوال ها تجزیه و تحلیل می شود و توسط متخصصان نوشته و ویرایش می شود.
- ۲- از آن جا که آزمون های میزان شده برای استفاده در مناطق مختلف و دانش آموزان مناطق گوناگون تهیه می شود هدف های آموزشی را باید در این آزمون ها منظور کرد.
- ۳- آزمون های میزان شده باید هنجار شوند در حالی که آزمون های معلم ساخته چنین نیازی ندارند.

موارد استفاده از آزمون های میزان شده و معلم ساخته

۱- تصمیمات آموزشی مبنی بر کیفیت و کم و کیف جریان آموزشی \Leftarrow آزمون های معلم ساخته

۲- تصمیمات مربوط به ارتقای دانش آموزان و دانشجویان \Leftarrow آزمون های معلم ساخته


۳- تصمیمات مربوط به مشاوره و راهنمایی \Leftarrow آزمون های میزان شده

۴- تصمیمات بر مبنای مقایسه افراد و گروه ها \Leftarrow آزمون های میزان شده

۵- تصمیمات مربوط به برنامه ریزی درسی \Leftarrow آزمون های میزان شده

۶- تصمیمات در مورد ارزشیابی از میزان موفقیت مدارس مختلف \Leftarrow آزمون های میزان

شده



دسته بندی ارزشیابیهای آموزشی

دسته بندی ارزشیابی از نظر ملاک مورد استفاده

۱- ارزشیابیهای وابسته به ملاک: (ملاک مطلق)

- ملاک یا معیار ارزشیاب از قبل تعیین میشود.

-

جدول مشخصات آزمون:

بهترین آزمون، آزمونی است که تمام مطالب (محتوی) و کلیه ی هدفهای آموزشی آن درس را شامل شود.

جدول مشخصات آزمون به ایجاد تعادل بین آموزش و سنجش کمک می کند. (پین ۲۰۰۳)

فصل ۴				فصل ۳			فصل ۲				فصل ۱				بعد محتوی
														بعد هدف	
															دانش
															فهمیدن
															کار بستن
															تعداد سوالها
															درصد سوالها

جدول و فرمول توزیع سؤالات به بخشهای مختلف:

برای تعیین درصد سؤالات به بخشهای مختلف باید به سؤالات ذیل پاسخ داد:

