



ارزشیابی از یادگیری

دکتر جواد وفوری

اندازه گیری (Measurement)

- عبارتست از فرایندی که تعیین می کند یک شخص یا یک شی چه مقدار از یک ویژگی برخوردار است. (می. ۱۹۹۱)
- اندازه گیری از کار بست ابزارهای پیشرفته و پیچیده الکترونیکی تا امتحانات کتبی، مقیاسهای درجه بندی و فرمهای دارایی را شامل می شود (پین، ۲۰۰۳)
- شیوه نسبت دادن یک عدد آن صفت یا ویژگی را که آن شخص دارا می باشد، نشان دهد. (کوهن، ۲۰۰۱) به یک صفت یا ویژگی معین به گونه ای که آن عدد، درجه ای از

اندازه گیری (Measurement)

➤ اندازه گیری فرایند یا شیوه ای دقیق و نظام دار است

▪

➤ اندازه گیری در بر گیرنده قواعدی است برای نسبت دادن اعداد به اشیا و یا افراد به گونه ای که صفت‌هایی از آنها را به صورت کمی نشان دهد .

➤ “قواعد اندازه گیری” خط مشی‌هایی هستند ، برای نشان دادن مقدار شیء مورد اندازه گیری

تعریف آزمون (Test)

- آزمون وسیله یا روشی نظامدار برای اندازه گیری نمونه‌ای از رفتار است (کمرانلاند، ۱۹۹۰)

- «آزمون به عنوان یک ابزار یا یک شیوه نظامدار مشاهده و توصیف ویژگیهای دانش آموزان تلقی می شود، که یک مقیاس عددی یا یک طرح طبقه‌بندی را به کار می‌بندد.» (نیتکو، ۲۰۰۱)

تعریف آزمون (Testing)

- وقتی برای اندازه‌گیری یکی از ویژگیهای روانی یا تربیتی یک فرد یا گروهی از افراد از آزمون استفاده می‌شود به این فعالیت آزمون، گویند.
- «اندازه‌گیری متغیرهای روانشناختی به وسیله تدابیر یا روشهایی که به منظور کسب نمونه‌هایی از رفتار طراحی شده‌اند، آزمون روانی یا روان آزمایی نام دارد»
(کوهن ۲۰۰۲)
- هر وقت آزمون صورت می‌گیرد، یعنی به کار می‌رود، نوعی اندازه‌گیری انجام می‌شود.
- اما هر نوع اندازه‌گیری الزاماً آزمون نیست.
- اندازه‌گیری از آزمون مفهوم گسترده‌تری دارد. چون آزمون نوع خاصی از اندازه‌گیری است.

سنجش Assessment

■ 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 آموزشی مورد استفاده قرار می گیرد

■ در سنجش وسایل و فنون مختلفی برای جمع آوری اطلاعات به کار می روند: از جمله آنها می توان بر آزمون، پرسشنامه، مقیاس درجه بندی واری، کار آزمایشگاهی، پرونده تحقیقی، امتحان شفاهی تکلیف درسی، مصاحبه و مشاهده عملکرد و رفتار دانش آموزان در موقعیتهای مختلف اشاره کرد

سنجش Assessment

■ سنجش به یک تحلیل جامع و چند وجهی از عملکرد گفته می شود (ویگینز، ۱۹۹۳)

■ سنجش شامل استفاده از فنون مختلف است و تأکید زیادی بر مشاهده عملکرد دارد ... سنجش نوعی تحلیل بالینی و پیش‌بینی عملکرد است، در حالی که آزمون یک ابزار است یک وسیله اندازه گیری است. (کرانباخ ۱۹۹۵)

■ سنجش تنها به جمع‌آوری داده‌ها یا اطلاعات خلاصه نمی‌شود بلکه به جمع‌آوری داده‌ها همراه با قصد یا منظور گفته می‌شود .

■ به دو منظور صورت می‌گیرد:

الف – تعیین و تشخیص مشکلات دانش آموزان

ب – تصمیم‌گیری درباره آنان

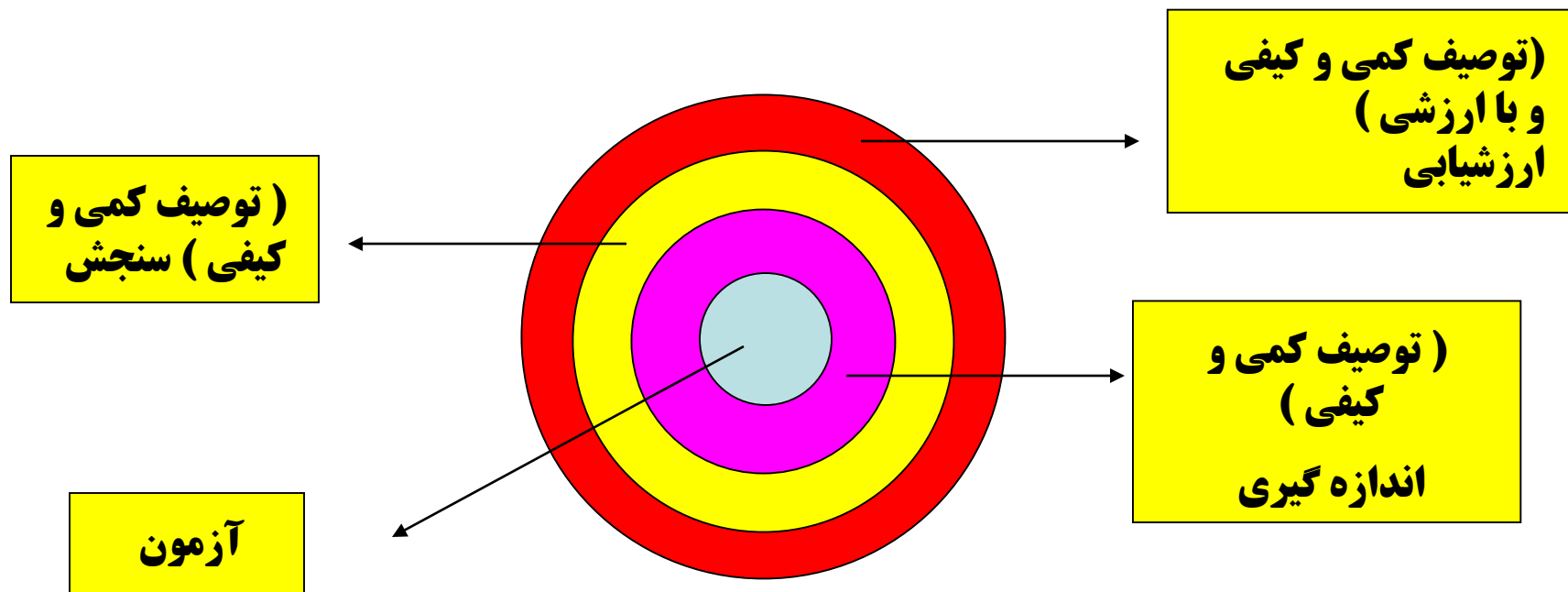
ارزشیابی Evaluation

■ ارزشیابی به یک فرایند نظامدار برای جمع آوری، تحلیل و تفسیر اطلاعات گفته می شود به این منظور که تعیین شود آیا هدفهای مورد نظر تحقق یافته اند یا در حال تحقق یافتن هستند و به چه میزان؟

■ یکی از ویژگیهای مهم ارزشیابی تعیین کیفیت است.

■ در فرایند ارزشیابی داوری ارزشی با توجه به کیفیت به عمل می آید.

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



رابطه بین آزمون، اندازه گیری، سنجش و ارزشیابی (از نگارنده)

<u>Dimension of Difference</u>	<i>Assessment</i>	<i>Evaluation</i>
Timing	Formative	Summative
Focus of Measurement	Process-Oriented	Product-Oriented
Relationship Between Administrator and Recipient	Reflective	Prescriptive
Findings, Uses Thereof	Diagnostic	Judgmental
Ongoing Modifiability of Criteria, Measures Thereof	Flexible	Fixed
Standards of Measurement	Absolute	Comparative
Relation Between Objects of A/E	Cooperative	Competitive

مقایسه‌های اندازه گیری

مقایسه‌های اندازه‌گیری به چهار نوع یا در چهار سطح دسته‌بندی شده‌اند. این طبقه‌بندی از ساده به پیچیده و به صورت سلسله مراتبی تنظیم شده است. یعنی هر طبقه دارای تمامی ویژگیهای طبقات قبل به اضافه ویژگیهای مخصوص آن طبقه است.

انواع داده ها (متغیرها)

نوع داده ها

کیفی

کمی

اسمی

Nominal

متغیرهایی که
فقط بوسیله
اسمشان از هم
جدا می شوند

ترتیبی

Ordinal

متغیرهایی هستند که هم
توسط اسمشان جدا می
شوند و هم ترتیبی نیز در
آنها وجود دارد

فاصله ای

Interval

دارای خصوصیات اسمی –
ترتیبی می باشد و با عدد
نشان داده می شود

نسبی

Ratio

دارای خصوصیات
اسمی – ترتیبی
– فاصله ای
است و نیز
دارای صفر
مطلق می باشد

مقیاس اسمی Nominal

■ پایین‌ترین سطح یا نوع اندازه‌گیری را معرفی می‌کند به دو صورت مورد استفاده قرار می‌گیرد ساده‌ترین.

■ **اسم‌گذاری** – از ارقام و اعداد برای اسم‌گذاری و تشخیص اشیاء و افراد استفاده می‌شود بدون اینکه این اعداد هیچ‌گونه مفهوم و معنای ریاضی داشته باشند.
شماره پیراهن بازیکنان – شماره خیابان‌های شهر

■ **طبقه‌بندی** – از ارقام و اعداد برای طبقه‌بندی استفاده می‌شود. اعداد هیچ‌گونه مفهوم و مبنایی ریاضی ندارد.
عدد ۱ به زن و عدد ۰ به مرد

مقیاس اسمی Nominal

عملیات مجاز آماری

شمارش فراوانی، تعیین تعداد موردهای درون هر طبقه، تعیین نما

عملیات مجاز ریاضی

هیچ یک از چهار عمل اصلی جمع، تفریق، ضرب، تقسیم میسر نیست.

مقیاس ترتیبی Ordinal

■ دسته‌ای از افراد یا اشیاء با توجه به یک صفت از بزرگ به کوچک (یا برعکس) مرتب می‌شوند.

■ معلوم نیست که هر یک از این افراد یا اشیاء به مفهوم مطلق چقدر از آن صفت را دارند.

■ مشخص نیست که این اشیاء از لحاظ صفت مورد اندازه‌گیری چه مقدار از یکدیگر فاصله دارند.

مثال : مرتب کردن دانش‌آموزان یک کلاس به ترتیب قد و شماره‌گذاری آنها از کوچک به بزرگ

مقیاس ترتیبی Ordinal

عملیات مجاز آماری :

شمارش فراوانی، تعیین نما، محاسبه میانی، محاسبه درصدها و محاسبه همبستگی رتبه‌ای اسپیرمن عملیات مجاز ریاضی

عملیات مجاز ریاضی :

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

مقیاس فاصله‌ای Interval

مقیاس فاصله‌ای مقیاسی است که در آن :

■ رتبه اشیاء یا افراد با توجه به یک صفت مشخص است .

■ معلوم است که اشیاء یا افراد از لحاظ صفت مورد اندازه‌گیری چه مقدار از یکدیگر فاصله دارند.

■ هیچ نوع اطلاعی راجع به مقدار مطلق صفت مورد نظر برای هیچ یک از اشیاء یا افراد در دست نیست مانند دمای هوا در مقیاس فارنهایت یا سانتی‌گراد.

مقیاس فاصله‌ای Interval

عملیات مجاز آماری :

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

عملیات مجاز ریاضی :

– جمع و تفریق مجاز است ولی ضرب و تقسیم مجاز نیست.

مقیاس نسبی Ratio

مقیاسی است که در آن:

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

۲- فاصله بین اشخاص معلوم است

۳- علاوه بر اینها فاصله حداقل یکی از اشخاص از یک صفر منطقی نیز معلوم است.

اندازه‌گیری طول، وزن، حجم و سایر صفات فیزیکی اشیاء تمامی عملیات مجاز آماری و ریاضی میسر است.

ویژگیها					مقیاس
عملیات مجاز ریاضی	عملیات مجاز آماری	وجود صفر مطلق	خواص مساوی طبقات	وجود تریب در طبقات	
هیچ یک از چهار عمل	شما، فراوانی	-	-	-	اسمی
هیچ یک از چهار عمل اصلی	شما، فروانی، میانه، ضریب همبستگی اسپریمن	-	-	+	ترتیبی
جمع، تفریق	همه عملیات آماری	-	+	+	فاصله‌ای
همه عملیات ریاضی	همه عملیات آماری	+		+	نسبی



مقیاس پیوسته و مقیاس گسسته

مقیاس گسسته (Discontinuous)

در این مقیاس متغیر مورد نظر تنها می‌تواند یکی از ارزشهای معین را به خود اختصاص دهد. ویژگی مهم این نوع مقیاس تساوی واحدهای شمارش آنها در طول مقیاس است.

انجام تمام اعمال ریاضی در مورد آن مجاز است.

