

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ

سبب و ارزشیابی در آموزش

استادنامداری

جلسه اول

تعریف یادگیری: ایجاد تغییرات نسبتاً پایدار در رفتار و تجربه ی فرد است و این رفتار و تجربه به دانش تعریف شده به عبارت دیگر در یادگیری یا میزان اطلاعات ما افزایش پیدا می کند یا مهارت ایجاد می شود و یا باعث نگرش مثبت یا منفی نسبت به چیزی میشود.

۱ - ۱ - برخی مفاهیم مربوط به ارزشیابی

اندازه گیری: Measurement

آزمون: Tes

آزمودن: Testing | سنجش: Assessment (گاهی ارزیابی نیز ترجمه شده)

ارزشیابی (ارزیابی): Evaluation

اندازه گیری: Measurement

در اندازه گیری، ویژگیها یا صفات اشیاء و افراد تعیین می گردند و مقدار آن ویژگیها یا صفات به صورت عدد یا رقم گزارش می شود. بنابراین می توان گفت:

اندازه گیری (تعریف گئی) عبارت است از فرایندی که تعیین می کند یک شخص یا یک شیء، چه مقدار از یک ویژگی برخوردار است.

اندازه گیری آموزشی (تعریف نیتکو): شیوه نسبت دادن یک عدد (که غالباً نمره نام می گیرد) 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 برخوردار است، نشان دهد.

مثال غیر اندازه گیری: حسن نسبت به سن و قدش لاغر است.

اندازه گیری: حسن ۱۶ سال سن، ۱۷۰ سانتی متر قد، ولی تنها ۴۰ کیلوگرم وزن دارد.

مثال غیر اندازه گیری: زهرا از پروین باهوش تر است.

اندازه گیری: هوشبهر پروین، ۱۰۲ و هوشبهر زهرا، ۱۱۱ است.

این اعداد حاصل اندازه گیری هایی است که از ویژگیهای جسمی و روانی آنان به دست آمده است

تأکید بر قواعد اندازه گیری (نانالی): اندازه گیری دربرگیرنده قواعدی است برای نسبت دادن اعداد به

اشیاء یا افراد به گونه ای که صفت‌هایی از آنها را به صورت کمیت نشان دهد.

قواعد اندازه گیری: خط مشی هایی برای نشان دادن مقدار (یا نوعی ویژگی دیگر) شیء مورد اندازه گیری (کوهن و سوردللیک)

بنابراین، روشهای نسبت دادن اعداد به ویژگیها باید به طور روشن بیان شود.

مثال:

در نمره گذاری نتایج یک آزمون حساب، جواب های درست دانش آموزان شمارش می شوند و به تعداد جوابهای درست به آنها نمره داده می شود. (قاعده اندازه گیری)

در بعضی موارد قواعد اندازه گیری به توضیح مفصل نیاز ندارد (مانند استفاده از متر برای اندازه گیری طول) چون فیزیکی و قابل مشاهده هستند؛ اما قواعد مربوط به اندازه گیری ویژگیهای روانی و تربیتی افراد، مانند هوش یا یادگیری دانش آموزان، به بیان دقیق نیاز دارند. چون پنهان و غیر قابل مشاهده هستند

فایده بیان قواعد اندازه گیری این است که کمک می کند تا افراد مختلفی که می خواهند وسیله اندازه گیری را مورد استفاده قرار دهند، آن را یکسان به کار بندند.

همچنین بیان قواعد اندازه گیری از جنبه‌های مهم استاندارد کردن یا میزان کردن آزمونهای روانی است در غیر اینصورت نتایج متعدد شده و وحدت رویه از بین می رود.

آزمون: Test

نکته: آزمون یکی از وسیله های اندازه گیری است و تنها ترین وسیله نیست

وسيله اندازه گیری: اندازه گیری، یک فرایند است و این فرایند نیاز به وسیله دارد. برای اندازه گیری ویژگیهای مختلف اشیاء و افراد از وسایل مختلفی استفاده میشود .

ویژگیها یا صفات روانی (مانند هوش، خلاقیت، انگیزش، نگرش، یادگیری و مانند اینها را باید بطور غیرمستقیم اندازه گیری کرد و همچنین، رفتارهای معرف ویژگیهای روانی بسیار متنوع اند و بنابر این نمی توان تمام آنها را اندازه گیری کرد.

معمول ترین وسیله اندازه گیری ویژگیها یا صفات روانی **آزمون** است:

آزمون، وسیله یا روش نظامدار (سیستماتیک برای اندازه گیری نمونه ای از رفتار است (گرانلاند و لین) و معمولاً شامل مجموعه ای از سؤالات است.

هرچقدر نمونه رفتارهای انتخاب شده برای آزمون، معرف ویژگی روانی مورد نظر باشد، به همان میزان نیز نتایج اندازه گیری دقیق و قابل اعتماد هستند

تعریف فوق از آزمون هم شامل آزمون یا همان تست های کلاسی می شود و هم شامل پرسشنامه که تبیین خواهد شد؛ میشود.

پرسشنامه، مجموعه ای از سؤالات است که هدف آن جمع آوری اطلاعات درباره ویژگی های افراد است.

آزمون و پرسشنامه هر دو مجموعه ای از سؤالات یا گویه ها را در اختیار فرد قرار میدهند و هر دو دربرگیرنده **نمونه** رفتارهای معرف یک ویژگی هستند و غالباً (نه همیشه به صورت کتبی اجرا می شوند. به علاوه هر دو، ابزار سنجش کمی هستند و نتیجه آنها به وسیله عدد و رقم (نمره) گزارش می شود.

تفاوت آزمون و پرسشنامه از لحاظ موضوع مورد اندازه گیری:

به کمک آزمون، توانایی ها اندازه گیری می شوند، اما پرسشنامه، ویژگیهای غیر توانایی را می سنجد. یعنی نگرش ها و علائق / مثلاً با آزمون توانایی علمی و سطح دانسته های یک دانش آموز از درس زبان را متوجه میشویم ولی با پرسشنامه نوع نگاه، علاقه مندی و سایر مسائل فهمیده میشود.

آزمون ابزار عملکرد پیشینه است، یعنی حداکثر توانایی فرد را در زمینه مورد سنجش نشان می دهد، ولی پرسشنامه **ابزار عملکرد نوعی** است، وضع معمول و طبیعی فرد را توضیح می دهد.

علیرغم وجود تفاوتها، آزمون و پرسشنامه در تعریف مشترک اند.

آزمودن: Testing

وقتی که برای اندازه گیری یکی از ویژگیهای روانی یا تربیتی یک فرد یا گروهی از افراد، از **آزمون** استفاده می شود، به این عمل یا فعالیت آزمودن می گوئیم. به طور خلاصه به عمل اجرای تست، تستینگ یا آزمودن گفته میشود.

آزمودن روانی (روان آزمایی): اندازه گیری متغیرهای روانشناختی به وسیله تدابیر یا روشهایی که به منظور کسب نمونه هایی از رفتار طراحی شده اند.

نکته: اندازه گیری از آزمودن گسترده تر است چون اندازه گیری با روش های مختلفی انجام می شود که یکی از این روشها آزمودن است.

سنجش: Assessment

سنجش مفهومی است گسترده تر از مفاهیم گذشته و اصطلاحی است کلی تر از اندازه گیری و آزمودن، و برای فرایندی به کار می رود که برای گردآوری اطلاعات معنادار و مورد نیاز **تصمیم گیری** (در آزمون هدف به دست آوردن عدد بود) درباره دانش آموزان، برنامه های درسی، سیاست های آموزشی و ... مورد استفاده قرار می گیرد.

سنجش یک تحلیل جامع و چندوجهی از عملکرد است و ممکن است از فنون مختلفی از جمله مشاهده عملکرد استفاده کند و الزاما با کمیت سر و کار ندارد.

اندازه گیری روانی و تربیتی صرفا از طریق آزمون یا پرسشنامه انجام می شود، اما سنجش الزاما به کمیت منحصر نمی شود و در آن وسایل و فنون مختلف (رسمی و غیررسمی برای گردآوری اطلاعات به کار می رود؛ از جمله: آزمون، پرسشنامه، مقیاس درجه بندی، فهرست واری، کار آزمایشگاهی، پروژه تحقیقی، امتحان شفاهی، تکلیف درسی، مصاحبه و مشاهده رفتار و عملکرد دانش آموزان در موقعیت های مختلف و همچنین نتایج می تواند به صورت غیر کمی (توصیفی) نیز گزارش شود.

به عنوان مثال در سنجش انگیزش دانش آموزان نسبت به یادگیری موضوع درسی، مشاهدات غیررسمی و پرسشهای مستقیم در گفتگوهای خصوصی نیز اطلاعات مفیدی دست می دهند.

سنجش کلاسی: تمامی فرایندهای مورد نیاز برای تصمیم گیری درباره یادگیری دانش آموزان.

اصطلاح سنجش، در حوزه روانشناسی بالینی و آموزش و پرورش بیشتر به کار می رود، زیرا در این حوزهها، برای تعیین رفتارها و ویژگیهای روانی و تربیتی افراد، علاوه بر اندازه گیری، از وسایل و روشهای دیگر نیز استفاده خواهد شد.

ارزشیابی: Evaluation

اصطلاح ارزشیابی (یا ارزیابی) به طور ساده به تعیین ارزش برای هر چیزی یا ارائه داوری ارزشی و تعیین کیفیت گفته می شود. (قضاوت کردن. قبول یا غیرقبول دانستن)

بنابراین:

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 میزانی (گی).

یکی از ویژگیهای مهم ارزشیابی، تعیین کیفیت است، زیرا داوری | ارزشی با توجه به کیفیت به عمل می آید. کیفیت در ارزشیابی یادگیری، دانش، مهارتها و توانایی هایی است که از دانش آموزان انتظار می رود (نیتکو)

نکته: ارزشیابی از سنجش نیز گسترده تر است

مقایسه مفاهیم:

ارزشیابی شامل داوری ارزشی درباره مطلوب بودن یا مطلوب نبودن ویژگی یا موضوع مورد ارزشیابی است، اما در اندازه گیری هیچ گونه داوری یا قضاوتی درباره نتایج حاصل انجام نمی گیرد. با وجود این ، اندازه گیری پایه و اساس ارزشیابی درست را می سازد زیرا بدون در دست داشتن نتایج یک اندازه گیری دقیق از ویژگی یا محصول مورد نظر نمی توان ارزشیابی درستی از آن به عمل آورد .

ارزشیابی آموزشی و پژوهش آموزشی

هر دو جزء فعالیتهای علمی هستند، زیرا هر دو قویا بر روشها و فنون بررسی تجربی متکی هستند. به عنوان مثال، بسیاری اوقات هر دو از آزمونهای کلاسی استفاده می کنند.

پژوهش ارزشیابی

ارزشیابی ای که بر پایه شواهد تجربی جمع آوری شده به راههایی که قابل تکرار باشند استوار است و در آن مسائل مربوط به روایی درونی، روایی بیرونی و روایی سازه منظور می شود، پژوهش ارزشیابی نام دارد (روسی). به عبارت دیگر، پژوهش ارزشیابی عبارت است از کاربرد روشهای پژوهش در ارزشیابی.

تعریف پژوهش ارزشیابی در ضمن مثال: مثلاً ما می خواهیم مهارت حل مسئله ریاضی در بخش معادلات را در کل کشور مورد ارزیابی قرار دهیم که در این صورت ما نمی توانیم کل کشور را مورد پژوهش قرار دهیم لذا کشور را به پنج منطقه و هر منطقه را به دو استان تقسیم می کنیم و از هر استان ما سه شهرستان را انتخاب می کنیم و در واقع با این کار یک نوع نمونه گیری کرده ایم و از آزمون های آماری استفاده شده.

ارزشیابی آموزشی و پژوهش ارزشیابی

تفاوتها:

پژوهش ارزشیابی	ارزشیابی آموزشی	ملاک
استنتاج	عمدتاً تصمیم گیری	هدف
مهم است	مهم نیست و متکی است بر یک پدیده آموزشی ویژه	تعمیم پذیری نتایج
خالی از ارزش است و به دنبال حقایق علمی است	در تصمیم گیری، ارزش معیار است	رابطه با ارزش
وابسته است، زیرا به دنبال دانش تازه و افزودن به نظریه ها است	الزاماً وابسته نیست	وابستگی به نظریه

مقیاس های اندازه گیری

تعریف و انواع مقیاس های اندازه گیری:

مقیاس، مجموعه ای از اعداد (یا نمادهای دیگر) است که ویژگیهای آنها بر ویژگی های تجربی اشیایی که اعداد به آنها نسبت داده می شوند منطبق است (کوهن و سوردلیک)

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

اسمی، ترتیبی، فاصله ای، نسبتی

مقیاس چیست؟

طبقه بندی مقیاسها به صورت اسمی، ترتیبی، فاصله ای و نسبتی، از ساده به پیچیده و به صورت سلسله مراتبی تنظیم شده است. یعنی هر طبقه ویژگیهای طبقات قبل را دارد.

مقیاس اسمی

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. مانند شماره روی پیراهن بازیکنان که هیچ ارتباطی بین آنها نیست و اعداد صرفاً برای تشخیص هستند.

۲- طبقه بندی

اعداد و ارقام برای طبقه بندی اشیاء و افراد استفاده می شود بدون این که مفهوم ریاضی داشته باشند. مانند طبقه بندی محل زندگی به روستایی شهری و عشایر و یا مانند طبقه بندی تحصیلات به ابتدائی راهنمایی و متوسطه

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 (رتبه ای)

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- مشخص نیست این اشیاء یا افراد از لحاظ صفت مورد اندازه گیری چقدر با هم فاصله دارند. (نانلی)

عملیات مجاز آماری:

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

عملیات مجاز ریاضی:

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

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 قدر مطلق مورد نظر برای هیچ یک از افراد در دست نیست (نانلی). مثلاً یک

دانش آموز در امتحان صفر میگیرد و باید بررسی شود صفر، صفر مطلق (مثلاً در این اتاق هیچ کتابی وجود

ندارد) است یا صفر قراردادی (مانند همان صفر گرفتن در امتحان) درست در این اینجا همان قراردادی است

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 (نانلی).

عملیات مجاز آماری و ریاضی:

در این مقیاس همه عملیات آماری و ریاضی مجاز است.

مقیاسها در یک نگاه:

مقیاس	ترتیب	فواصل یکسان	صفر مطلق	اعمال آماری	اعمال ریاضی
اسمی	-	-	-	۱	-
ترتیبی	√	-	-	۱ و ۲	-
فاصله ای	√	√	-	۱ و ۲ و ۳	الف
نسبتی	√	√	√	۱ و ۲ و ۳	الف و ب

عملیات آماری:

۱- نما (مد) و درصد

۲- میانه، ضریب همبستگی رتبه ای اسپیرمن

۳- میانگین، انحراف معیار، ضریب همبستگی گشتاوری پیرسون

عملیات ریاضی: الف- جمع و تفریق ب- ضرب و تقسیم

دسته بندی دیگر مقیاسها:

مقیاس پیوسته در مقابل مقیاس گسسته

در مقیاس گسسته (مجزا) متغیر مورد نظر می تواند یکی از ارزشهای معین (اسمی، ترتیبی یا تعداد و اعداد صحیح و غیر صحیح با تعداد ارزش محدود) را به خود اختصاص دهد. مانند تعداد اعضای خانواده، جنسیت، سطح تحصیلات و گردونه مضارب $۵/۰$ تا ۳

در مقیاس پیوسته متغیر مورد نظر می تواند در فاصله بین نقاط هر مقیاس هر تعدادی ارزش را به خود اختصاص دهد. مانند سرعت، زمان، طول، وزن، هوش، پیشرفت تحصیلی، استعداد ریاضی و نگرش.

اندازه های پیوسته همیشه تقریبی هستند. ویژگی مهم مقیاس پیوسته تساوی واحدهای اندازه گیری است.

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) مثلا حد واقعی عدد ۵۰ نیم واحد پایین تر و نیم واحد بالاتر از آن است.

$۵۰/۵$ ۵۰ $۴۹/۵$

جلسه سوم

انواع دسته بندی آزمون ها

انواع آزمونها

دسته بندی آزمونها از منظرهای مختلفی انجام شده است:

دسته بندی بر اساس ویژگی مورد اندازه گیری:

• آزمونهای توانایی:

. آزمونهای توانایی شناختی

. آزمونهای توانایی روانی حرکتی

• آزمونهای عاطفی

۱- آزمونهای توانایی

آزمونهای توانایی آموخته‌ها، استعدادها و مهارتهای افراد را می‌سنجد:

ویژگی مهم آزمونهای توانایی این است که در آنها از آزمون شونده خواسته می‌شود حداکثر تلاش خود را به کار بندد تا بهترین نمره ممکن را به دست آورد.

دسته بندی آزمونهای توانایی براساس محتوای سنجش:

دسته ای از این آزمونها توانایی های شناختی را می‌سنجد که برای فعالیت های ذهنی و فکری و تحصیلی کاربرد دارند.

دسته ای دیگر توانایی های روانی - حرکتی را می‌سنجد که برای فعالیتهای صنعتی، نظامی تربیت بدنی و ... به کار می‌روند.

انواع آزمونهای توانایی

دسته بندی بر اساس هدف سنجش:

آزمونهای استعداد: مقدار توانایی با آمادگی فرد برای انجام کارهایی که در پیش دارد و ظرفیت او برای یادگیری های مختلف (مثل آزمونهای هوش کلی و استعداد کلی برای پیش بینی موفقیت همه جانبه فرد، یا آزمونهای استعدادهای ویژه برای پیش بینی موفقیت فرد در زمینه فکری یا عملی مشخص)

آزمونهای پیشرفت: دانشها و مهارت هایی که فرد تا لحظه اجرای آزمون کسب کرده را اندازه می‌گیرد (مثل آزمون پایان ترم کلاسی، آزمون پیشرفت ریاضی، آزمون پیشرفت مهارتهای فنی و ...)

تذکر: در این دو نوع را نمی‌توان به روشنی از یکدیگر متمایز نمود؛ همیشه یک عامل پیش بینی یادگیری آینده، یادگیریهای قبلی یادگیرنده است.

ثورندایک و هیگن در کتاب معروف اندازه گیری و ارزشیابی در روان شناسی و آموزش و پرورش در رابطه با تفاوت آزمونهای استعداد و پیشرفت تحصیلی:

باید به خاطر داشت که همه آزمونهای توانایی، صرف نظر از نامی که دارند توانایی های موجود فرد را اندازه می‌گیرند. ما نمی‌توانیم مستقیماً توانایی «ذاتی» یا «ارثی» کسی را اندازه گیری کنیم. آن چه را که ما می‌توانیم در فرد اندازه بگیریم توانایی او برای انجام دادن اعمال در زمان حال است. بنابراین تمایز بین آزمونهای

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- آزمونهای عاطفی

آزمونهای عاطفی ویژگیهای عاطفی و شخصیتی افراد را می سنجند و بیشتر به امور مشاوره و حل مسایل روانی و تربیتی و حرفه ای کمک می کنند.

این آزمونها در صدد دستیابی به توصیفی دقیق از وضعیت فعلی فرد هستند و در آنها از فرد خواسته می شود وضعیت طبیعی خود را بروز دهد، نه حداکثر توانایی خود را.

آزمونهای عاطفی به اندازه آزمونهای شناختی عینی و دقیق نیستند؛ اما با وجود این، به قول آیکن، ' 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 آزمونهای عاطفی

آزمون شخصیت - آزمون علایق - آزمون نگرش و سایر آزمونهای عاطفی

آزمونهای شخصیت:

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

شخصیت مجموعه ای از رفتارهای فرد در موقعیتهای اجتماعی است (ایبل)
متغیرهای عمده مورد اندازه گیری در آزمونهای شخصیت:

- انگیزش Motivation

مزاج Temperament

سازگاری Adjustment

منش Character

ابزارهای اندازه گیری شخصیت

- پرسشنامه گزارش شخصی: نظر شخصی نسبت به خود

- پرسشنامه جامعه سنجی: نظر افراد دیگر نسبت به فرد

- فنون مشاهده ای: مشاهده رفتار فرد در یک موقعیت

انواع آزمونهای شخصیت:

قرار دادن آزمودنی در موقعیت های مبهم، مانند ارائه تصاویر مبهم (مثل لکه های نامنظم جوهر) و بررسی هرچه فرد در این تصاویر مشاهده می کند.

بر اساس روش مطالعه رفتار:

تک تعبیر (عینی): ارائه محرکها به صورت سازمان یافته و غیرقابل تعبیرپذیری های مختلف، مانند سؤال و پرسش و شکل‌های معلوم.

تعبیرپذیر (فرافکن): ابزارهایی هستند که در آنها محرکها به صورت سازمان نیافته و با قابلیت تفسیر و تعبیر متفاوتی از سوی آزمون شوندگان مختلف، ارائه می شوند. این آزمونها غالبا یک موقعیت مبهم مانند یک لکه جوهر یا تصاویر قابل تفسیر و یا داستانهای دارای قابلیت برداشتهای متفاوت هستند که به فرد داده میشود و از او خواسته می شود تفسیر خود را بگوید یا هرچه در آنها می بیند و یا برایش مهم است بیان نماید.

آزمونهای علاقه (رغبت و نگرش): به طور کلی علاقه و نگرش مشابه ویژگیهای شخصیت هستند، لکن این آزمونها به خاطر اهمیت در مشاوره شغلی و تربیتی از آزمونهای شخصیت جدا شده اند.

شباهت آزمونهای علاقه و نگرش:

هم علاقه و هم نگرش به دوست داشتنها و دوست نداشتنهای فرد اشاره می کند. هر دو به انتخاب فرد درباره فعالیت ها و نهادهای اجتماعی و یا گروه ها ارتباط دارند. هر دو شامل احساسهای شخصی نسبت به امور هستند.

تفاوت آزمونهای علاقه و نگرش:

آزمونهای علاقه احساس فرد را نسبت به فعالیتهایی که می تواند انتخاب کند می سنجند.

آزمونهای نگرش احساس فرد نسبت به نهادها، گروهها یا مفاهیم و اشیا را می سنجند (مهرنر و لهمان).

ابزار اندازه گیری این دو گروه معمولا پرسشنامه علائق و پرسشنامه نگرش است. در پرسشنامه علائق معمولا جمله هایی درباره مشاغل یا فعالیتهای مختلف داده میشود و پرسشنامه نگرش سنج معمولا حاوی گویه هایی درباره نهادهای اجتماعی، گروههای قومی و نژادی | با مفاهیم مختلف هستند.

دسته بندی آزمون ها با توجه به چگونگی تهیه آن ها

۱- آزمونهای میزان شده (استاندارد شده): توسط سازمانها و نهادهای آموزشی برای | ارزیابی و سنجش و پیشرفت در گروههای بزرگ استفاده میشود.

۲- آزمونهای معلم ساخته: آزمون هایی که معلم، مثلا برای سنجش و ارزیابی پیشرفت | تحصیلی دانش آموزان کلاس خود، می گیرد.