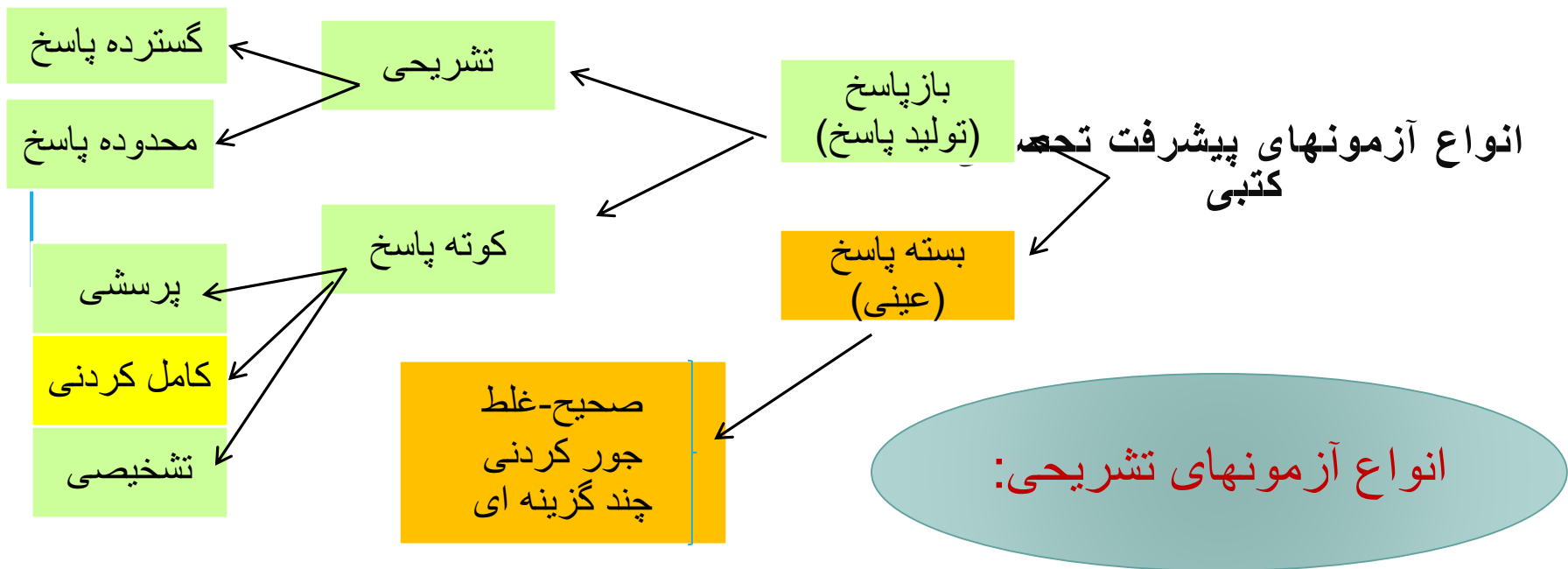
- ۱- کدام بخش از محتوا و هدفها دارای حجم بیشتری است و دقت آموزش را به خود اختصاص داد؟
 - ۲- کدام هدفها از نظر کاربرد و نگهداری حافظه اهمیت بیشتری دارد؟
 - ۳- کدام قسمت درسی از اهمیت بیشتری برخوردار است؟
- نحوه توزیع سؤالات باید متناسب با جواب این سؤالات باشد:
- چنانچه بخشهای مختلف از نظر اهمیت و کاربرد یکسان باشند برای توزیع سؤالات به بخشهای از فرمول زیر استفاده می کنیم:

تعداد ساعتهایی که صرف تدریس آن بخش شده

تعداد کل ساعتهای تدریس

کل سؤالات \times نسبت ساعت تدریس هر بخش = تعداد سؤالات هر بخش

= نسبت ساعت تدریس هر بخش



۱- آزمون گسترده پاسخ: هیچ محدودیتی برای آزمون شونده منظور نمی شود.

مثال: قانون جاذبه و تأثیر آن را بر علم زمین شناسی شرح دهید.

۲- آزمون محدود پاسخ: آزمون شونده ملزم است تا پاسخ خود را در چارچوب شرایط خاصی محدود کند.

مثال: توضیح دهید چرا آب در ۱۰۰ درجه به جوش می آید؟ پاسخ خود را در یک پاراگراف بنویسید.

قواعد تهیه سؤالات تشریحی:

- ۱- در نوشتن سؤالات آزمون، دقت کنید با رسم مشخصات آن درس، سؤالات به طور مستقیم به هدفهای آموزشی آن درس مرتبط باشد.
- ۲- سؤالهای تشریحی را تنها به هدفهایی محدود کنید که با سایر انواع آزمونها به خوبی قابل اندازه گیری نیستند.
- ۳- صورت سؤالات تشریحی را عباراتی کاملاً روشن و واضح بنویسید و از کلی گویی و ابهام در بیان پرهیزید.
- ۴- از کاربرد کلمات «چه کسی، چه وقت، کجا و...» پرهیزید.
- ۵- تا حد امکان از سؤالهای تازه و جدید استفاده کنید.
- ۶- در سؤالاتی که به ابراز عقیده مربوط می شود باید از آزمون شونده بخواهیم مستند پاسخ دهد و تنها عقیده شخصی خود را نگوید.
- ۷- به آزمون شوندگان حق انتخاب چند سؤال از میان سؤالات را ندهید.
- ۸- برای پاسخ به سؤالات وقت کافی در نظر بگیرید.