مثال:

- جنس (مذکر یا مونث)
- تعداد افراد خانوار
- وضعیت تأهل (مجرد، متأهل)
- سطح تحصیلات (۱ تا ۲۰ سال)



مقیاس پیوسته (Continuous)

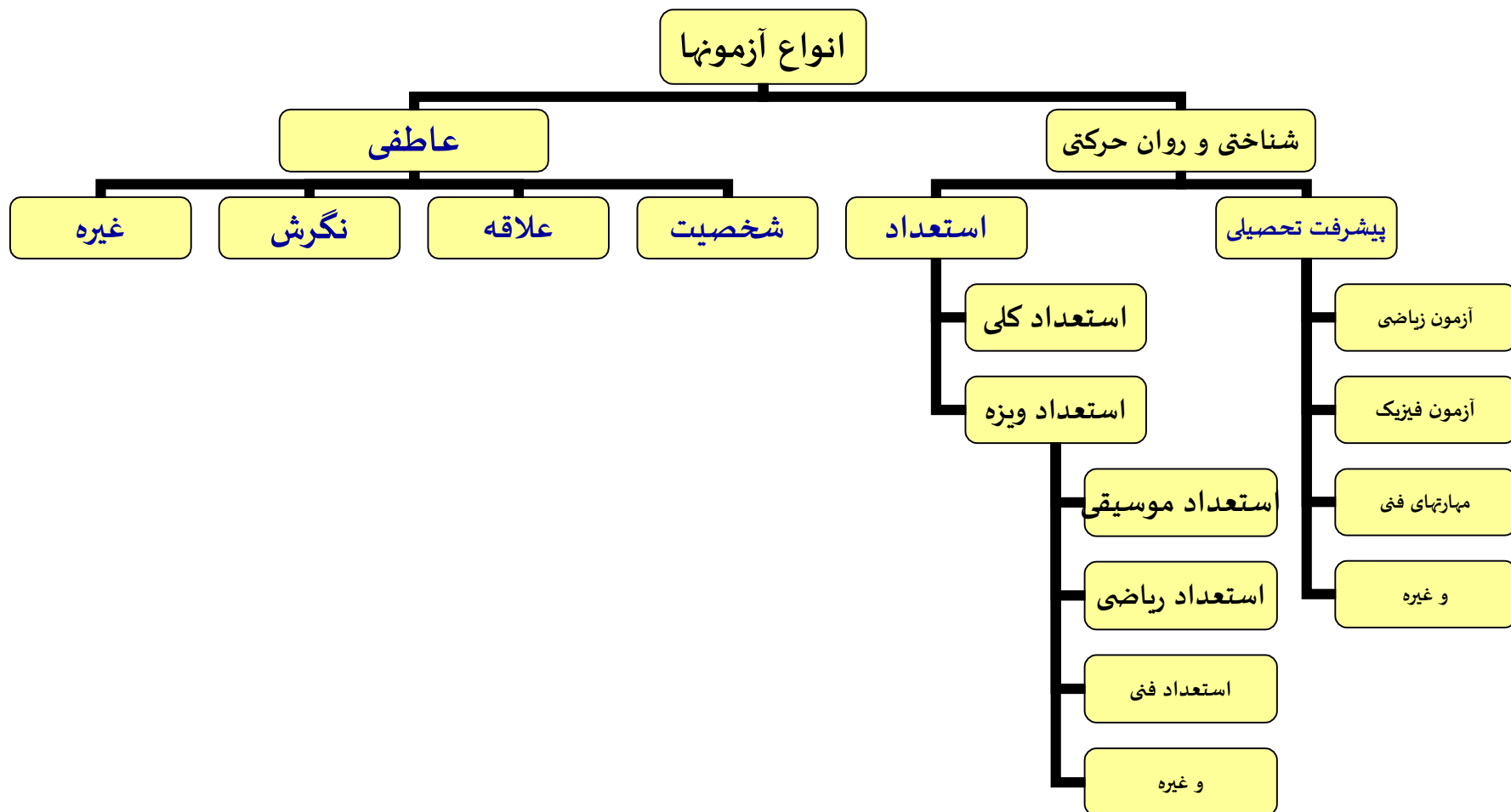
در مقیاس پیوسته متغیر مورد نظر در فاصله بین نقاط مقیاس، هر تعدادی ارزش را به خود اختصاص دهد.
مانند:

سرعت، زمان، طول، وزن، هوش، پیشرفت تحصیلی، استعداد ریاضی و نگرش

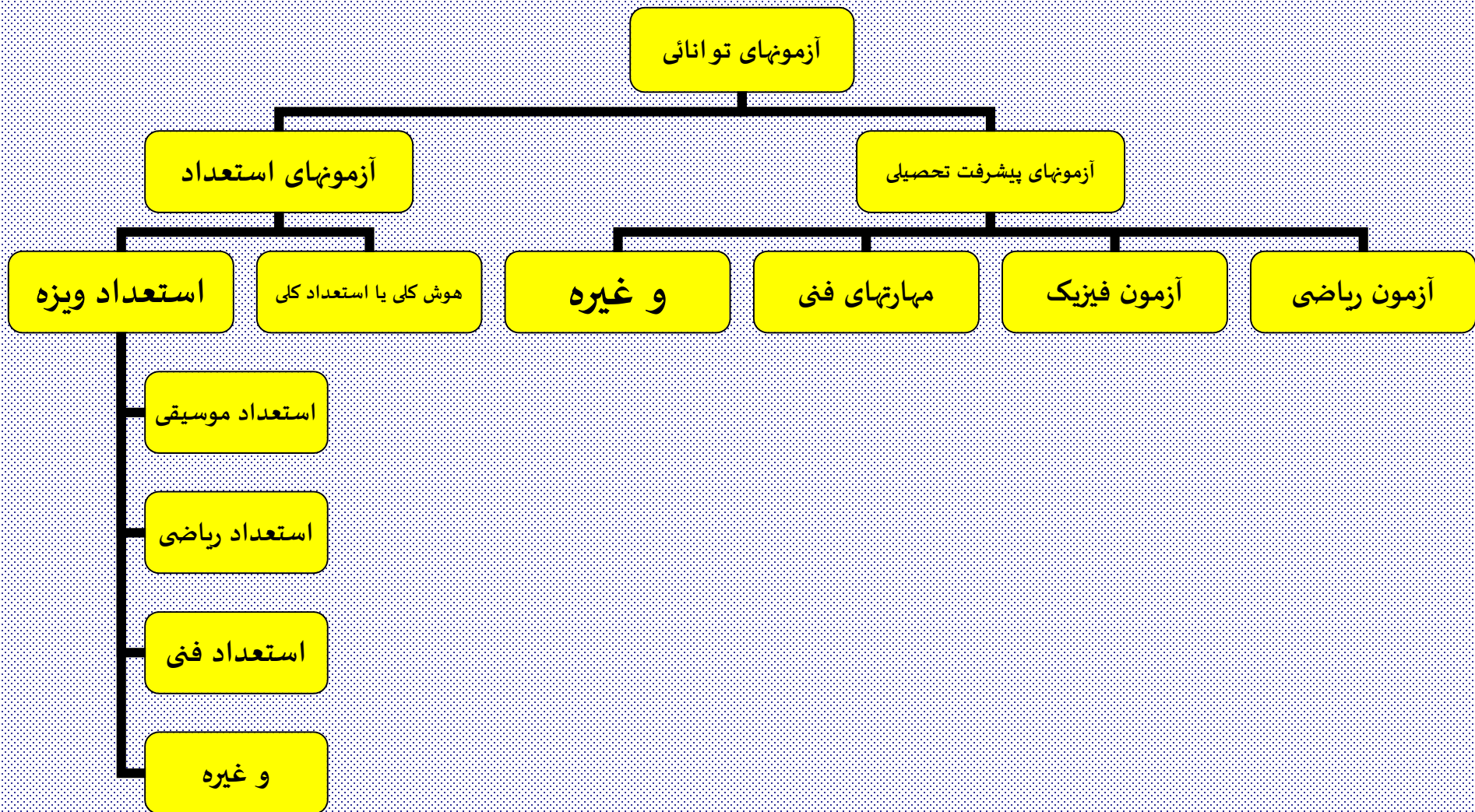
اندازه‌گیری از متغیرهای گسته یا مجزا همواره دقیق و قطعی است.

اگر بخواهیم قد دانش‌آموزان یک دبیرستان را اندازه بگیریم. هر گونه وسیله‌ی اندازه‌گیری که انتخاب کنیم، تقریبی خواهد بود زیرا همیشه تصور وسیله دقیق‌تری امکان‌پذیر است. اگر اندازه قد دانش‌آموزان را تا دهم اعشاری اندازه‌گیری کرده باشیم می‌توان مقیاس دیگری در نظر گرفت که قد این افراد را تا صدم اعشاری اندازه بگیرد.

انواع آزمونهای مورد استفاده در آموزش و پرورش



دسته بندی آزمون‌ها با توجه به ویژگیهای مورد اندازه گیری



دسته‌بندی آزمون‌ها با توجه به چگونگی تهیه آنها

❖ آزمونهای میزان شده

آزمون‌هایی هستند که با استفاده از نمونه‌های بزرگ آزمون شوندگان و غالباً به وسیله مؤسسه‌های دولتی یا خصوصی و یا مراکز آموزش و علمی و برای اندازه‌گیری پیشرفت تحصیلی یا هوش و استعداد گروه‌های بزرگ دانش‌آموزان و با مقاصد ارزشیابی از دوره‌های تحصیلی تهیه و توزیع می‌شوند.

❖ آزمونهای معلم ساخته

آزمون‌هایی هستند که معلمان برای سنجش پیشرفت تحصیلی دانش‌آموزان خود تهیه می‌کنند و در طول دوره آموزش یا در پایان دوره به اجرا در می‌آورند.



تفاوت‌های آزمون‌های میزان شده و معلم ساخته

۱- تهیه آزمون‌های میزان شده در مقایسه با آزمون‌های معلم ساخته به زمان و تخصص بیشتری نیاز دارند.

۲- آزمون‌های میزان شده ویژگی‌های اختصاصی کلاسها و مدارس بخصوص را نادیده می‌گیرند و بر وجوه مشترک محتوای درسی مناطق مختلف تکیه می‌کنند، ولی آزمون‌های معلم ساخته با توجه به محتوا و هدف‌های ویژه درس‌های بخصوص، برای مقایسه پیشرفت دانش‌آموزان مناطق مختلف مفید نیستند.

۳- اصطلاح آزمون میزان شده به وجود داده‌های هنجاری (Normative data) اشاره می‌کند. اما آزمون‌های معلم ساخته نیازی به هنجار و هنجاریابی ندارند.

دسته‌بندی ارزشیابی‌های آموزشی با توجه به ملاک مورد استفاده

ارزشیابی وابسته به ملاک Criterion referenced

در ارزشیابی وابسته به ملاک یا ارزشیابی ملاک مرجع، معیار یا ملاک از پیش تعیین می‌شود و عملکرد یادگیرنده با توجه به آن ملاک مورد قضاوت قرار می‌گیرد. در آن نوعی ملاک مطلق که همان هدفهای آموزشی از پیش تعیین شده هستند، مورد استفاده قرار می‌گیرند.

ارزشیابی وابسته به هنجار Norm referenced

آن نوع ارزشیابی است که به جای یک ملاک مطلق از پیش تعیین شده بر نوعی ملاک نسبی متکی است. در واقع عملکرد دانش‌آموزان با یکدیگر مورد مقایسه قرار می‌گیرد و در آن ملاک یا معیار ارزشیابی، نوعی هنجار گروهی نیست.



ارزشیابی تشخیصی
Diagnostic

ارزشیابی آغازین
Preassessment

انواع ارزشیابی
با توجه به زمان و هدف

ارزشیابی تراکمی
Summative

ارزشیابی تکوینی
Formative

ارزشیابی تکوینی

■ آنچه عمدتاً به منظور کمک به اصلاح موضوع مورد ارزشیابی یعنی برنامه یا روش آموزش مورد استفاده قرار می‌گیرد، ارزشیابی تکوینی نام دارد (پوفام، ۱۹۷۵)

■ هدف از اجرای آن آگاه ساختن تولیدکنندگان برنامه از نواقص برنامه خود و کمک به اصلاح آنهاست. این ارزشیابی در طول دوره آموزشی، یعنی زمانی که فعالیت آموزشی معلم هنوز در جریان است و یادگیری دانش‌آموزان در حال تکوین و شکل‌گیری است، انجام می‌شود.

■ برای بهره‌گیری کامل از نتایج ارزشیابی تکوینی، بیان هدفهای آموزشی به صورت دقیق و رفتاری ضروری است.



ارزشیابی تشخیصی

- ✱ این ارزشیابی با هدف تشخیص مشکلات یادگیری دانش آموزان به کار می رود.
- ✱ زمانی مورد استفاده قرار می گیرد که معلم با مشکلات مبرم و مکرری درمورد یک یا چند دانش آموز روبه رو می شود.
- ✱ که با روشهای اصلاحی معمول ارزشیابی تکوینی قابل رفع شدن نیست.
- ✱ آزمونهای تشخیصی توصیف کاملی از نقاط قوت و ضعف یادگیرندگان در زمینه های مهارتی خاص به دست می دهند. و روشهای مناسب رفع مشکلات را ارائه می کنند.



ارزشیابی تراکمی

- تمامی آموخته‌های دانش‌آموزان در طول یک دوره آموزشی تعیین می‌شوند و هدف آن نمره دادن به دانش‌آموزان و قضاوت درباره اثربخشی کار معلم و برنامه‌های مختلف درسی با یکدیگر است.
- هدف ارزشیابی تراکمی هم نمره دادن و اتخاذ تصمیمهای مختلف و هم تعیین اثربخشی طرح آموزشی است.
- این ارزشیابی در پایان برنامه یعنی زمانی که به طور کامل اجرا شده و به نتیجه رسیده است، انجام می‌پذیرد.

مرحله ارزشیابی	زمان اجرا	نوع تصمیمها	فعالیت‌های ویژه
مرحله طرح ریزی Planing Phase	پیش از اجرای آموزش ، برنامه، پروژه و غیره	۱- نتایج که قرار است به دست آیند ۲- اعمالی که باید انجام شوند.	۱- تحلیل موقعیت ۲- تعیین و توصیف هدف ۳- توصیف پیش‌نیازها ۴- التهاب یا تهیه وسایل اندازه‌گیری ۵- توصیف استراتژیها ۶- انتخاب طرح پژوهشی ۷- تدارک برنامه زمان‌بندی
مرحله فرایندی Process Phase	ضمن اجرای آموزش، برنامه پروژه و غیره	۱- به چه میزانی مراحل طرح‌ریزی شده به اجرا در می‌آیند؟ ۲- چه تغییراتی باید داده شوند تا پیشرفت لازم به دقت آید؟	۱- اجرای پیش‌آزمون ۲- اجرای آزمون رفتار ورودی ۳- سنجش مناسب بودن هدفها ۴- جمع‌آوری داده‌ها به طور دروهای ۵- تحلیل اثربخشی استراتژیها
مرحله فراورده‌ای Product Phase	پس از اجرای آموزش، برنامه، پروژه و غیره	۱- اثربخشی کلی آموزشی، پروژه، برنامه و غیره ۲- اقدامات آتی	۱- جمع‌آوری داده‌های مربوط به هدفها ۲- جمع‌آوری داده‌ای مربوط به بازده‌های پیش‌بینی نشده ۳- تحلیل و تغییرها ۴- تهیه گزارش



ارزشیابی پیشرفت تحصیلی

سنجش عملکرد یادگیرندگان و مقایسه نتایج حاصل با هدفهای آموزشی از پیش تعیین شده، به منظور تصمیم گیری در این باره که آیا فعالیتهای آموزشی معلم و کوششهای یادگیری دانش آموزان یا دانشجویان به چه میزان به نتایج مطلوب انجامیده اند .