استاندارد بودن یا میزان بودن آزمون (کران باخ):

به این صوت است که شیوههای اجرا، شیوههای مشاهده تجهیزات و مواد و قواعد نمره گذاری تا آن اندازه تثبیت شده اند که تا حد امکان در شرایط و زمانهای مختلف نتایج یکسان به دست آید.

تفاوت آزمونهای استاندارد (میزان شده) و معلم ساخته:

۱- آزمونهای میزان شده به زمان و تخصص بیشتری نیاز دارند چون برای گروههای زیاد و زمانهای مختلف است، و هزینه تخصصی بیشتری می طلبد و معمولا توسط متخصصان و مؤسسات تخصصی آزمون سازی تهیه می شوند. محتوای درسی و مهارتها نمونه گیری می شود، آزمون آزمایشی اجرا می شود، سوالها تجزیه و تحلیل می شود و توسط متخصصان نوشته و ویرایش می شود.

۲- از آن جا که آزمونهای میزان شده برای استفاده در مناطق مختلف و دانش آموزان مناطق گوناگون تهیه میشود هدفهای آموزشی مناطق مختلف را باید در این آزمونها منظور کرد.

۳- آزمونهای میزان شده باید هنجاریابی شوند تا بتوان یک فرد یا یک گروه را با افراد یا گروههای دیگر یا جنبه های مختلف عملکرد یک فرد را با گروههای دیگر مقایسه کرد، در حالی که آزمونهای معلم ساخته چنین نیازی ندارند.

موارد استفاده از آزمونهای میزان شده و معلم ساخته

۱- تصمیمات آموزشی مربوط به کم و کیف جریان آموزشی ← آزمونهای معلم ساخته

۲- تصمیمات مربوط به ارتقای دانش آموزان و دانشجویان ← آزمونهای معلم ساخته

۳- تصمیمات مربوط به مشاوره و راهنمایی ← آزمونهای میزان شده

۴- تصمیمات بر مبنای مقایسه افراد و گروهها ← آزمونهای میزان شده

۵- تصمیمات مربوط به برنامه ریزی درسی ← آزمونهای میزان شده

۶- تصمیمات در مورد ارزشیابی از میزان موفقیت مدارس مختلف ← آزمونهای میزان شده

دسته بندی روشهای اندازه گیری با توجه به نوع سؤال (ماده یا گویه) و تصحیح آن:

پاسخگزين (عینی): انتخاب پاسخ از میان پاسخهای معلوم (صحیحو غلط، جورکردنی و چند گزینه ای - استفاده برای سطوح شناختی پایین)

پاسخ ساز (غیرعینی): ساختن و ارائه پاسخ توسط آزمون شونده (تشریحی با انواعش و کامل کردنی - استفاده برای سطوح شناختی بالا و عمیق)

دسته بندی ارزشیابی های آموزشی

با توجه به موضوع ارزشیابی:

. ارزشیابی از دانشجویان و دانش آموزان

. ارزشیابی از برنامه های درسی و مواد آموزشی

. ارزشیابی از پروژه ها و برنامه های آموزشی ویژه (نو آوارانه)

. ارزشیابی از آموزشگاه

. ارزشیابی از کارکنان آموزشگاه

با توجه به ملاک مورد استفاده:

• ارزشیابی وابسته به ملاک (Criterion - Referenced Evaluation)

معیار یا ملاک ارزشیابی از پیش تعیین می شود و عملکرد یادگیرنده با توجه به آن ملاک مورد داوری قرار می گیرد. توزیع کجی منفی (معمولا همه در آن موفق میشوند) مانند امتحانات نهایی

• ارزشیابی وابسته به هنجار (Norm - Referenced Evaluation)

به نوعی ملاک نسبی وابسته است و افراد را با یکدیگر و یا با هنجار گروهی) مقایسه می کنیم. می توان تعیین نمود که فرد نسبت به دیگر آن چه وضعیتی دارد، اما نمی توان تعیین کرد فرد نسبت به هدفهای آموزشی و یا محتوای آموزش چه جایگاهی دارد. توزیع بهنجار مانند کنکور سراسری

با توجه به زمان و هدف استفاده از ارزشیابی:

. سنجش آغازین (سنجش رفتارهای ورودی - آزمون جایابی)

. سنجش تکوینی

. ارزشیابی تشخیصی

. ارزشیابی تراکمی

با توجه به ارزشیابان

. ارزشیابی درونی

توسط افراد درون برنامه (کارکنان و دست اندرکاران) تهیه و اجرا می شود.

. ارزشیابی بیرونی

توسط فردی که از بیرون برای ارزشیابی گمارده شده است و مشابه ارزشیابی تراکمی برای داوری درباره کیفیت با ارزش مواد آموزش روشهای آموزشی، برنامه های درسی، برنامه های آموزشی یا نظامهای آموزشی است.

سایر دسته بندی های ارزشیابی آموزشی

- . بر اساس اجرا: آزمونهای فردی، آزمونهای گروهی
- . بر اساس هدف: آزمون سرعت، آزمون قدرت
- . بر اساس پوشش اهداف (طبقه بندی اهداف، در بخشهای بعدی طرح خواهد شد)
- . بر اساس وابستگی به فرهنگ: آزمونهای وابسته به فرهنگ، آزمونهای نوابسته به فرهنگ
- . بر اساس فعالیت آزمودنی: آزمون شفاهی، آزمون مداد کاغذی، آزمون عملی

جلسه چهارم

مراحل و فعالیتهای ارزشیابی آموزشی

صرفنظر از

. الگوی ارزشیابی

. موضوع ارزشیابی

. نحوه کاربرد فرایند ارزشیابی

. ملاک های به کار رفته در ارزشیابی و

. نوع تصمیماتی که گرفته می شوند،

فرایند ارزشیابی همواره یکسان است و اصول و مراحل اساسی کار هیچ تغییر نمی کند.

مراحل اصلی ارزشیابی:

۱- مرحله طراحی

۲- مرحله فرایندی (اجرا و اعمال)

۳- مرحله فرآوری (خروجی و نتیجه)

مرحله طراحی

زمان اجرا: پیش از اجرای آموزش، برنامه، پروژه و غیره

نوع تصمیم ها:

۱- نتایجی که قرار است به دست آیند. ۲- اعمالی که باید انجام شوند.

مراحل طراحی

۱- تحلیل موقعیت: جمع آوری سوابق اطلاعاتی و برآورد و تعیین موانع موجود، مثال: برای ارزشیابی برنامه

آموزشی، مرور پژوهشهای مربوط به آن ضروری است تا بتوان رویکردهای

ارزشیابی مفید و همچنین ابزارهای موردنیاز را انتخاب کرد.

۲- **تعیین و توصیف هدفها:** اهداف به فعالیت های بعدی ارزشیابی جهت می دهند و نهایتاً در ارزشیابی میزان تحقق این هدف ها سنجش می شود.

هدف ها به فعالیت های بعدی ارزشیابی جهت می دهند و نهایتاً در ارزشیابی میزان تحقق این هدفها سنجش میشود. هدفهای آموزشی بر دو نوع هستند:

۱- **هدفهای کلی (غایتها)**، مثال: ارتقای سطح معلومات ریاضی دانش آموزان.

۲- **هدفهای جزئی** (یا دقیق که معمولاً هدف نامیده میشوند)، مثال: اعداد دو رقمی را با هم جمع کند.

هدفها(ی جزئی) به دو دسته کلی تقسیم میشوند:

. هدفهای فرایندی (کوششها و اقدامات ضمن اجرای برنامه)

. هدفهای فرآوردهای (بازدهها و نتایج مورد انتظار کوششها و اقدامات)

ارزشیابی هدف های آموزشی:

. ربط داشتن

. اندازه پذیر بودن

. مهم بودن

۳- **توصیف پیش نیازها یا رفتارهای ورودی:** رفتار های ورودی به پیش نیازهایی که یادگیرنده پیش از شروع به یادگیری هدف های تازه باید آنها را کسب کرده باشد و لذا به آمادگی فرد برای یادگیری هدفهای آموزشی اشاره دارد

۴- **انتخاب و تولید وسایل یا ابزارهای اندازه گیری:** اگر وسایل یا ابزار ها(مانند آزمون های پیشرفت، نگرش سنج، فرمهای ثبت مشاهدهای و ...) موجود باشند باید به بررسی و انتخاب مناسب ترین آنها اقدام کرد و اگر نه باید آنها را تهیه کرد که این امر به تخصص، دقت و صلاحیت نیاز دارد و در برنامه های سرنوشت ساز ارزشیابی باید آن را به متخصصین سپرد

۵- **توصیف راهبردها یا استراتژیها:** استراتژی یا راهبرد فعالیت یا تدبیری است که به منظور کمک به تحقق یافتن هدف های برنامه به کار می رود؛ شامل استراتژی های آموزشی و استراتژی های برنامه درسی توضیح بیشتر:

استراتژی با راهبرد، فعالیت یا تدبیری است که به منظور کمک به تحقق یافتن هدفهای برنامه به کار می رود؛ شامل استراتژی های آموزشی و استراتژی های برنامه درسی

مثال: موضوع چندجمله ای ها، با هدف های تشخیص مفاهیم یک جمله ای، دوجمله ای و چندجمله ای، تعیین درجه آنها و جمع کردن آنها.

برخی تصمیم های استراتژیک:

. توالی ارائه موضوعها

. توالی ارائه مفاهیم و روشها

. روش تعیین توالی ارائه هدفهای درس (تحلیل تکلیف و سلسله مراتب هدفهای آموزشی)

. مرور یا بازنگری و نیز تمرین و بازخورد

. دسته بندی یادگیرندگان و روش هایی مانند ایفای نقش، در هدفهای خاص

۶- انتخاب طرح پژوهشی در صورت انجام پژوهش ارزشیابی:

۷- تدارک برنامه زمانبندی: با برآورد واقع بینانه ای از جریان پیشرفت کار و دربرگیرنده فهرستی از فعالیت های اصلی مربوط به اقدامات پیشنهادی و زمان شروع و خاتمه هر یک از فعالیت ها با لحاظ کردن مقداری زمان اضافی و تاریخ آخرین اقدام قدری جلوتر از موعد رسمی.

یک پیشنهاد: روش چارت گانت (فهرست فعالیت ها در سمت راست کاغذ و زمان انجام تمامی پروژه در بالای صفحه)

مرحله فرایندی (اجرایی)

زمان اجرا: ضمن اجرای آموزش، برنامه، پروژه و غیره

نوع تصمیمها:

۱- به چه میزانی مراحل طرح ریزی شده به اجرا در می آیند؟

۲- چه تغییراتی باید داده شوند تا پیشرفت لازم به دست آید؟

نکته: توجه به تکوین ارزشیابی و تصمیم گیری درباره اقدامات عملی یا اجرایی (نوعی ارزشیابی تکوینی):

. اجرای پیش آزمون / آزمون رفتار ورودی و

. سنجش مناسب بودن هدفها / واری نتایج سنجش آغازین و در صورت لزوم تغییر برنامه

. اجرای استراتژیها و فعالیتهای طراحی شده به سبک تعیین شده و دریافت بازخوردهای لازم برای مراحل اجرایی و

. تحلیل اثربخشی راهبردها / تصمیم گیری درباره بخشهای مختلف برنامه تدریس و راهبردها و یا ارزشیابیها

بر حسب بازخوردهای سنجش و ارزشیابی

اجرای آزمونها، ارائه بار خورد و تقویت تدارک شده برای هر یادگیرنده و جمع آوری داده های دوره ای /
تصمیم گیری درباره پیشرفت فردی یا گروهی دانش آموزان با توجه به نوع ارزشیابی

مرحله فرآوردهای (تولیدی)

زمان اجرا: پس از اجرای آموزش، برنامه، پروژه و غیره

نوع تصمیمها:

۱- اثربخشی کلی آموزش، پروژه، برنامه و غیره ۲- اقدامات آتی

نکته: تصمیمات در پایان برنامه یا پروژه یا در پایان دوره یا مقطعی از آن، یعنی زمانی که فعالیتهای کامل می شوند (ارزشیابی تراکمی):

• جمع آوری و تنظیم دادههای مربوط به هدف ها مانند اجرای پس آزمون

• جمع آوری و تنظیم دادههای مربوط به بازدههای پیش بینی نشده

• تحلیل و تفسیر داده های مختلف

• تهیه گزارش شامل هدفها، روشها و بازدههای فعالیت های ارزشیابی

سه راه استفاده از نتایج مرحله فرآوردهای:

• تدارک بازخورد برای هر ذینفع (مانند فراگیران، مدرسان و مجریان برنامه)

• بازخورد برای تصمیم گیرندگان خارجی (مانند والدین، مدیران و منابع مالی)

• ارائه اطلاعات به سایر گروهها (مانند مشاوران و پژوهشگران)

فعالیت های ارزشیابی آموزشی

۱. تعیین، انتخاب، پالایش یا تغییر غایت های برنامه و هدف های ارزشیابی ۲- مشخص کردن استانداردها /

هدفها ۳. تهیه طرح ارزشیابی مناسب ۴. انتخاب یا تولید روشهای جمع آوری دادهها ۵. جمع آوری داده های

مناسب ۶. پردازش، خلاصه کردن و تحلیل دادهها ۷- مقایسه داده ها با استانداردها ۸- گزارش و بازخوردادن

از نتایج ۹. هزینه - سود / اثربخشی

ارزشیابی پیشرفت تحصیلی

تعریف :

سنجش عملکرد یادگیرندگان و مقایسه ی 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 به نتایج مطلوب

انجامیده اند و به چه میزانی

بنابراین، در ارزشیابی پیشرفت تحصیلی دو اقدام اساسی ضروری است: ۱- تعیین هدف های آموزشی ۲- سنجش یا اندازه گیری عملکرد یادگیرندگان

هدفهای کلی و هدف های دقیق

هدفهای کلی یا غایتها بیاناتی هستند کلی و آرمانی که قصد و منظور از ایجاد و اداره دوره های آموزشی یک جامعه (مانند دبستان، دبیرستان، دانشگاه و ...) را نشان می دهند و معمولاً به وسیله برنامه ریزان در سطح وزارت تعیین می شوند.

اما **هدفهای دقیق** را (که هدفهای آموزشی با هدف های یادگیری نامیده میشوند) غالباً مربیان و متخصصان آموزشی با توجه به غایتهای کلی و با استفاده از محتوای برنامه درسی تهیه می کنند و منظور و مقصودی هستند که معلم از آموزش خود و یادگیرنده از یادگیری خود دارد. در ارزشیابی آموزشی نیز عملکرد یادگیرندگان با توجه به این هدف های دقیق آموزشی سنجش میشوند و ملاک قضاوت درباره توفیق معلمان و یادگیرندگان به حساب می آیند.

استانداردهای تحصیلی و نشانگرها

. استانداردها یک فهرست، توصیف یا بازنمایی از کیفیت یا ویژگیهایی است که یک شیء باید داشته باشد. استانداردهای تحصیلی بیاناتی هستند که مشخص می کنند چه چیزی باید آموزش داده شود و دانش آموزان چه چیزی باید یاد بگیرند.

. معمولاً استانداردها به صورت کلی بیان می شوند و برای آنها یک یا چند نشانگر نوشته می شود تا آموزش پذیر و سنجش پذیر شوند.

. استانداردها به اهداف کلی و نشانگرها به هدفهای دقیق نزدیک هستند.

مثال:

دانش آموزان فرایند خواندن را به طور مؤثر مورد استفاده قرار می دهند.

(استاندارد)

دانش آموزان بر اساس عنوان و شکلهای یک متن، پیش بینی می کنند که آن متن درباره چه چیزی صحبت می کند. (نشانگر)

طبقه بندی اهداف آموزشی

طبقه بندی معروف اهداف آموزشی دربرگیرنده سه حوزه و حیطة است: **شناختی، عاطفی و روانی** -

حرکتی

به طور خلاصه حوزه شناختی با فکر، حوزه عاطفی با قلب و حوزه روانی حرکتی با دست و پا سر و کار دارد

حیطه شناختی:

هدفهای حوزه شناختی به جریان هایی که با فعالیت های ذهنی و فکری آدمی سروکار دارند مربوط می شوند. به طور کلی یادگیری به محتوای مطالب و کسب شناخت و معرفت درباره آنها مربوط می شود. نتایج حاصل از آموزش شناختی به مهارت های ذهنی از قبیل بازشناسی و یاد آوری، فهمیدن، کاربرد آموخته ها، تجزیه و تحلیل، ترکیب و ارزشیابی منتهی می شود. سلسله مراتب آموختنیهای حیطه شناختی از روند آسان به مشکل پیروی می کند. حیطه شناختی مهمترین حوزه یادگیری است، چون اکثریت فعالیت های تحصیلی آموزشگاهها و غالب موضوع های درسی و هدفهای آموزشی به این حوزه مربوط می شوند.

حوزه شناختی

طبقات حیطه شناختی بر اساس سلسله مراتب تنظیم شده و هر طبقه از آنها مستلزم مهارت ها و توانائی های طبقات پائین تر است:

ارزشیابی					
ترکیب	ترکیب				
تجزیه و تحلیل	تجزیه و تحلیل	تجزیه و تحلیل			
کاربرد	کاربرد	کاربرد	کاربرد		
درک و فهم	درک و فهم	درک و فهم	درک و فهم	درک و فهم	
دانش	دانش	دانش	دانش	دانش	دانش

طبقه بندی حوزه شناختی به جریانهایی که با فعالیتهای ذهنی و فکری آدمی یا به بیان دیگر با شناخت و اندیشه انسان سر و کار دارند مربوط است. این حوزه از شش طبقه اصلی به شرح زیر درست شده است.

۱- **دانش:** یادآوری امور جزئی و کلی، فرآیندها و روشها، الگوها، ساختها، یا موقعیتهای. این طبقه شامل حفظ و نگهداری موضوعهای قبلا آموخته شده است.

فعلهای نمونه: نام ببرید، یادآوری کنید، مشخص کنید، تعریف کنید.

هدف نمونه: یادگیرنده بتواند انواع آزمونهای عینی مورد استفاده معلم را در سنجش پیشرفت تحصیلی دانش آموزان و دانشجویان از حفظ نام ببرد.

۲ - **فهمیدن (درک):** درک مطالب آموخته شده. فهمیدن یک مرحله بالاتر از دانش است، زیرا در طبقه دانش فقط از یادگیرنده خواسته می شود تا مطالبی را که خوانده یا شنیده است، بدون تغییر زیاد، به یاد آورد، اما در طبقه فهمیدن، علاوه بر حفظ مطالب، باید آنها را بفهمد.

فعلهای نمونه: خلاصه بیان کنید، مثال بزنید، تفسیر کنید، ترجمه کنید، توضیح دهید.

هدف نمونه: یادگیرنده بتواند با ذکر مثال مفهوم همبستگی مثبت و منفی بین دو متغیر را توضیح دهد.

۳- کاربرد (کار بستن): استفاده از مطالب انتزاعی (اندیشه های کلی، قواعد اجرایی، روشهای کلی) در موقعیتهای ویژه و عینی. اصطلاح معمول آموزشی برای این طبقه حل مساله است.

فعلهای نمونه: نشان دهید، استفاده کنید، پیدا کنید، انجام دهید، اجرا کنید، بکار ببندید.

هدف نمونه: یادگیرنده بتواند با استفاده از اصول علم تغذیه که در آموزشگاه یاد گرفته است یک رژیم غذایی مناسب برای خودش تهیه کند.

۴- تحلیل: شکستن یک موضوع به اجزای تشکیل دهنده آن. این طبقه شامل یافتن عناصر و ارتباط میان عناصر یک کل پیچیده، مانند یک نظریه علمی، یک مقاله تحقیقی است.

فعلهای نمونه: اجزا را مشخص کنید، جدا کنید، تفاوتها و مشابهت ها را بیان کنید، تفکیک کنید، نشان دهید.

۵- ترکیب: پهلوی هم گذاشتن عناصر و اجزاء برای ایجاد یک اثر یا یک فرآورده تازه. ترکیب همان خلاقیت یا آفرینندگی است.

فعلهای نمونه: طراحی کنید، خلق کنید، ایجاد کنید، ارائه دهید، بنویسید، مرتب کنید.

هدف نمونه: یادگیرنده بتواند برای یک موقعیت جدید آموزشی یک طرح درس ابتکاری بنویسد.

۶- ارزشیابی: داوری یا قضاوت درباره ارزش یا اعتبار موضوعهای مختلف. اصطلاح دیگر مورد استفاده برای این طبقه تفکر انتقاد است. منظور از تفکر انتقادی این است که دانش آموز یا دانشجو یاد بگیرد که گفته ها، شنیدهها، و دیدهها را صرفا با توجه به اعتبار ظاهری آنها نپذیرد، بلکه پس از تفکر دقیق و تیزبینانه و واری درستی یا نادرستی آنها تصمیم بگیرد که آنها را بپذیرد یا رد کند.

فعل های نمونه: داوری کنید، مقایسه کنید، بسنجید، محاسن و معایب را مشخص کنید، نظر خود را بیان کنید.

هدف نمونه: یادگیرنده بتواند پس از شنیدن بیانات یک شخص یا خواندن یک مقاله دلایل غیر منطقی و سفسطه آمیز و نتیجه گیری های نامربوط آن را مشخص کند.

طبقه بندی حوزه شناختی این است که به صورت سلسله مراتبی درست شده و هدف های آن بنا به اصل پیچیدگی سازمان یافته اند؛ یعنی هدفهای طبقه دانش در پائین ترین سطح واقع اند، بعد از آن هدفهای طبقه فهمیدن، و بالاتر از همه هدفهای طبقه ارزشیابی قرار دارند.

طبقه بندی دو بعدی حوزه شناختی:

بعد دانش

بعد فرایند شناختی

طبقه بندی حوزه عاطفی

. دریافت کردن (توجه کردن)

. پاسخ دادن

. ارزش گزاردن

. سازمان دادن با ارزشها

. شخصیت پذیرفتن (سبک زندگی)

طبقه بندی دو بعدی حوزه روانی - حرکتی

. ادراک حسی . آمادگی . پاسخ هدایت شده . عادت (مکانیسم) . پاسخ پیچیده اشکار . انطباق . ابتکار

جدول مشخصات

یک آزمون پیشرفت تحصیلی زمانی خوب است که سؤالهای آن حاوی نمونه درستی از هدفها و محتوای درس باشد. یعنی آزمون به بهترین شکل منعکس کننده تمامی هدفهای آموزشی و همه محتوای برنامه درسی باشد. مهمترین گام در تهیه ی آزمونهای پیشرفت تحصیلی ، تهیه جدول مشخصات برای موضوعی است که قرار است آزمون آن تهیه شود. جدول مشخصات یک جدول دوبعدی است با تعدادی ردیف و ستون متناسب با محتوا و هدفهای آموزشی درس .

در ردیف بالای جدول که بعد محتوا نام دارد، اجزای مطالب درسی نوشته میشوند و در ستون کنار آن که بعد هدف نامیده می شود، هدفهای آموزشی قرار می گیرند.