قواعد تصحیح سؤالهای تشریحی:

- ۱- پاسخ سؤالها بر اساس هدفی که در سؤال گنجانده شده تصحیح کنید.
- ۲- با انتخاب یک کلید پاسخ از دخالت عوامل نامربوط جلوگیری کنید.
- ۳- پاسخ را سؤال به سؤال تصحیح کنید نه ورقه به ورقه.
- ۴- هنگام تصحیح ورقه های امتحانی از شناسایی نام صاحبان آنها خودداری کنید.
- ۵- در صورت امکان از یکی دو نفر از همکارانتان بخواهید که تصحیح شما را دوباره تصحیح کند.
- ۶- تمام پاسخهای آزمون شوندگان به یک سؤال را در یک نشست و بدون وقفه تصحیح کنید.
- ۷- به نمرات سؤالهای قبل نگاه نکنید.
- ۸- بر روی برگه آزمون اشتباهات دانش آموزان را تصحیح کنید و نظر خود را بنویسید.

تعریف آزمون کوتاه پاسخ:

مجموعه ای از سؤالهای مختصر که دارای پاسخهای بسیار کوتاه و برای سنجش و هدفهای آموزشی سطح پایین طرح می شود تشکیل شده است. (نیکتو ۱۹۸۳ و لهمان ۱۹۸۴)

انواع آزمونهاى کوتاه پاسخ:

الف = نوع پرسشی

- ۱- پایتخت عراق کجاست؟
- ۲- یک مگابایت چند کیلو بایت است؟

ب = نوع کامل کردنی

- ۱- یکی از روش های خوب مطالعه ----- از نکات مهم درس است.
- ۲- ----- $= 3 : (2 + 12)$

نوع تشخیصی

- ۱- نام گروه آزمایشی هر رشته را مقابل آن بنویسید.

----- مهندسی عمران

----- ژنتیک

----- حقوق

----- مهندسی پزشکی

- ۲- بعد از عناصر شیمیایی نام آنها را بنویسید.

----- سدیم

----- کلسیم

----- پتاسیم

----- باریم

کاربرد آزمونهای کوتاه پاسخ:

آزمونهای کوتاه پاسخ برای اندازه گیری هدفهای شناختی سطح پایین استفاده می شود.
(گرانلاندولین ۱۹۹۰)

مانند موارد ذیل:

- ۱- توانایی نشان دادن ارس مکانی اعداد، تشخیص اندامهای یک موجود زنده، تعریف یک مثلث و ...
- ۲- توانایی حل مسایل عددی در ریاضیات و علوم.
- ۳- توانایی حل کردن معادلات ریاضی و شیمیایی.

قواعد تهیه سؤالهای کوتاه پاسخ:

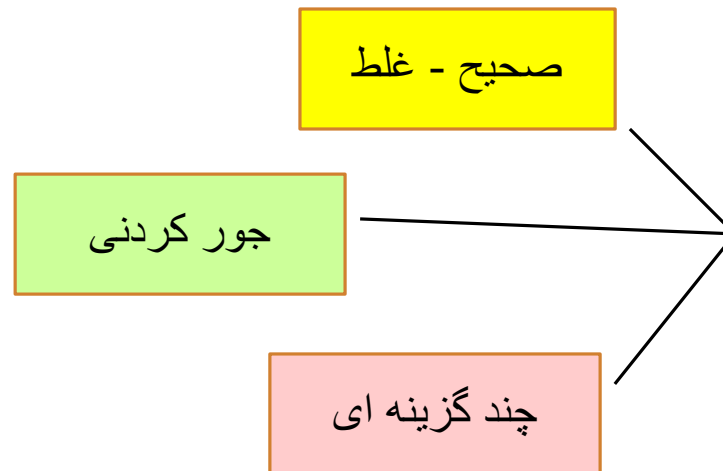
- ۱- هر سؤال باید موضوع مهمی را شامل شود.
- ۲- صورت سؤالات را کاملاً روشن بنویسید.
- ۳- صورت سؤال را از روی کتب عیناً نقل نکنید.
- ۴- در سؤالهایی که پاسخ آنها عدد است، واحد مقیاس را با دقت مشخص کنید.
- ۵- در سؤالات کامل کردنی تنها و عبارات مهم را حذف کنید.
- ۶- در سؤالات کامل کردنی تعداد زیادی جای خالی منظور نکنید.
- ۷- تا آنجایی که ممکن است جای خالی را در قسمت پایانی سؤال قرار دهید.
- ۸- تا حد امکان به جای سؤالهای کامل کردنی، از سؤالهای پرسشی استفاده کنید.
- ۹- در سؤالهای کامل کردنی، از کاربرد اشارات و مواردی که جواب سؤال را مشخص کند خوداری کنید.