ارزیابی عملکرد

در ارزشیابی پیشرفت تحصیل دو اقدام اساسی ضروری است :

- تعیین هدفهای آموزشی
- سنجش یا اندازه گیری عملکرد یادگیرندگان



هدفهای آموزشی یا هدفهای یادگیری

دو نوع هدف وجود دارد:

الف - کلی یا غائی یا آرمانی

ب - دقیق، عینی که برخاسته از هدفهای کلی هستند هدفهای آموزشی

ppt.۲



هدفهای رفتاری

ویژگیهای هدفهای رفتاری :

۱. این هدفها بر حسب رفتار اندازه پذیر دانش آموز نوشته می شوند.
۲. دارای شرایط یا وضعیت عملکرد هستند.
۳. دارای ملاک عملکرد هستند.

مثال : دانشجو باید بتواند از میان فهرستی از ماشینها و وسایل دیگری که معلم در اختیار او می گذارد، حداقل پنج ماشین را انتخاب کند که همزمان در دو راه یا بیشتر به انسان کمک می کنند.



طبقه بندي هدفهاي آموزشي

۱. هدفهاي حوزه شناختي (Cognitive domain)

به جريانهائي كه با فعاليتهاي ذهني و فكري آدمي سروكار دارند. مربوط ميشوند.

۲. هدفهاي حوزه عاطفي (Affective domain)

شامل هدفهايي هستند كه تغييرات حاصل در علاقه‌ها، نگرشها و ارزشها و نيز رشد ارج‌شناسي و سازگاري را نشان مي‌دهند.

۳. هدفهاي حوزه رواني حرکتي (Psychomotor domain)

به زمينه مهارتهاي حرکتي يا حرکات بدني مربوط مي شوند.

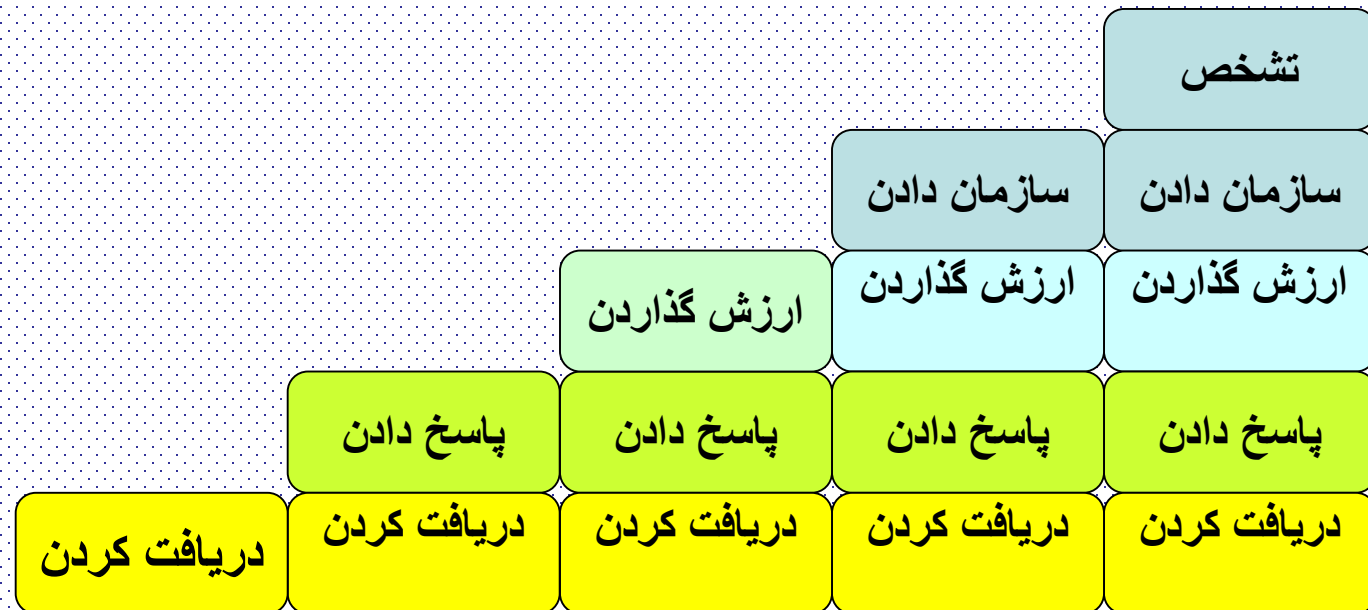
بعد دانشی

			دانش فراشناختی
		دانش روندی	دانش روندی
	دانش مفهومی	دانش مفهومی	دانش مفهومی
دانش امور واقعی	دانش امور واقعی	دانش امور واقعی	دانش امور واقعی

بعد فرایند شناختی

					آفریدن
				ارزشیابی	ارزشیابی
			تحلیل کردن	تحلیل کردن	تحلیل کردن
		به کار بردن	تحلیل کردن	تحلیل کردن	تحلیل کردن
	فهمیدن	فهمیدن	فهمیدن	فهمیدن	فهمیدن
به یاد آوردن	به یاد آوردن	به یاد آوردن	به یاد آوردن	به یاد آوردن	به یاد آوردن

حوزه عاطفی





ابتکار

انطباق

انطباق

پاسخ پیچیده
آشکار

پاسخ پیچیده
آشکار

پاسخ پیچیده
آشکار

مکانیسم یا
عادت

مکانیسم یا
عادت

مکانیسم یا
عادت

مکانیسم یا
عادت

پاسخ هدایت
شده

پاسخ هدایت
شده

پاسخ هدایت
شده

پاسخ هدایت
شده

پاسخ هدایت
شده

آمادگی

آمادگی

آمادگی

آمادگی

آمادگی

آمادگی

ادراک حسی

ادراک حسی

ادراک حسی

ادراک حسی

ادراک حسی

ادراک حسی

ادراک حسی

برای تعیین سئوالهای که به بخشهای مختلف محتوا و هدفهای تعلق می گیرند ، جواب سئوالهای زیر را باید تعیین کرد :

۱- کدام بخش از محتوا و کدام دسته از هدفها دارای حجم بیشتری است و وقت بیشتری صرف آموزش آنها شده است .

۲- کدام هدفها از لحاظ نگهداری در حافظه و کاربرد در آینده ، ارزش بیشتری دارند

۳- کدام قسمت یا قسمتهای درس یا موضوع درسی از اهمیت بیشتری برخوردارند.



■ آزمودنیهای پیشرفت تحصیلی متداول به آزمودنیهای کتبی یا آزمودنیهای مداد و کاغذی معروفند و به طور عمده با حوزه شناختی سر و کار دارند.

■ هدفهای آموزشی و بازدههای یادگیری متعلق به حوزههای غیر شناختی (از جمله نگرشها، علاقهها و ارج نهادنها) و مهارتهای مختلف به کمک وسائل دیگر سنجش از جمله:

■ مشاهده، مقیاس درجه بندی (Rating scale)، فهرست واریسی (Check list)، واقعہ نگاری (Anecdotal record)، ثبت حضور و غیاب (Inventory)، گزارشهای کار عملی، مورد سنجش قرار می گیرند.

ابزارهای سنجش هدفهای غیر شناختی

■ مشاهده

■ مقیاس درجه‌بندی (Rating scale)

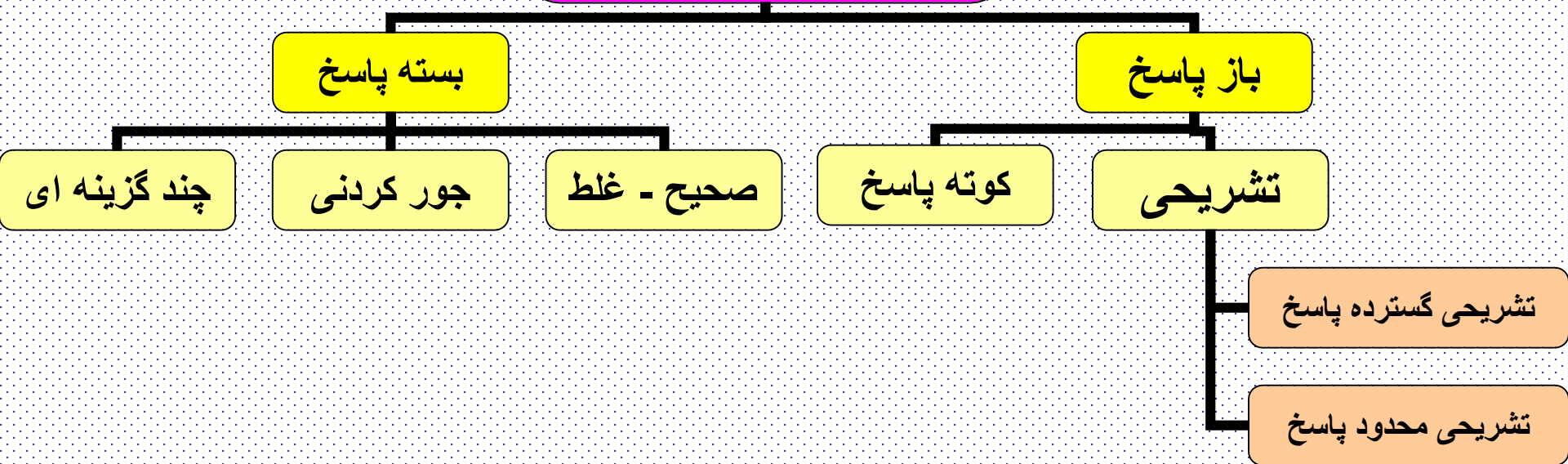
■ فهرست وارسی (Check list)

■ واقع‌نگاری (Anecdotal record)

■ ثبت حضور و غیاب (*Inventory*)

■ گزارشهای کار عملی

انواع آزمونهای پیشرفت تحصیلی



آزمونهای گسترده – پاسخ

هیچ گونه محدودیتی برای آزمون شونده منظور نمی شود و او آزاد است تا هر طور که مایل باشد، پاسخ خود را پیرو راند و سازمان دهد. آزمون شونده هم از لحاظ زمان پاسخدهی و هم از جهت مقدار پاسخ آزادی کامل دارد.
مثال: تاثیر قانون مند را بر تحول زیست شناسی به عنوان یک علم شرح دهید.

آزمونهای محدود – پاسخ

آزمون شونده در دادن پاسخ به سؤالات آزادی کامل ندارد. بلکه صورت سؤال او را ملزم می سازد تا پاسخ خود را در چارچوب شرایطی خاص محدود کند.
مثال: توضیح دهید که چرا فشار سنج یکی از ابزارهای مفید پیش بینی هواست؟
جواب خود را در یک پاراگراف بنویسید. برای

محاسن آزمونهای تشریحی :

۱- تهیه آزمونهای تشریحی از تهیه آزمونهای عینی آسانتر است.

۲- این نوع آزمون تنها وسیله موجود سنجش توانایی آزمون شونده در پروراندن جواب سؤالات و بیان آنهاست.

۳- این آزمونها توانایی پاسخ دادن به سؤالات را می سنجند نه توانایی انتخاب پاسخها را.

۴- آنها موقعیهای واقع تری را از آزمونهای عینی به آزمون شوندگان عرضه می کنند.

۵- آنها بر روش مطالعه یاد گیرندگان تاثیر مثبت به جای می گذارند.

معایب آزمونهای تشریحی :

۱- این آزمونها نمونه کوچکی از محتوای درس و هدفهای آموزش را اندازه می گیرند.

۲- تصحیح برگه های امتحانی این گونه آزمونها نمی تواند با دقت و به طور عینی انجام گیرد.

۳- تصحیح برگه های امتحانی این آزمونها بسیار وقت گیر است.

قواعد تهیه سؤالهای تشریحی :

۱- در نوشتن صورت سؤالهای آزمون یک درس، با رسم جدول مشخصات آن درس، دقت کنید که سؤالها به طور مستقیم به هدفهای آموزشی مربوط شوند.

۲- سؤالهای تشریحی را تنها به اندازه گیری هدفهایی محدود کنید که با سایر انواع آزمونها به خوبی قابل اندازه گیری نیستند.

۳- صورت سؤالهای تشریحی را با عبارت و کلمات واضح و روشن بنویسید و از کلی گویی و ابهام در بیان پرهیزید .

۴- از کاربرد کلمات «چه کسی» «چه وقت» «کجا» و چرائیها پرهیزید .

۵- تا حد امکان از سؤالهای تازه و موقعیتهای جدید استفاده کنید .

قواعد تهیه سؤالهای تشریحی :

۶- سؤالها مربوط به موضوعات و عقاید بحث انگیز باید طوری طرح شوند که از آزمون شونده بخواهند تا شواهد لازم برای مستند کردن عقیده انتخابی را بیان کند نه اینکه از او بخواهند تا صرفاً عقاید شخصی خود را شرح دهد.

۷- به آزمون شوندگان حق انتخاب چند سؤال از میان تعدادی سؤال را ندهند

۸- برای پاسخ دادن به سؤالها زمان کافی در نظر بگیرید و زمان هر سؤال را نیز به طور جداگانه مشخص کنید .

۹- با نوشتن سؤالهایی که به جواب کوتاه نیاز دارند تعداد آنها را افزایش دهید و از شکل ضعف نمونه گیری آزمونه‌های تشریحی بکاهید .

۱۰- عواملی را که در ارزشیابی تشریحی دخالت خواهید داد از پیش تعیین کنید و آنها را به اطلاع آزمون شوندگان برسانید.

قواعد تصحیح پاسخهای سؤالهای تشریحی :

- ۱- پاسخهای سؤالها را تنها بر اساس هدفی که در سؤال گنجانیده شده است , تصحیح کنید.
- ۲- با نوشتن یک پاسخ نمونه یا الگو برای هر سؤال به عنوان کلید از دخالت عوامل نا مربوط جلوگیری کنید.
- ۳- پاسخ را سؤال به سؤال تصحیح کنید , نه ورقه به ورقه .
- ۴- هنگام تصحیح ورقه‌های امتحانی از شناسایی نام صاحبان آنها خودداری کنید .
- ۵- در صورت امکان از یکی دو نفر از همکارانتان بخواهید تا سؤالهایی را که شما تصحیح کرده‌اید , تصحیح کنند .

قواعد تصحیح پاسخهای سؤالهای تشریحی :

۶- تمام پاسخهای آزمون شوندگان به یک سؤال را در یک نشست و بدون وقفه زمانی تصحیح کنید

۷- به نمرات سؤالهایی که قبلاً تصحیح کرده‌اید نگاه نکنید .