چگونگی توزیع سؤالها در بخشهای مختلف

برای تعیین درصد سوالها باید جواب سوال های زیر را تعیین کرد:

۱ - کدام بخش از محتوا و کدام دسته از هدفها دارای حجم بیشتری است و یا وقت بیشتری صرف آموزششان شده است ؟

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

۳ - کدام قسمت محتوای درس با موضوعات درسی از اهمیت بیشتری برخوردار است؟

جدول ۳-۶ الگوی کلی جدول مشخصات آزمونهای پیشرفت تحصیلی

تعداد هدف	بخش ۱			بخش ۲			بخش ۳			تعداد کل
	۱	۲	۳	۱	۲	۳	۱	۲	۳	
هدفهای حوزه دانش	دانش	فهمیدن	کار بستن	و غیره	تعداد سؤالات	تعداد سؤالات	تعداد سؤالات	تعداد سؤالات	تعداد سؤالات	تعداد سؤالات
هدفهای حوزه دانش	دریافت کردن	پاسخ دادن	ارزش گذاری	و غیره	تعداد سؤالات	تعداد سؤالات	تعداد سؤالات	تعداد سؤالات	تعداد سؤالات	تعداد سؤالات
هدفهای حوزه دانش	آزادی	آگاهی	پاسخ هدایت شده	و غیره	تعداد سؤالات	تعداد سؤالات	تعداد سؤالات	تعداد سؤالات	تعداد سؤالات	تعداد سؤالات

نکته: چنانچه بخشهای مختلف محتوا از اهمیت یکسانی برخوردار باشند و در توزیع سؤالات تنها بخواهیم مقدار ساعات تدریس را برای هر بخش منظور کنیم، از فرمول زیر استفاده می کنیم:

تعداد ساعت هایی که صرف تدریس بخش شده ۰۰۰ نسبت به ۰۰۰ تعداد کل ساعت های تدریس = درصد سؤالات هر بخش

سپس پاسخ به دست آمده را در تعداد کل سؤالات ضرب میکنیم تا تعداد سؤالات بخش مورد نظر به دست آید.

جلسه پنجم

مطالب مرح شده در این جلسه قبلا در جلسه چهارم بیان شده و عینا تکرار شده اند

جلسه ششم

یادآوری: انواع سوال در آزمون های کتبی: ۱- سوال های عینی (پاسخ گزین) ۲- سوال های غیر عینی (

ذهنی، پاسخ ساز) (تقسیم بندی بر اساس نوع تصحیح پاسخ ها)

۱- سؤالات کتبی عینی

سؤالاتی هستند که هم صورت سؤال و هم جوابهایی برای سؤال در اختیار آزمون شوندگان قرار میدهند و آزمونشوندگان درباره جوابهای داده شده اعمالی انجام میدهند یا تصمیمهایی اتخاذ می کنند.

سؤالهای عینی به سه نوع اصلی: صحیح غلط، جور کردنی و چند گزینه ای تقسیم میشوند.

صحیح غلط

در این سوال ها تعدادی پرسش یا جمله (گویه یا ماده) در اختیار آزمون شونده گذاشته می شود و صحیح یا غلط بودن آن ها را تعیین می کند

انواع سؤالهای صحیح - غلط:

: نوع صحیح - غلط ساده : نوع صحیح - غلط با توضیح : نوع بله - نه : نوع بله - نه با توضیح : نوع موافق - مخالف : نوع واقعیت - عقیده : نوع چند صحیح - غلط (نوع خوشه ای)

محاسن سؤالهای صحیح غلط:

۱. بیان موضوعات درسی در قالب عبارات ساده ۲. سادگی تهیه ۳. عینی بودن تصحیح جوابها
۴. امکان زیاد بودن تعداد آنها در هر جلسه امتحان

معایب سؤالهای صحیح غلط:

۱. اغلب برای سنجش اطلاعات جزئی و کم اهمیت به کار می رود. ۲. حدس کورکورانه را افزایش می دهد.
۳. کاربرد زیاد آنها حفظ طوطی وار را افزایش می دهد ۴. بیشتر، هدفهای سطح پایین را می سنجد.

قواعد تهیه سؤالهای صحیح - غلط:

۱. سؤالها را تا حد امکان مختصر، ساده و روشن بنویسید.
۲. جملاتی به کار ببرید که برای کسانی که موضوع سؤال را یاد گرفته اند به طور آشکار درست یا غلط باشد.
۳. حتی المقدور از کلماتی مثل همه، بعضی وقتها، غالبا، هرگز و بپرهیزید (چون این کلمات هدایتگر میشوند).
۴. تعداد سؤالهای صحیح و غلط تقریبا به یک اندازه باشد.
۵. سؤالهای صحیح غلط را از روی جملات کتاب ننویسیم
۶. هر سؤال به یک هدف آموزشی مهم مربوط باشد و مطلب بی اهمیت در سؤال قرار نگیرد.

سایر نکات:

: موضوعات مهم و در ارتباط با هدف های آموزشی را در سؤال های خود بیاورید

: سؤالهایی طرح کنید که درک و فهم را بسنجد نه یادآوری کلمه به کلمه را

: سؤالهایی طرح کنید که درست یا غلط بودن آنها به وسیله اشخاص صاحب نظر (قابل دفاع باشد

: از جمله های پیچیده و چند قسمتی استفاده نکنید

: در هر سؤال بیشتر از یک موضوع قرار ندهید

: در صورت امکان از زبان کمی و دقیق استفاده کنید نه از زبان کیفی و غیر دقیق .

: تا آنجا که ممکن است از طرح سؤالهای منفی پرهیز کنید

: از کلمات راهنما یا اشاره کننده به جواب درست استفاده کنید.

: سعی کنید تعداد سؤالهای غلط بیشتر از تعداد درست باشد..

: طول سؤالهای صحیح و غلط را تقریباً هم اندازه انتخاب کنید.

جوورکردنی

شامل تعدادی پرسش و تعدادی پاسخ می باشد و آزمودنی باید پرسشها و پاسخها را به نحوی به یکدیگر مربوط سازد؛ به عنوان مثال، یک ستون پرسشها و ستون دیگر پاسخها می باشد، و آزمون شونده پاسخها را با پرسشها جور می کند.

محاسن سؤالهای جوورکردنی

- (۱) برای آزمون دانش اصطلاحات، تعاریف، تاریخها و رویدادهاست.
- (۲) برای سنجش یادگیریهایی که با کلمات چه کسی، چه چیزی، چه زمانی و کجا مشخص میشوند، مفیدند.
- (۳) تعداد زیادی از آنها را میتوان در یک زمان بکار برد و نمونه نسبتاً بزرگی از محتوای درس و هدفهای آموزشی را دربرمی گیرد.

معایب سؤالهای جوورکردنی

- ۱- پیدا کردن مجموعه ای از پرسش ها و جواب ها که همگون باشند مشکل است.

۲- حفظ طوطی وار را افزایش می دهد

قواعد تهیه سؤال های جورکردنی:

- طول پرسشها و پاسخها را کوتاه کنید.
- به وسیله راهنمای سؤال، اطلاعات لازم را درباره نحوه جورکردن توضیح دهید.
- همه پرسشها و پاسخها را در یک صفحه قرار دهید
- در سؤالها از کاربرد جملات نیمه تمام خودداری کنید.
- پرسشها را با شماره و پاسخ ها را با حروف مشخص کنید.
- هر یک از پاسخ ها باید برای تمام پرسشها درست جلوه کند.
- پرسش و پاسخ از لحاظ تعداد یکسان نباشند.
- اطلاعات کامل درباره چگونگی جور کردن پرسش ها و پاسخ ها را در راهنمای سؤال بنویسید
- موضوع های مهم و در ارتباط با هدف های آموزشی را در سؤالها قرار دهید
- پرسش ها و پاسخ های هر تمرین را متجانس انتخاب کنید
- طول فهرست پرسش ها و پاسخ ها را کوتاه کنید
- توضیحات مفصل را در فهرست پرسش ها قرار دهید و در فهرست پاسخ ها مطالب کمتری بنویسید
- از نوشتن تمرین ها یا سؤالهایی که پرسش ها و پاسخ های آنها کاملاً جور می شوند پرهیز کنید
- هر یک از پاسخ ها را به گونه ای انتخاب کنید که برای تمام پرسش ها درست جلوه کند
- پرسشها را با شماره و پاسخ ها را با حرف مشخص کنید
- فهرست پرسش و فهرست پاسخ ها را به طور منطقی مرتب کنید
- درنوشتن صورت سؤالها ابتکاری عمل کنید
- در برگه های امتحانی همه پرسش ها و پاسخ های یک سؤال را در یک صفحه قرار دهید

جلسه هفتم

چند گزینه ای

سؤالهای عینی:

هر یک از سوال ها از یک قسمت اصلی و تعدادی گزینه(پاسخ) تشکیل می شود و آزمون شونده از میان گزینه های پیشنهادی، گزینه صحیح(پاسخ سوال) را انتخاب می کند.

قسمتهای مختلف سؤال چند گزینه ای :

۱- قسمت اصلی یا تنه سؤال (stem) این بخش از سؤال، متن اصلی سؤال را تشکیل میدهد و دربرگیرنده مسئله یا موضوعی است که سؤال باید آن را اندازه گیری کند.

۲- گزینه درست یا پاسخ سؤال (key) یکی از گزینه های پیشنهادی، پاسخ درست سؤال است که آزمون شونده باید آن را برگزیند. به این گزینه کلید گفته میشود.

۳- گزینه های انحرافی (distracters). به غیر از گزینه درست سؤال، تعدادی گزینه دیگر نیز برای هر سؤال تهیه میشود که به آنها گزینه های انحرافی می گویند. نقش گزینه های انحرافی، منحرف کردن آزمون شوندگانی است که پاسخ درست سؤال را نمی دانند.

متن اصلی یا تنه سؤال را یک جمله استفهامی یا یک جمله ناتمام تشکیل میدهد و گزینه های پیشنهادی، جواب احتمالی برای جمله استفهامی یا تکمیل کننده جمله ناقص هستند

نکته: تعداد گزینه های هر سال چند گزینه ای معمولاً از ۲ تا ۵ گزینه متغیر است از لحاظ نظری هرچه تعداد گزینه ها بیشتر باشد امکان حدس زدن کمتر است اما به سبب اینکه پیدا کردن بیشتر از سه پاسخ انحرافی مشکل است آزمونهای چهارگزینه ای مرسوم ترند

انواع سؤال های چند گزینه ای از نظر بدنه سوال:

پرسشی / ناتمام / تنها گزینه درست / بهترین گزینه / منفی / جایگزینی / ناتمام پیشنهادی / پاسخ ترکیبی

انواع سؤال های چند گزینه ای از نظر گزینه ها:

تنها گزینه درست / بهترین گزینه / منفی (بجز یکی صحیح است)

محاسن سؤالهای چندگزینه ای:

سؤالهای چندگزینه ای علاوه بر محاسن سؤالهای صحیح غلط و جورکردنی، محاسن زیر را نیز دارا میباشند:

۱. این سؤالها از سایر سؤالها انعطاف پذیری بیشتری دارند. این سؤالها علاوه بر دانش و یادآوری، سطوحی از استدلال و قضاوت و بازه های مهم یادگیری را می سنجند.
۲. این سؤالها در یک زمان محدود تعداد زیادی هدف آموزشی و بخش مهمی از محتوا را اندازه گیری می کنند.

۳. این سؤالها نسبت به سؤال های صحیح غلط امکان حدس زدن کمتری دارند.

۴. پاسخ این سؤاها به سادگی و با عینیت کامل تصحیح می شوند او تحلیل‌های بیشتری برای پاسخ‌های این سؤاها در دسترس است.]

۵. معمولا دانش آموزان و معلمان این آزمون ها را به سایر آزمون‌های عینی ترجیح می دهند.

۶. اگر گزینه های انحرافی این نوع سؤاها با توجه به اشتباهات و کج فهمی های متداول دانش آموزان تهیه شود، منبع بسیار مناسبی برای تشخیص مشکلات دانش آموزان خواهد بود.

معایب سؤا‌های چندگزینه ای:

۱. ساختن آنها دشوار است.

۲. تهیه کنندگان این سؤال ها، اغلب آنها را در حد اطلاعات جزئی و بی اهمیت یا کم اهمیت می نویسنده

۳. در مقایسه با سؤا‌های صحیح غلط، خواندن این سؤاها و پیدا کردن پاسخ وقت بیشتری احتیاج دارد.

۴. زمانی که نمره منفی وجود دارد، دانش آموزانی که ریسک می کنند گاهی از دانش آموزانی با توانایی مشابه خود نمره بهتری می گیرند.

۵. دانش آموزان قوی گاهی به دلیل تیزبینی و دقت زیاد، غلط جواب می دهند.

قواعد تهیه سؤا‌های چندگزینه ای :

- ۱- هر سؤال باید یک موضوع مهم یا یک هدف آموزشی را اندازه بگیرند. ۲- بیش از یک مسئله یا مطلب در یک سؤال قرار ندهید. ۳- سؤاها را کاملا روشن و واضح بنویسید و از عبارات و کلمات قابل فهم آزمون شوندگان استفاده کنید. ۴- از تکرار مطالب (مثلا در ابتدای همه 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 باشند. ۷- سؤال را طوری بنویسید که گزینه درست تنها پاسخ درست یا قطعا درست ترین پاسخ باشد. ۸- گزینه های انحرافی را طوری بنویسید که آزمون شوندگان بی اطلاع از موضوع سؤال را به خود جلب کند. ۹- گزینه های هر سؤال را طوری بنویسید که از نظر دستوری و جمله بندی به نحو درست مکمل متن سؤال باشند. ۱۰- در سؤال های منفی کلمات منفی را برجسته جلوه دهید. ۱۱- از نوشتن سؤا‌هایی که در آنها متن سؤال منفی و گزینه ها هم منفی باشند (منفی مضاعف) خودداری کنید. ۱۲- تا حد امکان از کاربرد عباراتی همچون «همه آنچه در بالا گفته شد» یا «هیچ یک از آنچه در بالا گفته شد» خودداری شود. ۱۳- سؤاها مستقل از یکدیگر نوشته شود، پاسخ یک مسئله نباید شرط لازم برای حل مسئله و یا تهیه سؤال بعدی باشد. ۱۴- حتی المقدور دو گزینه متضاد را که یکی از آنها درست است به کار نبرید. ۱۵- طول گزینه درست را در سؤا‌های مختلف تغییر دهید. ۱۶- محل گزینه درست را در میان گزینه های انحرافی به طور تصادفی انتخاب کنید. ۱۷- سؤا‌های آزمون طوری انتخاب شود که پاسخ درست هیچیک از پرسشها از روی متن و یا گزینه های پرسش های دیگر مستقیما یا به طور ضمنی استنباط نشود. ۱۸- در سؤا‌هایی که تنه آنها جمله ناتمام است، جای خالی را در قسمت آخر قرار دهید.

به طور خلاصه:

یک موضوع مهم یا یک هدف آموزشی را اندازه بگیرد / بیش از یک هدف یا مطالب در سؤال نباشد / سؤال کاملا روشن و در حد درک آزمون شوندگان باشد / از تکرار مطالب در گزینه ها خودداری شود. / مطلب اصلی سؤال را به طور کامل در تنه سؤال بنویسید / بکشید همه گزینه های یک سؤال متجانس و مربوط به موضوع واحد باشند. / پاسخ درست تنها پاسخ درست یا قطعا درست ترین پاسخ باشد. / برای اندازه گیری فرایندهای پیچیده ذهنی از موقعیتهای تازه استفاده کنید. / گزینه های انحرافی توجه آزمون شوندگان بی اطلاع از موضوع را به خود جلب کند. / جملات از نظر دستوری و جمله بندی به نحو درست مکمل متن سوال باشد / از به کار بردن اشاره های دستوری نامربوط خودداری کنید / در سؤالات منفی کلمات منفی را برجسته جلوه دهید. / از منفی مضاعف که هم متن سؤال و هم گزینه ها منفی باشد احتراز کنید / از عبارتهایی مانند (همه آنچه در بالا گفته شده با تمام موارد) خودداری کنید / تا حد امکان از گزینه «هیچیک از موارد» پرهیز کنید / سؤالها را مستقل از یکدیگر بنویسید / از طرح سؤالهای گمراه کننده بپرهیزید / دو گزینه متضاد که یکی از آنها درست است به کار نبرید / طول گزینه درست را در سؤالهای مختلف تغییر دهید / محل گزینه درست را در میان گزینه های انحرافی به طور تصادفی انتخاب کنید / برای هر متن سؤال بین ۳ تا ۵ گزینه در نظر بگیرید / در آزمونهای تکوینی از گزینه نمی دانم استفاده کنید. / در سؤالهایی که تنه سؤال ناتمام است، جای خالی را در قسمت آخر جمله قرار دهید / تا جایی که ممکن است در تنه سؤال به جای جمله ناتمام از جمله پرسشی استفاده کنید.

اعمال نمره منفی در آزمونهای عینی (چندگزینه‌ای)

$$S = R - \frac{W}{n - 1}$$

S	نمره واقعی
R	نمره خام، تعداد پاسخهای درست
W	تعداد پاسخهای غلط
n	تعداد گزینه‌های سؤالها

۲- سوال های غیر عینی (ذهنی ، پاسخ ساز)

♦ سوال های کوتاه پاسخ

♦ سوال های تشریحی (محدود پاسخ - گسترده پاسخ)

سوال های کوتاه پاسخ

در سؤالهای کوتاه پاسخ، کلمه، عبارت، جمله، عدد یا علامتی خواسته می شود. این سؤالها غالبا برای سنجش هدفهای آموزشی سطح پایین دانش و معلومات طرح می شوند.

انواع سؤالهای کوتاه پاسخ:

- نوع پرسشی (مؤسس حکومت ساسانیان چه کسی بود؟.....)
- نوع کامل کردنی (دیوار چین در دوره ساخته شد.)
- نوع تشخیصی یا تداعی (بنیانگذاران حکومت های زیر را بنویسید.
سفاریان آل بویه غزنویان طاهریان

محاسن سؤالهای کوتاه پاسخ:

۱. سؤالهای کوتاه پاسخ از نظر تهیه سؤال و تصحیح سهولت بیشتری دارد. ۲. اجرای آنها ساده تر است.
۳. نسبت به سؤال های عینی ثقل را کاهش می دهند. ۴. نسبت به سؤال های عینی چندگزینه ای، صحیح غلط و جور کردنی اطلاعات تشخیصی بیشتری را در اختیار معلمان می گذارد. ۵. احتمال حدس کورکورانه را کاهش میدهد.

معایب سؤالهای کوتاه پاسخ:

- ۱- برای سنجش هدفهای سطح بالا قابل استفاده نمی باشد. ۲- استفاده زیاد از این سؤالهای، دانش آموزان را به حفظ کردن تشویق می کند. ۳- تصحیح و نمره گذاری آنها به دقت و سرعت آزمونهای عینی نمی باشد.

قواعد تهیه سؤال های کوتاه پاسخ

- ۱- هر سؤال باید موضوع مهمی را شامل شود. ۲- صورت سؤال را طوری بنویسید که به پاسخ مشخص و واحدی نیاز داشته باشد. ۳- در سؤالهایی که پاسخ آنها اعداد هستند به واحد مقیاس و میزان دقتی که در محاسبات باید داشته باشند دقت کنید. ۴- کلمات و عبارات مهم را در سؤالهای کامل کردنی حذف کنید.

۵- در سؤال های کامل کردنی تعداد زیادی جای خالی قرار ندهید. سایر ۶- جای خالی سؤالهای کامل کردنی را تا حد امکان در قسمت آخر سؤال قرار دهید. ۷- تا حد امکان به جای سؤال های کامل کردنی از سؤالهای پرسشی استفاده کنید. ۸- در سؤالهای کامل کردنی از کاربرد اشارات دستوری و مواردی که جواب سؤال را مشخص می کند خودداری کنید. (مثلا: شاه اسماعیل و بنیانگذار چه حکومتی بود؟)

سایر نکات

هر سؤال را با توجه به یک موضوع مهم بنویسید / به صورت سؤال را کاملا روشن بنویسید به گونه ای که به پاسخ معینی نیاز داشته باشد / صورت سؤال را عینا از روی مطالب کتاب ننویسید / در سؤالهایی که پاسخ های آنها اعداد هستند واحد مقیاس و میزان دقتی را که در محاسبات باید رعایت شوند مشخص کنید / در سؤالهای کوتاه پاسخ کامل کردنی تنها کلمه ها و عبارت های مهم را حذف کنید / در سؤالهای کامل کردنی که برای ارزشیابی پیرفت تحصیلی به کار می روند تعداد زیادی جای خالی منظور نکنید / جای خالی سؤالهای کامل کردنی را تا آنجا که ممکن است در قسمت پایین سؤال قرار دهید / جای خالی سؤالهای کامل کردنی را همواره یک اندازه تعیین کنید / در سؤالهای کامل کردنی از کاربرد اشاره های دستوری و موارد دیگری که جواب سؤال را مشخص می کنند خودداری کنید / تا حد امکان به جای سؤال های کامل کردنی از سؤال های پرسشی استفاده کنید

قواعد تهیه سؤالهای کوتاه پاسخ ویژه مسایل عددی:

تاحدامکان از اعداد ساده استفاده کنید / تا حد امکان مسایلی را طرح کنید که دارای جواب صحیح باشد / میزان دقت مورد انتظار را تعیین کنید / اگر لازم است آزمون شوندگان در پاسخ خود واحد اندازه گیری را نشان دهند / در صورت امکان مسایل پیچیده و چند مرحله ای را به تعداد مسئله یک مرحله ای ساده تقسیم کنید / تا آنجا که امکان دارد مسائل عددی را به زبان ساده و به طور مختصر بیان کنید

سوال های تشریحی (محدود پاسخ - گسترده پاسخ)

سوال های تشریحی گسترده پاسخ : در این آزمونها هیچ گونه محدودیتی برای آزمون شونده وجود ندارد و عملا آزاد است تا پاسخ خود را بپروراند و سازمان دهد و از لحاظ زمان پاسخ دهی و مقدار پاسخ نیز آزادی کامل دارد.

این سؤالها بیشتر برای سطوح بالای شناختی، مانند ترکیب و ارزشیابی

سوال های تشریحی محدود پاسخ : در این آزمونهای آزمون شونده در دادن پاسخ به سؤاها آزادی کامل ندارد بلکه صورت سؤال او را ملزم می سازد تا پاسخ خود را در چارچوب خاصی محدود کند و از لحاظ زمان پاسخ دهی و مقدار پاسخ محدودیت هایی وجود دارد. این سؤاها برای سطوح فهمیدن و کار بستن و تحلیل مناسب است.

محاسن سؤاها تشریحی:

۱. طراحی سؤاها تشریحی از آزمونهای عینی آسان تر است. ۲. این نوع سؤال تنها وسیله موجود برای سنجش توانایی آزمون شونده در پروراندن جواب سؤاها و بیان آنهاست. ۳. این آزمونها توانایی پاسخ دادن به سؤال ها را می سنجند نه توانایی انتخاب پاسخها را.