تعریف آزمون های عینی:

آزمونهای عینی:

آزمونهای عینی (objective) آزمونهایی هستند که در آنها هم سوالات و جواب آنها در اختیار آزمون شوندگان قرار می گیرد و آزمون شوندگان درباره جوابها تصمیم می گیرند.

انواع آزمونهای عینی



انواع سؤال‌های صحیح غلط:

۱- نوه صحیح- غلط:

مثال: معلم در ابتدای کلاس باید با طرح پرسش‌هایی اطمینان یابد آیا دانش آموزان درس قبلی را فرا گرفته اند؟ صحیح - غلط

۲- نوع بله - نه:

مثال: آیا فلسفه یک علم است؟ بله- نه

۳- نوع اصلاحی:

مثال: جذر عدد ۶۴ ۹ است؟

دانش آموز باید عدد ۹ را خط بزند و صحیح آن که ۸ است را بنویسد.

۴- نوع خوشه ای:

در این روش تنه سؤال به صورت جمله ناتمام نوشته می شود و بدنبال آن چند جواب قرار می گیرد.

مثال: میانگین حسابی:

۱- یکی از اندازه های گرایش مرکزی است. صحیح- غلط

۲- کمتر از میانه است تحت تأثیر نمرات خیلی بزرگ یا خیلی کوچک قرار می گیرد. صحیح- غلط

۳- رابطه همبستگی بین دو متغیر را نشان می دهد. صحیح- غلط

قواعد تهیه سؤالهای صحیح - غلط

- ۱- سؤالها را تا حد امکان مختصر، ساده و روشن بنویسید.
- ۲- جمله هایی را بکار ببرید که درست یا غلط بودن آن ها برای کسانی که موضوع را یاد گرفته اند آشکار باشد.
- ۳- از کاربرد کلماتی چون همه، بعضی وقتها، غالباً، هرگز و از این قبیل خودداری کنید.
- ۴- سعی کنید تمام سؤالات صحیح- غلط هم اندازه باشند.
- ۵- سؤالهای صحیح- غلط را از روی جمله های کتب ننویسید.
- ۶- سعی کنید از مطالب بی اهمیت پرهیزید.

آزمون جور کردنی:

آزمونهای جور کردنی از دو ستون سؤالات و پاسخها تشکیل شده است که آزمون شونده باید پاسخ مربوط به هر سؤال را به آن سؤال مرتبط سازد.

مثال:

ستون سؤالات

(پ) ۱- پنی سیلین

(ث) ۲- پرتو ایکس

(الف) ۳- رادیو

(ج) ۴- واکسن فلج

(ب) ۵- قلب مصنوعی

(ت) ۶- گردش خون

ستون پاسخ

الف- ماری کوری

ب- دی باکی

پ- فلمینگ

ت- هاروی

ث- رونتگن

ج- سالک

چ- پاستور

قواعد تهیه سؤالات جور کردنی:

