

توضیحات:

- ویژه آزمون آموزش و پرورش
- خلاصه + نکات مهم
- حیطه اختصاصی
- با قابلیت پرینت

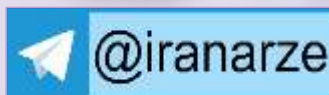
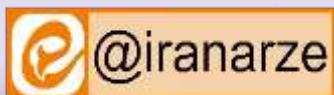
جزوه خلاصه و نکات مهم کتاب

اندازه گیری، سنجش و ارزشیابی آموزشی

برای دانلود رایگان جدیدترین سوالات استخدامی بخش های زیر، روی آن کلیک نمایید:

سنجش و اندازه گیری پیشرفت تحصیلی - راهنمای معلم در ارزشیابی توصیفی - سوالات آموزش و پرورش

« انتشار یا استفاده غیر تجاری از این فایل، بدون حذف لوگوی ایران عرضه، مجاز می باشد »



❖ فصل اول: خلاصه اندازه گیری، سنجش و ارزشیابی آموزشی تالیف ایران عرضه - صفحه ۲

❖ فصل دوم: نکات مهم اندازه گیری، سنجش و ارزشیابی آموزشی تالیف ایران عرضه - صفحه ۱۵

فصل اول: خلاصه اندازه گیری، سنجش و ارزشیابی آموزشی تالیف ایران عرضه

اندازه گیری عبارت است از فرایندی که تعیین می کند یک شخص یا یک شی چه مقدار از یک ویژگی برخوردار است. اندازه گیری فرایند یا شیوه دقیق و نظام دار است. اندازه گیری فرایند یا شیوه ای دقیق و نظام دار است. ناتالی در تعریف اندازه گیری به دست داده بر کاربرد قاعده تاکید کرده است: "اندازه گیری دربردارنده قواعدی است برای نسبت دادن اعداد به اشیا به گونه ای که صفتی از آنها را به صورت کمیت نشان دهد.

تعریف آزمون: گفتیم اندازه گیری یک فرایند است. این فرایند نیاز به وسیله ای دارد که به آن وسیله اندازه گیری میگویند. ویژگیها یا صفات روانی مانند هوش، خلاقیت، انگیزش، نگرش، یادگیری و مانند اینها را باید به طور غیرمستقیم اندازه گیری کرد زیرا هیچ کس به طور مستقیم به ویژگی ها دسترسی ندارد. معمول ترین وسیله برای اندازه گیری ویژگی ها یا صفات روانی آزمون است.

رفتارهای معرف ویژگی های روانی بسیار متنوع هستند و بنابراین نمی توان همه آنها را اندازه گیری کرد.

در بعضی موارد قاعده اندازه گیری تا حدی ساده است. فایده بیان قواعد اندازه گیری این هست که کمک می کند تا افرادی که می خواهند وسایل اندازه گیری را مورد استفاده قرار دهند آن را یکسان به کار بگیرند. چنانچه که گفتیم اندازه گیری یک فرایند است این فرایند به وسیله هایی نیاز دارد که به آن وسیله اندازه گیری می گویند.

بین آزمون و پرسشنامه تفاوت هایی وجود دارد که به وسیله آزمون می توانیم توانایی ها را اندازه گیری کنیم اما پرسشنامه ویژگی های غیر توانایی را ارزیابی می کند. اشتراک آزمون و پرسشنامه این هست که این دو ابزار کمی اند، یعنی اندازه گیری به وسیله آنها در قالب عدد و رقم بیان می شود. وقتی در اندازه گیری یکی از ویژگی های روانی یا تربیتی گروه یا فردی از آزمون استفاده می شود، این فعالیت آزمودن نامیده می شود

تعریف آزمودن: بنا بر تعریف، وقتی که برای اندازه گیری یکی از ویژگی های روانی یا تربیتی یک فرد یا گروهی از افراد از آزمون استفاده می شود به این فعالیت آزمودن می گویند.

تعریف سنجش: اصطلاح دیگر نزدیک به اندازه گیری و آزمودن اصطلاح سنجش است. سنجش از آزمودن و اندازه گیری مفهوم گسترده تری دارد. بنا بر گفته نیتکو سنجش یک اصطلاح کلی است و به صورت فرایندی تعریف میشود که برای گردآوری اطلاعات مورد نیاز تصمیم گیری درباره دانش آموزان، برنامه های درسی، و سیاست های آموزشی مورد استفاده قرار می گیرد.

وقتی اندازه گیری فقط با آزمون یا پرسشنامه انجام شود، آزمودن صورت می گیرد. سنجش یک اصطلاح کلی از آزمودن و اندازه گیری است.

تعریف ارزشیابی (یا ارزیابی): اصطلاح ارزشیابی به طور ساده به تعیین ارزش برای هر چیزی یا داوری ارزش کردن گفته میشود. ارزشیابی به یک فرایند نظام دار برای جمع آوری، تحلیل، و تفسیر اطلاعات گفته می شود. یکی از ویژگی های مه ارزشیابی تعیین کیفیت است.

اصطلاح سنجش بیشتر در روانشناسی بالینی و آموزش و پرورش کاربرد دارد. بین سنجش و اندازه گیری و آزمودن تفاوتی وجود دارد. اندازه گیری و آزمودن با کمیت سرکار دارند اما سنجش فقط محدود به کمیت نیست. ارزشیابی یعنی تعیین ارزش و داوری کردن.

یکی از ویژگی های مهم ارزشیابی تعیین کیفیت می باشد.

دو شخص از صاحب نظران در حوزه ارزشیابی ورتن و سندرز هستند. در ارزشیابی آموزشی تناسب داشتن موضوع با هدف مورد نظر، کیفیت را تعریف می کند.

ارزشیابی ها درست بر پایه اطلاعات دقیق استوار هستند. سنجش به جمع آوردن اطلاعات در مورد ویژگی عاطفی می پردازد. هدف ارزشیابی تصمیم گیری است ولی هدف پژوهش استنتاج است.

مقیاس به مجموعه‌ای از اعداد گفته می‌شود که ویژگی آنها به ویژگی‌های تجربی اشیایی که اعداد به آنها نسبت داده می‌شود، منطبق است. مقیاس اندازه‌گیری را به ۴ دسته تقسیم می‌کند. اسمی، ترتیبی، فاصله‌ای، نسبی. مقیاس اسمی ساده‌ترین نوع اندازه‌گیری است که به دو صورت اسم گذاری و طبقه بندی انجام می‌شود.

مقیاس اسمی دو عملیات مجاز دارد:

عملیات مجاز آماری، عملیات مجاز ریاضی.

مقیاس ترتیبی مجموعه‌ای از رتبه‌ها می‌باشد. دسته‌ای از اشیا یا افراد با توجه به یک صفتی که دارند از بزرگ به کوچک مرتب می‌شوند و بالعکس.

عملیات مجاز آماری به شمارش فراوانی محاسبه میانه محاسبه ضریب همبستگی و ... می پردازد. عملیات مجاز ریاضی هیچ کدام از چهار عمل جمع، تفریق، ضرب و تقسیم در این مقیاس میسر نیست.

مقیاس فاصله‌ای مقیاسی هست که رتبه اشیا با توجه به صفت مشخص است، اشیا از لحاظ صفت مورد اندازه‌گیری قرار می‌گیرند. تمایز مقیاس فاصله‌ای از ترتیبی این هست که در فاصله‌ای بین واحدها معلوم می‌باشد، اما در ترتیبی فاصله بین واحدها معلوم نیست. عملیات مجاز آماری در اینجا به محاسبه نما، میانگین، انحراف معیار و غیره می‌پردازد. عملیات مجاز ریاضی به جمع و تفریق مجاز است و ضرب و تقسیم در آن مجاز نیست.

مقیاس نسبی مقیاسی می‌باشد که رتبه اشخاص با توجه به یک صفت معلوم است فاصله بین اشخاص معلوم است و علاوه بر این فاصله حداقل یکی از اشخاص از یک صفر مطلق نیز معلوم است. یکی از مهمترین مقیاس‌های نسبتی همان مقیاس اعداد از صفر تا بی‌نهایت می‌باشد.

برای استفاده از مقیاس و تصمیم‌گیری از اینکه چه نوع مقیاسی را مورد استفاده قرار دهیم وابسته از به اینکه ما چگونه از مقیاس استفاده کرده باشیم دو نوع مقیاس دیگر داریم به نام:

مقیاس پیوسته و مقیاس گسسته.

مقیاس گسسته متغیر مورد نظر فقط می‌تواند یکی از ارزش‌های معین را به خود اختصاص دهد. یک ویژگی مهم این نوع مقیاس تساوی واحدهای شمارش آنها در طول مقیاس است. متغیر دیگر که بیانگر مقیاس گسسته هست عبارت است از جنس، وضعیت تاهل، سطح تحصیلات و ...

در مقیاس پیوسته متغیر مورد نظر می‌تواند در فاصله بین نقاط مقیاس هر تعدادی ارزش را به خود اختصاص دهد. متغیرهایی که بیانگر مقیاس پیوسته هستند شامل سرعت، زمان، طول، وزن و ... می‌باشند. اندازه‌گیری‌هایی که از متغیر گسسته به دست می‌آید دقیق و قطعی است اما اندازه‌گیری‌هایی که از متغیر پیوسته به دست می‌آید تقریبی می‌باشد. یک ویژگی مهم مقیاس پیوسته تساوی واحدهای اندازه‌گیری است.