معایب سؤاها تشریحی

۱- این آزمونها نمونه کوچکی از محتوای درس و هدفهای آموزش را اندازه می گیرند. ۲- تصحیح برگه های امتحانی این آزمونها نمیتواند با دقت و به طور عینی انجام شود. ۳- تصحیح برگه های امتحانی این آزمونها بسیار وقت گیر است.

قواعد تهیه سؤال های تشریحی:

۱- سؤاها تشریحی را تنها به اندازه گیری هدفهایی محدود کنید که با سایر انواع آزمونها به خوبی قابل اندازه گیری نیستند. ۲- صورت سؤاها تشریحی را با عبارات و کلمات واضح و روشن بنویسید و از کلی گویی و ابهام بپرهیزید. ۳- از به کاربردن کلمات چه کسی، چه وقت، کجا و جز اینها بپرهیزید و از کلماتی مثل چرا، چگونه و به چه دلیل استفاده کنید. (چون سؤاها تشریحی در سطح دانش نیست و دیگر اینکه به سمت کوتاه پاسخ می رود و جواب مشخص میشود) ۴- سؤاها مربوط به موضوعات و عقاید بحث انگیز باید طوری طرح شوند که از آزمون شونده بخواهند شواهد لازم برای مستند کردن عقیده انتخابی را بیان کند، نه اینکه از او بخواهند تا صرفا عقاید شخصی خود را شرح دهد. ۵- به آزمون شندگان حق انتخاب چند سؤال از میان تعدادی سؤال ندهید. ۶- برای پاسخ به سؤاها زمان کافی در نظر گرفته، زمان و بارم هر سؤال را مشخص کنید. ۷- با نوشتن سؤاها تشریحی که به جوابهای کوتاه تری احتیاج دارند تعداد سؤاها را افزایش دهید و از مشکل ضعف نمونه گیری سؤاها تشریحی بکاهید. (سؤاها گسترده پاسخ قسمت کمی را در برمی گیرند)

قواعد تصحیح سؤالهای تشریحی:

۱- پاسخ سؤالها را تنها بر اساس هدفی که در سؤال گنجانده شده، تصحیح کنیم ۲- با نوشتن یک پاسخ نمونه برای هر سؤال به عنوان کلید 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- تمام پاسخهای آزمون شوندگان به یک سؤال را در یک نشست و بدون وقفه زمانی تصحیح کند. ۶- بر روی برگه های آزمون اشتباهات دانش آموزان را تصحیح کرده و اظهارنظرهای خود را بنویسید.

سایر نکات:

دقت کنید که سؤالها به طور مستقیم به هدف های آموزشی مربوط شوند / استفاده از سؤالهای تشریحی را تنها به اندازه گیری هدف های محدود کنید که با سایر انواع آزمون ها به خوبی قابل اندازه گیری نیستند / صورت سؤالهای تشریحی را با عبارات و کلمات واضح و روشن بنویسید / از کاربرد کلمات چه کسی چه وقت کجا و جز اینها بپرهیزید / تا حد امکان از سؤالهای تازه و موقعیت های جدید استفاده کنید / سؤالهای مربوط به موضوعات و عقاید بحث برانگیز را طوری طرح کنید که آزمون شوندگان از مستندات استفاده کنند نه اینکه از عقاید شخصی خود استفاده کنند / به آزمون شوندگان حق انتخاب چند سؤال از میان تعدادی سؤال را ندهید / برای پاسخ دادن به سؤالها زمان کافی در نظر بگیرید با نوشتن سؤالهایی که به جواب کوتاه نیاز دارد تعداد سؤالها را افزایش دهید تا از مشکل ضعف نمونه گیری آزمون های تشریحی کم کنید / عواملی را که در ارزشیابی آزمون های تشریحی دخالت خواهید داد از پیش تعیین کنید و آنها را به اطلاع آزمون شوندگان برسانید / از روش کار و شش برای تکمیل آزمون های تشریحی استفاده کنید

قواعد تصحیح سؤالهای تشریحی:

پاسخ های سؤالهای تشریحی را تنها بر اساس هدفی که در سؤال گنجانیده شده است تصحیح کنید / پاسخ ها را سؤال به سؤال تصحیح کنید / هنگام تصحیح برگه های امتحانی از شناسایی نام صاحبان آنها خودداری کنید / در صورت امکان از یکی دو نفر از همکارانتان بخواهید تا سؤالهایی را که تصحیح کرده اید تصحیح کنند / تمام پاسخ های آزمون شوندگان به یک سؤال را در یک نشست و بدون وقفه زمانی تصحیح کنید / به نمرات سؤالهایی که قبلا تصحیح کرده اید نگاه نکنید / بر روی برگه های آزمون اشتباهات دانش آموزان را تصحیح کنید / با نوشتن یک نمونه با الگو برای هر سؤال به عنوان کلید از دخالت عوامل نامربوط جلوگیری کنید / از راهنمای نمره گذاری استفاده کنید

ترتیب سؤالها در یک آزمون

اگر بخواهیم سؤالهایی از انواع مختلف را در یک آزمون پیشرفت تحصیلی قرار دهیم، برای آنکه به آزمون شونده نظم فکری دهیم و پاسخدهی را آسان کنیم، ترتیب زیر پیشنهاد میشود:

۱. سؤالهای صحیح غلط

۲. سؤالهای جورکردنی

۳. سؤالهای چندگزینه ای

۴. سؤالهای کوتاه پاسخ

۵. سؤالهای تشریحی

علاوه بر این، ترتیب سؤالها را بر حسب موارد زیر نیز ما می توان دسته بندی کرد:

۱. سؤالهای مربوط به هر طبقه از هدف های آموزشی مانند: اصطلاحات، کاربرد اصول و... را به دنبال هم قرار دهیم.
۲. سؤالها را از ساده به دشوار مرتب کنیم.
۳. در تنظیم سؤالهای یک آزمون سعی کنیم سازمان اصلی مطالب درسی حفظ شود. ترکیب این اولویتها نیز مقدور و مفید است.

جلسه هشتم

موضوع این جلسه ارزشیابی برای و ارزشیابی به عنوان، است

ارزشیابی یادگیری، مقاصد ارزشیابی :

- پایش پیشرفت دانش آموزان و مستندسازی نتایج یادگیری آنان، نه تنها تمرکز بر دانش و مهارت، که طیفی از صفات و نگرشها اهمیت دارند. این ارزشیابی باید مسؤولانه و دارای معیارهایی شفاف باشد.
- حمایت از یادگیری و پیشرفت دانش آموز از طریق برنامه آموزشی؛ شکلی از داربست حمایتی و قادر شدن دانش آموزان به سمت درکی پیچیده تر. (نظریه ویگوتسکی)
- تصمیم گیری در مورد شیوه آموزش و بهبود برنامه ها
- گزارش به مراجع بالاتر یا سازمانهای بیرونی

ویژگیهای ارزیابی جدید:

- جمع آوری مداوم دادهها
- تلفیق ارزیابی با آموزش
- استفاده از طیف وسیعی از ابزارها (مشاهده، تکالیف عملکردی، پرسشهای کلاسی، پرسشهای تحلیلی متمرکز، پرسشهای باز پاسخ خودارزیابی، ارزیابی همسالان، یادداشت روزانه، پوشه کار و پروژه، آزمونهای قلم کاغذی و شفاهی، روبریکها و نقشه ها)
- توجه به مستندسازی به کمک ابزارهای بالا
- توجه به خودارزیابی دانش آموزان و ارزشیابی گروهی و چندبعدی. استفاده از شواهد متعدد
- گزارش دهی به صورتهای مختلف (کمی، کیفی، توصیفی و نتایج یادگیری)

- دادن مداوم بازخورد به دانش آموزان و مستندسازی سوابق آنان
 - ارزیابی از جنبه های پیچیده تر (مانند نگرش نسبت به درس، حل مساله و تفکر انتقادی، و به کارگیری در زندگی) همراه با جنبه های دانشی و مهارتی اولیه
- پرسشهای کلاسی :**

- لزوم مهارت یافتن در طرح پرسش به جای آموزش مستقیم و با هدف درگیر کردن دانش آموزان در تدریس
- پاسخ دانش آموزان به این پرسشها در برنامه ریزی معلم مفید است.
- توجه ویژه به «چه»، «چگونه» و «چرا»
- لزوم وجود طیفی از پرسشهای بسته و سطح پایین تا باز و سطح بالاتر
- حتی پرسشهای «بله خیر» برای کنترل جریان درس مفیدند.
- پرسشهای توجیه تفکر یا عمل خود، برای توجه به راه حل های مختلف یا تحلیلهای برگرفته از ذهن خود دانش آموزان، لزوم صبوری معلم در پاسخ این پرسشها
- توجه به طبقه بندی بلوم و سطوح بالای آن و تشویق تأمل، توجیه، استدلال و ... " چرا این درست است؟
- پرسشهای باز معمولاً راه حل آشکاری ندارند و پاسخهای متعددی دارند و با ایجاد تعامل، آموزگار نیز دسترسی بیشتری به عمق درک دانش آموزان می یابد.

ارزیابی تدریس خود :

چگونه متوجه شوم آیا در آموزش موفق بوده ام یا خیر؟

دو شکل ارزیابی آموزش:

- ارزیابی مداوم در طول آموزش از نوع تکوینی، شامل طرح پرسش، گوش دادن و پاسخ به دانش آموزان، و مشاهده عملکرد و نوشته ها و رفتار دانش آموزان در طول درس است.
- ارزیابی تراکمی در پایان تدریس با تأمل در تدریس، توجه به مواردی که مؤثر بوده یا نبوده و اینکه چه باید کرد و تدریس بعدی چه تغییری خواهد کرد.

عادتهای خوب:

- نوشتن نکات کوتاه تأمل بر کار خود در هر درس
- تحلیل عملکردها و شیوه های خود در فاصله های زمانی مختلف
- عادت به توجه مداوم به این که آیا تمام دانش آموزان در تمام فعالیتهای تدریس درگیر شده اند یا خیر؟

مواردی که در فرایند سنجش و ارزشیابی بااهمیت است.	مواردی که در فرایند سنجش و ارزشیابی کم اهمیت است.
آنچه ارزش اندازه گیری دارد.	آنچه به سادگی می شود اندازه گرفت.
دانش پایه و عمیق که به خوبی سازماندهی شده باشد.	دانش حافظه ای
توانایی تفکر، درک و استدلال علمی	دانستی های محض
تعیین حدودی که دانش آموزان درک کرده اند.	تعیین آنچه دانش آموزان نمی دانند.
توجه هم به فرایند و فرصت های یادگیری و هم به نتیجه ی نهایی	توجه صرف به نتیجه ی کار
ارزشیابی مستمر که دانش آموز را در فعالیت های روزمره و در فرایند آموزش می سنجد و نتایج حاصل مبنای طرح درس مرحله ی بعد می شود.	ارزشیابی پایان ماه، ترم یا پایان سال
ایجاد فرصت های مختلف برای معلم در انجام ارزشیابی های معتبر و متنوع	ارزشیابی هماهنگ منطقه ای یا استانی
استمرار ارزشیابی در فرایند آموزش	ارزشیابی انفرادی پایان دوره یا ترم
سنجش توان کارگروهی و انفرادی	سنجش توان کارهای انفرادی

جدول ۱: مقایسه ی موارد بااهمیت و کم اهمیت در سنجش و ارزشیابی

نکته: فراموش نکنیم که یکی از هدف های مهم آموزش، به خصوص در دوره های آموزش عمومی، آموختن برای زیستن، برای در جمع زیستن و برای بهتر زیستن است. امروزه دیدگاهی که دوران یادگیری را از بقیه ی عمر جدا می سازد، مطرود است. اگر آموزش در دوران تحصیل نتواند دانش آموز را در جهت حصول به ابزاری برای زندگی بهتر هدایت کند، در اعتبار آن آموزش باید شک کرد، زیرا یادگیری فرایندی است به اندازه ی تمام زندگی، هم در طول و هم در گوناگونی اش (پیام یونسکو)

اعتقاد داریم که فقط هر شبوهی ارزشیابی که با هدف فوق سازگار باشد، معتبر است، وگرنه مطرود است و باید اصلاح شود.

ارزشیابی معتبر به همراه آموزش معتبر امکان پذیر است بنابراین:

- آموزش باید با پژوهش همراه باشد
- آموزش باید سعی در تلفیق دانش داشته باشد
- ارزشیابی، ارزشی بیش از موفقیت در جلسات امتحان داشته باشد
- زندگی گروهی فرد را اصل قرار دهد
- افراد را به یک فراگیر مادام العمر تبدیل کنند

نکته: هدف ارزشیابی خدمت به آموزش است، نه در کمین نشستن برای غافلگیر کردن دانش آموز و محک زدن او با معیار آنچه " نمی داند". فرهنگ موفقیت را می توان با کاربرد روش صحیح در آموزش و ارزشیابی بر فضای کلاس غالب کرد.

ویژگی های سنجش معتبر:

بر فراشناخت و خودارزیابی تأکید دارد / آزمون عملکردی معنادار است / آموخته های قابل انتقال را می سنجد / معیارهای استاندارد، معتبر و شفاف برای موفقیت دارد / تعامل مثبت بین معلم و دانش آموز برقرار میکند / با تولید یک محصول و عملکرد کیفی همراه است

سنجش اصیل یا واقعی

روشهایی وجود دارد که به کمک آنها می توان یادگیری های روانی - حرکتی و عاطفی و همچنین فرایندها و فرآورده های فکری سطح بالای شناختی را نیز سنجید.

این روشها غالباً در موقعیتهای واقعی یا شبه واقعی انجام می شود: "سنجش واقعی یا اصیل مانند سنجش عملکردی"

سنجش عملکردی

روشی که در آن فرایندها و فرآورده های یادگیری دانش آموزان به طور مستقیم سنجیده می شود.

علت استقبال از این آزمونها:

- پیشرفتهای روانشناسی شناختی، و مطرح شدن موضوعاتی مثل تفکر انتقادی و خلاق
- اهمیت یافتن مهارتهای فکری پیچیده، خودنظم دهی
- رابطه بین انگیزش و بروز رفتار
- اعتراضات شدید نسبت به آزمونهای عینی

سنجش فر آیند و فر آورده

آزمون عملکردی: سنجش فرایند (نحوه انجام)، فرآورده (محصول تولیدشده) یا هر دو

ماهیت عملکرد مورد سنجش تعیین کننده نوع سنجش است

مثال: کار با وسایل آزمایشگاهی / سخنرانی کردن / انجام فعالیتهای ورزشی (شنا کردن، شوت کردن) "فرآیند"
داستان نویسی / انشا / شعر / نقاشی / "فرآورده" عیب یابی و تعمیر یک وسیله برقی. "هر دو"

نکته: سنجش عملکردی با مهارت سروکار دارد. مهارت در استفاده از فرایندها و شیوه های اجرایی و مهارت در تولید فرآورده ها.

در بیشتر دروس می توان از این آزمونها در کنار آزمونهای عینی بهره جسته

۳۴

انواع سنجش عملکردی (نیتکو)

تأکید بر کار بست: سنجش اینکه آیا دانش آموز علاوه بر دانستن دانش، می تواند آن را به کار بندد.

تأکید بر سنجش مستقیم: مشاهده مستقیم رفتار دانش آموز در یک هدف آموزشی

استفاده از مسائل واقعی: استفاده از مسائل و موقعیتهایی که در زندگی واقعی یافت می شوند یا شبیه آنها هستند

ترغیب و تشویق تفکر باز: هدایت دانش آموزان به یافتن راه حل‌های مختلف برای مسائل و تشویق آنان به کار کردن با هم.

انواع سنجش عملکردی (از نظر گرانلند)

• آزمون کتبی عملکردی

• آزمون شناسایی

• موقعیتهای شبیه سازی شده

• نمونه کار

آزمون کتبی عملکردی

در این نوع، موقعیت یادگیری به صورت کتبی سنجش می شود.

تفاوت با آزمون کتبی: این آزمونها بیشتر بر کاربرت دانش و مهارت در موقعیت شبیه سازی شده تاکید می شود. مثال:

کتبی تشریحی: ویژگیهای جدول مشخصات درس را تشریح کنید.

کتبی عملکردی: برای کتاب دین و زندگی پایه دهم، جدول مشخصات درس تهیه کنید.

• این آزمونها برای بازده های پایانی یادگیری یا مراحل میانی عملکردی که برای رسیدن به بازده های پایانی ضرورت دارند تأکید می شود.

• این آزمونها معمولاً مقدماتی هستند...

آزمون شناسایی

این آزمونها روشی برای سنجش تواناییهای یادگیرنده در تشخیص ویژگیهای محاسن و معایب و موارد استفاده امور مختلف است. مانند:

• شناسایی یک ابزار و توصیف آن

• مواجهه دانش آموز با یک مشکل و شناسایی ابزار و روشهای حل

• گوش دادن به صدای موتور اتومبیل و شناسایی علتها

• شناسایی سنگهای مختلف

• تشخیص تلفظ درست کلمات در زبان انگلیسی

• نمونه خیلی بارز در آموزش پزشکی: آزمون آسکی (توصیف شرح حال بیمار)

عملکرد در موقعیت شبیه سازی شده (سنجش قیاس پذیر)

از دانش آموز خواسته می شود در یک موقعیت شبیه سازی شده یا تصنعی همان اعمالی را انجام دهد که در موقعیتهای واقعی ضروری به نظر می رسد. مانند: انجام حرکت شنا بدون آب / نحوه معلمی کردن بدون دانش آموز / رانندگی بدون ماشین

• هدف از این آزمون: پیش بینی وقوع رفتارهای مورد سنجش در آینده است.

نمونه کار

از یادگیرنده خواسته می شود اعمالی را انجام دهد که معرف عملکرد واقعی مورد سنجش هستند. در روش نمونه کار، اعمالی که می خواهیم دانش آموز انجام دهد باید شامل عناصر مهم عملکرد کلی باشد. مثال: از دانشجوی تربیت معلم خواسته می شود. با استفاده از یک روش نام نویسی موضوع را به دانش آموزان آموزش بدهد

انواع روشهای سنجش عملکردی (نیتکو)

• سنجش ساختارمند (کنترل شده)

• سنجش در موقعیتهای طبیعی

• پروژه های طولانی مدت

• کارپوشه

سنجش ساختارمند

• در این نوع معلم هم بر تکالیف سنجش و هم بر شرایطی که در آن دانش آموز باید پاسخ دهد کنترل اعمال می کند.

• معلم تصمیم می گیرد که از چه موادی استفاده کند و چه زمانی آنها را مورد استفاده قرار دهد.

• معلم باید راهنمای عملکرد دانش آموز و نوع بازدههای یادگیری را تعیین کند.

• فنون سنجش ساختارمند: ۱- تکالیف مداد - کاغذی (آزمونهای تشریحی یا انشایی) ۲- تکالیف مبتنی بر تجهیزات و منابع.

سنجش در موقعیت طبیعی

در این نوع به جای اینکه تکلیفی برای دانش آموز تعیین شود و موقعیت سنجش زیر کنترل معلم قرار گیرد، معلم منتظر می ماند تا عملکرد در شرایط طبیعی رخ دهد و آنگاه سنجش صورت می گیرد.

به این روش، روش تکالیف عملکرد نوعی هم می گویند. زیرا رفتارهای معمولی یا نوعی دانش آموز سنجیده میشود. مثال: مشاهده نحوه عمل دانش آموز در زمین فوتبال .

پروژه های طولانی مدت

بر دو بخش است: پروژه فردی و پروژه گروهی

پروژه فردی یک فعالیت طولانی مدت گفته میشود که به نوعی محصول یا فراورده مانند یک الگوی محصول، وسیله یا مجموعه می انجامد. ویژگی این نوع آن است که چندین فعالیت مختلف ترکیب می شود تا یک محصول تولید شود

پروژه گروهی دو یا چند نفر به صورت مشارکتی در تولید یک محصول یا فراورده می پردازند. مانند تهیه روزنامه دیواری نوشتن مقاله ای درباره کهکشان

کارپوشه (پوشه کار = پورتفولیو)

- مجموعه محدودی از کارهای دانش آموز است که برای سنجش عملکرد او استفاده میشود.
- مهم ترین مزیت این روش، فراهم آوردن اطلاعات کافی از میزان و چگونگی پیشرفت دانش آموز است.
- پوشه کار، علاوه بر سنجش عملکرد، می تواند مهارتهایی مانند تفکر، تحلیل انتقادی و راهبردهای فراشناختی را در دانش آموزان تقویت کند.

تعریف : پوشه کار، مجموعه ای از نمونه کارهای دانش آموزان که در جریان یاددهی-یادگیری، به گونه ای هدفمند گردآوری شده، آگاهانه انتخاب گردیده و منظم سازماندهی شده است و بیانگر میزان تلاش، پیشرفت و موفقیت دانش آموزان است.

قواعد تهیه آزمونهای عملکردی

بازدههای مورد نظر عملکرد را مشخص کنید. / از تکالیف معنی دار استفاده کنید / موقعیت آزمون را واقع بینانه انتخاب کنید. / دستورالعمل روشن و واضح داشته باشید / راهنمای نمره گذاری (روبریک) داشته باشید (کلی و تحلیلی). / از روشهای مختلف سنجش عملکردی بهره بگیرید.

مزایا و محدودیتهای آزمونهای عملکردی

مزایا:

بوسیله آن می توان یادگیری هایی را سنجید که به روش مداد کاغذی نمی توان سنجید (دانستن چگونگی انجام) / هم فراورده و هم فرایند را می توان با آن سنجید. / مبتنی بر نظریه های یادگیری جدید مانند سازنده گرایی است و انسان گراست

محدودیتهای آنها :

کمبود پایایی با دقت نمره گذاری (دخاله ذهنیت) / صرف وقت زیاد

سنجش مشاهده ای

سنجش و نمره گذاری همه عملکردهای فرایندها و فرآوردهها) دانش آموزان مبتنی بر مشاهده است.

روشهای سنجش مشاهده ای :

• فهرست واریسی

• مقیاس درجه بندی

• واقعه نگاری

فهرست واریسی (چک لیست)

• مجموعه پرسشهایی که به معلم امکان می دهد تا حضور یا عدم حضور رفتارها، اندیشه ها، رویدادها یا کیفیتها را تعیین کند.

• معلم با استفاده از این ابزار، هم فرایند [شیوه انجام عمل] و هم فرآورده [محصول عمل] را سنجش نماید.

• در این روش فقط بود و نبود یا کمیت ویژگی مورد نظر ثبت می شود.

• عملکردها: سخن گفتن شرکت در بحث / اجرای یک آزمایش

• فرآورده ها: نقاشی / رسم / نقشه

ویژگیهای فهرست واریسی

• هم برای سنجش فرایند و هم سنجش فرآورده قابل استفاده است؛

• برای ارزشیابی فرایند یا فرآورده بر جنبه های قابل مشاهده تأکید می شود.

• ویژگیهای مطلوب (شرکت در بحث / همکاری / سازگاری) یا نامطلوب رفتار یا فرآورده مورد نظر

(پرخاشگری / انزوا) سنجش قرار می گیرد.