- ۱- پرسشها و پاسخهای متجانس و همگون انتخاب کنید.
- ۲- طول سؤال و پاسخ را کوتاه انتخاب کنید.
- ۳- در راهنمای سؤالات، اطلاعات لازم را درباره نحوه مقایسه و جور کردن پرسشها و پاسخ ها در اختیار دانش آموزان قرار دهید.
- ۴- همه پرسشها و پاسخهای یک سؤال را در یک صفحه قرار دهید.
- ۵- در صورت امکان فهرست پاسخها را به طور منطقی مرتب کنید.
- ۶- در فهرست پرسشها از جملات نیمه تمام استفاده نکنید.
- ۷- پرسشها را با شماره و پاسخ ها را با حروف مشخص کنید.
- ۸- هر یک از پاسخها باید برای تمام پرسشهای یک سؤال درست جلوه کند.
- ۹- از نوشتن سؤالهایی که پرسشها و پاسخهای آن ها کاملاً با هم جور می شوند پرهیزید.

آزمون چند گزینه ای:

قسمتهای مختلف سؤال چند گزینه ای

- ۱- قسمت اصلی یا تنه سؤال stem
- ۲- گزینه درست key
- ۳- گزینه های انحرافی distractor

انواع آزمونهای چند گزینه ای

۱- نوع تنها گزینه درست:

مثال: اگر $R=27$ ، R برابر است با:

- الف) ۳ ب) ۹ ج) ۱۱ د) ۳۰

۲- نوع بهترین گزینه:

مثال: کدام مورد تعریف کامل تری از یادگیری است؟

الف) کسب اطلاعات تازه در آموزش

ب) ایجاد تغییرات نسبتاً پایدار در توان رفتاری یاد گیرنده

ج) ایجاد تغییرات مطلوب و مورد نر در رفتار یادگیرنده

د) توانایی به خاطر سپردن و یادآوری تجارب پیشین

۳- نوع منفی:

مثال: کدامیک از عبارتهای زیر درباره ویروس درست نیست؟

- الف) ب) ج) د)

قواعد تهیه سؤالهای چند گزینه ای:

- ۱- هر سؤال باید یک موضوع یا یک هدف آموزشی را در بر گیرد.
(نیکتو ۱۹۸۳)
- ۲- بیشتر از یک مسأله یا یک مطلب در هر سؤال قرار ندهید.
- ۳- سؤالات را کاملاً روشن و واضح بنویسید.
- ۴- از تکرار مطالب در گزینه ها خودداری کنید.
- ۵- مطالب اصلی سؤال را به طور کامل در تنه سؤال بنویسید.
- ۶- همه گزینه های یک سؤال باید متجانس و به موضوع واحدی مربوط باشد.
- ۷- سؤال را طوری بنویسید که پاسخ درست تنها پاسخ درست یا درست ترین پاسخ باشد.
- ۹- از به کار بردن اشاره های دستوری نامربوط خودداری کنید.

۱۰- در سؤالات منفی کلمات منفی را برجسته کنید.

۱۱- از نوشتن سؤالاتی که هم متن سؤال منفی و هم پاسخ منفی خودداری کنید.

۱۲- تا حد امکان تاز گزینه همه موارد فوق خودداری کنید.

۱۳- از طرح سؤالهای گمراه کننده پرهیزید.

۱۴- دو گزینه متضاد که یکی از آنها درست است بکار نبرید.

۱۵- محل گزینه درست را میان گزینه های انحرافی تصادفی انتخاب کنید.

۱۶- برای هر سؤال بین ۳ تا ۵ گزینه در نظر بگیرید.

۱۷- عبارتهای هیچکدام یا هیچ یک از موارد فوق را زیاد بکار نبرید.

۱۸- تنها در صورت لزوم از آزمونهای چند گزینه ای استفاده کنید.