حدود واقعی ارزش عدد یک متغیر پیوسته برابر می‌باشد با همان عدد به اضافه و منهای نصف واحد اندازه‌گیری.

یک مشکلی که متخصص‌های اندازه‌گیری در کار با اعداد با آن روبرو می‌شوند تعداد زیاد ارقام اعشاری می‌باشد. برای حل این مشکل از روش گرد کردن استفاده می‌شود.

روش اندازه‌گیری با توجه به ویژگی‌های مختلف به دو دسته آزمون‌های سنجش توانایی و پرسشنامه‌های سنجش عاطفی تقسیم می‌شود. آزمون‌های اندازه‌گیری توانایی به دو دسته توانایی‌های شناختی و توانایی‌های روانی-حرکتی تقسیم می‌شود. ویژگی مهم آزمون توانایی این هست که هنگام اجرای آن از آزمون شونده خواسته می‌شود تا حداکثر

سعی خود را بکند تا بهترین نمره ممکن را به دست آورد. آزمون توانایی به دو دسته آزمون‌های استعداد و آزمون‌های پیشرفت تحصیلی تقسیم می‌شود. آزمون استعداد توانایی یا آمادگی فرد را برای انجام دادن کارهایی که در پیش داره می‌سنجد، اما آزمون پیشرفت تحصیلی دانش و مهارت را که فرد تا لحظه اجرای آزمون کسب کرده، اندازه می‌گیرد.

این آزمون‌های استعداد و پیشرفته تحصیلی را نمی‌توان از همدیگر جدا کرد زیرا آن چیزی را که فرد در گذشته آموخته است می‌تواند برای پیش‌بینی موفقیت خود در آینده مورد استفاده قرار دهد. این آزمون‌های استعداد را برای پیش‌بینی موفقیت‌های آینده می‌توان به کار برد در حالی که پیشرفت تحصیلی برای تعیین مقدار آموخته‌ها یا

توانایی‌های موجود مورد استفاده قرار می‌دهند. معروف‌ترین آزمون استعداد آزمون هوش یا استعداد کلی هست. اهداف بیشتر آزمون‌های هوش اندازه‌گیری آن دسته از توانایی ذهنی است که در همه فعالیت‌های فکری و شناختی ضروری‌اند.

هدف از کاربرد آزمون هوش پیش بینی موفقیت‌های همه جانبه افراد است در صورتی که آزمون‌های استعداد های ویژه برای پیش بینی موفقیت افراد در زمینه فکری و عملی مشخص به کار می‌رود. آزمون‌های شناختی و روانی توانایی فکری عملی افراد را اندازه می‌گیرند اما پرسشنامه‌های عاطفی ویژگی‌های عاطفی شخصیتی را می‌سنجند. پرسشنامه‌های سنجش عاطفی دارای تنوع زیادی هستند اما مهم‌ترین آنها پرسشنامه‌های شخصیت، نگرش علاقه است.

پرسشنامه شخصیت در واقع هست که شخصیت را به صورت مجموعه رفتارهای فرد در موقعیت‌های اجتماعی می‌توان تعریف کرد. یعنی علاوه بر رفتارهای آشکار ویژگی‌های غیر رفتاری مثل هوش پیشرفت قد و ... به این تعریف شخصیت اضافه می‌شود. پرسشنامه شخصیت به اندازه‌گیری متغیرهای انگیزش، مزاج، سازگاری و از این قبیل اختصاص دارد. پرسشنامه‌ها و ابزارهای اندازه‌گیری شخصیت به سه دسته تقسیم می‌شود. پرسشنامه‌های گزارش شخصی که در آن فرد نظر شخصی خود را بیان می‌کند. پرسشنامه جامعه‌سنجی که در آن افراد دیگر نظرشان را نسبت به فرد مورد سنجش بیان می‌کنند. فنون مشاهده رفتار فرد در یک موقعیت خاص مورد مشاهده قرار می‌گیرد. ابزارهای سنجش شخصیت را می‌توان به دو دسته تقسیم کرد. محرک سازمان یافته و محرک سازمان نیافته.

دو بخش دیگر شخصیت علاقه و نگرش است. علاقه و نگرش به دوست داشتن‌ها و دوست نداشتن‌های افراد اشاره می‌کند. علاقه به احساس فرد نسبت به یک فعالیت اشاره دارد در حالی که نگرش حاکی از احساس فرد نسبت به یک شیء یک نهاد اجتماعی یا یک گروه است. علاقه به ترجیح یک فعالیت به سایر فعالیت‌ها اشاره دارد در حالی که نگرش به معنی ترجیح دادن گروه‌ها نهادها یا اشیا است. برای سنجیدن علاقه و نگرش افراد روش‌های مختلفی وجود دارد معروف‌ترین وسیله اندازه‌گیری در این دو زمینه پرسشنامه‌های (خلاصه شده توسط ایران‌عرشه) علاقه و نگرش است.

ابزارهای اندازه‌گیری به دو دسته ابزارهای اندازه‌گیری میزان شده یا استاندارد شده و ابزارهای اندازه‌گیری معلم تقسیم می‌شوند. ابزار اندازه‌گیری میزان شده ابزارهایی اند که با استفاده از نمونه‌های بزرگ افراد و قالباً به وسیله موسسه‌ها یا آزمون‌سازی دولتی یا خصوصی یا مراکز آموزشی و برای اندازه‌گیری پیشرفت تحصیلی و هوش و استعداد و ... تهیه می‌شوند. در آزمون‌های میزان شده ملاک نسبی اندازه‌گیری مورد استفاده قرار می‌گیرد. اصطلاح استاندارد شده به این معنی است که آزمون دارای هنجار می‌باشد. آزمون‌های معلم ساخته آزمون‌های اند که معلمان برای سنجیدن پیشرفت تحصیلی دانش‌آموزان تهیه می‌کنند و در طول دوره آموزش و یا در پایان دوره به اجرا در می‌آورند. این آزمون‌ها برای تعیین میزان دقیق یادگیرندگان در رسیدن به هدف‌های دقیق آموزشی طرح‌ریزی می‌شوند که در آنها ملاک مطلق اندازه‌گیری به کار گرفته می‌شود. تهیه آزمون‌های میزان شده در مقایسه با آزمون‌های معلم ساخته به زمان و تشخیص بیشتر نیاز دارد

آزمون‌های میزان شده برای استفاده در مناطق مختلف و مقایسه بین دانش‌آموزان مدارس تهیه می‌شود. آزمون میزان شده به وجود داده‌های هنجاری اشاره می‌کند. هدف اصلی آزمون‌های میزان شده مقایسه است مقایسه یک فرد یا یک گروه با فرد یا گروه دیگر.

روش‌های مختلف اندازه‌گیری و سنجش در آزمون‌ها را می‌توان به نوع سوال یا ماده دسته‌بندی کرد. در اندازه‌گیری به سوال آزمون یا هر تکلیف یا تمرینی که به آزمون شونده داده می‌شود، ماده می‌گویند. وسایل مختلف اندازه‌گیری با توجه به نوع سوال یا ماده به بسته پاسخ و باز پاسخ تقسیم می‌شوند. در بسته پاسخ هم سوال و هم پاسخ تهیه می‌شود و در اختیار و از آن می‌خواهد تا در مورد آن تصمیم بگیرد. در باز پاسخ دانش آموز خود پاسخ سوال را تولید می‌کند. آزمون‌های بسته پاسخ و باز پاسخ به ترتیب عینی و غیرعینی نامیده می‌شود. به جای اصطلاحات بسته پاسخ و باز پاسخ از آزمون‌های گزینه ثابت و سنجش‌های عملکرد پیچیده استفاده کرده‌اند

گونه دیگر دسته‌بندی وسایل اندازه‌گیری تقسیم آن‌ها به دو گروه روش انتخاب پاسخ و روش ساختن پاسخ است. اصطلاح دیگری که در سال‌های اخیر به وسیله صاحب نظران به کار رفته است اصطلاح سنجش قیاس پذیر می‌باشد سنجش قیاس پذیر رفتار آزمون شونده در موقعیت‌های شبیه‌سازی شده با موقعیت زندگی واقعی مورد سنجش قرار می‌گیرد.

صاحب نظران ورتن و سندرز ارزشیابی تمام رویکردهای ارزشیابی آموزشی را به ۶ دسته تقسیم می‌کنند. رویکرد مبتنی بر هدف - رویکرد مبتنی بر مدیریت- رویکرد مبتنی بر مصرف‌کننده- رویکرد مبتنی بر نظر متخصصان- رویکرد مبتنی بر اختلاف نظر متخصصان- رویکرد طبیعت‌گرایانه و مبتنی بر مشارکت کنندگان.

این تقسیم بندی ششگانه بر تقسیم بندی دوگانه هاوس از ارزشیابی‌های آموزشی استوار است. این تقسیم بندی بیشتر مبنای فلسفی و اپیدئولوژی دارد. رویکرد ارزشیابی پیرو فلسفه عینیت‌گرایی است.