۸- بر روی برگه‌های آزمون اشتباهات دانش آموزان را تصحیح کنید و اظهار نظرهای خود را بنویسید .

۹- از روشهای متنوع نمره گذاری استفاده کنید .

روشهای مختلف نمره گذاری :

- روش تحلیلی Analytical
- روش کلی Global
- روش ویژگیهای اصلی Primery Traits

آزمون کوتاه پاسخ (Short answer)

از مجموعه‌ای سؤال مختصر که غالباً برای سنجش هدفهای آموزشی سطح پایین طرح می‌شوند، تشکیل می‌یابد. این سؤالات از آزمون شونده می‌خواهند تا کلمه، عبارت، جمله، عدد یا علامتی را در پاسخ به یک سؤال، یا تکمیل یک جمله بنویسند.

انواع آزمونهای کوتاه پاسخ :

نوع پرسشی — مرکز استان مازندران کجاست ؟
نوع کامل کردنی — بزرگترین دریاچه دنیا دریاچه است
نوع تشخیصی یا تداعی
— بعد از نام هر استان نام مرکز آن استان را بنویسید .

خوزستان :

خراسان :

مرکزی :



کاربرد آزمونهای کوتاه پاسخ

✓ برای اندازه گیری هدفهای شناختی سطوح پایین به ویژه دانش یا معلومات مفیدند .

✓ این سؤالات به طور عمده با کلمات و اعداد و اسامی اشخاص ، مکانها ، اشیاء ، فرایندها ، رنگها ، اعداد معرف تاریخها ، فاصله ها ، هزینه ها و جمعیتها سرو کار دارند .



محاسن آزمونهای کوتاه پاسخ

۱ – سهولت در تهیه سوالات و تصحیح پاسخها

۲ – اجرای آسان

۳ – تقلب را کاهش می دهد

۴ – نسبت به آزمونهای عینی اطلاعات تشخیصی بیشتری در اختیار معلمان می گذارد.

۵ – در آن احتمال حدس کورکورانه وجود ندارد .

معایب آزمونهای کوتاه پاسخ

۱- سؤالهای این آزمون بسیار مختصرند، لذا از آنها نمی‌توان برای سنجش هدفهای سطوح بالا استفاده کرد.

۲- کاربرد زیاد این گونه آزمونها سبب تشویق یادگیرندگان به حفظ کردن اطلاعات جزئی و بی اهمیت خواهد شد.

۳- تصحیح و نمره گذاری پاسخهای این آزمونها به دقت و سرعت آزمونهای عینی عملی نیست.

قواعد تهیه سؤالهای کوتاه پاسخ

۱- برای آزمون خود جدول مشخصات تهیه کنید و سؤالهای آزمون را با توجه به جدول مشخصات بنویسید .

۲- هر سؤال باید موضوع مهمی را شامل شود .
در سال ۱۹۵۰ در کشور آمریکا چند مورد بیماری آبله مشاهده شده است ؟

۳- صورت سؤال را کاملاً روشن بنویسید به گونه‌ای که به پاسخ مشخص معینی نیاز داشته باشد .

ذغال سنگ چیست ؟ منظور طراح سؤال نوعی سوخت
پاسخهای دیگر : یک ماده قابل اشتعال - کربن نا خالص ، بقایای سخت شده مواد گیاهی .

ذغال سنگ به چه منظوری به کار می‌رود ؟
ذغال سنگ از چه موادی درست شده است ؟
ذغال سنگ از چه عنصر شیمیایی درست شده است ؟

قواعد تهیه سؤالهای کوتاه پاسخ

۴- صورت سؤال را از روی مطالب کتاب عیناً نقل نکنید .

۵- در سؤالهایی که پاسخهای آنها اعداد هستند واحد مقیاس و میزان دقتی را که در محاسبات باید رعایت شوند مشخص کنید .
- مقدار عدد پی چقدر است ؟ $1/3$ $14/3$ $141/3$

۶- در سؤالهای کوتاه پاسخ کامل کردنی تنها کلمات و عبارات مهم را حذف کنید .
کبد گلوکز اضافی را به صورت گلیکوژن در خود می کند .

۷- در سؤالهای کامل کردنی که برای ارزشیابی پیشرفت تحصیلی به کار می روند تعداد زیادی جای خالی منظور نکنید .

۸- جای خالی سؤالهای کامل کردنی را تا آنجا که ممکن است در قسمت پایانی سؤال قرار دهید .

..... در آزمایشهای شرطی کردن پاولفی به پاسخی گفته می شود که حیوان از آن طریق به صدای زنگ واکنش نشان دهد .

قواعد تهیه سؤالهای کوتاه پاسخ

۹- جای خالی سؤالهای کامل کردنی را همواره یک اندازه تعیین کنید

۱۰- تا حد امکان به جای سؤالهای کامل کردنی از سؤالهای پرسشی استفاده کنید .

۱۱- در سؤالهای کامل موارد دیگری کردنی از کاربرد اشارات دستوری و که جواب سؤال را مشخص می کنند ، خودداری کنید .

موتور کشتی بخار با نیروی کار می کند



تعریف آزمون عینی

✓ آزمونهای عینی آزمونهایی هستند که در آنها هم سئوالها و هم جواب سئوالها در اختیار آزمون شوندگان قرار داده می‌شوند و آزمون شوندگان درباره جوابهای داده شده، تصمیمهایی می‌گیرند یا اعمالی انجام می‌دهند.

از سه نوع تشکیل شده است :

- صحیح - غلط (Tru – False)

- جور کردنی (Matching)

- چند گزینه‌ای (Multiple – choice)

✓ آزمونهای عینی توانایی بازشناسی (Recognition) ولی آزمونهای تشریحی و کوتاه پاسخ توانایی باز خوانی (Recall) را در آزمودنیها اندازه می‌گیرند.

انواع سئوالهاي صحيح - غلط

نوع صحيح - غلط :

نخستين سئوالي كه معلم هنگام انتخاب تقويت كننده براي يك كودك بايد بپرسد , اين است كه آيا آن تقويت كننده به اندازه كافي نيرومند است يا نه ؟

نوع اصلاحي:

جذر عدد ۶۴ ، ۹ است .

نوع خوشه‌اي

ميانگين حسابي :

الف- يكي از اندازه‌هاي گرايش مركزي است

ب- كمتر از ميانه تحت تاثير نمرات خيلي بزرگ يا خيلي كوچك قراغر مي‌گيرد .

ج- رابطه همبستگي بين دو متغير را نشان مي‌دهد .



محاسن

۱- بیان موضوعهای درسی در قالب عبارتهای ساده

۲- سادگی تهیه

۳- عینی بودن تصحیح جوابها

۴- زیادی تعداد آنها در هر جلسه امتحان

معایب



۱- این آزمونها اغلب براي سنجش اطلاعات جزئي و کم اهمیت به کار می‌روند

۲- حدس زدن کورکورانه در آنها خیلی زیاد دخالت دارد .

۳- کاربرد زیاد آنها سبب تاکید بی مورد بر حفظ طوطی وار اطلاعات کم اهمیت می‌شود .

۴- بیشتر براي سنجش هدفهای سطح پایین مفیدند و تواناییهای پیچیده سطح

قواعد تهیه سئوالهاي صحيح - غلط

۱- سئوالها را تا حد امکان مختصر ، ساده و روشن بنویسید

سئوال ضعیف : نهنگها جزو پستانداران هستند و به استثنای انسانها ، آنها با هوش ترین پستانداران هستند . ص غ
سئوال بهتر : نهنگها با هوش ترین پستانداران هستند . ص غ

۲- جمله‌هایی به کار ببرید که درست یا غلط بودن آنها برای کسانی که موضوع سئوال را یاد گرفته‌اند آشکار باشد .

سئوال ضعیف : اگر چه يك سئوال واحد يك اندازه ناپایاست و اگر چه نوشتن سئوالهاي اضافی دشوار است ، پایانی يك آزمون را می‌توان با افزایش طول آزمون و در نتیجه بالا بردن تعداد سئوالهاي که رفتار دانش آموزان را می‌سنجد افزایش داد. ص غ

سئوال بهتر : آزمونهایی که دارای سئوالهاي بیشتری هستند در مقایسه با آزمونهایی که سئوالهاي کمتری دارند پایاترند . ص غ



۳- از کاربرد کلماتی چون همه ، بعضی وقتها ، غالباً ، هرگز و از این قبیل پرهیزید .

۴- سعی کنید طول سئوالهای صحیح و سئوالهای غلط هم اندازه باشند .

۵- سعی کنید تعداد سئوالهای صحیح و سئوالهای غلط به يك اندازه اشند .

۶- سئوالهای صحیح – غلط را از روی جمله‌های کتاب ننویسید .

۷- سعی کنید هر سئوال به يك هدف آموزشی مهم مربوط باشد و از گنجاندن مطالب بی اهمیت در سئوال پرهیز کنید .

آزمونهای جور کردنی

➤ از تعدادی سؤال تشکیل می‌شوند که هر یک از آنها شامل دو ستون است. یک ستون معرف پرسشها و ستون دیگر نشان دهنده پاسخهاست. وظیفه آزمون شوندگان آن است که پاسخها را با پرسشهای مربوط جور کند.

مثال : در ستون سمت چپ اسامی مخترعین و در ستون سمت راست اختراعات آنها فهرست شده اند، با نوشتن حرف مربوطه در داخل () آنها را به هم مربوط کنید

الف- ماری کوری	(پ) پنی سیلین
ب- دی باکی	(ث) پرتو ایکس
پ- فلمینگ	(الف) رادیوم
ت- هاروی	(ج) واکسن فلج
ث- رونتگن	(ب) قلب مصنوعی
ج- سالک	(ت) گردش خون
چ- پاستور	



قواعد تهیه سئوالهاي جور کردني

۱- پرسشها و پاسخهاي متجانس يا همگون انتخاب کنید .

- | | |
|---------------|---|
| الف-آرمسترانگ | (پ) ۱-پني سيلين را كشف كرد |
| ب-كريستف كلمب | (ب) ۲-آمریکا را كشف كرد |
| پ-فلمینگ | (الف) ۳-اولين فضا نوري كه بر ماه قدم نهاد |
| ت-گلن | (ت) ۴-اولين كاشف كه دماغه اميدنيك را دور زد |
| ث-ماژلان | |

-
- | | |
|-----------------------|---|
| الف-الكساندر گراهامبل | ۱-مخترع تلفن |
| ب-كريستوفر كولومبوس | ۲-كاشف آمریکا |
| پ-جان گلن | ۳-اولين فضانورد در آمریکا كه زمين را دور زد |
| ت-ابراهام لينكزن | ۴-اولين رئيس جمهور آمریکا |
| ث-فرديناند مگان | |
| ج-جرج واشنگتن | |
| چ-الي ويتني | |

۲- طول فهرستهای پرسش و پاسخ کوتاه انتخاب کنید و مطالب کمتری را در هر يك از پاسخها قرار دهید .

مثال : در جاي خالي هر پرسش در ستون سمت راست خود مربوط به ضريب پاياني يكي از موارد ستون سمت چپ را بنويسيد . هر يك از حروف ستون سمت چپ را مي توان يك بار يا چند بار به كار بريد يا هر گز از آن استفاده نكنيد .

الف-ثبات	(ث) ۱-نمرات استانفورد بينه با نمرات آزمون وكسلر همبستگي دارند .
ب - تعادل	(ت) ۲-روشي كه غالباً براي تخمين پاياني آزمونهاي معلم ساخته به كار مي رود .
ج-ثبات و تعادل	(ت) ۳-معمولاً كمترين برآورد پاياني را به دست مي دهد .
ت-دونيمة كردن	(ت) ۴-نتايج تحت تاثير آزمونهاي سرعت قرار دارند .
ث-هيچ يك از موارد بالا	(الف) ۵-آزمون استعداد تحصيلي يك بار به دانش آموزان سال اول دبيرانستان و بار دوم به همان دانش آموزان در سال آخر دبيرانستان داده مي شود

۳- در راهنمایی سؤال اطلاعات لازم را درباره نحوه مقایسه و جور کردن پرسشها و پاسخها در اختیار دانش آموزان قرار دهید

۴- همه پرسشها و پاسخهای يك سؤال را در يك صفحه قرار دهید

۵- در صورت امکان فهرست پاسخها را به صورت منطقی مرتب کنید .

مثال : در جاهای خالی مقابل آثار ستون سمت راست ، حروف مربوط به نویسندگان آن آثار را که در ستون سمت چپ نوشته شده‌اند قرار دهید .

- | | |
|------------------|-----------------|
| ۱- شاهنامه | الف- بوعلی سینا |
| ۲ - بوستان | ب- حافظ |
| ۳ - کیمیای سعادت | پ- سعدی |
| ۴- شفا | ت- عراقی |
| ۵- چهار مقاله | ج- فردوسی |



۶- در فهرست پرسشهای سئوالهای جور کردنی از کاربرد جملات نیمه تمام خودداری کنید .

۷- پرسشها را با شماره و پاسخها را با حروف مشخص کنید

۸- هر يك از پاسخها بايد براي تمام پرسشهاي يك سؤال درست جلوه كند .

۹- از نوشتن سئوالهایی که پرسشها و پاسخهای آن کاملاً با هم



آزمون چند گزینه‌ای

(Stem)

۱- قسمت اصلی یا تنه سؤال

(گزینه کلید = Key)

۲- گزینه درست یا پاسخ سؤال

(Distracters)

۳- گزینه‌های انحرافی

انواع سئوالات چند گزینه‌ای :

۱- نوع تنها گزینه درست

- اگر $27 = R$ باشد R برابر است با :

الف- ۳

ب- ۹

ج- ۱۱

د- ۳۰

۲- نوع بهترین گزینه

- می‌توان گفت که یادگیری عبارت است از :

الف- کسب اطلاعات تازه در نتیجه آموزش

ب- ایجاد تغییرات نسبتاً پایدار توان رفتاری یادگیرنده

ج- ایجاد تغییرات مطلوب و مورد نظر در رفتار یادگیرنده

د- توانایی به خاطر سپردن و یاد آوری تجارب پیشین

۳- نوع منفي

- کدام يك از عبارتهای زیر درباره ویروس درست نیست .