نمره گذاری آزمون های عینی:

برای نمره گذاری آزمونهای عینی دو روش عمده وجود دارد:

۱- محاسبه کلید پاسخهای درست بدون کسر نمره
برای حدس زدن

۲- کسر مقداری از نمره برای جبران حدس زدن

نمره گذاری روش دوم:

$$\text{نمره اصلاح} = R - \frac{W}{N-1}$$

R=تعداد پاسخهای درست

W=تعداد پاسخهای غلط

N =تعداد گزینه ها

مثال: در یک آزمون چهار گزینه ای ($N=4$) که دارای ۱۰۰ سؤال می باشد. دانش آموزی به ۸۰ سؤال پاسخ داده و از این تعداد ۱۲ سؤال غلط ($W=14$) و ۶۸ سؤال درسا است ($R=68$) نمره اصلاح شده این دانش آموز برابر است با:

$$\text{نمره اصلاح شده} = 68 - 12 = 56$$

۴-۱

تحلیل سؤال های آزمون:

هدف از تحلیل، واریسی تک تک سؤالهای آزمون و تعیین میزان دقت و نارسایی آن است.

مراحل تحلیل سؤال های آزمون:

برای تحلیل سؤال های آزمون باید تعیین کنیم چه تعداد افراد گزینه درست را انتخاب کردند و چه تعدادی گزینه های انحرافی را برگزیدند و چند نفر آن را بی جواب گذاشتند.

برای این منظور، ابتدا دو گروه بالا(قوی) و پایین(ضعیف) را انتخاب می کنیم.

گروه بالا از بین (۲۵% تا ۳۳%) دانش آموزانی که بیشترین نمره را آوردند متخصصان آزمون سازی (از جمله وتینی و سیبرنر ۱۹۷۰ و نیکتو ۱۹۸۰) پیشنهاد کردند اگر تعداد کل دانش آموز شرکت کننده بیشتر از ۲۰ یا ۴۰ نفر هستند ۱۰ نفر گروه بالا و ۱۰ نفر گروه پایین انتخاب شوند.

- بعد از تهیه گروههای بالا و پایین، لازم است برای هر سوال اطلاعات زیر تهیه شود:
- ۱- تعداد افراد گروه بالا که هر یک از گزینه ها را انتخاب کرده یا آن را بی جواب گذاشتند.
 - ۲- تعداد افراد گروه پایین که هر یک از گزینه ها را انتخاب یا آن را بی جواب گذاشتند.
- مثال: تهیه کارت و جدول زیر (جدول ۲-۱۳ کتاب)

گروهها	الف	ب	ج	د	بدون پاسخ	جمع
۲۵% بالا	۰	۵	۳	۰	۲	۱۰
۲۵% پایین	۵	۲	۳	۰	۰	۱۰

محاسبه ضریب دشواری سؤال:

درصد کل آزمون شوندگانی که به یک سؤال جواب درست می دهند. ضریب دشواری (difficulty index) آن سؤال است که با حرف P نمایش داده می شود.

ضریب دشواری عبارتست از:

R= تعداد کسانی که به سؤال جواب درست دادند

T= تعداد کل آزمون شوندگان

$$P = \frac{R}{T}$$

مثال: اگر از تعداد کل ۸۰ نفر شرکت کننده ۶۵ نفر به سؤال جواب درست داده باشند ضریب دشواری:

$$P = \frac{65}{80} = 81$$

در مواردی که تعداد شرکت کنندگان زیاد است و اطلاعات محدود است لازم است از فرمول زیر استفاده کنیم:

$$P = \frac{\text{انتخابات درست} + \text{انتخابهای درست گروه بالا}}{\text{تعداد افراد گروه بالا} + \text{تعداد افراد گروه پایین}} \times 100$$

$$\text{برای مثال جدول قبل: } 35 = 100 \times \frac{7}{20} \times 100 = \frac{5+2}{10+10}$$

تفسیر ضریب دشواری:

در آزمونهای وابسته به هنجار برای تفسیر ضریب دشواری از آمار استفاده می کنیم. انتظار ما این است که برای آزمون شوندگان نمراتی که بدست می آید در طول یک پیوستار باشد و هر چه پراکندگی بیشتر باشد بهتر است به عبارت دیگر هر چه واریانس نمرات حاصل از یک آزمون بزرگتر باشد آن آزمون، آزمون بهتری است.