ویژگیهای فهرست واریسی خوب: نسبتا کوتاه است / هر ماده فهرست (سنجه = گویه) به طور کاملا روشن، موضوع مورد ارزشیابی را بیان می کند / تنها رفتارهای مهم در فهرست منظور می شوند / قابل استفاده در موقعیت های مختلف است؛ / موقعیتهایی را که مستلزم انجام فعالیت مرحله ای است را می تواند مورد سنجش قرار دهد (مانند مشاهده اجسام توسط میکروسکوپ)

مراحل تهیه فهرست واریسی

۱- نشانه های هریک از انتظارات تهیه شود.

۲- فعالیتها و شواهد تحقق هدف مورد نظر به صورت عبارات یا جملات کوتاه نوشته شود (سنجه = گویه)

۳- اگر هدف به صورت فرایندی حاصل می شود، ترتیب زمانی آنها رعایت شود (مانند مشاهده اجسام توسط میکروسکوپ)؛

۴- جدول مناسبی برای آن تهیه شود؛

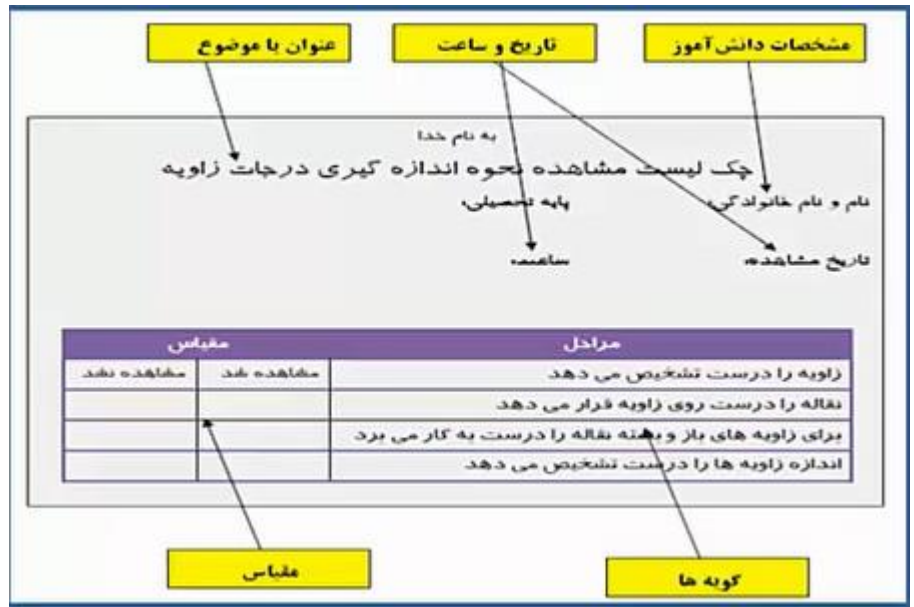
۵- یک بار فهرست واریسی را اجرا کنید [رفع نارساییها]

۶- از دوستان نقاد برای اصلاح نهایی فهرست کمک بگیرید.

اجزا و عناصر فهرست واریسی

۱- مشخصات دانش آموز ۲- تاریخ برای قضاوت مهم است. نشاندهنده پروسه زمانی ۳- عنوان با موضوع ۴- سنجه ها ۵- مقیاسها: مجموعه ای عدد یا نمادها که ویژگیهای آنها بر ویژگیهای تجربی اشیایی که اعداد به آنها نسبت داده میشود منطبق است. مانند: بلی خیر / مشاهده شد مشاهده نشد / دارد ندارد / هست نیست / زیاد تا حدودی کم.

نمونه فهرست واریسی



ارزشیابی فهرست وارسی

- ۱- نگاه کلی و اجمالی به پاسخ ها و پیوند بین این یافته ها با مشاهدات دیگر؛ یعنی اظهار نظر به صورت کیفی - توصیفی ۲- تجزیه و تحلیل کمی: کدگذاری مقیاسها (مثلا بصورت ۰، ۱ یا ۲) و محاسبه میانگین ۰. مثال: ۷ گویه با سه مقیاس کم، زیاد، تاحدودی (۰، ۱ و ۲)

مقیاس درجه بندی

- ۰ مقیاس درجه بندی در واقع همان فهرست وارسی است با این تفاوت که علاوه بر نشان دادن حضور یا عدم حضور یک ویژگی یا رفتار، میزان و سطح آن نیز معین می شود.
- ۰ هم بود یا نبود رفتار و هم فراوانی یا شدت و ضعف رفتار مورد بررسی و مشاهده قرار می گیرد.

انواع مقیاس درجه بندی

الف - مقیاس درجه بندی عددی (Numerical)

در این مقیاس، معلم دور هر عددی که درجه یا میزان مناسب بودن، توافق یا فراوانی با یک بیان را مشخص می کند با ترسیم خط بسته ای یا با گذاشتن علامت مشخص می کند. مثلا:

به چه میزان دانش آموز به درس ریاضی علاقه نشان می دهد؟

۵ ۴ ۳ ۲ ۱

۱=ضعیف، ۲=زیر متوسط، ۳=متوسط، ۴=بالاتر از متوسط، ۵=عالی

تعداد درجات: ۳ تا ۷ درجه گفته شده ولی بهترین همان بین ۴ تا ۵ درجه است.

ب - مقیاس درجه بندی نگاره ای (Graphic)

ج - مقیاس درجه بندی نگاره ای توصیفی (Descriptive Graphic)

بهترین نوع مقیاس برای مقاصد تحصیلی است. در این مقیاس عبارتها و توصیفهایی نگاشته می شوند که به شکل رفتاری، چگونگی عملکرد دانش آموز را نشان می دهد. مثلا:

۱- کیفیت کار دانش آموز را هنگام انجام تکالیف درسی چگونه ارزیابی می کنید؟

کار خوب است و نیازی به نظارت ندارد	کار خوب است و به ندرت به نظارت نیاز دارد	کار خوب است و گهگاه به نظارت نیاز دارد	کار قابل قبول اما احتیاج فراوان به نظارت دارد	کار ضعیف است است. غلطهای زیادی دارد
------------------------------------	--	--	---	-------------------------------------

نکته: مقیاس درجه بندی همان لیست واری است با طول و تفصیل.

قواعد تهیه مقیاس درجه بندی

ویژگیهای مهم آموزشی را انتخاب کند / ویژگیهای مورد ارزشیابی قابل مشاهده مستقیم باشد. / اصطلاحات و نقاط روی مقیاس را به روشنی بیان کند. / تعداد درجات را ۳ تا ۷ برگزیند.

واقعہ نگاری (رویداد نگاری)

- توصیفی واقعی از رویدادها و اتفاقات معنی داری که معلم در نتیجه مشاهده از زندگی دانش آموزان به دست می آورد.
- معلم باید رویدادی که در زندگی دانش آموز رخ می دهد و آن را مهم می داند، بلافاصله بعد از وقوع ثبت نماید، بنابراین این روش بیشتر جنبه تحلیلی دارد تا ارزشیابی.
- روش کیفی در مشاهده به شمار می رود.
- ابزاری برای ثبت رفتارهای دانش آموز در حیطه های عاطفی، اجتماعی، روانی حرکتی و عملکردی است.

• معلم رفتار یا رویداد مشخصی را به طور کامل گزارش می کند.

ویژگیهای رویداد نگاری

روش مشاهده مستقیم رفتار است؛ / بیشتر در حیطه های عاطفی و اجتماعی کاربرد دارد / بلافاصله حین یا پس از یک رویداد نوشته می شود؛ / اطلاعات را در موقعیتی واقعی نشان می دهد.

قواعد برنامه ریزی و اجرای رویداد نگاری

برای رویداد مورد نظر طرحریزی صورت گیرد؛ / رویدادها را پس از وقوع ثبت کنید، نیاز به ثبت تمامی جزئیات نیست فقط نکات مهم نوشته شود؛ / توصیف را با تفسیر نیامیزید؛ / هر رفتار باید دارای تاریخ، و مشخصات باشد؛

اجزا و عناصر رویدادنگاری

۱- مشخصات دانش آموز ۲- تاریخ، زمان و مکان رویداد(جغرافیای رفتار) ۳- عنوان یا موضوع ۴- توصیف عینی رویداد ۵- تفسیر و تحلیل رویداد ۶- ارائه راهکار

خودسنجی

زمانی که دانش آموز، عملکرد و فعالیتهای یادگیری خودش را ارزشیابی می کند، می گوئیم خود سنجی انجام گرفته است؛ در صورتی که در همسال سنجی دانش آموز به ارزشیابی از عملکرد و فعالیتهای یادگیری همکلاسی هایش می پردازد.

هدف اساسی: ایجاد احساس مسئولیت بیشتر نسبت به یادگیری خود و نیز خود آگاهی و افزایش حس مراقبت از فرایند یادگیری

نقش معلم در ارتباط با خودسنجی: هر چند که خودسنجی توسط دانش آموزان انجام می گیرد ولی آماده نمودن شرایط برای خودارزیابی دانش آموزان به عهده معلم خواهد بود؛ به عبارت دیگر معلم است که باید فرصت این خودارزیابی را فراهم آورد.

نکات مهمی در خودسنجی

۱. معلم در فرایند یاددهی یادگیری و در زمانهای مناسب، حتما باید فرصت خودسنجی را برای دانش آموزان فراهم کند.

۲. آگاهی دانش آموزان از معیارها و ملاکهای فعالیت مورد ارزشیابی به اجرای بهتر خودسنجی منجر می شود.

۳. معلم می تواند در تهیه معیارها و ملاکهای ارزشیابی، دانش آموزان را سهیم نماید. این عمل می تواند به گونه ای مناسب بر مسئولیت پذیری دانش آموزان بیافزاید.

۴. شرایط باید به گونه ای فراهم شود که دانش آموزان در جریان خودسنجی، در عملکردشان احساس آزادی کنند.

۵. چنانچه در جریان ارزشیابی از عملکرد فعالیت دانش آموز بازخورد توصیفی به دانش آموز باد داده شود تاثیر موثر تری خواهد داشت.

روشها:

۱. استفاده از فهرست وارسی: در این فهرست دانش آموز به کمک معلم خود انتظاراتی را که باید در جریان ارزشیابی از خود برابر بسازد، فهرست و مرتب می کند.

۲. استفاده از جملات ناقص: در این روش معلم به دانش آموز می آموزد که در قالب یک سری از جملات ناقص که کامل کردن آن به عهده ی خود دانش آموز است به ارزشیابی از خود با همکلاسی اش بپردازد.

۳. استفاده از توصیف کامل عملکرد: آنگونه که از نام روش پیداست معلم انتظار دارد دانش آموز به کمک بیان عباراتی کامل، ارزشیابی از فعالیت و عملکرد یادگیری خود را بیان نماید.

گروه سنجی

• اندازه گیری الگوی تعامل در گروهها

• با گروه سنجی می توان پذیرش اجتماعی فرد دانش آموزان و ساختار اجتماعی گروه را سنجش کرد.

• نظریه میدان روانی (لوین): لوین معتقد است که هر گروهی در هر لحظه از زمان دارای یک "میدان روانی" خاصی است که بی شباهت به میدان الکتریکی در علم فیزیک نیست.

• این موضوع حائز اهمیت است که همه عناصر و موقعیتهای موجود گروه در وضع موجودش، به عنوان یک کل مورد نظر قرار می گیرد و این خود همچون یک روش، به تجزیه و تحلیل نیروهای گروه پرداخته و قوانین حاکم بر آن را کشف می کند.

روش های گروه سنجی

۱- روش نگاره گروه

۲- روش حدس بزن چه کسی

روش نگاره گروه

تعریف: یک بازنمایی نموداری از روابط اجتماعی درون گروه یا ساختار گروه است که با بررسی آن می توان میزان تعامل یا نفوذ افراد را بررسی کرد.

• این روش مبتنی بر پاسخگویی به برخی سوالات است که گاه به صورت محرمانه مطرح می شود مانند: کسانی را که دوست دارید با آنها همکاری کنید را با توضیح نام ببرید. / با چه کسانی دوست ندارید کار کنید؟ چرا؟

• بعد از پاسخگویی به این سؤاها باید نمودار آن را ترسیم کرد. انتخابهای دوجانبه با فلش دوسویه و انتخابهای یک جانبه با فلش یکسویه. دوست نداشتنها با نقطه چین.

روش حدس بزن چه کسی

تعریف: روشی است برای سنجش روابط عاطفی میان اعضای یک گروه کلاسی مبتنی بر یافتن سرنخهایی از عاداتها، علایق دوستیها، سرگرمی ها و مشکلات دانش آموزان.

در این روش به دانش آموزان تعدادی جمله توصیف کننده رفتار داده می شود و از آنان خواسته می شود تا دانش آموزانی که دارای آن رفتارها یا ویژگی ها هستند را نام ببرند. مانند: این شخص خواندن را دوست دارد. / این شخص همیشه نگران است. / این شخص دوستان زیادی دارد.

جلسه دهم

شرایط اجرای آزمونهای کتبی

- نخستین گام در اجرای آزمونهای پیشرفت تحصیلی، آماده سازی و فراهم آوردن شرایط مناسب برای اجرا
- تهیه دفترچه آزمون و پاسخنامه، ترتیب سؤالات، در نظر گرفتن زمان و مکان مناسب برای اجرا

• فراهم آوردن حداکثر امکانات برای نشان دادن بهترین عملکرد

ترتیب سؤالات در یک آزمون

اگر بخواهیم سؤال های مختلفی را در آزمون پیشرفت تحصیلی قرار دهیم، برای آنکه به آزمون شونده نظم فکری دهیم و پاسخ دهی را آسان کنیم ترتیب زیر پیشنهاد میشود: ۱. سؤالهای صحیح غلط ۲. سؤالهای جورکردنی ۳. سؤالهای چندگزینه ای ۴. سؤال های کوتاه پاسخ ۵. سؤال های تشریحی

همچنین:

۱. سؤال های مربوط به هر طبقه از هدفهای آموزشی مانند: اصطلاحات، کاربرد اصول و... را به دنبال هم قرار دهیم.

۲. سؤال ها را از ساده به دشوار مرتب کنیم.

۳. در تنظیم سؤالهای یک آزمون سعی کنیم سازمان اصلی مطالب درسی حفظ شود.

۴. خوش خطی، خوانایی، فاصله مناسب برای جواب ها، قرار دادن مناسب شکل ها و نمودارها

دستورالعمل (راهنمای آزمون باید مختصر و مفید و آشکارا به دانش آموز بگوید که چه کاری باید انجام دهد و باید شامل اطلاعات ذیل باشد:

: زمان لازم برای هر قسمت - ارزش هر سؤال : مجاز یا غیرمجاز بودن حدس

دستورالعمل (راهنمای) آزمون

• باید مختصر، مفید و آشکارا به به دانش آموز بگوید که چه کاری باید انجام دهد و باید شاما اطلاعات ذیل

باشد: زمان لازم برای هر قسمت / ارزش هر سوال / مجاز یا غیر مجاز بودن حدس (نمره منفی)

• اگر افراد نیاز به توضیح شفاهی داشته باشند، یعنی راهنما روشن نیست و باید بازنویسی شود.

• راهنمایی کلی برای تمامی آزمون و راهنمایی برای هر بخش

• در مسائل عددی میزان دقت مشخص شود و واحد هم بیان شود.

• اگر خوش خطی، املاي کلمات و جمله بندی امتیاز دارد، مشخص شود.

• نحوه نمره گذاری پاسخ ها مشخص شود.

اجرای آزمون

• شرایط به گونه ای باشد که از هر لحاظ آسایش جسمی و فکری افراد فراهم شود.

• شرایط روانی بسیار مهم است. عامل اضطراب مهم است (میزان).

نمره گذاری

پاسخنامه کلید: یک پاسخنامه که در آن محل گزینه های صحیح سوراخ شده است و با قرار گرفتن روی

پاسخنامه ها به فوریت می توان آنها را تصحیح کرد.

دو روش عمده نمره گذاری:

۱- محاسبه کلیه پاسخهای درست بدون کسر نمره حدس زدن

۲- کسر مقداری از نمره برای جبران حدس زدن

نمره گذاری

نمره اصلاح شد (برای جبران حدس زدن و اعمال نمره منفی):

فرمول:

$$R = \frac{W-1}{n}$$

R = تعداد پاسخ‌های صحیح

W = تعداد پاسخ‌های غلط

n = تعداد گزینه‌های هر سؤال

مثال:

در یک آزمون چهارگزینه‌ای، که دارای ۴۰ سؤال است، دانش‌آموزی ۳۸ سؤال پاسخ داده که از این تعداد ۵ سؤال نادرست و بقیه درست بوده است. نمره اصلاح شده او را حساب کنید.

$$33 = R$$

$$5 = W$$

$$4 = n$$

$$33 - \frac{5-1}{4} = 32$$

تحلیل سؤال‌های آزمون

هدف:

- واریسی تک تک سؤالها و تعیین میزان دقت و نارسایی‌های آنها
- تعیین نقاط قوت و ضعف یک آزمون و کیفیت همه سؤالهای آن
- تجدید نظر در آزمون و بهبود کیفیت سؤالها برای استفاده‌های بعدی

مراحل تحلیل سؤال

اطلاعات مورد نیاز برای تحلیل سؤال

پاسخهایی که دانش آموزان به هر سؤال داده اند. (تعیین اینکه در هر سؤال چند نفر گزینه درست، چند نفر پاسخ نادرست و چند نفر بدون پاسخ / تعیین گروه بالا و پایین (۲۵ تا ۳۳ درصد) و تعداد گروه بالا که به سؤال پاسخ درست داده اند تعداد گروه پایین که به سؤال پاسخ نادرست داده ان.

ضریب دشواری یا سهولت سؤال (P):

درصد کل آزمون شوندگان که به یک سؤال جواب درست داده اند.

$$P = \text{ضریب دشواری}$$

انتخابهای درست گروه پایین + انتخابهای درست گروه بالا

تعداد افراد گروه پایین + تعداد افراد گروه بالا

ضریب دشواری بین صفر و یک است. (هر چه به صفر نزدیکتر باشد ضریب دشواری بالا و سخت می باشد و هر چه به یک نزدیک تر باشد ضریب دشواری پایین تر و آسان می باشد)

مثال:

در یک آزمون با ۴۰ شرکت کننده، ۱۸ نفر از گروه بالا و ۱۲ نفر از گروه پایین

به سؤال ۱۰ جواب درست داده اند. ضریب دشواری را حساب کنید.

$$\frac{18 + 12}{20 + 20}$$

$$= 0,75$$

تفسیر ضریب دشواری

• اگر آزمون ما هنجاری باشد، انتظار بر این است که برای دانش آموزان مختلف نمرات متفاوتی به دست بدهد (پراکندگی = واریانس)

• واریانس یکی از راههای قضاوت درباره مفید بودن سؤالهاست ۱- واریانس سؤالها ۲- همبستگی بین سؤالها

• هر سؤالی که واریانس بزرگتری دارد و با سایر سؤال های آزمون همبستگی بیشتری دارد واریانس کل آزمون را بالا می برد. و مفید بودن آزمون را ثابت می کند

واریانس چیست؟

$$\text{واریانس سؤال} = P \times (1 - P)$$

$P =$ درصد کسانی که به سوال جواب درست داده اند. (ضریب دشواری)

مثال : اگر ضریب دشواری سؤالی ۰/۳۰ باشد، واریانس آن چقدر است؟

$$\text{واریانس سؤال} \approx 0.21 \approx 0.30 \times (1 - 0.30)$$

چه سوالی واریانس بیشتری دارد؟

سوالی بهتر است که ضریب دشواری آن از ۱ کمتر و از صفر بیشتر باشد و به ۰/۵ نزدیک باشد.

• ضریب دشواری بین ۰/۳ تا ۰/۷ حداکثر اطلاع را بین افراد نشان می دهد(نرمال)

ضریب تمیز سؤال (D):

میزان ایجاد تفاوت بین دو گروه قوی و ضعیف کلاس.

$$D = \text{ضریب تمیز}$$

انتخاب‌های درست گروه پایین – انتخاب‌های درست گروه بالا
تعداد افراد یک گروه (بالا یا پایین)

ضریب تمیز عددی بین ۱- و ۱+ است.

مثال

در یک آزمون با ۴۰ شرکت‌کننده، ۱۸ نفر از گروه بالا و ۱۲ نفر از گروه پایین به

سؤال ۱۰ جواب درست داده‌اند. ضریب تمیز را حساب کنید.

$$\frac{18-12}{20} = 0.30$$

همبستگی سؤال با کل آزمون

ضریب تمیز \approx ضریب همبستگی سؤال با کل آزمون

• اگر فرد در کل آزمون نمره بالایی بگیرد در یک سوال خاص بالایی بگیرد... ضریب تمیز سوال بالا

ضریب همبستگی دو رشته ای نقطه ای :

۱- اگر ضریب همبستگی سؤال با آزمون بیشتر باشد یعنی سؤال با کل آزمون توافق بیشتری دارد .

۲- اگر ضریب همبستگی سؤال با گل آزمون منفی باشد یعنی سؤال با آزمون ناهمخوانی دارد .

۳- اگر ضریب همبستگی سؤال با کل آزمون صفر باشد سرال از ضریب تمیز برخوردار نیست.

مزیت: از لحاظ آماری می توان آن را قضاوت کرد. اگر ضریب همبستگی $0/40$ باشد، 16 درصد از واریانس نمرات را به خود اختصاص می دهد.

نکته: علاوه بر ضریب همبستگی سؤال با کل آزمون، ضریب همبستگی هر جفت سؤال با 0 هم نیز مفید است. چون:

۱- با در دست داشتن تعداد سؤالهای آزمون و میانگین ضرایب همبستگی بین سوالها ضریب پایایی آزمون به دست می آید.