الف- تنها در حیوانات و گیاهان زندگی می‌کنند

ب- تولید مثل می‌کنند

ج - از سلولهای زنده خیلی بزرگ تشکیل می‌شوند

د- می‌توانند سبب بیماری شوند

محاسن آزمونهاي چند گزينه‌اي



- ۱- از ساير آزمونهاي عيني انعطاف پذيرترند ، علاوه بر دانش، توانايي استدلال ، قصاوت...را مي‌سنجند .
- ۲- با استفاده از آنها مي‌توان در يك زمان محدود تعداد زيادي از هدفهاي آموزشي و محتواي درس را اندازه گرفت .
- ۳- نسبت به آزمونهاي صحيح – غلط كمتر امكان حدس زدن كوركورانه را به آزمون مي‌دهند .
- ۴- پاسخهاي آن به سادگي و با عينييت كامل قابل تصحيح هستند .
- ۵- دانش آموزان و معلمان اين نوع آزمونها را ترجيح مي‌دهند .
- ۶- اگر گزينه‌هاي انحرافي اين آزمونها با توجه به اشتباهات و كج فهمي‌هاي متداول دانش آموزان تهيه شوند ، منبع بسيار مناسبی براي تشخيص مشكلات دانش آموزان خواهد بود .

معایب آزمونهای چند گزینه‌ای

۱- ساختن این آزمونها بسیار دشوار است . (انتخاب گزینه‌های انحرافی خوب)

۲- معلمان اغلب این سئوالها را در حد سنجش اطلاعات جزئی و کم اهمیت می‌نویسند .

۳- در مقایسه با آزمونهای صحیح – غلط , خواندن این آزمونها و پیدا کردن گزینه درست , مستلزم صرف وقت بیشتری است .

۴- در صورت محاسبه نمره منفی , دانش آموزانی که خطر می‌کنند , از سایر دانش آموزان نمره بهتری می‌گیرند .

۵- دانش آموزان قوی بیش از دانش آموزان معمولی قادر به پیدا کردن اشکالات , پیچیدگیها و دیدگاههای واگرا در سئوالها هستند , اما از آنجا که تنها یک پاسخ صحیح در سئوال وجود دارد , بابت این تیزبینی‌های خود نه تنها تشویق نمی‌شوند بلکه ممکن است تنبیه هم بشوند .

قواعد تهیه سئوالهاي چند گزینه اي

۱- هر سؤال باید يك موضوع مهم يا يك هدف آموزشي را اندازه بگیرد .

۲- بیشتر از يك مسئله يا يك مطلب در هر سؤال قرار ندهید .

۳- سئوالها را کاملاً روشن و واضح بنویسید و از عبارات و کلمات قابل آزمون شنوندگان
کنید .

۴- از تکرار مطالب در گزینه ها خودداری کنید .

سؤال ضعیف : در آزمونهایی عینی اصطلاح عینی :

الف- به روش مشخص کردن نتایج یادگیری مربوط است

ب- به روش انتخاب محتوای آزمون مربوط است .

ج- به روش تصحیح جوابها مربوط است .

د- به روش ارائه مطالب مربوط است .

سؤال بهتر : در آزمونهایی عینی به روشمربوط است .

الف- مشخص کردن نتایج یادگیری

ب- انتخاب محتوای آزمون

ج- تصحیح جوابها

د- ارائه مطالب

استفاده





۵- مطالب اصلي سؤال را به طور كامل در تنه سؤال بنويسيد .

سؤال ضعيف : تصعيد:

- الف: به معني تبديل جامد به مايع است
- ب: به معني جامد به گاز است
- ج: نشان دهنده تبديل مايع به گاز است.
- د : نشان دهنده تبديل گاز به مايع است.

۶- همه گزينه هاي سؤال بايد متجانس و به موضوع واحد مربوط باشند.

- سؤال ضعيف: کدام يك از مطالب زير درست است.
- الف- تهران مركز ايران است.
 - ب: كيوي يك از انواع سبزيهاست.
 - ج: خواجه نظام الملک وزير قاجار به بود.
 - و: به تبديل گاز به مايع تبخير مي گويند.

۷- سؤال را طوري بنويسيد كه پاسخ درست تنها پاسخ درست يا قطعاً درست ترين پاسخ باشد.

سؤال ضعيف : يك پسر بچه كوررنگ، كوررنگي خود را از چه كسي به ارث مي برد؟

- الف: پدر
- ب: مادر
- ج: پدر و مادر
- د: پدر پدر

۸- برای اندازه‌گیری فرایندهای پیچیده ذهنی از موقعیتهای تازه استفاده کنید.

سوال ضعیف: کدام يك از موارد زیر وقتی كه در بدن مصرف می‌شود. بیشترین كالري را تولید می‌کند.
الف: يك گرم چربی ب: يك گرم شكر ج: يك گرم نشاسته د: يك گرم پروتئين
سؤال بهتر : کدام يك از موارد زیر اگر از غذاهای روزانه حذف شود. بیشترین مقدار کاهش كالري را سبب می‌شود؟
. الف: يك قاشق کره ب: يك قاشق شكر ج: يك قرص نان د: يك تخم‌مرغ آب پز

۹- گزینه‌های انحرافی را طوری بنویسید که توجه آزمون شوندگان بی‌اطلاع از موضوع سؤال را به خود جلب کنند.

سؤال ضعیف: - امیر کبیر صدر اعظم کدام يك از شاهان زیر بود؟
الف: شاه عباس ب: سلطان محمود ج: ناصرالدین شاه د: نادر شاه
سؤال بهتر: امیر کبیر صدر اعظم کدام يك از شاهان زیر بود؟
الف: مظفرالدین شاه ب: محمد علي شاه ج: ناصرالدین شاه د: احمد شاه

۱۰- گزینه‌هایی هر سؤال را طوری بنویسید که از نظر دستوری و جمله‌بندی به نحو درست مکمل متن سؤال باشد.

سؤال ضعیف: چرا زندگی شهری و روستایی در ناحیه کنار خزر به هم وابسته است ؟

الف: شهرها بازار فروش محصولات روستایی است.

ب: وابستگی کارخانه‌ها به محصولات کشاورزی

ج: وجود بازارهای هفتگی

د: همه موارد بالا

۱۱- از به کار بردن اشاره‌هایی دستوری نامربوط و غیره خودداری کنید.

سؤال ضعیف: اسناد مالکیت املاک مردم در کجا ثبت می‌شوند؟

الف: اداره ثبت اسناد و املاک

ب: اداره کل ثبت احوال

ج: وزارت امور اقتصادی و دارایی

د: وزارت فرهنگ و آموزش عالی

۱۲- در سؤالات منفي كلمات منفي را برجسته جلوه دهید.

۱۳- از نوشتن سؤالهایی که در آنها متن سؤال منفي و گزینه‌ها هم منفي هستند، يعني

سوال ضعیف :

در رابطه با استفاده از سوال های انشائی و ارزشیابی هدفهای آموزشی نمی توان گفت که ه سوال های انشائی

الف: برای ارزشیابی هدفهای طبقات ترکیب و ارزشیابی بهترین نوع نیستند.

ب: برای ارزشیابی هدفهای طبقات ترکیب و ارزشیابی بهترین نوع نیستند.

ج: و عینی تفاوت زیادی با یکدیگر ندارند.

د: برای هدفهای دانش همیشه صلاح نیست.

۱۴- تا حد امکان از کاربرد عبارتهایی مانند «همه آنچه در بالا گفته شد»، «تمامی موارد فوق» و غیره خودداری کنید.

۱۵- سؤاها را مستقل از یکدیگر بنویسید.

۱۶- از طرح سؤاها گمراه کننده پرهیزید.

۱۷ - از به کار بردن دو گزینه متضاد در یک سوال پرهیز کنید

سوال ضعیف : در نظریه شرطی سازی سوال پاولفی خاموشی عبارت است از

- الف - بی اثر شدن محرک شرطی ب - بی اثر شدن محرک غیر شرطی
ج - تغییر شکل دادن پاسخ د - خسته شدن جاندار در اثر پاسخ زیاد

سوال بهتر : در نظریه شرطی سازی پاولفی خاموشی عبارت است از

- الف - بی اثر شدن محرک شرطی ب - بی اثر شدن محرک غیر شرطی
ج - تغییر شکل دادن محرک طبیعی د - تغییر شکل دادن محرک غیر شرطی

۱۸ - طول گزینه درست را در سوالهای مختلف تغییر دهید .

سوال ضعیف : کدام یک از موارد زیر تعریف درست یادگیری به شمار می رود /

- الف - افزایش اطلاعات ب - کسب رفتار مطلوب
ج - اجتناب از رفتار نا مطلوب د - ایجاد تغییر نسبتا پایدار در رفتار بالقوه فراگیر

۱۹ - محل گزینه درست را در میان گزینه های انحرافی به طور تصادفی انتخاب کنید .



۲۰- برای هر سوال ۳ تا ۵ گزینه در نظر بگیرید

۲۱- عبارت هائی نظیر “ هیچ کدام از مواد بالا “ “ هیچ کدام از آنها “ و از این قبیل را زیاد به کار نبرید.

۲۲- در آزمونهای تکوینی از گزینه نمی دانم استفاده کنید

۲۳- در سوال هائی که تنه آنها جمله نا تمام است جای خالی را در قسمت آخر جمله قرار دهید.

سوال ضعیف : طبق نظر دیباکی عامل اصلی بیماری قلبی است .
الف - کلسترول ب - فربهی ج - سیگار کشیدن د - استرس

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

ترتیب قرار دادن سؤالهای مختلف یک آزمون به دنبال یکدیگر :

- ۱- سؤالهای صحیح - غلط
- ۲- سؤالهای جور کردنی
- ۳- سؤالهای چند گزینه ای
- ۴- سؤالهای کوتاه پاسخ
- ۵- سؤالهای تشریحی

علاوه بر گروه بندی سؤالهای بر حسب نوع ، لازم است سؤالهای مختلف هر نوع را نیز بر حسب موارد زیر دسته بندی کنید !

- ۱- سؤالهای مربوط به هر طبقه از هدفهای آموزشی مانند : دانش ، کاربرد ، و ... به دنبال هم قرار دهید .**
- ۲- سؤالها را از ساده به دشوار مرتب کنید .**
- ۳- در تنظیم سؤالهای یک آزمون سازمان اصلی مطالب را حفظ کنید .**

نمره گذاری

دو روش وجود دارد:

- ۱- محاسبه کلیه پاسخهای درست بدون کسر نمره برای حدس زدن
- ۲- کسر مقداری از نمره برای جبران حدس زدن

$$R - \frac{W}{N-1}$$

$R =$ تعداد پاسخهای درست

$W =$ تعداد پاسخهای غلط

$N =$ تعداد گزینه‌های هر سؤال

هدف از تحلیل سؤالهای آزمون , واریسی تکتک سؤالها و تعیین میزان دقت و نارساییهای آنهاست

P - محاسبه ضریب دشواری سؤال

D - محاسبه ضریب تمییز سؤال

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

$$D = \frac{\text{انتخاب های درست گروه پایین} - \text{انتخاب های درست گروه بالا}}{\text{تعداد افراد یک گروه (بالا یا پایین)}}$$

سوال: کدام یک از ارقام زیر معرف ضریب هبستگی بیشتری است؟

الف - ۵۵/۰ - ب - ۶۱/۰ - ج - ۴۹/۰ - د - ۲۳/۰

$$D = ۳/۰ \quad P = ۳۵$$

گروه ها	الف	ب	ج	د	بدون پاسخ	مجموع
۲۵/۰ بالا	♦	۵	۳	♦	۲	۱۰
۲۵/۰ پایین	۵	۲	۳	♦	♦	۱۰

تفسیر ضریب دشواری (Difficulty Index)

هر اندازه ضریب دشواری یک سؤال بزرگتر باشد، آن سؤال آسان تر است و هر اندازه این ضریب کوچکتر باشد، سؤال دشوارتر است.

هر اندازه ضریب دشواری یک سؤال بزرگتر (به ۱۰۰ نزدیکتر) باشد، آن سؤال آسان تر خواهد بود و هر اندازه که این ضریب کوچکتر (به صفر نزدیکتر) باشد سؤال دشوارتر است.

تفسیر ضریب تمیز (discriminative index)

هر قدر ضریب تمیز بزرگتر باشد، قوه تمیز آن سؤال بیشتر و هر قدر این ضریب کوچکتر باشد، قوه تمیز آن کمتر است. اگر قوه تمیز سوآلی ۹۰ / ۰ باشد، آن سوال آزمودنی های قوی و ضعیف را خیلی خوب از هم جدا کرده است.

گاه اتفاق می افتد که ضریب تمیز سوالی منفی است، این ضریب تمیز نشان می دهد که در آن سوال گروه قوی بدتر از گروه ضعیف عمل کرده است. دلیل چنین ضریب تمیزی می تواند این باشد که افراد گروه بالا یا به طور کامل یاد نگرفته اند یا آن را به غلط آموخته اند.

تحلیل گزینه های انحرافی

تفسیر ضریب دشواری

اگر در تفسیر ضریب دشواری یا ضریب سهولت سوال با آزمونهای وابسته به هنجار سر و کار داشته باشیم، می توانیم این را از دیدگاه آماری مورد بررسی قرار دهیم. یکی از انتظارات ما از آزمونهای وابسته به هنجار این است که برای آزمون شوندگان مختلف نمراتی به دست بدهد که در طول یک پیوستار پراکنده باشند. هر چه این پراکندگی بیشتر باشد، بهتر است.

هر چه واریانس نمرات حاصل از یک آزمون وابسته به هنجار بزرگتر باشد، آن آزمون، آزمون بهتری است

واریانس نمرات گروهی پاسخ دهندگان به یک آزمون از دو قسمت تشکیل می شود:

۱ - واریانس سوالها

۲ - همبستگیهای بین سوالها

هر سوالی که دارای واریانس بزرگتری است و با سایر سوالهای آزمون همبستگی بیشتری دارد به واریانس کل آزمون کمک بیشتری می کند

سوالی که با سایر سوالها همبستگی ندارد و دارای واریانس کوچکی است، کمک زیادی به واریانس کل آزمون نمی کند. اگر در تحلیل آزمون به انتخاب سوالهایی بپردازیم که همبستگی زیادی با یکدیگر دارند و در ضمن دارای واریانسهای بزرگتری هستند، در مجموع آزمون بهتری را تدارک دیده ایم.