رویکرد شهود گرایي - کثرت گرایي بر این اندیشه استوار است که ارزش به تاثیر برنامه بر فرد افراد وابسته است نه بر اکثریت آنها. در رویکرد کثرت گرایي ارزشیابی آموزشی احساسات دریافت‌های فردی ملاک خوب و درست به حساب می‌آید. متعادل کردن و وزن دادن داورها و ملاک‌های مختلف در این رویکرد عمدتاً جنبه شهودی دارد که به آن عنوان شهودگرایانه دادند. به لحاظ اجرایی در ارزشیابی‌های شهودگرا و کثرت گرا تمام افراد و گروه‌هایی که تحت تاثیر برنامه آموزش بودند به عنوان داوران استفاده می‌شوند.

در رویکرد ارزشیابی مبتنی بر هدف ابتدا مقاصد یک فعالیت آموزشی تعیین می‌شود و بعد از آن در ارزشیابی کوشش انجام می‌گیرد تا نشان داده شود که آن مقاصد تا چه اندازه تحقق یافته‌اند. الگوی ارزشیابی تایلر اولین الگوی ارزشیابی آموزشی در تاریخ پیدایش آن به دهه ۱۹۳۰ میلادی برمی‌گردد. مفهوم تایلر از ارزشیابی آموزشی این هست که تعیین کند هدف آموزشی برنامه آموزشگاه با برنامه درسی تا چه میزان تحقق یافته است.

الگوی ارزشیابی آموزشی تایلر حوزه برنامه‌ریزی درسی را تحت تاثیر قرار داده است. الگوهای ارزیابی که از رویکرد مبتنی بر مدیریت آمدند برای کمک به مدیران و تصمیم گیرندگان به عمل آمده‌اند. پدید آوردندگان این الگوهای ارزشیابی بر یک رویکرد سیستمی تاکید دارند.

ورتن و سندرز گفتم موسسات یا افراد مستقلی که مسئولیت قبول کردند تا درباره فرآورده‌های آموزشی اطلاعات جمع آوری کنند یا دیگران را یاری دهند رویکرد ارزشیابی مبتنی بر مصرف کننده را ترویج کرده‌اند. فرآورده‌های آموزشی عبارتند از بسته‌های برنامه درسی، کارگاه‌های آموزشی، رسانه‌های آموزشی، مواد آموزشی و حتی خدماتی که به مدارس داده می‌شود. خلاصه وار می‌توان اشاره کرد که رویکرد مبتنی بر مصرف کننده برابر است با تعیین میزان مطلوب بودن یک فرآورده آموزشی. مایکل اسکریون معروف‌ترین شخصی است که در زمینه ارزشیابی آموزشی برای کمک به مصرف کننده کار کرده است. ورتن و سندرز الگوی ارزشیابی که برخاسته از نظر متخصصان است را به چهار طبقه تقسیم بندی می‌کند. نظام‌های بررسی تخصصی رسمی - نظام‌های بررسی تخصصی غیر رسمی - بررسی توسط جلسات میز گرد ویژه - بررسی توسط افراد چیره.

الگوی برخاسته از اختلاف نظر متخصصان به این فرض استوار است که تمام متخصصان ارزشیابی خواسته و ناخواسته تعصبات خود را در جریان ارزشیابی دخالت می‌دهند و برای به دست آوردن نتایج درست‌تر باید نظرهای موافق و مخالف متخصصان صاحب نظران را مورد استفاده قرار دهیم. رویکرد طبیعت گرایانه مبتنی بر مشارکت به تجارب دست اول در موقعیت‌ها و فعالیت‌های تربیتی آموزشی تاکید دارد.

منظور از ارزشیابی از دانش آموزان و دانشجویان عمدتاً این است که پیشرفت تحصیلی و میزان یادگیری آنها مورد ارزیابی واقع شود. عمده‌ترین ملاک در ارزشیابی آموزشی، هدف‌های آموزشی از پیش تعیین شده، هستند. ارزشیابی پیشرفت تحصیلی مهم‌ترین نوع ارزشیابی آموزشی است که تعیین می‌کند دانش‌آموزان یا دانشجویان تا چه میزان به هدف‌های آموزشی خود که از پیش تعیین شده بودند، رسیدند. ارزشیابی آموزشی بعضی مواقع برای تعیین اثربخشی برنامه درسی یا مواد آموزشی انجام می‌شود. ارزشیابی از برنامه‌های درسی و مواد آموزشی عمده‌ترین ملاک پیشرفته تحصیلی میزان یادگیری دانش‌آموزان و دانشجویان می‌باشد. ملاک دیگر ارزشیابی از برنامه درسی و مواد آموزشی هزینه آن‌هاست.

منظور از ارزشیابی از آموزشگاه این هست که تعیین شود که اهدافی که پیش بینی شده برای آموزشگاه به چه میزان تحقق پیدا کرده و نقاط قوت و ضعف آن چیست. مانند دیگر ارزشیابی‌ها در ارزشیابی از آموزشگاه پیشرفت تحصیلی دانش‌آموزان و دانشجویان مهم‌ترین بخش برنامه آزمون و منبع کسب اطلاعات است.

ارزیابی از کارکنان آموزشگاه یک ارزشیابی است که از همه ساکنانی که به طور مستقیم و غیر مستقیم مسئول بازده‌های آموزشی موسسه آموزشی هستند، انجام می‌شود.

مهم‌ترین نوع ارزشیابی از کارکنان آموزشگاه ارزشیابی معلم است. پیچیده‌ترین نوع ارزشیابی، ارزشیابی از کارکنان به ویژه از کار معلمان است.

آن چیزی که برای کمک به اصطلاح ارزشیابی یعنی برنامه یا روش آموزش مورد استفاده قرار می‌گیرد ارزشیابی تکوینی می‌باشد.

ارزشیابی آموزشی به تصمیم گیری برای مسائل مختلف کمک می‌کند این تصمیم گیری‌ها در رابطه با هدف هستند. این تصمیم گیری‌ها به سه دسته تقسیم می‌شوند. مرحله طراحی - مرحله فرایندی - مرحله فرآورده‌ای. مرحله طراحی شامل یک سری از تصمیم‌هایی است درباره فعالیت‌هایی که قرار است انجام بگیرد. در این مرحله اقداماتی از قبیل

تحلیل موقعیت تعیین هدفها توصیف پیش نیازها انتخاب یا تولید وسایل و ابزار مورد نیاز انتخاب طرح و ... انجام می شود. در اولین گام مرحله طرح ریزی موقعیت موجود را تحلیل می کنیم که از جمله کارهایی که برای این باید انجام دهیم جمع آوری سوابق اطلاعاتی و برآورد و تعیین موانع موجود است.

در گام اول همانطور که گفتیم تشخیص دادن موانع و محدودیتها برای تعیین هدف ارزشیابی بسیار ضروری است. در ارزشیابی پیشرفت تحصیلی برای اینکه موقعیت موجود را تحلیل کنیم، معلم باید از آموخته‌های قبلی دانش آموزان اطلاعات کافی کسب کرده باشد. هدف آموزشی در دو سطح می باشد، کلی و جزئی. هدفهای کلی را غایت و هدفهای جزئی را هدف می نامند. هدفها به دو دسته کلی هدف های فرایندی و هدفهای فرآورده‌ای یا تولیدی تقسیم می شوند. در ارزشیابی پیشرفت تحصیلی یادگیرندگان غالباً هدفهای فرآورده‌ای مورد استفاده قرار می گیرند.

این هدفهای فرآورده‌ای تر بازده‌های مورد انتظار را وصف می کنند و معمولاً به صورت تاثیرات مطلوبی در رفتار یادگیرندگان رخ می دهند و بیان می شوند. هدفهای فرایندی کمتر با رفتار یادگیرنده و بیشتر با راهبردها دیده می شود و به اجرا در می آید. تحقق هدفهای آموزشی به وجود رفتارهای ورودی وابسته می باشد. رفتارهای ورودی همان پیش نیازها هستند که یادگیرنده قبل از شروع به یادگیری هدفهای تازه باید آنها را کسب کرده باشد.

برای جمع آوری اطلاعات مهم و لازم در ارزشیابی آموزشی به ابزارهایی نیاز هست از جمله آنها می توان به آزمونهای پیشرفته تحصیلی، آزمونهای نگرش سنجش، یا فرمهای ثبت مشاهده‌ای را نام برد. تهیه وسایل اندازه گیری مانند آزمونها و پرسشنامهها کاری هست که به دقت و صلاحیت نیاز استراتژی یا راهبرد فعالیتی است که برای کمک به تحقق یافتن هدفهای برنامه به کار برده می شود. اولین تصمیم استراتژی این هست که معین کنیم توالی ارائه این موضوعها چگونه باشد.