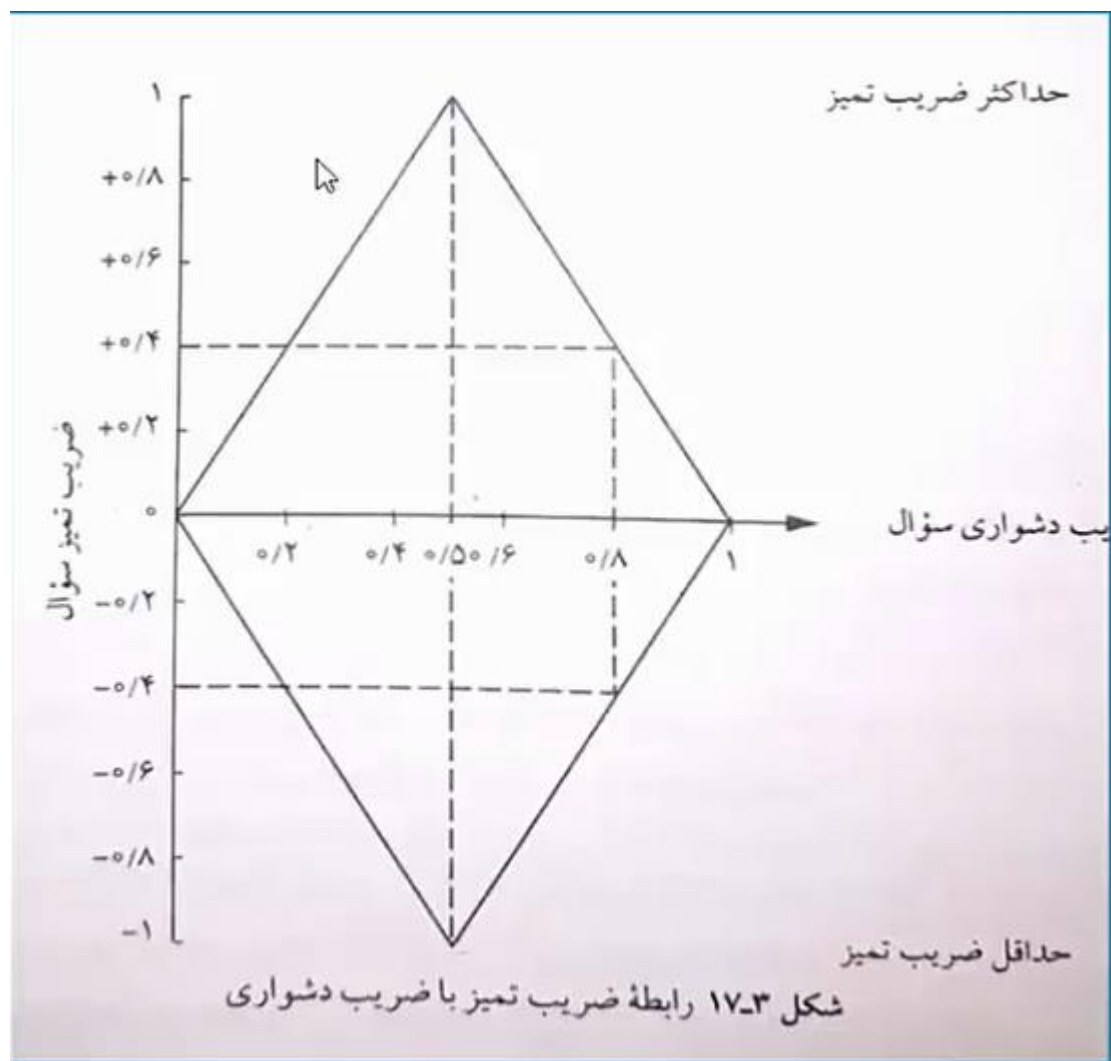
۲- با این محاسبه، ضریب تمیز به دست می آید

نکته: اگر همبستگی بین یک سوال و کل آزمون کم باشد دو دلیل وجود دارد:

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

دوم: ممکن است این سوال با بعضی سوال ها همبستگی زیاد داشته باشد و با برخی سوالات همبسته نباشد و موجب آن می شود سوالات آزمون چیز های مختلفی را اندازه می گیرند و پایایی آن پایین است.

رابطه بین ضریب دشواری و ضریب تمیز



تحلیل گزینه های انحرافی

- هدف گزینه های انحرافی: منحرف کردن افرادی که جواب سؤال را نمی دانند.
- این گزینه ها باید افراد گروه پایین را به خود جذب کند و این افراد آنها را انتخاب کنند.
- اگر از گروه قوی فرد یا افرادی را جذب کرد، افراد گروه پایین بیشتر باشد.

اصلاح و تجدید نظر سؤال

پس از تعیین ضریبهای دشواری و تمیز همه سوال های آزمون و بررسی گزینه های هر سؤال، باید به اصلاح سؤال ها و بازنویسی آنهايي که نیاز به تغییر دارند اقدام کرد. سؤالهایی که ضریب دشواری آنها بسیار بالا یا بسیار پایین است و نیز سؤالاتی که ضریب تمیز آنها خیلی کم است، باید مورد تجدید نظر قرار گیرند. همچنین گزینه های انحرافی معیوب نیز باید اصلاح یا تعویض شوند.

تحلیل سؤالهای تشریحی و آزمونهای عملکردی

روش ویتنی و سیبرز

مراحل: ۱- انتخاب دو گروه ۲۵ درصد بالا و پایین، مانند قبل

۲- محاسبه مجموع نمرات گروه بالا و پایین

۳- محاسبه ضریب دشواری و تمیز نتایج با فرمول های زیر

$$P = \text{ضریب دشواری}$$

$$P = \frac{\text{پایین ترین نمره ممکن سؤال} \times (\text{تعداد یک گروه} \times 2) - \text{مجموع نمره های گروه پایین} + \text{مجموع نمره های گروه بالا}}{\text{پایین ترین نمره ممکن سؤال} - \text{بالا ترین نمره ممکن سؤال}} \times \text{تعداد یک گروه} \times 2$$

$$D = \text{ضریب تمیز}$$

$$D = \frac{\text{مجموع نمره های گروه پایین} - \text{مجموع نمره های گروه بالا}}{\text{پایین ترین نمره ممکن سؤال} - \text{بالا ترین نمره ممکن سؤال}} \times \text{تعداد یک گروه}$$

روش نیتکو

مراحل روش نیتکو با مراحل روش ویتنی و سیبرز یکسان است و تفاوت در فرمول هاست

توجه: روش نیتکو در مواردی که کمترین نمره سوال برابر صفر باشد با روش ویتنی و سیبیل برابر خواهد بود
فرمولهای آن به شرح ذیل است:

$P =$ ضریب دشواری

میانگین نمرات سؤال در گروه پایین – میانگین نمرات سؤال در گروه بالا

پایین‌ترین نمره ممکن سؤال – بالاترین نمره ممکن سؤال

تحلیل سؤال های آزمونهای وابسته به ملاک

تمیز بین پیش آزمون و پس آزمون در یادگیری در حد تسلط

مراحل: ۱- اجرای آزمون پیش از آموزش و پس از آن

۲- محاسبه ضریب دشواری هر سوال برای دو آزمون

۳- محاسبه تفاوت ضرایب دشواری سوال های دو آزمون

تمیز بین پیش آزمون و پس آزمون در یادگیری در حد تسلط

یا

$S =$ ضریب حساسیت نسبت به آثار آموزش

تعداد پاسخ‌های درست پیش از آموزش – تعداد پاسخ‌های درست پس از آموزش

تعداد کل افراد

• ایده آل این است که ضریب تمیز پیش آزمون و پس آزمون برابر یک باشد، یعنی پیش از آموزش هیچ یک از افراد آن را پاسخ نداده باشند، و پس از آموزش تمام افراد پاسخ درست داده باشند.

جلسه یازدهم

آمار چیست؟

آمار چارچوب روشهای علمی را که در تجزیه و تحلیل دادههای مقداری به کار برده می شوند بنیانگذاری می کند.

روشهای آماری دارای دو وظیفه مهم هستند:

۱- به پژوهشگر در طبقه بندی ، خلاصه کردن، توصیف و تفسیر و برقراری ارتباط از طریق اطلاعات جمع آوری شده کمک می کند.

۲- به پژوهشگران امکان می دهد که با استفاده از اطلاعات جمع آوری شده از نمونه کوچکی از آزمودنیها ویژگیهای جامعه ای را که نمونه از آن انتخاب شده است برآورد یا استنباط کنند.

انواع روشهای آماری که برای وظایف اول و دوم به کار برده می شوند:

۱- **آمار توصیفی:** مجموعه روشهایی است که به خلاصه کردن، طبقه بندی ، توصیف و تفسیر داده ها می پردازد.

۲- **آمار استنباطی:** مجموعه روشهایی است که معمولاً برای بیان رابطه بین دو یا چند متغیر و تعمیم ویژگیهای نمونه آماری به جامعه آماری به کار برده می شوند.

نکته: اولین و مفیدترین قدم در سازمان دادن به داده ها ، مرتب کردن آنها بر اساس یک ملاک منطقی است.

نکته: با استفاده از روشهای آمار توصیفی می توان دقیقاً ویژگیهای یک دسته از اطلاعات را بیان کرد.

نکته: روشهای آمار توصیفی همیشه برای تعیین و بیان ویژگیهای اطلاعاتی به کار برده می شود که به وسیله پژوهشگران جمع آوری شده اند.

نکته: هدف نهایی آمار استنباطی برآورد ویژگیهای جامعه است.

علائم مربوط به پارامتر و آمار

ویژگی شاخص	میانگین	واریانس	انحراف استاندارد	نسبت	همبستگی	تعداد مشاهدات
آمار	\bar{X} یا M	S^2	s	P	r	n
پارامتر	μ	σ^2	σ	-		N

ویژگی‌هایی که پژوهشگران مشاهده و اندازه‌گیری می‌کنند متغیر نامیده می‌شود. واژه متغیر به ویژگی اطلاق می‌شود که بیش از یک ارزش به آن اختصاص داده می‌شود و تغییرات را از فردی به فردی یا از شیئی به شیئی دیگر نشان می‌دهد. مانند: قد، سن، وزن، بهره هوشی و...

• ویژگی‌هایی که در یک پژوهش به عنوان متغیر اندازه‌گیری یا مشاهده می‌شود ممکن است در پژوهش دیگر ثابت نگاه داشته شوند. واژه ثابت به ویژگی‌هایی که مقدار یا ارزش آنها در نزد افراد مختلف یکسان است اطلاق می‌شود. مثلاً در پژوهشی که در مورد دانش آموزان کلاس چهارم است، درجه تحصیلی ثابت است.

انواع متغیره

• الف: متغیرها از نظر ماهیت به دو دسته کمی و کیفی تقسیم می‌شوند:

۱- **متغیر کمی:** به متغیرهایی اطلاق می‌شوند که از نظر مقدار یا ارزش متفاوت هستند و به صورت عدد نوشته می‌شوند مانند سن، نمرات درسی و ...

۲- **متغیر کیفی:** هر متغیری که نتوان آن را بصورت عددی نمایش داد مانند: جنس، رنگ مو، مذهب.

• ب: انواع متغیر از نظر نقشی که در پژوهش دارد

۱- **متغیر مستقل:** متغیری است که بر متغیرهای دیگر اثر می‌گذارد متغیر پیشفرض است و از طریق آن متغیر وابسته اندازه‌گیری و تعیین می‌شود در تحقیق آزمایش متغیر مستقل متغیری است که توسط محقق دستکاری می‌شود تا تأثیرش بر متغیر وابسته مشخص شود

۲- **متغیر وابسته:** متغیری است که ارزش یا مقدار آن به متغیر مستقل بستگی دارد متغیر وابسته در اختیار محقق نیست و محقق نمی‌تواند در آن بخت و تصرف و دستکاری به عمل آورد

• ج: انواع متغیرها از نظر اینکه فاصله بین اعداد در نظر گرفته می‌شود یا خیر. داده‌ها دو نوع هستند:

۱- **داده‌های گسسته:** متغیری که فاصله بین اعداد را در نظر نمی‌گیرد و ارزشهای موجود بین دو مقدار دارای معنی نیست. مانند: تعداد دانشجویان، تعداد معلمان...

۲- داده های پیوسته: متغیری که هر ارزش یا مقداری (کسری - اعشاری) را می توان به آن اختصاص داد مانند: قد، وزن و...

نکته: در عمل تشخیص بین متغیر پیوسته و گسسته به صورت نظری امکان پذیر نیست. دلیل این امر فقدان وسایل اندازه گیری دقیق و مناسب است. در پژوهش غالباً متغیرهایی که ذاتاً پیوسته هستند به صورت گسسته مورد بحث قرار می گیرند مثلاً سن (پیوسته) به دلیل طبقه بندی کردن افراد به متغیر گسسته تبدیل می شود.

مقیاس های اندازه گیری:

اساس فعالیت در هر پژوهشی اندازه گیری است به طور کلی اندازه گیری عبارت است از نسبت عددی دادن به یک صفت یا حادثه بر اساس یک قانون معین.

- چهار نوع مقیاس اندازه گیری وجود دارد که به صورت سلسله مراتب هستند: ۱- اسمی ۲- ترتیبی ۳- فاصله ای ۴- نسبی

مقیاس اسمی

ساده ترین مقیاس اندازه گیری و سنجش میباشد که فقط به نام گذاری و طبقه بندی داده ها می پردازد مانند: جنس که به دو طبقه مرد (۱) و زن (۲) تقسیم می شود.

نکته: مقیاس اسمی باید کامل باشد به این معنی که باید در هر طبقه کلیه افراد هم نوع جایگزین شود.

مقیاس ترتیبی

این مقیاس علاوه بر اینکه داده ها را نامگذاری و طبقه بندی می کند به هر طبقه رتبه ای هم میدهد یعنی افراد یا اشیاء بر اساس ویژگیهای مورد اندازه گیری از بالاترین به پایین ترین مرتبه می شوند. در این مقیاس فاصله ها یکسان نیست.

مقیاس فاصله ای

این مقیاس علاوه بر طبقه بندی و نام گذاری دادهها، نسبت دو تفاضل یا دو فاصله را حفظ می کند. مانند درجه حرارت و ...

• در این مقیاس صفر مطلق وجود ندارد. مثلاً دانش آموزی که در یک آزمون هوش صفر می‌گیرد و به این معنی نیست که اصلاً هوشی ندارد.

مقیاس نسبی

این مقیاس کامل‌ترین نوع مقیاس سنجش می‌باشد و تمام خصوصیات مقیاس‌های قبلی را دارا است و می‌توان هر گونه عملیات و کاری با آن انجام داد. در این مقیاس صفر مطلق است به این معنی که صفر به معنی وجود نداشتن می‌باشد مثلاً اگر در آمد فردی صفر باشد یعنی هیچگونه درآمدی ندارد.

پژوهشگران غالباً با توده‌ای از اطلاعات که نیاز به تفسیر دارند، روبرو هستند که برای معنی بخشیدن به اطلاعات باید آنها را خلاصه و سازمان بندی کنند. یکی از کارآمدترین روشها برای خلاصه و سازمان بندی اطلاعات توزیع فراوانی می‌باشد.

توزیع‌های فراوانی و نمودارها

• پژوهشگران غالباً با توده‌ای از اطلاعات که نیاز به تفسیر دارد روبرو هستند که برای معنی بخشیدن به اطلاعات باید آنها را خلاصه و سازمان بندی کنند یکی از کارآمدترین روشها برای خلاصه و سازمان بندی اطلاعات توزیع فراوانی می‌باشد

توزیع فراوانی

عبارت است از سازمان دادن اندازه‌ها یا مشاهدات به صورت طبقات همراه با فراوانی هر طبقه. توزیع فراوانی داده‌ها را بصورت خلاصه و مرتب، به نحوی که تفسیر آنها آسان شود، نمایش می‌دهد.

مراحل ساخت جدول توزیع فراوانی

۱- مرتب کردن اعداد از کوچک به بزرگ یا برعکس.

۲- مشخص کردن تعداد دفعاتی که هر عدد تکرار شده است (تعداد فراوانی)

+ زمانی که همه اعداد تک تک در جدول آورده شوند جدول توزیع فراوانی منفرد یا طبقه بندی نشده گفته می‌شود. اما زمانی که نمرهها یا اعداد دارای دامنه گسترده‌ای هستند و تنظیم اعداد بصورت توزیع فراوانی طبقه بندی نشده وقت گیر و طاقت فرسا است، اعداد را طبقه بندی می‌کنیم و از جدول توزیع فراوانی طبقه بندی شده استفاده می‌کنیم.

نکته: در جدول فراوانی ، ستون دادهها (طبقات) را با X نشان می دهند.

نکته: فراوانی مطلق (f) برابر است با مقدار دفعات تکرار هر داده در هر طبقه.

مثال :

**در توزیع فراوانی درس ارزشیابی یک کلاس،
نمرات به شرح ذیل می باشد.
جدول فراوانی مربوط به توزیع را فراهم کنید؟**

۱۰-۱۵-۱۳-۱۰-۱۲-۱۱-۱۰-۱۱-۱۲-۱۰

X	f
15	1
13	1
12	2
11	2
10	4

جواب:

نکته : با توجه به جدول فوق ، عدد ۴ در ستون f بیانگر اینست که عدد ۱۰ چهار بار تکرار شده است.

نکته: اگر داده های ستون فراوانی (f) را با هم جمع کنیم تعداد کل داده ها بدست می آید .

توزیع فراوانی طبقه بندی شده

زمانی که تعداد اعداد یک توزیع و همچنین فاصله بین آنها خیلی زیاد باشد، از توزیع فراوانی طبقه بندی شده استفاده می شود.

نکته: زمانی که تفاضل بین بزرگترین و کوچکترین نمره یا عدد مساوی با بزرگتر از ۲۰ باشد از توزیع فراوانی طبقه بندی شده استفاده می شود

نکته: طبقات بایستی ناسازگار باشند. یعنی یک عدد معین فقط در یک طبقه قرار داده شود.

۱- تعیین دامنه تغییرات

$$R = X_H - X_L$$

۲- تعیین تعداد طبقات با استفاده از قانون استرژ

$$K = 1 + 3 / 3 \text{Log}_N$$

۳- تعیین اندازه یا حجم هر طبقه (فاصله طبقات)

$$i = \frac{R}{K}$$

۴- نوشتن طبقات

۵- نوشتن فراوانی طبقات

نکته: اگر تعداد طبقات بزرگتر از ۲۰ باشد تهیه و تنظیم جدول نیاز به وقت و کار بیشتر است

نکته: اگر تعداد طبقات کوچکتر از ۱۰ باشد اندازه طبقات بزرگ میشود و اطلاعات بیشتری از دست می رود.

نماینده طبقات (نقاط وسط طبقات)

نماینده طبقات یا نقاط میانی را با X' نمایش می دهند و از طریق فرمول زیر به دست می آید:

$$X' = \frac{\text{حد بالای طبقه} + \text{حد پایین طبقه}}{2}$$

Handwritten example calculation:

$$X' = \frac{15 + 12}{2} = 13.5$$

توزیع فراوانی تراکمی

اگر پژوهشگری علاقه‌مند به دانستن تعداد افراد یا نمره‌هایی باشد که در پایین نمره یا عدد خاصی وجود دارند که نیاز به توزیع فراوانی تراکمی دارد فراوانی تراکمی با (cf) نشان داده می‌شود که از جمع کردن فراوانی‌های ساده هر طبقه به طبقه بزرگ‌تر به دست می‌آید.

نکته: فراوانی تراکمی کوچکترین طبقه همیشه برابر با فراوانی ساده یا مطلق آن طبقه است.

نکته: فراوانی تراکمی بزرگترین طبقه همیشه برابر با مجموع داده‌ها (EF) یا N می‌باشد.

نمودارهای فراوانی

اطلاعات جدول به سرعت قابل درک نیست و برای این کار باید جدول بطور تفکیک و جزء به جزء مورد مطالعه قرار گیرد به همین خاطر از نمودار استفاده می‌شود که سرعت انتقال اطلاعات در آن بالا است. + نمودار ابزاری است تصویری که برای توصیف و نمایش داده‌های جمع‌آوری شده به کار برده می‌شود.

انواع نمودارهای فراوانی

الف) هیستوگرام (ب) ستونی (ج) چندضلعی (د) چندضلعی تراکمی (و) دایره‌ای

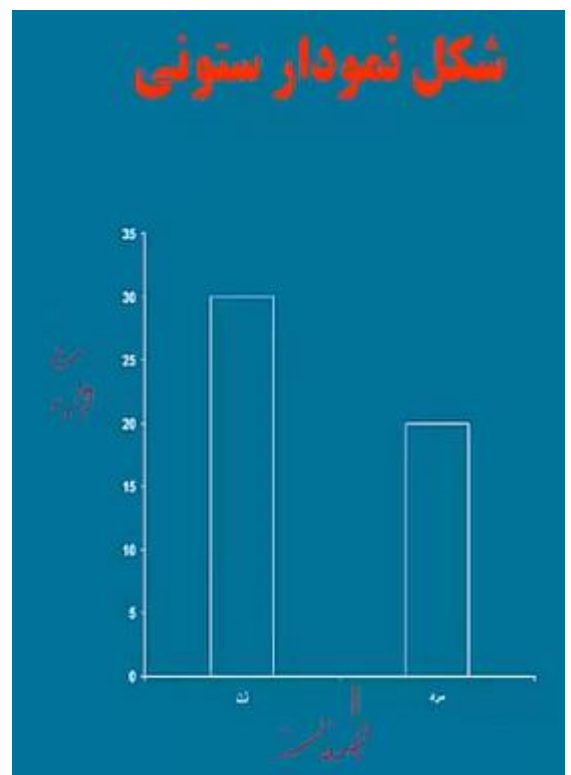
نمودار هیستوگرام

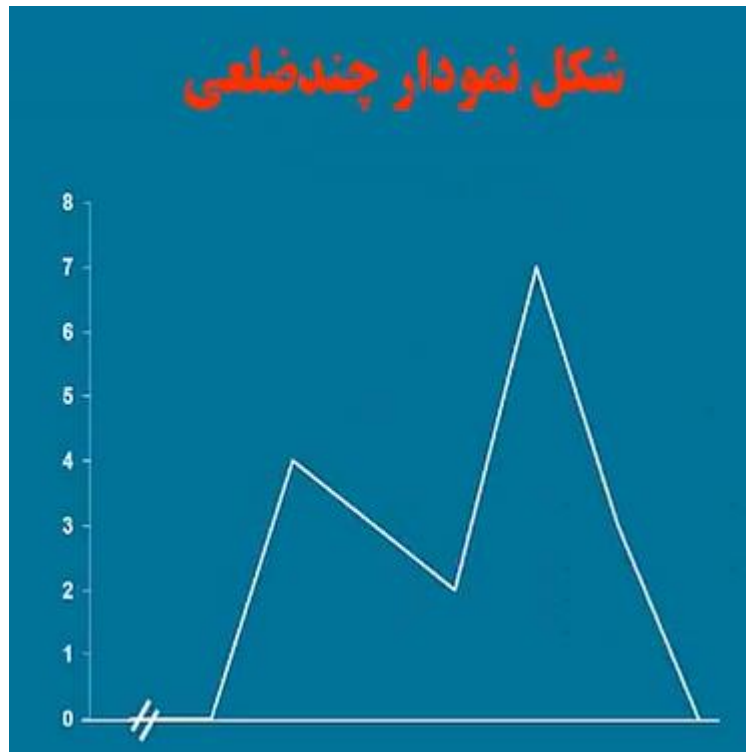
این نمودار از ستون‌هایی که به هم چسبیده شده‌اند تشکیل شده است و وسیله مناسبی برای نمایش داده‌های پیوسته و متغیرهایی در سطح مقیاس فاصله‌ای و نسبی می‌باشد. در این نمودار در محور عمودی (Y) فراوانی مطلق و در محور افقی (X) حدود واقعی طبقات نوشته می‌شود.



نمودار ستونی

این نمودار همانند نمودار هیستوگرام است اما ستون‌ها مجزا از یکدیگر هستند و زمانی استفاده می‌شود که داده‌ها گسسته یا در سطح مقیاس اسمی باشند در سطح مقیاس فاصله‌ای و نسبی می‌باشد. در این نمودار در محور عمودی (Y) فراوانی مطلق و در محور افقی (X) حدود واقعی طبقات نوشته می‌شود.