محاسبه واریانس سوال :

$$\text{واریانس سوال} = P (1 - P)$$

P در فرمول بالا برابر است با درصد کسانی که به سوال جواب درست داده اند
یک سوال زمانی دارای حداقل واریانس است که ضریب دشواری آن ۰ یا ۱ باشد
یعنی هیچیک از آزمودنی ها به آن جواب درست نداده باشند یا اینکه همه به آن
پاسخ داده باشند.

$$\bullet (1 - 0) = 0$$

$$\bullet (1 - 1) = 0$$

اما زمانی که ضریب دشواری سوال معادل ۰/۵ باشد , واریانس سوال حداکثر
خواهد بود.

سوال هایی بهتر هستند که ضریب دشواری آنها از ۱ کمتر و از صفر بیشتر است و
به ۰/۵ نزدیک باشد.

برای آزمون های چند گزینه ای سطح بهینه دشواری اندکی کمتر از وسط فاصله ی بین و سطح موفقیت با حدس زدن است منظور از سطح موفقیت با حدس زدن عبارت است از نسبت پاسخهای درستی که با حدس زدن به دست می آید. برای سوالهای چهار گزینه ای سطح موفقیت با حدس زدن حدود ۲۵٪ است، لذا سطح بهینه ی دشواری در وسط ۲۵٪ و ۱ یعنی در حدود ۶٪ خواهد بود

$$۱/۰۰ - - ۰/۲۵ \quad ۰/۷۵$$

$$۰/۳۷۵ = \frac{\quad}{۲} = \frac{\quad}{۲} = \text{نقطه میانی}$$

ضریب دشواری بهینه = عملکرد با حدس زدن + نقطه میانی

$$۰/۶۲۵ = ۰/۲۵ + ۰/۳۷۵$$

راه ساده تر این است که عدد ۱ را با عملکرد مبتنی بر حدس زدن جمع کنیم و نتیجه را بر ۲ تقسیم کنیم

$$\frac{۱ + ۰/۲۵}{۲} = ۰/۶۲۵$$

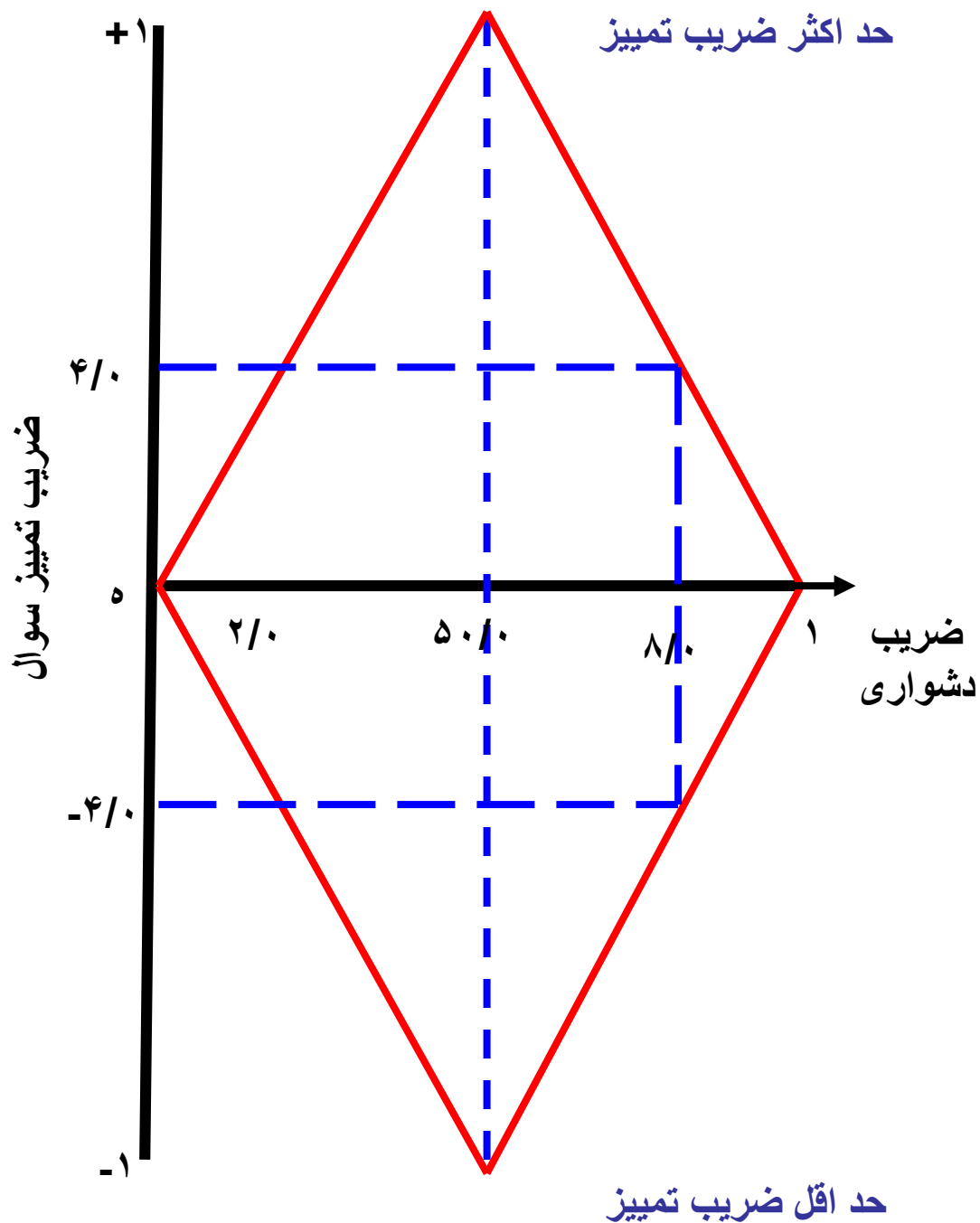
تعیین گروه بالا و گروه پایین :

متخصصان آزمون سازی (ویتنی و سیرز و نیکو) پیشنهاد کرده اند اگر تعداد کل برگه های آزمون بین ۲۰ تا ۴۰ نفر هستند ۱۰ برگه بالا و ۱۰ برگه پایین را انتخاب کنید . اگر تعداد کل برگه های شما ۲۰ عدد باشد ، همه آنها به دو دسته بالا و پایین تقسیم خواهند شد . اگر تعداد کل برگه ها کمتر از ۲۰ باشد به دو دسته بالا و پایین تقسیم کنید . در شرایطی که تعداد کل آزمون شوندگان بیش از ۴۰ نفر است ، بهترین رقم برای گروه بالا و پایین ۲۷٪ کل برگه هاست .

همبستگی سوال با کل آزمون

ضریب تمیز سوال نوعی ضریب توافق سوال با کل آزمون است
سوالهای دارای قوه تمیز خوب با کل آزمون همبسته اند، اما سوالهای
دارای قوه تمیز صفر یا منفی با کل آزمون همبسته نیستند یا با آن
همبستگی منفی دارند.

رابطه ضریب دشواری با ضریب تمییز



هٲنٲارها و نيمرخها

Norms and profiles

■ پس از تصحیح برگه های آزمون ، برای هر یک از دانش آموزان نمره ای به نام نمره خام به دست می آید .

■ نمرات خام به خود خودی قابل تغییر و استفاده نیستند و باید در رابطه با آنها اقداماتی انجام داد تا قابل استفاده گردند .

■ پاره ای از اقدامات عبارتند از : تعیین وضع تمرکز یا حالت پراکندگی نمرات ، یعنی محاسبه شاخصهای گرایش مرکزی و شاخصهای پراکندگی .

موضوع مورد بحث در این فصل هنجارها و نیمرخها است . نیمرخها و هنجارها روشهایی هستند که به ما کمک می کنند تا سطح موفقیت دانش آموزان را در درسهای مختلف با هم مقایسه کنیم و میزان توفیق آنان را نسبت به یکدیگر بسنجیم .

تعریف هنجار :

اگر برای تفسیر نمره یک دانش آموز نمره او را با نمره یک گروه مرجع مقایسه کنیم به آن گروه مرجع هنجار گویند . گروه مرجع از کسانی تشکیل می یابد که به گونه ای شبیه به دانش آموز مورد نظر هستند .

انواع نمرات هنجار

نمرات هنجار که به آنها نمرات مشتق هم گفته می شود دارای انواع مختلفی هستند که معروفترین آنها عبارتند از هنجارهای سنی یا معادل های سنی ، هنجارهای کلاسی یا معادل های کلاسی ، هنجارهای درصدی و نمرات معیار (استاندارد) .

هنجارهای سنی

هنجارهای سنی بر میانگین یا میانه نمراتی مبتنی هستند که دانش آموزان در سنین مختلف کسب می کنند و به صورت معادلهای سنی نشان داده می شوند . برای مثال اگر دانش آموزانی که ۱۰ سال و ده ماه سن دارند ، در یک آزمون نمره ۱۸ بگیرند ، به این نمره معادل سن ۲-۱۰ متعلق می گیرد .

بسیاری از آزمون های میزان شده پیشرفت تحصیلی دارای هنجارهای سنی هستند . در این آزمونها معادل های سنی با عنوان موضوعهای درسی نامگذاری می شوند . برای مثال سن خواندن ، سن حساب ، سن املاء و جز اینها . استفاده از هنجارهای سنی دارای محدودیت هایی است . اما در سطح دبستان بهترین مرحله کاربرد این نوع هنجارها است .

هنجارهای کلاسی :

هنجارهای کلاسی شبیه به هنجارهای سنی هستند ، اما در هنجارهای کلاسی از گروههای کلاسی به عنوان گروههای مرجع استفاده می کنند .
برای مثال : اگر دانش آموزان گروه نمونه ای که کلاس پنجم را شروع کرده اند ، در یک آزمون نمره خام ۱۹ بگیرند ، به این نمره خام معادل کلاسی ۵/۰ داده می شود .

هنجارهای سنی نمایی :

برای بدست آوردن شاخص درست تری از نمره متوسط دانش آموزان در یک کلاس معین ، نمرات دانش آموزان خیلی بزرگ یا خیلی کوچک را کنار می گذارند ، و نمره هنجار را تنها برای دانش آموزان سن مورد نظر برای آن کلاس محاسبه می کنند . این نوع هنجار کلاسی محدود شده را هنجار سنی نمایی می نامند .

هنجار سن ذهنی :

هنجار سن ذهنی که اولین بار بوسیله آلفرد بینه معرفی شد نوعی هنجار سنی است که در آزمونهای هوشی به کار می رود . زوش محاسبه سن ذهنی یک دانش آموز شبیه به محاسبه معادلهای کلاسی است . بنابراین ، اگر نمره خام یک دانش آموز در یک آزمون هوشی برابر با میانه نمرات خام کودکان ۹ ساله باشد . سن ذهنی او ۹ خواهد بود .

هوشبهر (بهره هوشی) IQ

هوشبهر (بهره هوشی) یک خارج قسمت است که از فرمول زیر به دست می آید .

$$IQ = \frac{A.M \text{ (سن عقلی)}}{C.A \text{ (سن تقدیمی)}} \times$$

هنجارهای درصدی

در هنجارهای درصدی ، با گروه سنی یا کلاسی خودش یعنی گروهی که منطق می توان او را عضوی از آن دانست مقایسه می شود . در هنجارهای درصدی از رتبه درصدی یا صدکها استفاده می شود .

انواع مختلف نمرات معیار :

معروف ترین آنها عبارتند از :

۱- نمرات Z

۲- نمرات T

۳- هوشبهر انحرافی

۴- نمرات ۹ بخشی

نمره Z

آن دسته از نمرات معیار که دارای میانگین صفر و انحراف معیار واحد (۱) هستند به نمرات Z شهرت دارند. برای تبدیل نمرات خام به نمرات Z از فرمول زیر استفاده می شود.

$$Z = \frac{X - \bar{X}}{SD}$$

در این فرمول :

نمره خام = X ، میانگین نمرات = $\bar{X}$ ، انحراف معیار = SD

مثال : اگر نمره خام دانش آموزی در یک آزمون که دارای میانگین $(X=56)$ و انحراف معیار $(SD=4)$ است ، ۵۸ باشد نمره Z او برابر خواهد بود با :

$$Z = \frac{58-56}{4} = 0.5$$

همچنین نمره Z دانش آموز دیگری که دارای نمره خام ۴۰ است برابر خواهد بود با :

$$Z = \frac{40-56}{4} = -4$$

نمره Z به ما نشان می دهند که یک نمره خام به اندازه چه تعداد از واحدهای انحراف معیار بالاتر یا پایین تر از میانگین است . بنابراین دانش آموزی که نمره خام او به اندازه یک انحراف معیار بالای میانگین باشد دارای نمره Z برابر با $+1$ و دانش آموزی که نمره خام او به اندازه دو انحراف معیار پایین تر از میانگین باشد دارای نمره Z برابر با -2 خواهد بود و الی آخر .

نمره T

نمرات معیار دیگری که به نمرات T شهرت دارند برای رفع مشکلات نمرات Z ابداع شده اند. مقیاس نمرات T با ضرب نمرات Z در ۱۰ و جمع کردن نمرات حاصل با ۵۰ به دست می آید. بنابراین می توان فرمول محاسبه نمرات T را به صورت زیر نوشت:

$$T = 10(Z) + 50$$

مثال : اگر نمرات دو دانش آموز در مقیاس Z برابر $5/0$ و $-5/1$ باشند نمرات آنها در مقیاس T به شرح زیر خواهد بود .

$$T = 10(5/0) + 50 \qquad T = 10(-5/1) + 50$$

$$= 5 + 50 = 55$$

$$= -15 + 50 = 35$$

از آنجا که همه نمرات T دارای میانگین ۵۰ و انحراف معیار ۱۰ هستند ، هر نمره T به تنهایی و به طور مستقیم قابل تغییر و استفاده است . برای مثال ، نمره $T = 55$ حاکی از این است که آن نمره به اندازه نصف انحراف معیار بالای میانگین است ، و نمره $T = 35$ نشان می دهد که آن نمره به اندازه یک و نیم انحراف معیار پایین میانگین قرار دارد و الی آخر .

نمرات ۹ بخشی

نوع دیگر هنجار به نام نمرات ۹ بخش وجود دارد که به صورت نمرات معیار یک رقمی نشان داده می شود. علت نامگذاری این مقیاس نمرات به ۹ بخشی این است که توزیع این نمرات به ۹ بخش محدود می شود. این مقیاس دارای میانگین ۵ و انحراف معیار ۲ است.