نمونه دیگر از استراتژی آموزشی مرور یا بازنگری مطالب است. بسیاری از پژوهشها نشان داده که مرور مطالب درسی باعث نگهداری بیشتر آنها در حافظه می شود. طرحهای پژوهشی بسیار متفاوتند و به دو گونه طرحهای آزمایشی و طرحهای شبه آزمایشی تقسیم می شوند. گی برای تدارک جدول زمانی روش چارتری گانت را پیشنهاد داده است. دومین مرحله ارزشیابی مرحله فرایندی است. در این مرحله تصمیمهایی که مرتبط با اقدامات عملی و اجرایی گرفته می شود. هدف مهم و اصلی ارزشیابی در مرحله فرایندی هدایت فعالیت‌های آتی برای بهبود فعالیت های آموزش و یادگیری است. گی گفته است که مرحله فرایندی ارزشیابی همان چیزی است که ارزشیابی تکوینی نامیده می شود و هدف ارزشیابی تکوینی را کمک به اصلاح برنامه آموزشی می داند.

در مرحله فرآورده‌ای یا تولیدی برعکس مرحله فرایندی تصمیمات در پایان برنامه اتخاذ می شود. هدف این مرحله جمع آوری دادههایی است که متخصص ارزشیابی را در تصمیم گیری هایش درباره اثربخشی کلیه دانش آموزشی برنامه درسی و ... کمک نماید. در این مرحله مشخص می شود که هدفهایی که از پیش تعیین شده اند تا چه مقدار تحقق یافته اند. به طور کلی در مرحله فرآورده‌ای ارزشیابی تصمیمات مربوط به اثربخشی موضوع مورد ارزشیابی اتخاذ می گردند.

ارزشیابی به یک فرایند نظامدار برای جمع آوری و تحلیل و تفسیر اطلاعات گفته می شود برای اینکه معین شود آیا هدفهای مورد نظر تحقق پیدا کرده اند یا در حال تحقق یافتن هستند. غایتها له بیاناتی هستند که به صورت کلی و آرمانی می باشند که قصد و منظور از ایجاد و اداره دوره‌های آموزشی یک جامعه را نشان می دهند. در مقابل هدف کلی نوع دیگری هدف وجود دارد که دقیق تر است و معلمان و متخصصان آموزشی با توجه به هدفهای کلی و با استفاده از محتوای درسهای مختلف آنها را تهیه می کنند. هدف آموزشی یا هدف یادگیری یک درس مقصودی است که معلم از آموزش خود و یادگیرنده از یادگیری خود دارد. برای آموزش و هم برای ارزشیابی نیاز است هدفهای کلی به هدفهای دقیق تجزیه شوند.

هدف کلی کسب مهارت‌های مورد نیاز استفاده از وسایل معمولی برای اندازه گیری طول، حجم و وزن است. در تجزیه کردن هدفهای کلی به هدفهای جزئی نباید زیاده روی کرد زیرا هدفهای خیلی جزئی مفید نیستند. درجه اندازه پذیری کلید شناسایی هدفهای آموزشی خوب است. هدفی که اندازه پذیر باشد هرچه گسترده تر باشد بهتر است. استاندارد تحصیلی از جمله بیاناتی هست که مشخص می کند چه چیزی باید آموزش داده شود و دانش آموزان چه چیزی را باید یاد بگیرند. استانداردها معمولاً به شکل کلی (طراحی شده توسط ایران عرضه) بیان می شوند. برای هر استاندارد یک یا چند نشانگر می نویسند تا آموزش پذیر و سنجش پذیر بشود.

استانداردهای تحصیلی به دو گونه اند. استانداردهای محتوایی و استانداردهای عملکردی. نشانگرها در عمل همان استانداردهای عملکردی هستند زیرا با استفاده از آنها می توان نشان داد که یادگیرندگان به استانداردهای محتوایی رسیده اند یا نه. استاندارد و هدف بسیار به هم شبیه اند. استانداردها به هدفهای کلی و نشانگرها به هدفهای دقیق آموزشی مانند هستند.

هدفهای آموزشی ابتدا به سه طبقه کلی با نامهای حوزه شناختی، حوزه عاطفی و حوزه روانی-حرکتی تقسیم شده‌اند. هدفهای حوزه شناختی به جریان‌هایی که با فعالیت ذهنی و فکر آدم سر و کار دارد، مربوط می‌شود. حوزه عاطفی شامل هدفهایی است که تغییرات حاصل در علاقه‌ها نگرش‌ها و ارزش‌ها را نشان می‌دهد. حوزه روانی-حرکتی به مهارت‌های حرکتی یا حرکات بدنی مربوط است. هدفهای حوزه عاطفی و حوزه روانی-حرکتی با آزمون کتبی به خوبی قابل سنجش نیستند.

بلوم و همکارانش هدفهای حوزه شناختی را به ۶ طبقه کلی طبقه بندی کردند. دانش- فهمیدن- کار بستن- تحلیل- ترکیب- ارزشیابی. هر کدام از این طبقه‌ها یک نوع فرایند شناختی را توصیف می‌کند.

در طبقه‌بندی تازه حوزه شناختی یک بعد دانش و یک بعد فرایند شناختی وجود دارد. بعد دانش شامل دانش امور واقعی، دانش مفهومی، دانش فراشناختی و ... است. بعد فرایند شناختی دربرگیرنده به یاد آوردن، فهمیدن و کار بستن و آفریدن است.

بلوم و همکاران هدفهای حوزه عاطفی را طبقه بندی کرده‌اند. این طبقه‌بندی شامل ۵ طبقه است. دریافت کردن- پاسخ دادن- ارزش گذاشتن- سازمان دادن به ارزش‌ها- شخصیت پذیرفتن.

برای حوزه روانی حرکتی تعداد طبقه‌بندی وجود دارد که در اینجا طبقه‌بندی سیمپسون مورد نظر می‌باشد که شامل ۷ طبقه اصلی و تعدادی خورده طبقه است. ادراک- آمادگی- پاسخ هدایت شده- مکانیسم- پاسخ پیچیده آشکار- انطباق- ابتکار.

مهمترین گام در تهیه آزمون‌های پیشرفت تحصیلی تهیه جدول مشخصات برای موضوعی می‌باشد که قرار است آزمون از آن تهیه شود. جدول مشخصات یک جدول دوبعدی است با تعدادی ردیف و ستون متناسب با محتوا و هدفهای آموزش درس.

یک آزمون خوب رفته تحصیلی آزمون می‌باشد که به بهترین شکل منعکس کننده تمام هدفهای آموزشی و محتوای برنامه درسی باشد. یک آزمون خوب یا به اصطلاح یک آزمون دارای بی خواهد بود که سوال‌های آن حاوی نمونه درستی از هدفها و محتوای درس باشد.

در آزمون‌هایی که صحیح غلط نامیده می‌شوند هر سوال از یک جمله تشکیل شده است که آزمون شونده آن را به یکی از صورت‌های صحیح-غلط، درست-نادرست، بله-خیر و ... پاسخ می‌دهد. سوال‌های صحیح-غلط شبیه به سوال‌های چند گزینه‌ای هستند اما تفاوتی هم دارند که عبارت است از سوال‌های چند گزینه‌ای آزمون شونده از میان گزینه‌های پیشنهادی فقط یکی را انتخاب می‌کند، اما در سوال‌های چند صحیح-غلط لازم است به همه گزینه‌ها جواب دهد.

نکته‌ای که در مورد سوال‌های چند صحیح غلط به یاد داشته باشیم این هست که در هر مجموعه همه انتخاب‌های صحیح یا غلط باید به موضوع یکسانی مربوط باشد. برای تهیه سوال‌های صحیح - غلط از سوی متخصصان قواعدی پیشنهاد شده است.

موضوع‌های مهم در ارتباط با هدفهای آموزشی را در سوال‌های خود بیاوریم.

سوال‌های طرح شود که درک و فهم را بسنجد نه یادآوری کلمه به کلمه مطالب را.

سوالاتی طرح شود که درست یا غلط بودن آن قابل دفاع باشد.

از جمله‌های پیچیده استفاده نشود..

در هر یک سوال بیشتر از یک موضوع قرار داده نشود.

تا جایی که امکان دارد از زبان کمی و دقیق استفاده شود نه از زبان کیفی و غیر دقیق.

تا جایی که ممکن است از طرح سوال‌های منفی پرهیز شود. از کلمات راهنما یا اشاره کننده به جواب درست استفاده در طرح سوال سعی شود تعداد سوال‌های غلط بیشتر از تعداد سوال‌های درست باشد. طول سوال‌های صحیح و سال‌های غلط تقریباً باید با هم، هم اندازه انتخاب شود.

از کسانی که این آزمون‌ها را مفید تشخیص داده‌اند می‌توان به ایبل اشاره کرد. از جمله محاسن سوال‌های صحیح-غلط به این شرح است: تهیه این سوال‌ها آسان است. تصحیح

جواب آزمون شوندگان به سادگی و سرعت انجام می‌گیرد. به این خاطر می‌توان تعداد زیادی سوال صحیح غلط در جلسه امتحان مورد استفاده قرار داد. ایراداتی نیز به

سوال‌های صحیح-غلط وارد شده است که از جمله آن می‌توان به اینکه همه موضوع‌های درسی را نمی‌توان به صورت عبارت‌های صحیح یا غلط بیان کرد و اینکه میزان بالای حدس پذیری آنها یکی دیگر از این مشکلات می‌باشد.