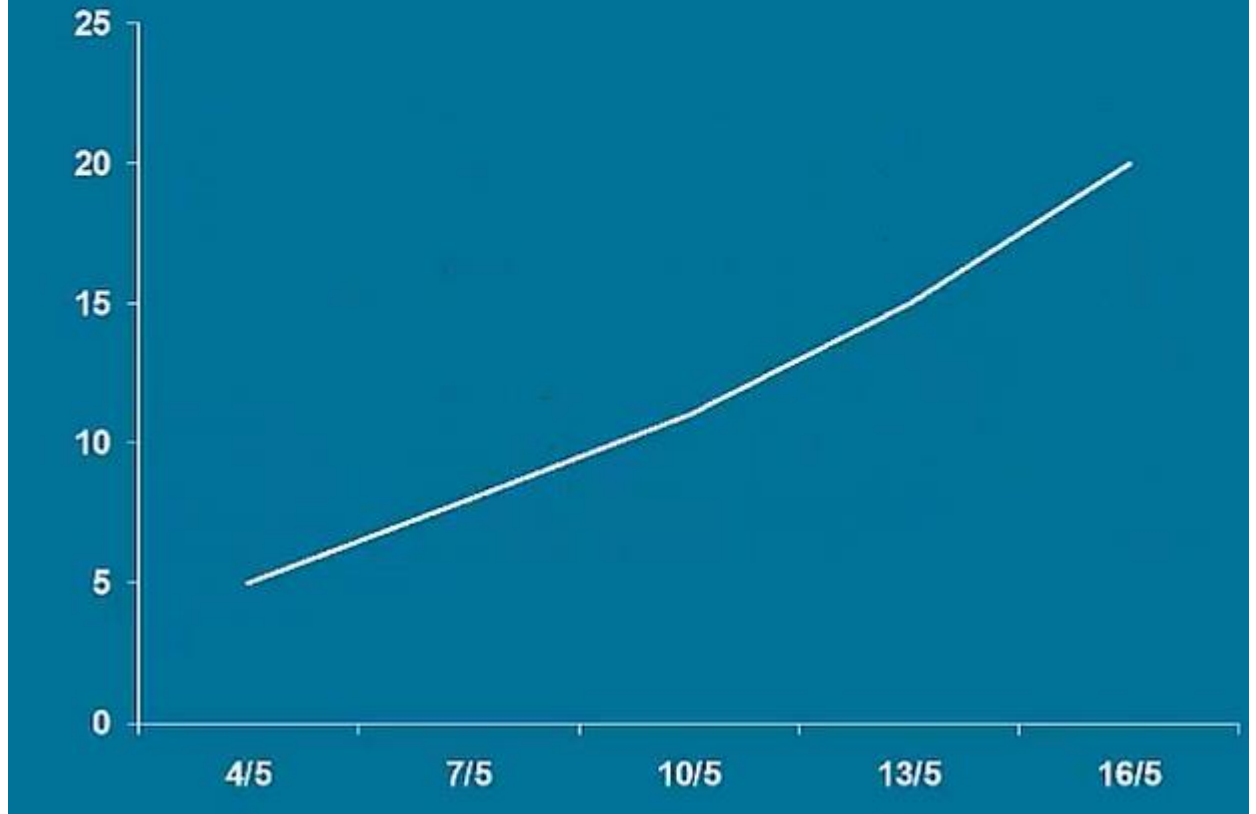


نکته: در صورتی که دو دسته فراوانی مختلف وجود داشته باشد، هنگام ترسیم نمودار چندضلعی احتمال اینکه خطوط بر روی هم قرار گیرد زیاد است و در نتیجه، مقایسه نمودارها به صورت مستقیم امکان پذیر نیست. در چنین شرایطی یکی از راهها، تبدیل فراوانی ها به درصد یا نسبت می باشد.

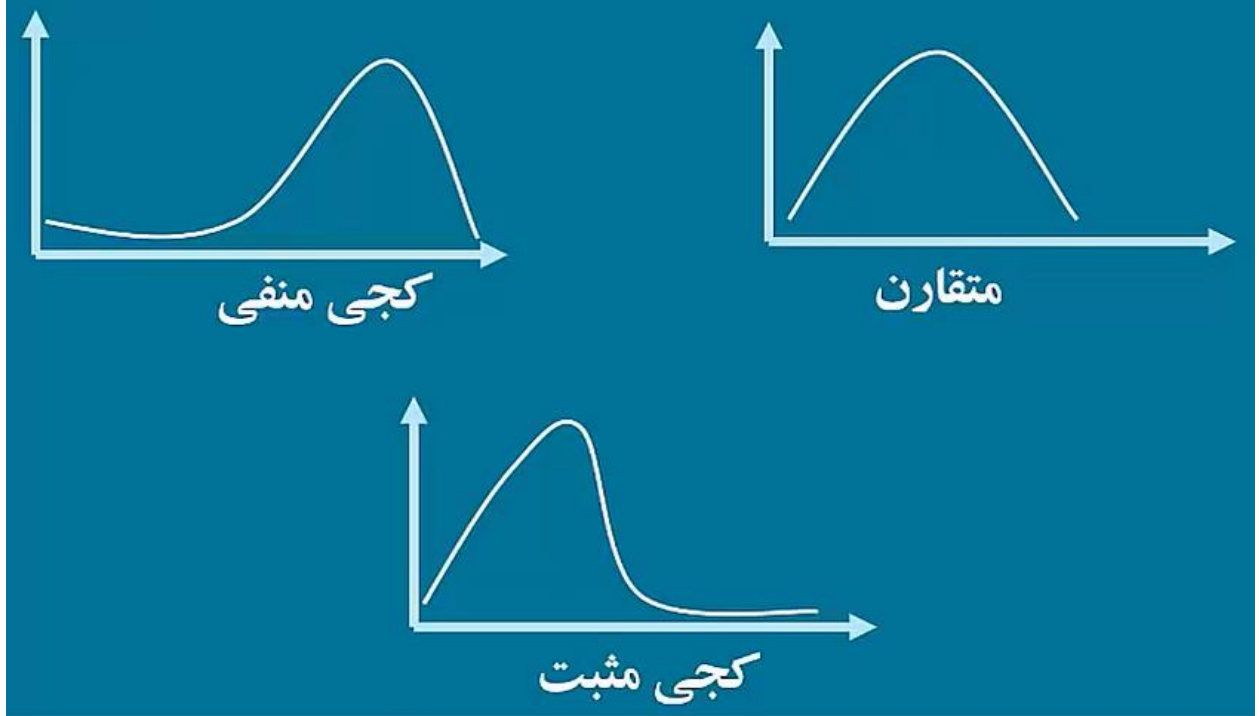
نمودار چندضلعی تراکمی

این نمودار وقتی مفید است که پژوهشگر علاقمند باشد وضعیت یک نمره یا یک فرد را نسبت به بقیه نمرهها یا افراد مشخص کند. برای ترسیم این نمودار در روی محور عمودی (Y) درصد فراوانی تراکمی و در روی محور افقی (X) حدود واقعی طبقات قرار می گیرد.

شکل نمودار



شکل‌های مختلف نمودار چندضلعی



اندازه های گرایش مرکزی

برای طبقه بندی و خلاصه کردن اطلاعات روشهای دقیق تری از جدول توزیع فراوانی نیاز می باشد. یکی از این روش ها تعیین جایگاه و موقعیت کلی نمره ها است. سه شاخص گرایش مرکزی به نام **نما، میانه و میانگین** وجود دارد.

نما (مد)

+ نما ساده ترین شاخص گرایش مرکزی است که عبارت است از عدد یا نمره ای که در توزیع فراوانی دارای بیشترین فراوانی است.

نکته: نما همیشه در مرکز توزیع فراوانی قرار ندارد به همین دلیل نمی توان به عنوان یک شاخص مرکزی به آن اطمینان داشت.

نکته: یک توزیع ممکن است تک نمایی، دو نمایی، چند نمایی باشد.

ویژگیهای نما

۱- زمانی مورد استفاده قرار می گیرد که نیاز به تعیین شاخص مرکزی به صورتی تقریبی و سریع باشد.

۲- به سهولت محاسبه می شود.

۳- با داده های اسمی به کار برده می شود

محاسبه نما در داده های طبقه بندی شده

$$M_o = L + i \left(\frac{d_1}{d_1 + d_2} \right)$$

▪ حد واقعی پایین طبقه ای که دارای بیشترین فراوانی است $L =$

▪ طول یا فاصله طبقات $i =$

▪ تفاضل فراوانی ساده طبقه ای که دارای بیشترین فراوانی است
با فراوانی ساده طبقه کوچکتر $d_1 =$

▪ تفاضل فراوانی ساده طبقه نما با طبقه بزرگتر $d_2 =$

مثال:

X	F
۵۴-۵۶	۱
۵۷-۵۹	۳
۶۰-۶۲	۶
۶۳-۶۵	۸
۶۶-۶۸	۲

$MO = L + i \left(\frac{d_1}{d_1 + d_2} \right)$

$d_1 = 8 - 6 = 2$

طبقه نما $d_2 = 8 - 2 = 6$

$MO = 62/5 + 3 \left(\frac{2}{2+6} \right) = 63/5$

میانه

میانه نقطه وسط در توزیع نمرهها است. به عبارت دیگر میانه نقطه ای است که نیمی از نمرهها در بالای آن و نیم دیگر در پایین آن قرار دارند. میانه را با m نشان میدهند.

مراحل محاسبه میانه

- ۱- مرتب کردن اعداد. ۲- تعیین نقطه ای که نیمی از داده ها بالاتر و نیمی دیگر پایین تر از آن هستند.
- نکته: در صورتی که تعداد دادهها فرد باشد میانه عددی است که در وسط قرار دارد و اما در صورتی که تعداد دادهها زوج باشد، میانه عبارت است از معدل دو نمره ای که در وسط واقع می شوند.

مثال:



محاسبه میانه در داده های طبقه بندی شده (میانه در جدول فراوانی)

- ۱- تقسیم تعداد کل فراوانی ها (N) بر دو
- ۲- مشخص کردن طبقه میانه، یعنی اولین طبقه ای که فراوانی تراکمی آن مساوی یا بزرگتر از $\frac{N}{2}$ باشد.
- ۳- جایگزینی مقادیر در فرمول.

فرمول محاسبه میانه در داده های طبقه بندی شده

$$m = L + \left(\frac{\frac{N}{2} - CF}{f} \right) \times i$$

L = حد پایین واقعی طبقه‌ای که میانه در آن قرار دارد

N = تعداد کل فراوانی‌ها

CF = فراوانی تراکمی طبقه کوچکتر از میانه

f = فراوانی ساده طبقه میانه

i = فاصله طبقات

میانگین

مشهورترین و معتبرترین شاخص گرایش مرکزی میانگین است. میانگین معدل حسابی گروهی از نمره هاست.

$$\text{میانگین} = \frac{\text{حاصل جمع کل نمره‌ها}}{\text{تعداد کل نمره‌ها}} \quad \bar{X} = \frac{\sum X}{N}$$

محاسبه میانگین در جدول توزیع فراوانی

X	f	fx
18	1	18
17	2	34
15	2	30
12	3	36
11	5	55
	13	173

$$\bar{X} = \frac{\sum fX}{N}$$

$$\bar{X} = \frac{173}{13} = 13 / 3$$

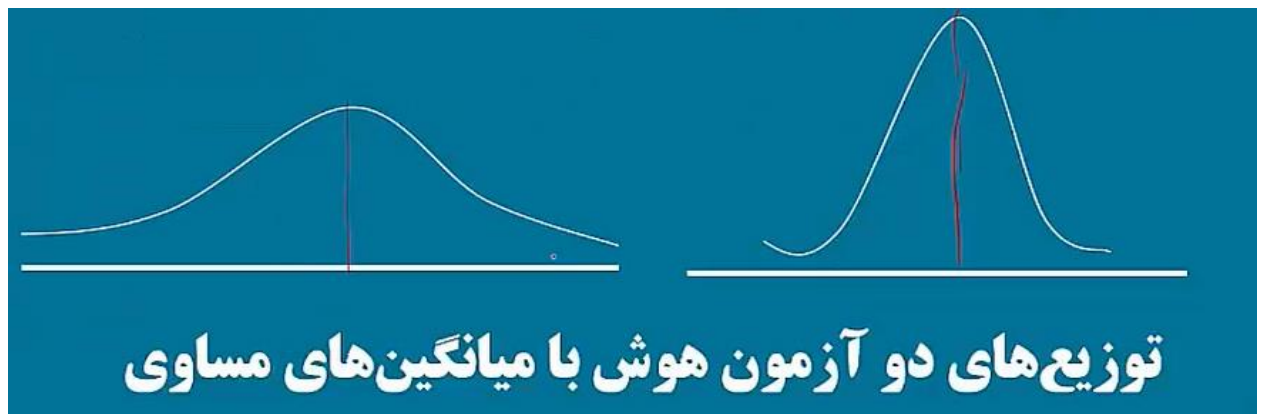
ویژگیهای میانگین

۱- به تک تک اعداد توزیع فراوانی حساس است. ۲- دارای ثبات بیشتری نسبت به میانه و نما است ۳- میانگین در نمونه های مختلف یک جامعه بیشتر از نما و میانه به همدیگر نزدیک می باشند. ۴- اگر تمام اعداد با داده ها با عدد ثابتی جمع یا تفریق یا ضرب یا تقسیم شوند، میانگین در آن عدد جمع، تفریق، ضرب و تقسیم می شود. ۵- مجموع انحراف نمرهها از میانگین برابر صفر است.

جلسه دوازدهم

شاخص های پراکندگی

شاخصهای پراکندگی، میزان پراکندگی یا تغییراتی که در بین داده های یک توزیع وجود دارد را نشان می دهند. ممکن است توزیعی وجود داشته باشند که میانگین های آنها مساوی باشد ولی پراکندگی آنها در اطراف میانگین اختلاف داشته باشد.



انواع شاخصهای پراکندگی

الف: دامنه تغییرات ب: انحراف چارگی ج: واریانس د: انحراف استاندارد

دامنه تغییرات:

بزرگترین عدد منهای کوچکترین عدد می شود.

نکته: وجود یا فقدان پراکندگی ضرورتاً نه خوب است نه بد، بلکه بستگی به هدف سنجش یا پژوهش دارد. مثلاً در صورتی که تعیین میزان معلومات یک گروه مدنظر باشد، پراکندگی زیاد مطلوب است، ولی در صورتی که میزان یادگیری مطالب تدریس شده مدنظر باشد پراکندگی کم مطلوب است.

نکته: به دلیل اینکه دامنه تغییرات فقط پراکندگی بین بزرگترین و کوچکترین نمره ها را تعیین می کند قادر به توصیف توزیع نمره ها به صورت حقیقی نیست.

نکته: دامنه تغییرات یک شاخص پایدار پراکندگی نیست، زیرا مقدار آن با تغییر یک نمره (کوچکترین یا بزرگترین) ممکن است به شدت تغییر کند، اما با تغییر سایر اعداد میانی گروه، ممکن است هیچ تغییری در دامنه تغییرات ایجاد نشود.

انحراف چارگی

این شاخص میزان پراکندگی را در اطراف مرکز توزیع نمرهها نشان می دهد.

$$Q = \frac{Q_3 - Q_1}{2}$$

نقطه ۷۵/۰ \Rightarrow چارک سوم = Q_3

انحراف چارگی = Q

نقطه ۲۵/۰ \Rightarrow چارک اول = Q_1

نکته: انحراف چارگی همانند میانه تحت تأثیر نمره های خیلی بزرگ یا کوچک قرار نمی گیرد.

نکته: بهترین مورد استفاده آن زمانی است که نمره های خیلی بزرگ یا خیلی کوچک شکل توزیع نمره ها را از بین می برد.

واریانس

واریانس عبارتست از میانگین انحراف نمره ها از میانگین یا مجموع مجذورات انحراف نمره ها از میانگین تقسیم بر تعداد نمره ها.

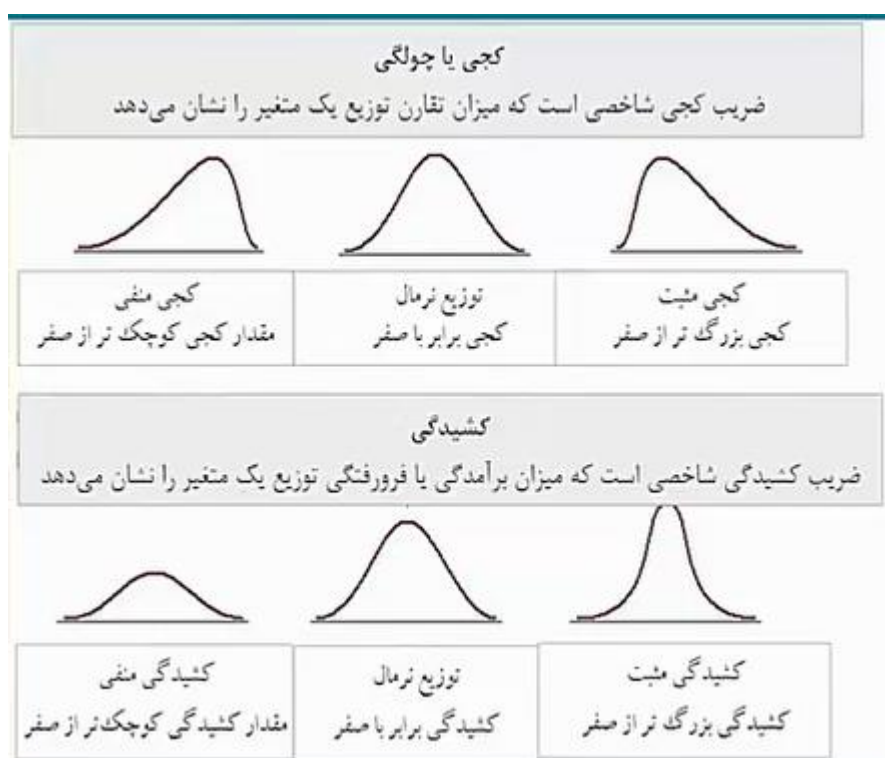
انحراف استاندارد

جذر واریانس، انحراف استاندارد می باشد

+ تمام روشهایی که برای محاسبه واریانس به کار برده می شوند برای به دست آوردن انحراف استاندارد هم کاربرد دارند.

نکته: در صورتی که عدد ثابتی به همه اعداد یک توزیع اضافه شود یا از آنها کم شود، واریانس و انحراف استاندارد تغییری نخواهد کرد، اما اگر عدد ثابتی در اعداد خام ضرب شود، همان عدد در انحراف استاندارد ضرب می شود و توان دوی آن در واریانس ضرب می شود.

ویژگیهای انحراف استاندارد: ۱- انحراف استاندارد معتبرترین شاخص پراکندگی است. ۲- اگر پراکندگی نمرات بالا باشد، انحراف استاندارد بزرگ خواهد شد. ۳- در محاسبه ی انحراف استاندارد تمامی اعداد دخالت دارند.



همبستگی

حوادث متعددی در طبیعت اتفاق می افتد که بین آنها همبستگی یا رابطه وجود دارد. منظور از رابطه بین متغیرها وجود رابطه علت معلولی نیست.

نکته: همبستگی رابطه بین دو متغیر در جامعه را توصیف می کند که متغیرها را یکی را X و دیگری را Y می نامند.

نکته: ضریب همبستگی با (rxy) نشان داده میشود. ضریب همبستگی هم جهت رابطه و هم شدت رابطه را بیان می کند.

نکته: جهت همبستگی با علامت +/- نشان داده می شود و شدت همبستگی با قدر مطلق عددی.

نکته : بالاترین ضریب همبستگی +1 می باشد.

محاسبه ضریب همبستگی

برای محاسبه ضریب همبستگی شاخصهای متعددی تدوین شده است. معروفترین و در عین حال پرمصرفترین آنها زمانی به کار برده می شود که متغیرهای مورد مطالعه با استفاده از مقیاس فاصله ای یا نسبی اندازه گیری شده باشند. این روش توسط کارل پیرسون تهیه و تنظیم گردیده است و می شود. ضریب همبستگی گشتاوری پیرسون نامیده

ضریب همبستگی اسپیرمن


ضریب همبستگی اسپیرمن صورتی از ضریب همبستگی پیرسون است و زمانی به کار برده می شود که نمره ها رتبه بندی شده باشند یا به جای اعداد رتبه های آنها در دست باشد.

ضریب همبستگی رتبه ای اسپیرمن

ضریب همبستگی اسپیرمن صورتی از ضریب همبستگی پیرسون است و زمانی به کار برده می شود که نمره ها رتبه بندی شده باشند یا به جای اعداد رتبه های آنها در دست باشد.

$$P = 1 - \frac{6\sum D^2}{N(N^2 - 1)}$$

P = ضریب همبستگی رتبه ای اسپیرمن
 D^2 = مجذور تفاوت رتبه ها
N = تعداد داده ها یا نمره ها



مراحل محاسبه ضریب همبستگی رتبه ای اسپیرمن

۱- اعداد متغیر X به ترتیب از بزرگ به کوچک مرتب می کنیم (RX) و بعد از آن اعداد Y مربوط به هر عدد X را در ستون Y ها قرار می دهیم و به آنها رتبه می دهیم (RY)

۲- RY را از RX کم می کنیم و اختلاف آنها را D می نامیم

۳- D ها را به توان دو می رسانیم و مجموع آنها را محاسبه می کنیم

عواملی که بر ضریب همبستگی تأثیر می گذارند:

۱- اساس رابطه از جامعه ای به جامعه دیگر فرق می کند. مثلا در افراد بشر در سنین ۱۶-۱۰ سالگی بین سن تقویمی و توانایی فیزیولوژی همبستگی بالایی وجود دارد. ولی بین این دو متغیر در سنین ۲۶-۲۰ سالگی همبستگی وجود ندارد.

۲- پراکندگی متغیرها در جوامع مختلف متفاوت است بدین معنی که هر چه تجانس بیشتر باشد همبستگی کمتر است. به عنوان مثال، در یک دانشکده (جامعه) بین قد و موفقیت در بازی بسکتبال همبستگی مثبت و بالای وجود دارد. اما در تیم بسکتبال یک کشور چنین رابطه ای وجود ندارد.

۳- همبستگی بین دو متغیر تحت تأثیر همبستگی آنها با متغیر ثالثی قرار دارد به عنوان مثال ، همبستگی بین فیزیک و ریاضی ممکن است به دلیل همبستگی این متغیرها با هوش باشد.

نکته: تفسیر ضریب همبستگی نباید بر حسب درصد و نسبت باشد. مثلا $r_{XY} = ۷۰ / ۰$ هفتاد درصد رابطه بین متغیرها را تبیین نمی کند و $r_{XY} = ۰۹ / ۰$ دقیقا دو برابر $r_{XY} = ۴۵ / ۰$ نیست.