نیتکو نمرات ۹ بخشی را برای نمره گذاری آزمون های کلی و تشریحی که ارزشیابی آنها جنبه کیفی به شرح زیر پیشنهاد داده است

- ۹- بسیار عالی (۴٪ بالا)
- ۸- عالی (۷٪ بعدی)
- ۷- خیلی خوب (۱۲٪ بعدی)
- ۶- خوب (۱۷٪ بعدی)
- ۵- متوسط (۲۰٪ وسط)
- ۴- نسبتاً ضعیف (۱۷٪ بعدی)
- ۳- ضعیف (۱۲٪ بعدی)
- ۲- خیلی ضعیف (۷٪ بعدی)
- ۱- بیش از حد ضعیف (۴٪) آخر

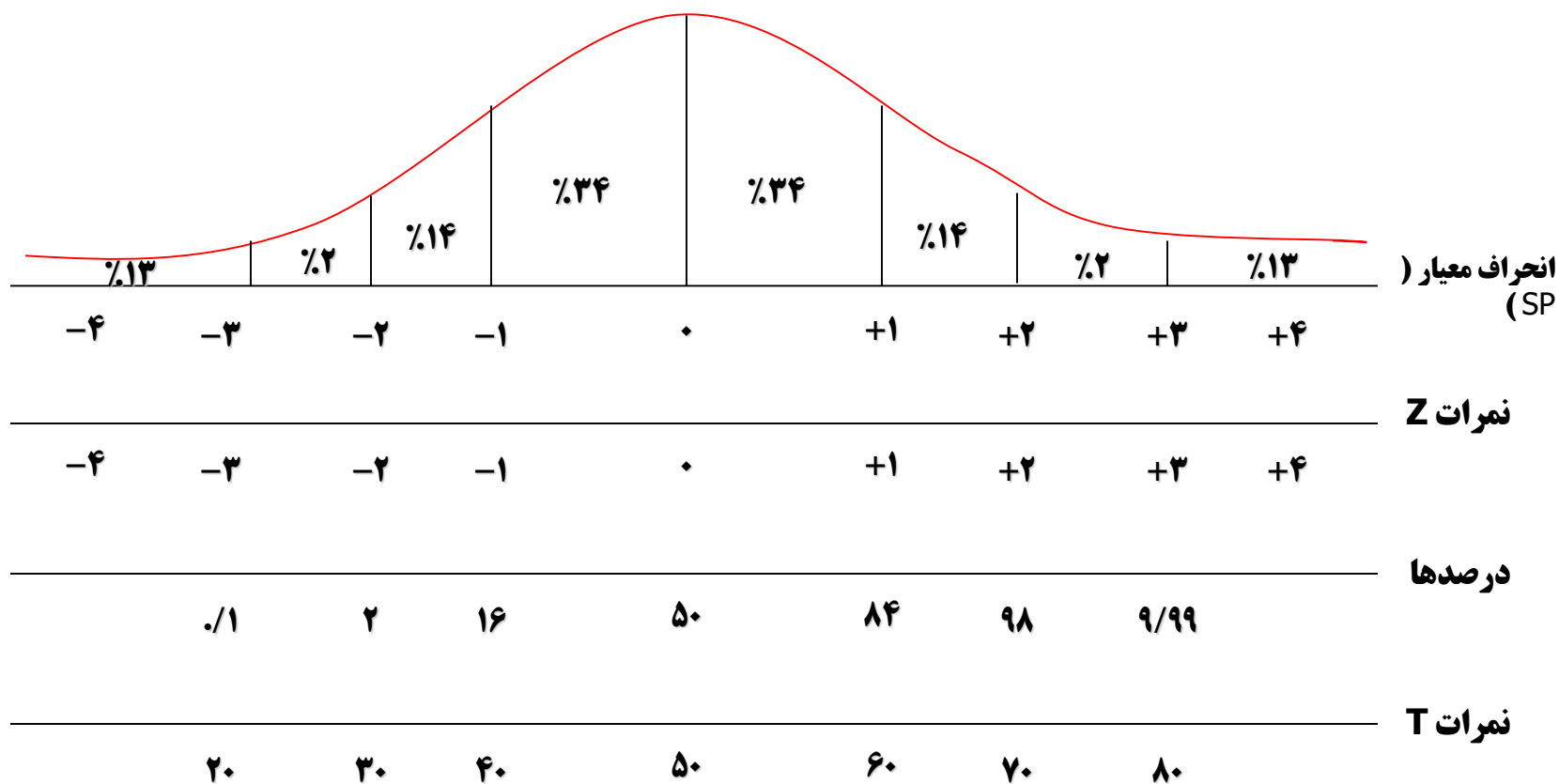
هوشبهر انحرافی

هوشبهر انحرافی یک هنجار مربوط به نمرات معیار است که دارای میانگین ۱۰۰ و انحراف معیار ۱۵ یا ۱۶ است. بنابراین، اگر انحراف معیار ۱۵ باشد معدل هوشبهر انحرافی به قرار زیر است:

$$D IQ = 15Z + 100$$

مقایسه هنجارهای مختلف با یکدیگر .

رابطه میان هنجارهای مختلف در شکل زیر نشان داده شده است . این شکل رابطه هنجارها را برای نمراتی که توزیع بهنجار دارند نشان می دهد . برای مثال ، نمره ای که یک انحراف معیار پایین میانگین است ($SQ = -1$) دارای -1 - $Z =$ ، ملاک 16 ، $T = 40$ ، هوشبهر انحرافی 85 و نمره 9 بخشی 3 است .



مقایسه هنجارها و نمره های معیار در توزیع بهنجار (از گران لاند، ۱۹۷۱)

توزیع بهنجار :

توزیع بهنجار نوعی توزیع است که بر منحنی بهنجار منطبق است .
منحنی بهنجار یک منحنی زنگوله شکل است که ویژگیهای زیر
را داراست :

۱- دو قسمت منحنی بهنجار و نسبت به خط وسط منحنی قرینه
هستند .

۲- در توزیعی که این منحنی معرف آن است اکثر نمره ها در وسط
توزیع انباشته شده اند .

۳- دو امتداد منحنی بهنجار در بینهایت محور افقی را قطع می کنند .

۴- در منحنی بهنجار ، میانگین ، میانه ، و نمابر بر روی هم قرار دارند.

۵- در روی این منحنی دو نقطه وجود دارند که در آنها جهت منحنی تغییر می کند . این دو نقطه در فواصل $1 \pm$ انحراف معیار از میانگین قرار دارند .

۶- در منحنی بهنجار ، بین انحراف معیارها و نسبت افراد یک رابطه دقیق ریاضی وجود دارد و در هر واحد انحراف معیار ، همیشه یک نسبت ثابت از افراد موجود است .

نیمرخ

نیمرخ وسیله ای است که به کمک آن می توان نمرات یک دانش آموز (یا گروهی دانش آموز) در دو یا چند درس را با هم مقایسه کرد . این مقایسه به کمک اعداد خام درست نخواهد بود . هدف از ترسیم نیمرخ این است که نتایج را به صورت تصویر یا نمودار نشان دهیم .

نیمرخ یعنی بیان تصویری یا نموداری نتایج ، نیمرخ
نقاط قوت و ضعف عملکرد یک فرد را نشان می
دهد ، درست همانطور که تصویر یک فرد وضع
ظاهری او را نشان می دهد ، نیمرخ نمرات وضع
تحصیلی و روانی او را نشان می دهد .

مثال :

برای مقایسه نمره یک دانش آموز در آزمون درس علوم که تنها به ۱۰ سؤال آن جواب داده است با نمره همین دانش آموز در آزمون درسی ریاضی که به ۱۶ سؤال آن پاسخ گفته است ، باید این دو نمره را به نوعی هنجار کلاسی ، سنی یا درصدی تبدیل کنیم تا مقایسه آنها امکان پذیر شود .

بنابراین ، هرگاه تعدادی نمره حاصل از آزمونهای مختلف با آزمونهای فرعی یک آزمون کلی را برای یک دانش آموز به یکی از هنجارها تبدیل کنیم . و این نمرات را در یک نمودار قرار دهیم نیمرخ عملکرد آن دانش آموز را بدست داده ایم .

شکل ۲-۹: کتاب در صفحه ۲۴۳ نمونه ای از نیمرخ
نمره های یک دانش آموز کلاس نهم را در
درسهای علوم اجتماعی ، انشاء و ... نشان می دهد .

باید توجه داشت :

تفاوت نمره های یک دانش آموز در آزمونهای مختلف ممکن است ناشی از خطای اندازه گیری آزمونها باشد نه یک تفاوت واقعی . بنابراین تفاوت میان نمرات آزمونها را با احتیاط تفسیر کنیم .

روش دیگر ترسیم نیمرخ نمرات که برای رفع مشکل خطای اندازه گیری پیشنهاد شده است این است که ، به عوض نشان دادن نمرات آزمون به صورت نقطه هایی در روی نیمرخ نمرات ، عملکرد دانش آموز به صورت نوارهایی ترسیم شود که هر نوار به اندازه یک خطای معیار اندازه گیری بالای نمره دانش آموز و یک خطای اندازه گیری پایین آن نمره را شامل شود .

دروایی آزمون

روایی آزمون

متخصصان اندازه گیری و ارزشیابی برای آزمونهای مختلف ویژگیهای زیادی را برشمرده اند ، از جمله می توان عینیت ، سهولت اجرا ، سهولت نمره گذاری ، عملی بودن ، سهولت تعبیر و تفسیر ، روایی و پایایی را نام برد . از میان این ویژگیها آنچه از بقیه مهمترند دو ویژگی روایی و پایایی هستند .

تعریف روایی

روایی اصطلاحی است که به هدفی که آزمون برای تحقق بخشیدن به آن درست شده است اشاره می کند به سخن دیگر ، آزمونی دارای روایی است که برای اندازه گیری آنچه مورد نظر است مناسب باشد . برای مثال ، یک آزمون پیشرفت ریاضی کلاس پنجم دبستان در صورتی یک آزمون روا است که محتوا و هدفهای آن درس را بخوبی اندازه گیری کند و بجز یادگیری آن درس چیز دیگری را شامل نباشد .

سه هدف عمده از آزمونهای مورد استفاده در آموزش و پرورش

۱- آیا سئوالهای آزمون از محتوای درس یا موضوع موردنظر یک نمونه خوب ارائه می دهند ؟

۲- آیا نمرات آزمون عملکرد فعلی یا آتی دانش آموزان را پیش بینی می کنند ؟

۳- آیا نمرات با مفاهیم نظری یا سازه های که آزمون برای سنجش آن درست شده است مربوط می شوند ؟

سه مورد بالا به ترتیب سه نوع روایی محتوا ، پیش بینی و سازه را نشان می دهند .

روایی محتوایی

روایی محتوایی به این مطلب اشاره می کند که نمونه سئوالهای مورد استفاده در یک آزمون تا چه حد معرف کل جامعه سئوالهای ممکن است که می توان از محتوا یا موضوع موردنظر تهیه کرد . هرچه آزمون از این لحاظ بهتر باشد ، دارای روایی بیشتر است . بنابراین ، اگر معلم فیزیک بخواهد برای درس خود یک آزمون پیشرفت تحصیلی روا بسازد ، آزمون او باید در برگیرنده نمونه ای درست و دقیق از مطالب درس فیزیک باشد .

روایی صوری

روایی صوری یکی از مشتقات روایی محتوایی است .
روایی صوری به این مطلب اشاره می کند ، که سئوالهای
آزمون تا چه حد در ظاهر شبیه به موضوعی هستند که
برای اندازه گیری آنها تهیه شده اند .

روایی پیش بینی

روایی پیش بینی ویژه آزمونهای است که برای پیش بینی موفقیت افراد در امور تحصیلی یا شغلی بکار می روند . آزمونهایی که برای مقاصد پیش بینی بکار می روند مختلف هستند . گاه از آزمونهای پیشرفت تحصیلی فعلی یا قبلی برای پیش بینی پیشرفت تحصیلی استفاده می شود . مثلاً استفاده از آزمونهای ورودی دانشگاهها برای پیش بینی پیشرفت تحصیلی داوطلبان در دوره تحصیلی در دانشگاه .

روایی ملاکی

به فن تجربی مطالعه رابطه بین نمرات آزمون (پیش بینی کننده) و نوعی اندازه بیرونی مستقل (ملاک) مربوط می شود. روایی ملاکی را می توان با توجه به درجه ای که عملکرد آزمون شونده در یک آزمون (آزمون پیش بینی) به عملکرد او در آزمون دیگر (آزمون ملاک) ارتباط دارد تعریف کرد.

نحوه تعیین روایی پیش بینی . یک آزمون

برای تعیین روایی پیش بینی یک آزمون ، دو آزمون ضروری هستند . ضریب همبستگی بین نمرات حاصل از این دو آزمون شاخص روایی پیش بینی آزمون مورد نظر است ، هرچه این ضریب بزرگتر باشد ، آزمون دارای روایی بیشتری خواهد بود و هرچه این ضریب کوچکتر باشد آزمون مورد نظر روایی کمتری خواهد داشت .

روایی همزمان

روایی همزمان ، نوعی روایی پیش بینی است که در آن به جای تعیین رابطه بین دو آزمون در یک فاصله زمانی ، رابطه دو آزمون به طور همزمان تعیین می شود . هدف از تعیین روایی همزمان بین دو آزمون این است که معلوم سازیم آیا می توان یک آزمون را به جای آزمون دیگری استفاده قرار داد یا نه ؛ در تعیین روایی همزمان ، هنگام اندازه گیری متغیر پیش بینی ، علاوه بر این متغیر ، متغیر ملاک نیز موجود است .

متخصصان روانسنجی دو نوع روایی پیش بینی و همزمان را با اصطلاح روایی ملاکی معرفی می کنند و بین این دو نوع روایی تمایز چندانی قائل نمی شوند . اما با وجود شباهت زیاد بین این دو نوع روایی ، می توان بین آن دو نوعی تمایز منطقی قائل شد .

تفاوت روایی همزمان و روایی پیش بینی

روایی پیش بینی ، نمرات آزمون واقعاً برای انجام نوعی پیش بینی بکار می روند ، اما چون در روایی همزمان داده های پیش بینی کننده و پیش بینی شوند همزمان جمع اوری می شوند ، هدف آزمون پیش بینی ، پیش بینی ملاک نیست ، بلکه هدف آن جانشین ساختن آزمون پیش بینی با آزمون ملاک است .