هر آزمون جور کردنی از تعداد تمرین تشکیل می‌شود و هر تمرین سه بخش دارد. راهنمای انجام دادن تمرین - فهرست محرک‌ها یا پرسش‌ها - فهرست پاسخ‌ها. سوال‌های جور کردنی بندی از نوع بالا با سوال‌های جور کردنی معمولی از لحاظ تعداد ماده‌های فهرست پرسش و پاسخ فرق دارند. یکی از امتیاز سوال‌های جور کردنی از نوع رده‌بندی یا ماده فهرست این است که در تصحیح جواب‌ها می‌توان از پاسخنامه‌های جداگانه استفاده کرد. نکته‌ای که در رابطه با سوال‌های جور کردنی م این هست که فهرست پرسش‌ها و فهرست پاسخ‌ها را هم می‌توان در ستون موازی در کنار هم قرار داد و هم می‌توان آنها را پشت سر هم نوشت.

به طور کل در راهنمای سوال سه دسته اطلاع را باید قرار داد. اساس رابطه بین پرسش‌ها و پاسخ‌ها یا مبنای جور کردن آنها - روش مشخص کردن پاسخ‌ها یا جور کردن پرسش‌ها و پاسخ‌ها - و اینکه از یک پاسخ تنها یک بار استفاده شود یا بیشتر از یک بار.

تجانس یا همگونی یک امر نسبی هست طراحان آزمون ممکن است در نوشتن سوال‌های آزمون‌های جور کردنی با مشکل مواجه شوند زیرا آنچه برای یک گروه از آزمون شوندگان متجانس است ممکن است برای گروه دیگر نامتجانس باشد.

آزمون‌های جور کردنی در قیاس با آزمون‌های صحیح-غلط برای سنجش دامنه وسیع‌تری از هدف‌های آموزشی قابل استفاده است. عمده‌ترین مورد استفاده آزمون‌های جور کردنی سنجش هدف‌هایی است که روابط بین امور را شامل می‌شود.

از آزمون‌های جور کردنی می‌توان برای سنجش توانایی تشخیص روابط در شکل‌ها و نمودارها استفاده کرد. علاوه بر سنجش توانایی تشخیص روابط بین امور از آزمون‌های جور کردنی نوع رده‌بندی یا ماده فهرست اصلی می‌توان برای سنجش درک مفاهیم استفاده کرد. همچنین از آزمون‌های جور کردنی می‌توان برای سنجش توانایی تحلیل استفاده کرد.

آزمون‌های چند گزینه‌ای از تعداد سوال درست می‌شوند و هر سوال چند گزینه‌ای از سه قسمت اصلی تشکیل می‌شود: تنه سوال - گزینه کلید - گزینه‌های انحرافی. گزینه‌های انحرافی پیشنهاد‌های غلطی هستند که نقش آنها منحرف کردن آزمون‌شونده از موضوع سوال است. این اصطلاحات متداول‌ترین اصلاحات مورد استفاده در سوال‌های چند گزینه‌ای هستند. تعداد گزینه سوال‌های چند گزینه‌ای از ۲ تا ۵ تغییر می‌کند. به لحاظ نظری هر چه تعداد گزینه‌ها بیشتر باشد امکان حدس زدن کمتر است.

قواعدی که برای تهیه سوال‌های چند گزینه‌ای پیشنهاد می‌شود بر اساس رز فکر تدوین شده است که یک سوال چند گزینه‌ای باید آزمون‌شونده را با بلا تکلیفی که هم مهم و هم کاملاً برای قابل فهم است مواجه کند هر سال باید طوری نوشته شود که تنها به وسیله کسانی که هدف مورد اندازه‌گیری را می‌دانند به درستی پاسخ داده شود و کسانی که تسلط بر مطلب ندارند نتوانند به آن پاسخ درست بدهند. زیرا هر سوال باید به گونه‌ای انتخاب شود که یک موضوع مهم با یک هدف آموزشی را اندازه بگیرد. بیشتر از یک هدف یا یک مطلب در هر سوال قرار داده نشود.

سوال‌ها کاملاً روشن و در حد درک آزمون‌شوندگان نوشته شود.

از تکرار مطلب‌ها در گزینه‌ها جلوگیری شود.

مطالب اصلی سوال به طور کامل در تنه سوال نوشته شود.

پکوشید که همه گزینه‌های یک سوال متجانس و به یک موضوع واحد مربوط باشند. تجانس گزینه‌ها یکی از راه‌های کنترل سطح دشواری سوال‌ها است که کلی این هست که هر چه گزینه‌ها متجانس‌تر باشند به همان نسبت سوال دشوارتر است.

سوال را طوری بنویسید که پاسخ درست تنها پاسخ درست یا قطعاً درست‌ترین پاسخ باشد.

برای اندازه‌گیری فرایندهای پیچیده ذهنی از موقعیت‌های تازه استفاده کنید.

گزینه‌های انحرافی را طوری بنویسید که توجه آزمون‌شوندگان بی‌اطلاع از موضوع سوال را به خود جلب کنند.

برای نوشتن گزینه‌های انحرافی پیشنهاداتی بیان شده است: از غلط‌های معمول دانش‌آموزان استفاده کنید. از کلماتی که با تنه سوال تداعی کلامی دارند استفاده کنید. از

گزینه‌های انحرافی که از لحاظ محتوا متجانس و شبیه گزینه درست است استفاده کنید و ...

گزینه هر سوال را طوری بنویسید که از نظر دستوری و جمله‌بندی به نحو درست مکمل متن سوال باشد.

از به کار بردن اشاره‌های دستوری نامرتبط خودداری شود.

در سوالات منفی کلمات منفی را برجسته جلوه دهید تا متوجه منفی بودن سوال باشند مثلاً خط بکشید.

از نوشتن سوال‌هایی که در آنها متن سوال منفی و گزینه‌ها هم منفی هستند جلوگیری کنید.

از کاربرد عبارت‌هایی مانند همه آنچه در بالا گفته شد و تمامی موارد بالا و غیره خودداری کنید.

تا حد ممکن از کاربرد گزینه هیچ یک از موارد بالا پرهیز کنید این گزینه در ارتباط با سوال‌های درس‌های ریاضیات مورد استفاده قرار می‌گیرد.

سوال‌ها را مستقل از یکدیگر بنویسید پاسخ دادن به یک سوال نباید شرط لازم برای پاسخ دادن به سوال‌های بعدی باشد.

از طرح سوال گمراه کننده پرهیز کنید. استفاده از سوال‌های گمراه کننده در آزمون‌های پیشرفت تحصیلی به هیچ وجه جایز نیست.

دو گزینه متضاد را که یکی از آنها درست است به کار نبرید بهتر است به طور کلی از کاربرد متضادها صرف نظر شود یا دو زوج متضاد برای سوال‌های چهار گزینه‌ای مورد استفاده قرار گیرد.

طول گزینه درست را در سوال مختلف تغییر دهید. گزینه درست در مقایسه با گزینه انحرافی باید جوری نوشته شود که به هیچ وجه نمایان نباشد.

محل گزینه درست را در میان گزینه‌های انحرافی به طور تصادفی انتخاب کنید. ساده‌ترین راه برای انتخاب محل گزینه درست استفاده از شماره صفحات یک کتاب است.

برای هر سوال بین ۳ تا ۵ گزینه در نظر بگیرید. مهترز و لهمان نهاد کردند که بهتر است برای اندازه‌گیری پیشرفت تحصیلی دانش آموزان از سوال‌های دارای گزینه کمتر استفاده شود.

در آزمون‌های تکوینی از گزینه نمی‌دانم استفاده کنید.

در سال‌هایی که تنه آنها جمله ناتمام است جای خالی را در قسمت آخر جمله قرار دهید.

تا آنجا که ممکن است در تنه سال به جای جمله ناتمام از جمله پرسشی استفاده کنید. جمله پرسشی هدف سوال را به طور آشکار بیان می‌کند تا جمله ناتمام.

یکی از کاربردهای ابتکاری سوال‌های چند گزینه‌ای استفاده از رایانه است. این آزمون‌ها تاکنون در حوزه آموزش پزشکی بیشترین مورد استفاده را داشتند و معلوم شده که وقتی با آزمون‌های سنتی به کار می‌رود روش بسیار مفید و معتبری است. به طور کلی برای امتیازهای آزمون‌های چند گزینه‌ای می‌توان به مواردی اشاره کرد. این آزمون‌ها از بقیه آزمون‌های عینی انعطاف پذیرتر هستند. آزمون‌های چند گزینه‌ای نسبت به آزمون‌های صحیح_غلط کمتر امکان حدس زدن کورکورانه را می‌دهند. پاسخ این سوال‌ها به سادگی و با عینیت کامل قابل تصحیح هستند. معلمان و دانش آموزان این نوع آزمون را به سایر آزمون‌های عینی ترجیح می‌دهند.

محدودیت‌های آزمون‌های چند گزینه‌ای شامل: ساختن این آزمون‌ها بسیار دشوار است. در مقایسه با بعضی آزمون‌های دیگر خواندن این آزمون و پیدا کردن گزینه درست وقت زیادی را می‌گیرد. معلمان اغلب سوال‌های چند گزینه‌ای را در حد سنجش اطلاعات جزئی می‌نویسند.