ضریب تعیین

ضریب تعیین نشان دهنده میزان تأثیری است که متغیر X (مستقل) در متغیر (وابسته) ایجاد می کند یا به عبارتی با محاسبه این ضریب می توان تعیین کرد چند درصد از کل واریانس X ناشی از واریانس Y است.

$$V = (r_{xy})^2 (100)$$

نکته: ضریب تعیین هیچ وقت منفی نخواهد شد زیرا برای محاسبه آن ضریب همبستگی مجذور می شود.

ضریب همبستگی دورشته ای نقطه ای

اگر یکی از متغیرها پیوسته و دیگری دو ارزشی واقعی باشد.

مانند ضریب همبستگی بین نمرات ریاضی و جنسیت

ضریب همبستگی بین نمرات کل آزمون و نمره یک سؤال (۰ و ۱)

ضریب همبستگی دورشته ای

اگر یکی از متغیرها پیوسته و دیگری دو ارزشی ساختگی باشد.

مانند ضریب همبستگی بین نمرات ریاضی و قبولی / ردی در آزمون

ضریب همبستگی فی

اگر هر دو متغیر از نوع دوارزشی واقعی باشند. مانند ضریب همبستگی بین جنسیت و موفقیت / شکست

ضریب همبستگی تراکوریک (چهارخانه ای)

اگر هر دو متغیر از نوع دوارزشی ساختگی باشند.

جلسه سیزدهم چهاردهم

ویژگی های یک سؤال خوب در امتحان کتبی

۱- سؤال نمونه درستی از هدفها و محتوای درس باشد.

۲- سؤال به جنبه های کاربردی موضوع در زندگی واقعی آدمی توجه داشته باشد. یک سؤال خوب می تواند با آموزش جنبه های صنعتی و اقتصادی و زیست محیطی پیرامون یک هدف درسی همراه باشد.

۳- به تحقیق پیرامون طبیعت اهمیت بدهد. به کنترل متغیرها و طراحی روش های مناسب برای فرضیه سازی و آزمون ها توجه کند.

۴- از حافظه آزمایی و پرسش های مستقیم اجتناب کند.

۵- از شکل نمودار و جدول برای طرح سؤال با پاسخگویی استفاده کند.

۶- به نظرهای گوناگون دانش آموزان توجه کند. تفکر واگرا را تقویت کند و به جای پاسخهای بسته و محدود به پاسخ های گسترده و گوناگون توجه داشته باشد.

۷- دانش آموزان را به تجزیه و تحلیل و قضاوت و ارزیابی فراخواند.

۸- به تقویت تفکر نقاد توجه کند.

۹- به افزایش انعطاف پذیری و جامع نگری یاری برساند و از یک سو نگری بکاهد.

۱۰- دانش آموزان را به دسته بندی سازمان دهی و مقایسه تطبیقی واقعیتها اصطلاحات و قوانین فراخواند.

۱۱- فاقد شکل و محتوای کلیشه ای و تکراری باشد.

۱۲- به تلفیق رشته های مختلف دانش توجه کند.

۱۳- به جنبه های عاطفی دانش آموزان توجه کند.

۱۴- دانش آموزان را به تفسیر رویدادها و اطلاعات فراهم آمده فرا خواند.

ویژگیهای یک آزمون خوب

• عینیت (Objectivity)

• عملی بودن (Usability)

• سهولت اجرا، نمره گذاری و تعبیر و تفسیر

• روایی (Validity)

• پایایی (Reliability)

عینیت (Objectivity)

آزاد بودن آزمون از خطای شخصی

+ دو جنبه مهم عینیت آزمون: ۱. عینیت نمره گذاری (مصحح متعدد) ۲. عینیت در تفسیر سؤال های آزمون
به وسیله آزمودنی ها (دانش آموزان)

عملی بودن (Usability)

ارزش عملی آزمون از لحاظ زمان، صرفه اقتصادی و اجرا

+ پنج جنبه مهم عملی بودن آزمون: آسانی اجرا / زمان لازم برای اجرا / آسانی تفسیر و کاربرد آزمون /
در دسترس بودن فرم های معادل / هزینه آزمون

روایی (Validity)

یک آزمون باید برای هدفی که برای تحقق بخشیدن آن درست شده است، مناسب باشد. روایی، یعنی آزمون دقیقا آنچه را که باید بسنجد، مورد ارزیابی قرار دهد.

+ روایی یک ارزشیابی از کفایت و مناسبت تفسیرها و استفاده های نتایج سنجش است (لین و گرانلاند).

+ روایی را می توان به صورت توافق بین نمره آزمون و کیفیتی که قرار است اندازه بگیرد تعریف کرد (کاپلان و ساکوزو).

+ در حقیقت، **شواهد** به دست آمده از نمرات آزمون و نشانه های دیگر، **استنباط هایی** برای معنا، تفسیر و کاربرد نمره های آزمون در عمل در پی دارد.

+ بررسی روایی (یا روایی Validation)، جمع آوری این شواهد و استنباط هاست.

نکته: مهم ترین ویژگی یک آزمون روایی آن است

در حقیقت می توان گفت تنها ویژگی یک آزمون خوب روایی آن است و سایر ویژگی ها (مانند طراحی خوب سوالها و عینیت، عملی بودن و پایایی آزمون) جزئی از شواهد روایی هستند و شواهد نقض آن ها شواهدی بر نقض روایی محسوب می شوند.

+ شواهد روایی، شامل شواهد همگرا و واگرا است.

شواهد همگرا، به این قرینه بر می گردند که آزمون باید فقط ویژگی مورد نظر را بسنجد؛ مثلا آزمون باید با آزمون های دیگری که همین ویژگی را می سنجند، همبستگی بالایی داشته باشد.

شواهد واگرا (تشخیصی یا تمیزی)، از این قبیل اند که آزمون نباید ویژگی های دیگر را به جای این ویژگی بسنجد؛ مثلا آزمون باید با آزمون های ویژگی های دیگر همبستگی پایینی داشته باشد.

در کل، هر آنچه به درستی نشان دهد شما آنچه باید بسنجید، آزمون کرده اید و موارد غیر آن شما را از هدف دور نکرده است، شواهد روایی محسوب میشود.

انواع روایی؛ سه سؤال و سه نوع روایی:

(۱) آیا سؤال های آزمون از محتوای موضوع مورد نظر یک نمونه خوب ارائه میدهند؟

۲) آیا نمرات آزمون، عملکرد فعلی یا آتی دانش آموزان را به درستی پیش بینی می کنند؟

۳) آیا نمرات آزمون به مفاهیم نظری (یا سازه هایی) که آزمون برای سنجش آنها درست شده است مربوط اند؟

• پاسخ سؤال یک، روایی محتوایی (منطقی)

• پاسخ سؤال دو، روایی ملاکی

• و پاسخ سؤال سه، روایی سازه را نشان میدهد.

جلسه پانزدهم

پایایی (Reliability)

آزمون یک محتوا در دفعات متعدد با شرایط مختلف اجرا برای یک گروه واحد، نتایج نزدیک به هم داشته باشد. یعنی فاصله میانگین نمره آزمون در دفعات متعدد اجرا یا در فاصله زمانی کوتاه زیاد نباشد.

توجه: برای آنکه یک آزمون روا باشد باید دارای پایایی باشد.

روشهای بررسی پایایی:

۱- بررسی پایایی مصححانه

پایایی بین مصححان

۲- بررسی پایایی آزمون :

الف) دو بار اجرای آزمون:

• بازآزمایی

• فرمهای موازی با هم ارز

ب) یک بار اجرای آزمون (روشهای همسانی درونی):

• دو نیمه کردن و اسپیرمن براون

• گودر ریچاردسون

• ضریب آلفای گورنباخ (برای آزمون های دارای چند بخش)

روش باز آزمایی

اجرای آزمون در دو نوبت، در فاصله حدود دو هفته، با گروه واحدی از آزمون شوندگان.

ضریب پایایی \approx با ضریب همبستگی بین نمره های دو اجرای آزمون

این روش ساده و قابل اجراست، اما:

- در اجرای اول آزمون شوندگان با آزمون آشنا می شوند و در اجرای دوم بهتر عمل می کنند.

- نمونه سوال های دو آزمون واحد است و لذا اطلاعاتی درباره نمونه سوال های دیگر نمی دهد.

- اگر فاصله دو اجرا زیاد باشد، ممکن است تغییرات واقعی در توانایی آزمون شوندگان با نتایج یادگیری در هم آمیزد.

- اجرای دوباره یک آزمون مورد علاقه دانش آموزان و معلمان نیست.

توجه: روش باز آزمایی بیشتر با آزمونهای هوش و شخصیت و یا آزمونهای حرکتی که در طول زمانی کوتاه چندان تغییر نمی کنند مفید است.

روش فرم های موازی یا هم ارز

در آزمون معادل با موازی برای یک موضوع تهیه می کنند و آنها را در فاصله زمانی کوتاهی (حدود دو هفته) به یک گروه واحد از آزمون شوندگان می دهند.

دو فرم آزمون زمانی موازی هستند که میانگین و واریانس دو آزمون یکسان باشد و پایایی آزمونها و رابطه میان نمرات آزمونها و متغیرهای دیگر نیز باید برابر باشد. در این صورت، ضریب پایایی برابر است با ضریب همبستگی بین نمرات حاصل از اجرای دو فرم موازی. آزمون

+ مشکل عمده این روش این است که آزمون پیشرفت معلم ساخته غالبا در یک فرم واحد تهیه میشود و تهیه فرمهای موازی برای آنها به سادگی میسر نیست.

+ مشکل بی علاقهگی دانش آموزان و معلمان به اجرای دو فرم از یک آزمون بر گروه واحد باقی است

+ همچنین با جواب دادن به فرم اول با اصولی آشنا می شوند که بر عملکرد آنان در فرم دوم تاثیر می گذارد

+ همچنین مدلی از اجرای فرمهای موازی وجود دارد که ابتدا دو گروه نمونه به طور تصادفی یا معادل انتخاب می شود و سپس هر یک از فرمها برای یکی از این گروه ها به اجرا در می آید.

روش های همسانی درونی

روشهای بازآزمایی و فرمهای موازی برای تعیین ضریب پایایی آزمون های ویژگی هایی که نسبتاً ثابت هستند کار آمد هستند، اما برای تعیین پایایی ابزارهای سنجش ویژگی های ناپایدار (متغیرهای گذرا) مناسب نیستند. آن روشها، تغییرپذیری ذاتی ویژگیهای ناپایدار را به عنوان خطا در نظر میگیرند و در نتیجه پایایی را کمتر از واقعیت برآورد می کنند.

+ در حقیقت روش بازآزمایی بر ثبات نتایج آزمون و روش فرمهای هم ارز بر تراز نتایج آزمون تأکید دارند.
 + در روش همسانی درونی، بر همسانی و یکنواختی ماده ها یا اجزای تشکیل دهنده یک آزمون تأکید می شود.

روش دونیمه کردن (تصرف یا دونیمه سازی)

در این روش آزمون یک بار و برای یک گروه اجرا می شود و پس از اجرا آن را به دو نیمه تقسیم می کنیم. سپس ضریب همبستگی در نیمه (r) را محاسبه نموده و سپس با فرمول اسپیرمن - براون ضریب پایایی (rtt) را به دست می آوریم:

مثال: اگر ضریب همبستگی دو نیمه آزمون، 0.7 باشد، ضریب پایایی آن برابر است با 0.82

$$r_{tt} = \frac{2r}{1+r}$$

$$r_{tt} = \frac{2 \times 0.7}{1 + 0.7} = 0.82$$

روش دو نیمه کردن

روشهای تقسیم کردن آزمون به دو نیمه:

• نیمه اول از اولین سؤال تا سؤال وسط و باقی سؤالات در نیمه دوم. به دلیل تفاوت حالت شرکت کنندگان در دو نیمه آزمون و احتمال خستگی در نیمه دوم روش مناسبی نیست

• سؤالات فرد در یک نیمه و سؤالات زوج در نیمه دیگر، که اشکال قبلی را رفع می کند و روش مناسبی است، لکن ممکن است ویژگیهای سؤالات دو نیمه یکسان نباشد.

• دونیمه سازی تصادفی که در این روش سؤالات به طور تصادفی به دو نیمه تقسیم می شوند. این روش نیز اشکال روش قبلی را دارد.

• دونیمه سازی تصادفی که در این روش سؤالات به طور تصادفی به دو نیمه تقسیم میشوند.

این روش نیز اشکال روش قبلی را دارد: دونیمه سازی بر اساس ویژگیهای آماری، که در این روش بر اساس ویژگیهایی مشابه تمیز و واریانس سؤال، سؤالات را به "جفتهای جور شده" دسته بندی می کنند و از هر جفت یک سؤال در نیمه اول و یک جفت در نیمه دوم آزمون قرار می گیرد.

روش دو نیمه کردن افزایش سؤال های آزمون: چنانچه پایین بودن پایایی آزمون در اثر کم بودن سؤالات آن باشد، می توان با افزایش سؤال تا حدودی پایایی را بالا برد. در این شرایط، می توان از فرمول اسپیرمن براون برای محاسبه ضریب پایایی جدید آزمون وقتی به آزمون تعدادی سؤال با همان کیفیت سؤال های قبلی افزوده شود، استفاده نمود. در این صورت اگر آزمون اولیه ضریب پایایی ttv داشته باشد و در اثر افزودن سؤال های جدید، تعداد سؤالات k برابر شود، ضریب پایایی برابر است با

+ افزایش سؤال های آزمون: چنانچه پایین بودن پایایی آزمون در اثر کم بودن سؤال های آن باشد، می توان با افزایش سؤال تا حدودی پایایی را بالا برد. در این شرایط، می توان از فرمول اسپیرمن براون برای محاسبه ضریب پایایی جدید آزمون وقتی به آزمون تعدادی سؤال با همان کیفیت سؤالات قبلی افزوده شود، استفاده نمود.

پایایی بین مصححان (برای پایایی تصحیح)

پایایی بین مصححان (بین ارزیابان، بین مشاهده گران یا بین درجه بندی کنندگان) برای تعیین پایایی نمرات دو یا چند ارزیاب است که مستقلاً پاسخ های تشریحی با عملکردی آزمون شوندگان را تصحیح کرده با عمل آنها را مشاهده و نمره گذاری کرده اند.

• شاخص پایایی بین مصححان برابر است با همبستگی بین نمرات نمره گذاران مختلف. هرچه توافق ارزیابان بیشتر باشد، این نوع پایایی نیز بیشتر است.

• ضریب پایایی مصححان برای ضریب پایایی نمرات حاصل از اجرای آزمون یک سقف به حساب می آید، به این معنی که ضریب پایایی نمرات یک آزمون که توسط دو نفر تصحیح شده است، احتمالاً از پایایی مصححان در آزمون کمتر خواهد بود.

جلسه شانزدهم

پایایی آزمون های سرعت

یاد آوری: در آزمونهای قدرت هدف اصلی سنجش سطح عملکرد یا میزان دانش و توانایی آزمون شوندگان است و شامل سؤالهای چالش انگیز هستند، اما آزمون های سرعت از سؤالهایی آسان تشکیل شده و هدف آن صرفاً سنجش سرعت عمل آزمون شوندگان در پاسخ دادن به سؤالها است و برای پاسخگویی محدودیت زمانی وجود دارد و اکثر افراد نمی توانند در محدوده زمانی تعیین شده آزمون را به اتمام برسانند.

+ در محاسبه این مورد از ضریب پایایی، روشهای پیشین مناسب نیستند؛

• مثلاً در روش دونیمه کردن اگر بر حسب نیمه ابتدایی و نیمه انتهایی سؤالها نیمه بندی کنیم، دچار کم برآورد کردن پایایی و با تقسیم سؤالهای زوج و فرد، دچار بیش برآورد کردن خواهیم شد.

• از آنجا که سؤالها دشوار نیستند، بازآزمایی نیز سبب افزایش تمرین و افزایش تفاوت بین عملکرد آنان در دو اجرای آزمون خواهد شد.

• در این شرایط روش فرم های موازی که بلافاصله بعد از یکدیگر اجرا می شوند، مناسب تر است.

پایایی نمرات اختلاف (نمرات افزایش یا تغییر)

فرض کنید بخواهیم اختلاف بین نمره های دانش آموزان در یک آزمون واحد یا در دو فرم موازی یک آزمون را که در دو زمان مختلف اجرا شده اند، یا اختلاف بین عملکرد دو گروه دانش آموزی در یک آزمون یا تفاوت نمرات یک دسته دانش آموز در آزمونهای مختلف را بررسی کنیم.

در این مثال ها به دنبال میزان تغییرات یادگیری هستیم. در مثال اول سؤال اصلی این بوده است که دانش آموزان در طول زمان مشخص چه مقدار یادگیری جدید داشته اند.

$$r_d = \frac{\frac{r_{11} + r_{22}}{2} - r_{12}}{1 - r_{12}}$$

عوامل مؤثر در افزایش پایایی آزمونهای وابسته به هنجار

- ۱- با افزودن تعداد سؤالهای یک آزمون
 - ۲- متجانس تر و همگون تر کردن سؤالهای یک آزمون
 - ۳- افزودن تعدادی سؤال با ضریب تمیز زیاد
 - ۴- افزودن تعدادی سؤال با ضریب دشواری متوسط
 - ۵- اجرای آزمون در یک گروه با تجانس توانایی کمتر
- همچنین معمولاً ضریب پایایی آزمون های سرعت بیشتر از آزمون های قدرت است.

خطای معیار اندازه گیری

Standard Error of Measurement (SEM)

وقتی از شخصی بخواهیم توپی را ۲۰۰ مرتبه تا دورترین فاصله ممکن پرتاب کند، میانگین فواصل را می توان شاخصی از توانایی او در پرتاب توپ قلمداد کرد. همچنین، انحراف معیار اختلاف های پرتابهای او از میانگین را انحراف معیار توزیع نمرات خطا یا خطای معیار اندازه گیری می نامند. در حقیقت میزان فقدان دقت نمرات (که از میزان ناپایایی نمرات ناشی می شود) در خطای معیار اندازه گیری انعکاس می یابد.

+ در متغیرهای روانی و پرورشی، که تکرار مداوم اندازه گیری مطرح نیست، خطای معیار اندازه گیری با استفاده از ضریب پایایی آزمون به دست می آید.

رابطه روایی و پایایی: فرمول اصلاح کاهش

قبلاً گفتیم پایایی یک شرط روایی است، بنابراین هر اندازه از میزان پایایی یک آزمون کاسته شود، به همان نسبت از روایی آزمون نیز کاسته خواهد شد.

پایایی روش های درجه بندی و رتبه بندی

معروف ترین روش: روش کلیفورد

این روش اساساً با داده‌های مربوط به آرایش رتبه‌های به دست آمده از دوران مختلف به کار می‌رود این داوران گروهی از افراد را از بالا به پایین یا از زیاد به کم رتبه بندی یا درجه بندی می‌کنند.

نحوه محاسبه:

$$\bar{r} = 1 - \frac{M(4N + 2)}{(M - 1)(N - 1)} + \frac{12 \sum S^2}{M(M - 1)N(N^2 - 1)}$$

تعداد داوران: M
تعداد افراد مورد داوری: N
مجموع رتبه‌های داوران برای هر نفر: $\sum S^2$

پایایی روش درجه‌بندی یک داوری (میانگین همبستگی‌های داوران)

پایایی آزمونهای وابسته به ملاک

در آزمونهای وابسته به ملاک، ما بیشتر به ثبات نتایج آزمون در ارتباط با تصمیم‌های مربوط به حد تسلط علاقمندیم (نه به پایایی نمره‌های آزمون)، تا تعیین شود وضعیت تعدادی از دانش‌آموزان در دستیابی یا عدم دستیابی به حد تسلط در دو موقعیت مختلف با شرایط مشابه (با فاصله‌ای کوتاه بدون هیچ‌گونه آموزش یا بازخورد اصلاحی جدید) تا چه اندازه از ثبات برخوردار بوده است. اگر این وضعیت اکثریت آزمون‌شوندگان در دو بار اجرای آزمون یکسان بوده است، آزمون مورد نظر از میزان بالایی از پایایی برخوردار است.

+ در آزمونهای وابسته به ملاک، روش باز آزمایی و روش فرمهای معادل (هم‌ارز) به کار می‌رود و روش فرمهای هم‌ارز ترجیح داده می‌شود.

- اما برای تعیین پایایی ثبات تصمیم‌گیری (در مقابل ثبات نمره) روشهای دیگری نیز به کار رفته است. روش‌های درصد توافق، کاپا، لیوینگ استون و نیز روش لیبرمن و مرندا در این زمینه مطرح شده، که روش اول را بررسی می‌کنیم

روش درصد توافق

به روش درصد توافق، روش تعیین پایایی ثبات تصمیم گیری با ضریب ثبات تصمیم گیری نیز می گویند و زمانی که دو فرم معادل آزمون برای یک گروه به کار رفته باشد، اجرا میشود.

نحوه محاسبه:

$$\text{ضریب پایایی در روش درصد توافق} = \frac{\text{تعداد افرادی که در هر دو آزمون به حد تسلط نرسیده اند} + \text{تعداد افرادی که در هر دو آزمون به حد تسلط نرسیده اند}}{\text{تعداد کل افراد}} \times 100$$

عوامل مؤثر در افزایش پایایی آزمونهای وابسته به ملاک:

- ۱- در صورت امکان، آزمونی بسازید که بین میانگین نمرات آن و نمره حد تسلط تفاوت محسوسی وجود داشته باشد.
- ۲- برای هر هدف آموزشی با هدف یادگیری تا حد امکان سؤالهای زیادتری بنویسید.
- ۳- هدفهای خود را تا آنجا که ممکن است دقیق و مشخص بنویسید.
- ۴- از روش عینی نمره گذاری استفاده کنید.
- ۵- سؤال های نمونه تمرینی شبیه به سؤال های آزمون اصلی در اختیار آزمون شوندگان قرار دهید.
- ۶- سؤال های آزمون را سازمان یافته و روشن در اختیار آزمون شوندگان قرار دهید.
- ۷- به فرض این پیش بینی که بین نمرات اختلاف دیده خواهد شد، از روش هایی که پایایی آزمونهای وابسته به هنجار را بالا می برد استفاده کنید.

رابطه روایی و پایایی

• اگر آزمون پایا نباشد، تلاش برای توصیف روایی آن بیهوده خواهد بود.

• ملاک اصلی ارزشیابی هر روش اندازه گیری را روایی آن تشکیل می دهد و پایایی از آن جهت مهم است که شرط لازم برای روایی است.

• ضریب پایایی بیشینه مقدار ممکن روایی را مشخص می کند. در واقع بیشینه مقدار نظری روایی هر آزمون برابر است با جذر ضریب پایایی (که به آن شاخص پایایی گفته می شود).

• همچنین هرگاه ضریب پایایی یک آزمون کم باشد (مثلا ۰/۳۰) هیچ آزمونی نمی تواند با ملاک مذکور همبستگی قابل توجهی داشته باشد؛ زیرا آزمون از ثبات لازم برخوردار نیست.

پایان