ویژگیهای مطلوب ملاک

شوند ایک وهیگن برای داده های ملاکی چهار ویژگی مطلوب ذکر کرده اند :

این ویژگیها به ترتیب اهمیت عبارتند از :

۱- ربط داشتن

۲- بی طرفی

۳- پایایی

۴- در دسترس بودن

رابط داشتن

رابط داشتن داده های یک آزمون ملاکی ، با توجه به قابلیت انطباق آن با موفقیت واقعی شخص مورد قضاوت قرار می گیرد . برای اینکار هیچ گونه شواهد تجربی وجود ندارد و باید بر اساس ملاک منطقی تصمیم گرفت .

بی طرفی

منظور از بی طرفی این است که آزمون باید به گونه ای باشد که همه افراد در کسب نمره خوب فرصت مساوی داشته باشند . مثالهای مربوط به عدم بی طرفی ، تفاوت بین تجهیزات و شرایط کار برای کارگران یک کارخانه یا تفاوت بین کیفیت آموزش داده شده به دانش آموزان مختلف است .

پایایی

یک آزمون یا یک اندازه پایایی مربوط به موفقیت شناسی باید با ثبات و قابل تکرار باشد . اگر این نمره از روزی به روز دیگر تغییر کند ، به طوری که کارگر از روزی به روز دیگر یا از هفته ای به هفته دیگر نمرات متفاوتی بگیرد . در این صورت هیچ آزمونی وجود نخواهد داشت که نمره ملاک را پیش بینی کند .

در دسترس بودن

در دسترس بودن به این معنی است که دستیابی به اندازه ملاک عملی و میسر باشد. این ویژگی را با توجه به زمان ، هزینه و امکانات مورد نیاز مورد قضاوت قرار می دهند . بنابراین ملاکی که به سهولت بیشتر در چهار چوب امکانات سهل تر قابل دسترس باشد ملاک بهتری است .

روایی سازه

یک آزمون در صورتی دارای روایی سازه که نمرات حاصل از اجرای شان به مفاهیم سازه های نظریه مورد نظر مربوط باشند . برای مثال ، یک آزمون یا پرسشنامه اضطراب در صورتی دارای روایی سازه است که نمرات حاصل از آن به سازه هایی که در نظریه های اضطراب آمده اند ارتباط داشته باشد .

تعریف سازه

معنی سازه به معنی مفهوم بسیار نزدیک است . کرلینجر سازه را اینگونه تعریف کرده است :

در سازه یک مفهوم است ، اما سازه یک معنی اضافی بر مفهوم دارد و آن این است که سازه برای مقاصد ویژه ی علمی بطور عمد و از روی آگاهی ابداع می شود . مثالهای سازه : هوش - انگیزش - اضطراب - برون گرایی - درون گرایی

بار علمی + مفهوم = سازه

روشهای تعیین روایی سازه

- ۱- روش تعیین همبستگی آزمون با سایر آزمونهای روا
- ۲- روش تمایز سنی
- ۳- روش تحلیل عوامل
- ۴- روش همسانی درونی

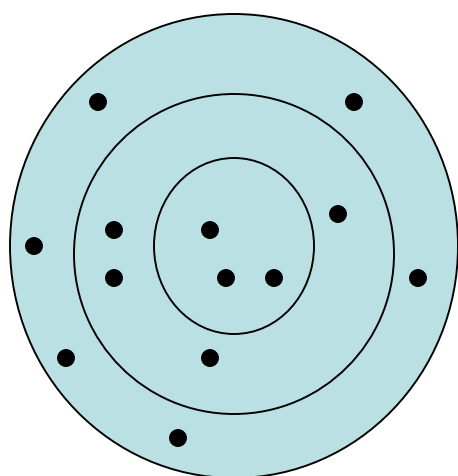
پایایی آزمون

تعریف پایایی

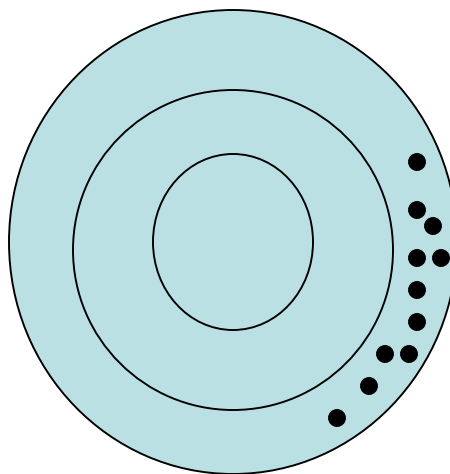
پایایی یک وسیله اندازه گیری به دقت آن اشاره می کند. یک آزمون در صورتی دارای پایایی است که اگر آن را در یک فاصله زمانی کوتاه چندین بار به گروه واحدی از افراد بدهیم نمرات حاصل از این چندین بار اجرا نزدیک به هم باشند.

رابطه بین پایایی و روایی

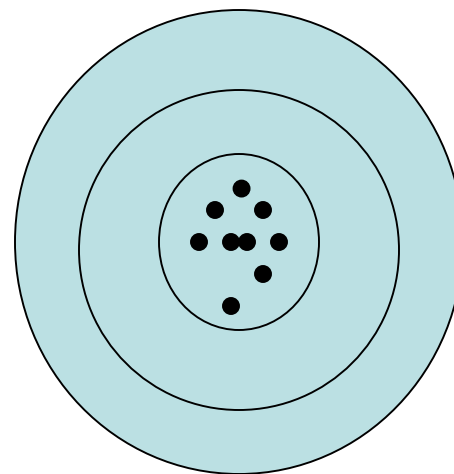
رابطه بین روایی و پایایی از این قرار است که یک آزمون باید پایا باشد تا بتواند روا باشد. اگر آزمونی در هر بار اجرا در مورد تعدادی دانش آموز نتایج مختلفی بدست بدهد آن آزمون یک آزمون پایا نخواهد بود. پس برای اینکه یک آزمون روا باشد باید پایا باشد، یعنی پایایی شرط روایی است، اما روایی بر پایایی ضروری نیست.



ج



ب



الف

شکل ۱۰-۱ رابطه بین روایی و پایایی (نتایج تیر اندازی با سه تفنگ مختلف)

برای مقایسه پایایی و روایی در شکل ۱۰-۱ نتایج تیراندازی با سه تفنگ مختلف نشان داده است. چنانکه مشاهده می نمایند نتایج تیراندازی با تفنگ الف هم پایاست و هم روا، زیرا همه تیرها به هدف خورده اند. نتایج تیراندازی با تفنگ ب پایا است اما روا نیست، زیرا تیرها همه تقریباً به یک نقطه اثابت کرده اند اما آن نقطه مورد نظر تیر انداز نیست. نتایج تیراندازی با تفنگ پ نه پایا است و نه روا.

روشهای تعیین پایایی

ایبل پنج روش تعیین پایایی را معرفی کرده است . این روشها عبارتند از :

- ۱- روش پایایی مصحح
- ۲- روش باز آزمایی
- ۳- روش فرمهای موازی یا هم ارز
- ۴- روش دو نیم کردن آزمون
- ۵- روش کودر - ریچاردسون

روش پایایی مصحح

برای تعیین پایایی آزمونهای تشریحی یا انشایی که نمرات آنها تحت تأثیر قضاوت محصلان برگه های آزمون قرار می گیرد باید از دو یا چند مصحح که مستقلاً برگه های امتحان را تصحیح می کنند استفاده کرد . همبستگی بین نمرات این مصححان شاخص پایایی به حساب می آید .

روش باز آزمایی

ساده ترین روش تعیین پایایی یک آزمون باز آزمایی است در این روش ، آزمون را در دو نوبت به گروه واحدی از آزمون شوندگان می دهند و نمرات حاصل را با هم مقاسه می کنند . ضریب همبستگی بین نمرات حاصل از دوبار اجرای آزمون ضریب پایایی آزمون است . به این نوع پایایی ، پایایی باز آزمایی می گویند .

روش فرمهای هم ارز

روش دیگری که بای رفع مشکلات روش باز آزمایی به کار می رود . روش فرمهای هم ارز است . در این روش دو آزمون معادل یا همارز برای یک مطلب یا موضوع تهیه می کنند و آنها را در فاصله زمانی کوتاهی به یک گروه واحد از آزمون شوندگان می دهند . ضریب همبستگی بین نمرات این دو فرم آزمون ، ضریب پایانی آن آزمون به حساب می آید .

روش دو نیمه کردن

در روش دو نیمه کردن آزمون یک آزمون واحد یک بار به یک گروه از آزمون شوندگان داده می شود و پس از اجرا ، آن را به دو نیمه مساوی تقسیم می کنند . در این روش ، بهترین راه دو نیم کردن آزمون ، این است که همه سئوالهای فرد را یک آزمون به حساب آوریم ، و سئوالهای زوج را نیز آزمون دیگری بدانیم ضریب همبستگی حاصل از نمرات دو نیمه آزمون ضریب پایایی هر یک از دو نیمه خواهد بود .

نحوه محاسبه ضریب پایایی به روش دو نیمه کردن

برای محاسبه ضریب پایایی کل آزمون ، ضریب
همبستگی بین نیمه ها را در فرمول زیر که به فرمول
الپیرمن براون شهرت دارد قرار می دهند .

ضریب پایایی آزمون r_{tt}

$$R_{tt} = \frac{2r^{1/2} \cdot 1/2}{1+r^{1/2} \cdot 1/2}$$

ضریب همبستگی بین نیمه های آزمون = $\frac{R_1}{2} \cdot \frac{1}{2}$

مثال :

اگر ضریب همبستگی بین نیمه های یک آزمون برابر ۵۰٪ باشد ، ضریب پایایی کل آزمون با استفاده از فرمول الپیرمن – براون به نحو زیر محاسبه خواهد شد .

$$r_{tt} = \frac{2 \times 5/10}{1 + 5/10} = \frac{1}{5/10} = 60/100$$

روش کودر ریچاردسون

در روش کودر ریچاردسون نیز آزمون تنها یک بار اجرا می شود . در این روش همسانی درونی کل آزمون بررسی میشود و برای این منظور همه ماده های آزمون تحلیل می شوند برای بررسی همسانی درونی آزمون و پایایی آن دو فرمول معروف به کار برده اند که $kr20$ و $kr21$ مشهورند . ما در اینجا این دو فرمول و نحوه استفاده از آنها را به ترتب توضیح می دهیم .

فرمول kr20 :

$$R_{tt} = \left(\frac{K}{K-1} \right) \left(\frac{\Sigma t^2 - \Sigma p_i q_i}{\sigma_t^2} \right)$$

r_{tt} = ضریب پایانی آزمون

k = تعداد سئوالهای آزمون

σ_t = واریانس یا مربع انحراف معیار آزمون

P_i = نسبت دانش آموزانی که به سئوال پاسخ داده اند

$g_i = 1 - p$ = نسبت دانش آموزانی که به سئوال پاسخ غلط داده اند

$\Sigma p_i g_i$ = مجموع $p \times g$ ها برای تمام سئوال

..

مثال :

$$P_{tt} = \left(\frac{8}{8-1} \right) \left(\frac{25/2 - 63/1}{25/2} \right)$$

$$= 1/14 \frac{25/2 - 63/1}{25/2}$$

$$= 1/14 \frac{62/1}{25/2}$$

$$= 1/14 \times 0/27$$

$$= 0/31$$

فرمول کودر ریچاردسون ۲۱ برابر محاسبه پایایی

$$R_{tt} = \frac{K}{K-1} \frac{\sigma_t^2 - kpq}{\sigma_t^2}$$

مثال

$$r_{tt} = \left(\frac{8}{7} \right) \left(\frac{25/2 - 8 \times 56/10 \times 44/10}{25/2} \right)$$

$$= 1/14 \frac{25/2 - 97/1}{25/2}$$

$$= 1/14 \frac{28/10}{25/2}$$

$$= 0/14$$

$$K = 8$$

$$\sigma^2 = 2/25$$

$$\bar{P} = 0/56$$

$$\bar{q} = 0/44$$

عوامل مؤثر بر پایایی آزمون

- ۱- با افزودن تعداد سؤالات یک آزمون ، پایایی آن نیز افزایش می یابد
- ۲- با متناسب تر و همگون تر کردن سؤالات یک آزمون ، پایایی آزمون افزایش می یابد .
- ۳- افزودن تعدادی سؤال دارای ضریب دشواری متوسط ، پایایی آزمون افزایش می یابد .
- ۴- آزمونهای سرعت معمولاً از آزمونهای دارای سرعت کمتر ضریب پایایی بیشتری نشان می دهند

خطای معیار اندازه گیری

یکی از فواید ضریب پایایی آزمون این است که با استفاده از آن می توان خطای معیار اندازه گیری آزمون را تعیین کرد . خطای معیار اندازه گیری عبارت است از انحراف معیار توزیع نمرات خطا .

فرمول محاسبه خطای معیار اندازه گیری

$$S_m = s_{dt} \sqrt{1 - r_{tt}}$$

در این - فرمول :

s_m = خطای معیار اندازه گیری بر حسب نمرات خام

s_{dt} = انحراف معیار آزمون

r_{tt} = ضریب پایانی آزمون

مثال

فرض کنید ضریب پایانی یک آزمون که با روش دو نیم کردن آزمون محاسبه شده ۹۱/۰ و انحراف معیار آن ۱۰ باشد. طبق فرمول بالا، مقدار خطای معیار اندازه گیری این آزمون برابر است با:

$$\begin{aligned} S_m &= 10 \sqrt{1-0/91} \\ &= 10 \sqrt{0/09} \\ &= 10 \times 0/13 \\ &= 3 \end{aligned}$$

مثال دیگر از خطای معیار اندازه گیری

یک مثال واقعی از محاسبه خطای معیار اندازه گیری آزمونهای گروهی استعداد که در مورد کودکان سالهای آخر دبستان اجرا می شوند به دست داده اند. برای این آزمونها ضریب پایانی ۸۵/۰ و انحراف معیار ۱۵ گزارش شده است. با استفاده از فرمول خطای معیار اندازه گیری، خطای معیار این آزمونها برابر است با:

$$\begin{aligned} S_m &= 15 \sqrt{1-0/85} \\ &= 15 \sqrt{0/15} \\ &= 15 \times 0/387 \\ &= 5/8 \end{aligned}$$

نتیجه گیری :

بر اساس خطای $۸/۵$ می توان گفت که احتمال ۱ در ۲ وجود دارد که نمره به دست آمده برای دانش آموز از نمره واقعی او $۸/۵$ (یک خطای معیار اندازه گیری فاصله داشته باشد . یا احتمال ۱ در ۱۹ وجود دارد که نمره به دست آمده از نمره واقعی فرد $۶/۱۱$ (دو خطای معیار اندازه گیری) فرق کند .

موفق باشید