آزمون‌های چند گزینه‌ای بهترین ابزارهای سنجش گزینش پاسخ یا بهترین نوع آزمون‌های بسته پاسخ هستند.

آزمون تشریحی پیش همه ما به خوبی شناخته شده است. این آزمون چند ویژگی مهم دارد: هیچ پاسخ واحدی برای سوال نمی‌توان در نظر گرفت که صد درصد کامل و درست باشد. پاسخ‌ها از لحاظ درجه کیفیت درستی با هم فرق دارند. در این سنجش به آزمون شونده آزادی عمل داده می‌شود. در آزمون تشریحی هیچگونه محدودیتی برای آزمون شونده منظور نمی‌شود. آزمون شونده در پاسخ دادن به سوال این آزمون هم از لحاظ زمان و هم مقدار پاسخ آزادی کامل دارد.

سوال‌های آزمون‌های تشریحی گسترده پاسخ برای سنجش هدف‌های تحلیل و آفرینندگی مناسب‌ترین سوال است. در آزمون تشریحی محدود پاسخ آزمون شونده در دادن پاسخ به سوال آزادی کامل ندارد و باید پاسخ خود را در چهارچوب شرایط خاصی محدود کند. بنابراین سوال‌های آزمون محدود پاسخ برای اندازه‌گیری توانایی یادگیرندگان در سطوح فهمیدن، کار بستن و تحلیل مناسب اند. ولی برای سنجش توانایی ترکیب و ارزشیابی مفید نیستند.

تهیه سوال‌های تشریحی قواعدی دارد که از جمله آن می‌توان به این موارد اشاره کرد.

دقت کنید که سوال‌ها به طور مستقیم و به هدف‌های آموزشی مرتبط باشند.

استفاده از آزمون‌های تشریحی تنها به اندازه‌گیری هدف‌های محدود کنید که با سایر انواع آزمون‌ها به خوبی قابل اندازه‌گیری نیستند.

صورت سوال‌های تشریحی را با عبارات و کلمات واضح و روشن بنویسید.

تا حد امکان از سوال‌های تازه و موقعیت‌های جدید استفاده کنید.

به آزمون شوندگان حق انتخاب چند سوال از میان تعداد سوال را ندهید. از جمله مشکلات این مورد می‌توان به اینکه تهیه سوال‌هایی که از نظر دشواری مانند هم باشند کار آسانی نیست. آزمون شونده توانایی انتخاب سوال‌هایی را که جواب آنها بهتر می‌داند، ندارد. آزمون شوندگان قوی‌تر بیشتر به سراغ سوال‌های دشوار می‌روند و ممکن است داشتن حق انتخاب در نهایت به ضرر این افراد تمام شود.

دلیل عمده مخالفت برخی از صاحب نظرها با به کار بردن این آزمون به خاطر بی‌ثباتی نمره گذاشتن آنها می‌باشد.

در روش تحلیلی نمره گذاری که به آن روش امتیازبندی گفته می‌شود پاسخ نمونه خود را به اجزای کوچک‌تری تقسیم کنید و برای هر جز مشخص نمره یا امتیاز جداگانه در نظر بگیرید. در روش کلی نمره گذاری که به آن روش درجه‌بندی گفته می‌شود پاسخ نمونه به اجزا و قسمت‌های کوچک تقسیم نمی‌شود بلکه تنها به صورت یک معیار به کار می‌رود. ایبل درباره نمره گذاری به روش کلی گفته است که در مجموع این روش نمره گذاری آسان‌تر و سریع‌تر از روش تحلیلی است.

روش دیگر نمره گذاری آزمون تشریحی روش ویژگی‌های اصلی نام دارد در این روش معلم ویژگی اصلی پاسخ دانش آموزان و هر سوال را سنجش می‌کند و برای هر یک از آنها نمره‌های ۰ تا ۴ که معرف عالی تا غیر قابل قبول‌اند را منظور می‌کند. این نوع آزمون‌ها امتیازهای آنها به این شرح است: تهیه آزمون‌های تشریحی آسان است. این آزمون توانایی تولید پاسخ را می‌سنجد نه توانایی انتخاب پاسخ را. این آزمون‌ها موقعیت‌های واقعی‌تری را در اختیار آزمون شوندگان می‌گذارد. این آزمون‌ها بر روش مطالعه یادگیرندگان تاثیر مثبت به جا می‌گذارد.

در نتیجه آزمون‌های تشریحی نمونه ضعیفی از محتوا و هدف‌های درس را اندازه می‌گیرند. مشکل دیگر آزمون‌های تشریحی ذهنی بودن نمره گذاری مصححان و یکنواخت نبودن ارزشیابی آنان است.

یکی از تدابیر متداول معلمان در ارزشیابی غیر رسمی از یادگیری دانش‌آموزان و رفع مشکلات یادگیری آنان استفاده از پرسش شفاهی است. مہرنز و مهمان کاربرد پرسش‌های شفاهی را برای مقاصد تشخیصی پیشنهاد می‌کند.

آزمون کوتاه پاسخ در بین آزمون‌های عینی و تشریحی قرار دارد. آزمون کوتاه پاسخ تعدادی سؤال مختصر را در بر می‌گیرد که پاسخ‌های آنها به یک عدد، علامت، کلمه، عبارت یا جمله محدود می‌شود. آزمون‌های کوتاه پاسخ را به سه دسته پرسشی، کامل کردنی و تشخیصی طبقه بندی می‌کنند.

در آزمون کوتاه پاسخ نوع پرسشی سؤال به صورت یک جمله پرسشی طرح میشود و آزمون شونده پاسخ کوتاهی به آن میدهد. در آزمون کوتاه پاسخ نوع کامل کردنی سؤال به صورت یک جمله ناقص نوشته میشود که در آن یک جای خالی وجود دارد و آزمون شونده با پر کردن جای خالی به سؤال جواب میدهد و در آزمون کوتاه پاسخ نوع تشخیصی یا تداعی صورت سؤال ارائه می‌شود و آزمون شونده باید جواب آن را تداعی کند.

علاوه بر این‌ها نوع دیگری از آزمون‌های کوتاه پاسخ وجود دارد که به آزمون‌های کوتاه پاسخ تشکیل یافته از مسائل عددی شهرت دارند. اینگونه آزمون در ریاضیات، علوم، حسابداری و زمینه‌های دیگری که با کاربرد اعداد سر و کار دارند قابل استفاده است.

صاحب نظران سنجش و اندازه‌گیری برای تهیه سوال‌های کوتاه پاسخ قواعدی را به دست دادند.

برای آزمون‌های کوتاه پاسخ امتیازها و محدودیت‌هایی را قائل شده‌اند. مهم‌ترین امتیازهای این سال‌ها به شرح زیر می‌باشد: سهولت تهیه و اجرا. عینی بودن تصحیح پاسخ سوال‌ها در مقایسه با آزمون‌های تشریحی. قابلیت استفاده با مسائل ساده عددی. کاهش دادن تقلب. فراهم آوردن اطلاعات تشخیصی بیشتر از آزمون‌های عینی.

در مقابل امتیازهای بالا برای آزمون‌های محدود پاسخ محدودیت‌هایی را نیز ذکر کردند در زیر به مهم‌ترین آنها توجه کنید: یادگیری سطح پایین را می‌سنجد. استفاده زیاد از

آنها سبب تشویق دانش آموزان به یادگیری اطلاعات جزئی می‌شود. در مقایسه با آزمون‌های عینی پاسخ سوال‌ها با سرعت کمتری قابل تشخیص است. ص ۲۵۵

در رابطه با موارد استفاده آزمون‌های عینی باید گفت که این آزمون‌ها عمدتاً برای سنجش هدف‌های سطوح پایین یادگیری مفیدند. می‌توان این گونه سوال‌ها را طوری تهیه کرد که هدف‌های طبقه‌های فهمیدن و کار بستن را نیز سنجش کنند. به ویژه از آزمون‌های تشکیل یافته از مسائل ساده عددی به خوبی می‌توان در سطح کار بستن استفاده کرد.

یک دسته از روش‌های سنجش واقعی یا اصیل سنجش عملکردی نام دارد.

آزمون‌های عملکردی یا فرایند یا شیوه اجرا، فراورده با محصول کار و یا ترکیبی از این دو سر و کار دارند. ماهیت عملکرد مورد سنجش مشخص می‌کند که تاکید بر چه قسمتی باید باشد. فراورده چیزی است که یادگیرنده تولید می‌کند و فرایند جریانی است که او مورد استفاده قرار می‌دهد یا به فراورده دست یابد.

ویژگی آنها به چهار دسته تقسیم می‌شود. تاکید بر کاربرست- تاکید بر سنجش مستقیم- استفاده از مسائل واقعی- ترغیب و تشویق تفکر باز.

منظور از آزمون شناسی یک روشی می‌باشد برای سنجش توانایی یادگیرنده در تشخیص ویژگی‌ها محاسن معایب و موارد مورد استفاده امور مختلف.