واریانس نمرات گروهی

۱- واریانس سؤالات

پاسخ دهند به یک آزمون

از دو قسمت تشکیل شده است

۲- همبستگی بین سؤالات

هر سؤالی که دارای واریانس بزرگتری است و با سایر سؤالات آزمون همبستگی بیشتر دارد آن سؤال قوی تر و بهتر است.



محاسبه واریانس سؤال:

ضریب دشواری P واریانس $S = P \times (1 - P)$

یک سؤال زمانی دارای حداقل واریانس است که ضریب دشواری آن ۰ یا ۱ باشد یعنی یا همه به آن سؤال پاسخ درست داده باشند. یا هیچکدام پاسخ درست نداده باشند.

$$S = 0 \times (1 - 0) = 0$$

$$S = 1 \times (1 - 1) = 0$$

اما زمانی که ضریب دشواری ۵۰٪ باشد واریانس حداکثر است.

$$S = 50\% \times (1 - 50\%) = 25\%$$

نتیجه: برای انتخاب یک سؤال برای آزمون، سؤالهایی بهتر هستند که ضریب دشواری آنان بین (۰ و ۱) باشد و به ۵۰٪ نزدیک تر باشد.

محاسبه ضریب تمیز، سؤال:

ضریب تمیز برخلاف ضریب دشواری که سختی یا آسانی سؤال را مشخص می کند. قدرت سؤال را در تمایز یا تشخیص گروه قوی و ضعیف آزمون شوندگان مشخص می کند، یعنی مشخص می کند سؤال تا چه اندازه می تواند گروه قوی و ضعیف را از هم جدا سازد:

انتخابهای درست گروه پایین- انتخابهای درست گروه بالا = (ضریب تمیز سؤال)

تعداد افراد یک گروه (بالا یا پایین)

5-2

10

در مثال جدول قبل:
 $d = \frac{3}{10} = 30\%$

- تفسیر ضریب تمیز: هر قدر ضریب تمیز بزرگ باشد، قوه تمیز آن سؤال بیشتر است
- ۱- مثلاً اگر قوه تمیز ۹% باشد، آن سؤال آزمون شوندگان قوی و ضعیف را خوب از هم جدا می کند.
 - ۲- اما اگر قوه تمیز ۱۰% باشد آن سؤال از عهده جداسازی قوی و ضعیف بر نمی آید.
 - ۳- ضریب تمیز صفر یعنی این که سؤال بهیچ وجه نتوانسته بین گروه قوی و ضعیف تمایز ایجاد کند.
 - ۴- گاهی اتفاق می افتد ضریب تمیز منفی است یعنی در این سؤال گروه قوی برتر از گروه ضعیف عمل کرده پس این سؤال دارای عیب و اشکال اساسی است و باید تجدید نظر شود.

همبستگی سؤال با کل آزمون:

ضریب تمیز سؤال نوعی ضریب توافق سؤال با کل آزمون است سؤالی که ضریب تمیز آن بالاست همبستگی آن با کل آزمون بیشتر است یعنی اگر در سؤالی افراد قوی (بالا) به آن پاسخ درست داده اند یعنی کسانی که نمره شان در کل آزمون بالاست بنابراین آن سؤال با کل آزمون همبسته است.

پس یکی از راههای محاسبه ضریب تمیز سؤالهای یک آزمون محاسبه ضریب همبستگی بین نمره آزمون شوندگان در آن سؤال و نمرات آنان در کل آزمون است.

این ضریب همبستگی به صورت ضریب همبستگی دو رشته ای نقطه ای محاسبه می شود.

- ۱- اگر ضریب همبستگی سؤال با آزمون بیشتر باشد یعنی سؤال با کل آزمون توافق بیشتری دارد.
- ۲- اگر ضریب همبستگی سؤال با کل آزمون منفی باشد یعنی سؤال با آزمون ناهمخوانی دارد.
- ۳- اگر ضریب همبستگی سؤال با کل آزمون صفر باشد سؤال از ضریب تمیز برخوردار نیست.

همبستگی میان سؤالهای آزمون:

علاوه بر ضریب همبستگی میان سؤال آزمون با کل آزمون ضریب همبستگی تک تک سؤالها را با هم تعیین می کنیم.

۱- با در دست داشتن میانگین ضرایب همبستگی بین سؤالها می توان پایایی آزمون را محاسبه کرد

فواید این کار

۲- با در دست داشتن این نوع همبستگی می توان به مفهوم ضریب تمیز نیز دست یافت.

اگر همبستگی بین یک سؤال

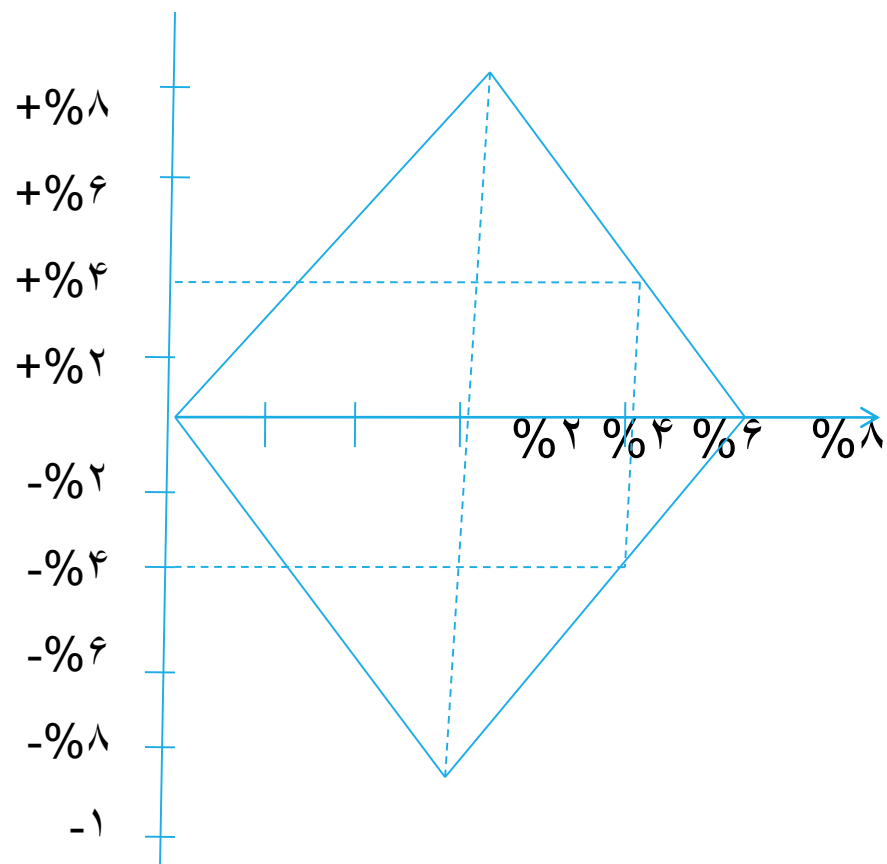
و کل آزمون کم باشد

دو دلیل وجود دارد

اول: ممکن است سؤال مورد نظر با هیچ کدام از سؤالهای دیگر همبسته نباشد که باید کنار گذاشت.

دوم: ممکن است این سؤال با بعضی سؤالها همبستگی زیاد داشته باشد. با برخی سؤالات همبسته نباشد و موجب آن می شود. سؤالات آزمون چیزهای مختلفی را اندازه می گیرند و پایایی آن پایین است.

رابطه بین ضریب دشواری و ضریب تمیز:



سؤال خوب، سؤالی است با ضریب دشواری متوسط و ضریب تمیز بزرگ.