در روش نمونه کار از یادگیرنده درخواست می‌شود تا کاری را انجام دهد که معرف عملکرد واقعی مورد سنجش باشد. در سنجش ساختارمند معلم هم بر تکالیف سنجش و هم بر شرایطی که در آن دانش آموز باید پاسخ دهد کنترل دارد.

پروژه‌های طولانی مدت به دو قسمت پروژه‌های فردی و پروژه‌های گروهی تقسیم بندی می‌شوند.

از محدودیت‌ها و امتیازهای روش عملکردی می‌توان به وسیله آن یادگیری‌ها را سنجید. با استفاده از آن می‌توان عمل یادگیرنده و جریان کار او را مورد سنجش قرار داد. در این نوع سنجش به نظریه‌های جدید یادگیری تاکید می‌شود.

کار پوشه به یک مجموعه از کارهای یادگیرنده گفته می‌شود که نشان دهنده کوشش و موقعیت و پیشرفت او در یک زمینه خاص می‌باشد.

کار پوشه حاوی کارهایی است که در یک مدت معین انتخاب شده است نه تمام کارها.

طبق گفته سانتروک به طور کل چهار دسته مدرک را می‌توان در کار پوشه دانش آموز قرار داد. کارهای انجام شده توسط دانش آموز- کارهای تهیه شده توسط دانش آموز- اظهار نظر دیگران درباره دانش آموز- تولیدات دانش آموز.

برای تهیه کار پوشه به چهار نکته باید توجه داشت. استفاده‌هایی که از کار پوشه به عمل خواهد آمد- کسانی که به آن دسترسی دارند- تناسب داشتن هدف تهیه کار پوشه با کارهایی که قرار است در کار پوشه قرار بگیرد- ملاک‌هایی که در ارزشیابی از کار پوشه به کار خواهد رفت.

از امتیازهای خاص این روش می‌توان به مواردی اشاره کرد. با این روش هم می‌توان سنجش کرد و هم آموزش داد- به کمک این روش می‌توان موفقیت یادگیرندگان را سنجش کرد و هم جریان رشد موفقیت آنها را- به یادگیرندگان کمک می‌کند تا مسئولیت انتخاب هدف خود را بپذیرند و بر جریان پیشرفت نظارت داشته باشند.

معروف‌ترین روش‌های سنجش مشاهده‌ای عبارتند از فهرست واری، مقیاس درجه‌بندی و واقعه‌نگاری. منظور از فهرست واری فهرستی است از رفتارها، ویژگی‌ها و فعالیت‌های یادگیرنده.

در روش فهرست واری تنها بود و نبود یا کمیت ویژگی مورد نظر ثبت می‌شود. در تهیه فهرست واری برای ارزشیابی از محصول رفتار بر جنبه‌ها یا ویژگی‌های قابل مشاهده محصول مورد سنجش تاکید می‌شود.

ویژگی‌های یک فهرست واری خوب عبارت است از فهرست نسبتاً کوتاه ده بر یک رفتار با ویژگی قابل مشاهده تاکید می‌کند تنها رفتارها یا ویژگی‌های مهم منظور می‌شود و ...

ساده‌ترین نوع مقیاس درجه‌بندی مقیاس درجه‌بندی عددی است. متخصص‌ها برای هر سه روش مشاهده‌ای راهنماهایی ارائه داده‌اند. روش‌های سنجش مشاهده‌ای دارای امتیازها و محدودیت‌هایی هستند که از جمله امتیازها می‌توان به این موارد اشاره کرد. آسان بودن استفاده با تکالیف و موقعیت‌های گوناگون- بالا بردن قدرت پیش بینی نتایج سنجش- به دست آوردن داده‌های کمی و اطلاعات کیفی و ... از جمله محدودیت‌هایی که برای روش سنجش مشاهده‌ای وجود دارد شامل این موارد می‌باشد. امکان نتیجه‌گیری‌های غلط- نشان ندادن علت‌های رفتار- دشواری ثبت مشاهدات همزمان- قطعیت غیر موجه.

رفتار هم دارای ویژگی‌های ظاهری و هم دارای ویژگی‌های زمانی است. به مجموعه این ویژگی‌ها بعدها رفتاری می‌گویند. شکل ظاهری رفتار یعنی اینکه رفتار در چه شکل یا به چه صورت رخ می‌دهد.

در سنجش رفتاری علاوه بر رفتار مورد نظر معمولاً محرک‌های پیش آیندی و محرک‌های پس آیندی مورد مشاهده قرار گرفته و ثبت می‌شوند زیرا در توصیف و تحلیل رفتارها برای مقاصد آموزشی بالینی مفیدند.

اگر دفعات وقوع یک رفتار را در یک فاصله زمانی معین شمارش کنیم از روش ثبت رویداد استفاده کردیم گاهی به این روش، روش ثبت فراوانی هم گفته می‌شود. روش ثبت فراوانی برای اندازه‌گیری رفتارهایی که مشخصاً بسیار مفید است.

با استفاده از روش ثبت دوره نرفتگی می‌توان مشخص کرد که چه زمانی طول می‌کشد تا شخص رفتاری را آغاز کند. روش ثبت دوره نرفتگی شبیه به روش ثبت طول مدت رفتار است. به روش ثبت تمام موارد یک رفتار در یک مدت مشخص روش ثبت مداوم می‌گویند.

در روش ثبت نمونه‌گیری زمان وقوع یا عدم وقوع یک رفتار در فاصله‌های کوتاه مشاهده و ثبت می‌شود. تنها تفاوت بین این دو روش در ثبت فاصله رفتار مورد نظر در یک مدت زمان مشخص است.

سنجش رفتاری یک سری امتیازها و محدودیت‌هایی دارد. با استفاده از سنجش رفتاری می‌توان ویژگی‌های رفتار مورد نظر را شناسایی و سنجش کرد یا استفاده از روش‌های سنجش رفتاری می‌توان داده‌های مورد نیاز پژوهش‌های آموزشی را گردآوری کرد. از این روش برای سنجش ویژگی‌های دانش آموزان می‌توان بهره برد. محدودیت‌های روش سنجش رفتاری عبارتند از استفاده از این روش وقتگیر و لازمه صرف نیروی زیاد است. در پاره مواقع به علت وجود پدیده واکنش پذیری این روش‌ها نیز از روایی کامل برخوردار نیستند و ...

ویژگی عاطفی باید ملاک‌هایی را دارا باشد که از جمله این ملاک‌ها، دربرگیرنده احساس یا هیجان باشد. ویژگی معمولی یا نوع عاطفی فرد باشد. هر ویژگی عاطفی باید دارای سه ملاک خاص شدت، جهت و هدف باشد. شدت درجه یا نیروی احساس یا عاطفه است. جهت به جنبه مثبت یا منفی احساس گفته می‌شود. هدف به موضوعی که عاطفه یا احساس متوجه آن است، اشاره می‌کند.

در سنجش عاطفی سعی می‌شود تا اطلاعات مربوط به ویژگی‌های عاطفی فرد از طریق سنجش رفتارهای به دست آید که جز رفتارهای نوعی او هست و به طور معمول از او سر می‌زند.

به طور کل روش‌های سنجش عاطفی را می‌توان به دو دسته تقسیم کرد. مشاهده گزارش شخصی. روش مشاهده یعنی جمع آوری اطلاعات از راه دیدن. گزارش شخصی یعنی جمع آوری اطلاعات از راه پاسخ‌هایی که شخص مورد سنجش به سوال‌هایی که از او پرسیده می‌شود، بدهد. پرسشنامه بر دانش واقعیت‌ها تاکید می‌کند و برای سنجش متغیرهای شناختی مفید است.

در سنجش هدف‌های حوزه عاطفی روش پاسخنامه‌های باز پاسخ با روش آزمون‌های تشریحی و کوت پاسخ در سنجش هدف‌های حوزه شناختی شبیه است. ص ۳۸۸ پرسشنامه باز پاسخ از پرسشنامه با بست پاسخ برای استفاده در کلاس درس مفیدتر است زیرا سوال‌های اینگونه پرسشنامه‌ها به معلم اجازه می‌دهد تا دامنه گسترده از موضوع‌های گوناگون را به صورت کارآمد مورد استفاده قرار دهد. برای تهیه و اجرای ابزارهای سنجش عاطفه قواعدی وجود دارد. بعد از آنکه تصمیم به تهیه پرسشنامه گرفتید جمله ساده و واضح بنویسید. جمله‌ها را در زمان حال بنویسید نه در زمان گذشته. جمله‌هایی بنویسید که معرف عواطف باشند نه واقعیت در هر جمله یا سوال پرسشنامه بیشتر از یک اندیشه قرار ندهید.

روش سنجش عاطفی دارای امتیازهایی می‌باشد. اجرای این روش در مقایسه با بعضی روش‌های دیگر نیاز به صرف وقت زیادی ندارد. این روش با گروه‌های کوچک و بزرگ قابل اجرا است. استفاده از این روش نسبتاً کم هزینه و آسان است. این روش‌ها اما محدودیت‌هایی هم دارند. چندان قابل اعتماد نیستند یعنی داده‌ها حاصل از عینیت لازم نیستند. ذهنیت زیادی در آنها دخالت ندارد. ساختن پرسش‌های مفید و موثر دشوار است و از عهده هر کسی بر نمی‌آید. ص ۴۰۶

گام اول در اجرای آزمون پیشرفت تحصیلی آماده سازی و فراهم کردن شرایط مطلوب برای اجرای آن می‌باشد. در رابطه با قرار دادن سوال‌های آزمون به دنبال یکدیگر چند پیشنهاد ارائه شده است. سوال با توجه به نوع آنها دسته‌بندی شود و پشت سر هم قرار بگیرند. سوال‌ها از ساده به دشوار مرتب شود. سال‌ها به ترتیب سازمان اصلی مطالب به دنبال هم مرتب شوند.

گفته شده است که راهنمای آزمون باید یک سری اطلاعات را در اختیار آزمائش شوندگان قرار دهد که از جمله آن زمان لازم برای هر قسمت، ارزش هر سوال و مجاز یا غیر مجاز بودن حدس زدن است.

شرایط اجرای آزمون پیشرفته تحصیلی باید به صورتی باشد که از لحاظ آسایش جسمی و فکری آزمون شوندگان را تامین کند این نکته در رابطه با آزمون شوندگان خردسال با دقت بیشتری باید رعایت شود.

بعد از تهیه و اجرای آزمون پیشرفت تحصیلی نوبت به تصحیح برگه‌های آزمون و نمره گذاری آنها می‌رسد. هدف از تحلیل سوال‌های آزمون بررسی تک تک سوال‌ها و تعیین میزان دقت و نارسایی آنهاست. اطلاعات مورد نیاز برای تحلیل سوال‌های یک آزمون پاسخ‌هایی هستند که آزمون شوندگان به هر سوال دادند. درصد کل آزمون شوندگانی که به یک سوال جواب درست می‌دهند ضریب دشواری آن سوال است که با حرف P نشان داده می‌شود. واریانس نمرات گروهی پاسخ دهنده‌ها به یک آزمون از دو قسمت تشکیل می‌شود که شامل واریانس سوال‌ها و همبستگی بین سوال‌ها می‌باشد. هر چقدر ضریب تمیز بزرگتر باشد قوه تمیز آن سال بیشتر و هر قدر این ضریب کوچکتر باشد قوه تمیز آن کمتر است. ضریب تمیز سوال نوعی ضریب توافق سوال با کل آزمون است.

بهتر است برای تحلیل آزمون علاوه بر محاسبه ضریب همبستگی بین هر سوال و کل آزمون ضریب همبستگی تک تک سوال‌ها با یکدیگر را تعیین کنیم یکی از روش‌های تحلیل سوال‌های آزمون‌های تشریحی و عملکردی روش ویتنی و سیرز است.

نتیجه آرمانی آزمون‌های وابسته به ملاک این است که در آنها اکثر آزمون شوندگان اکثر سوال‌ها را درست جواب دهند که در چنین حالتی دشواری سال‌ها بسیار بزرگ و ضریب تمیز بسیار کوچک خواهد بود.

در این فصل نظریه صفت مکنون و الگوی سوال پاسخ را توضیح دادیم در این نظریه‌ها با استفاده از منحنی ویژگی سوال ضرایب دشواری و تمیز سوال تعیین می‌شود.

به هرگونه ویژگی یا صفتی از اشیا، اشخاص که بتواند مقادیر مختلفی به خود بگیرد متغیر گفته می‌شود. قد، وزن، هوش، انگیزش و ... نمونه‌هایی از متغیرها هستند. بعد از اینکه آزمونی را به اجرا درآوردیم و نتایج را به دست آوردیم با دسته‌ای از اعداد به نام نمرات خام روبرو هستیم که به خودی خود قابل استفاده نیستند. هر وقت تعداد نمرات زیاد باشد پشت سر هم قرار دادن آنها برای تهیه جدول توزیع فراوانی یک ستون طولانی از نمرات به دست می‌دهد که خلاصه خوبی از آنها نیست.

اطلاعات مندرج در توزیع فراوانی را می‌توان به صورت نمودار نشان داد. در همه نمودارها از جمله هیستوگرام توزیع فراوانی نمرات با استفاده از محورهای مختصات نشان داده می‌شود. برای ترسیم نمودار هیستوگرام فاصله طبقات را بر روی محور افقی و فراوانی‌ها را بر روی محور عمودی مشخص می‌کنیم. نمودار چند ضلعی توزیع فراوانی نام دارد.

نمونه دیگری از نمودارها نمودار ستونی یا میله‌ای است برای رسم نمودار ستونی بسیار شبیه به رسم هیستوگرام عمل می‌کنیم یعنی بر بالای هر طبقه از اعداد یک ستون یا یک میله عمودی به گونه‌ای می‌سازیم که عرض آن معرف فاصله طبقه و ارتفاع آن نشان دهنده فراوانی طبقه باشد. برای داده‌های مربوط به مقیاس ترتیبی نیز از نمودار ستونی استفاده می‌شود.

نمودار دایره‌ای یک منحنی دایره شکل است که نسبت‌های مختلف یک کل را نشان می‌دهد. نمودارهای تصویری تعداد تصویر یا شکل مورد استفاده قرار می‌گیرند به گونه‌ای که هر کدام معرف فراوانی با درصد وقوع یک رویداد است.

نمای هر دسته از نمرات عبارت است از نمره‌ای که بیشترین فراوانی را دارد مورد استفاده نما زمانی است که معلم بخواهد با یک نگاه سطحی نمره‌ای را به عنوان نماینده نمرات گروه نشان دهد.

میان نقطه‌ای است که روی مقیاس نمرات یا توزیع نمرات که نصف بالا و نصف پایین گروه نمرات را از یک جدا می‌کند.

اندازه‌های معروف تغییر پذیری عبارتند از دامنه تغییر، واریانس یا پراکنش و انحراف معیار. ساده‌ترین اندازه تغییرپذیری دامنه تغییر نمرات است. واریانس و انحراف معیار بهترین اندازه‌گیری‌های تغییر پذیری هستند.

با داشتن هنجارهای مناسب می‌توان نمرات آزمون شوندگان به یک آزمون یا پاسخ دهندگان به یک پرسشنامه را تفسیر کرد و تعیین کرد که عملکرد آنان در چه وضعی است.

در هنجار سنی و هنجار کلاسی عملکرد فرد با عملکرد گروه مرجعی که فرد منطقاً عضوی از آن به حساب می‌آید مقایسه می‌شود اما در هنجار درصدی موقعیت نسبی فرد نسبت به گروهی که او خود عضو از آن است نشان داده می‌شود. هنجار سنی یا معادل سنی بر میانگین یا میان نمراتی مبتنی است که کودکان و نوجوانان در سن‌های مختلف

کسب می‌کنند و به صورت معادل‌های سنی نشان داده می‌شود. حنجره کلاسی شبیه به هنجار سنی است تنها تفاوت بین این دو نوع هنجار است که در هنجار سنی گروه‌های مرجع گروه‌های سنی هستند اما در هنجار کلاسی از گروه‌های کلاسی به عنوان گروه‌های مرجع استفاده می‌شود.

هنجار سن ذهنی یا سن عقلی اولین بار به وسیله آلفرد معرفی شد نوعی هنجار سنی است که در آزمون‌های هوش به کار می‌رود. در حنجار درصدی فرد یا گروه کلاسی خودش یعنی گروهی که منطقاً می‌توان او را عضوی از آن دانست مقایسه می‌شود. رتبه درصدی و صدک دو اصطلاح مکمل یکدیگرند.

آن دسته از نمرات معیار که دارای توضیح یا میانگین صفر و انحراف معیار واحد هستند به نمرات Z شهرت دارند. نمرات Z به ما نشان می‌دهد که یک نمره خام به اندازه چه تعداد از واحدهای انحراف معیار بالاتر یا پایین‌تر از میانگین است. مجموعه نمرات معیار دیگری که به نمرات T شهرت دارند برای رفع مشکلات نمرات Z ابداع شده‌اند.

نمره معیار دیگری که در آزمون‌های میزان شده توانایی کاربرد زیادی دارد نمره سنی معیار است در این مقیاس استاندارد میانگین ۱۰۰ و انحراف معیار ۱۶ است.

توضیح بهنجار توزیعی است که بر منحنی بهنجار منطبق و دارای یک سری ویژگی‌ها می‌باشد. دو قسمت منحنی بهنجار نسبت به خط وسط آن قرینه دو امتداد منحنی بهنجار در بی‌نهایت محور افقی را قطع می‌کنند در منحنی بهنجار میانگین، میانه و نما بر روی هم قرار دارند و

روایی اصطلاحی است که به هدفی که آزمون برای تحقق بخشیدن به آن درست شده اشاره می‌کند. لین و گرانلاند روایی را به صورت یک ارزشیابی از کفایت و مناسبت تفسیرها و استفاده‌های نتایج سنجش تعریف کردند. آزمون دارای روایی که برای اندازه‌گیری آنچه مورد نظر است کافی و مناسب باشد. آزمون‌های مورد استفاده در آموزش و پرورش و روانشناسی دارای سه هدف عمده هستند

روایی محتوایی به این مطلب اشاره می‌کند که نمونه سوال‌های مورد استفاده در یک آزمون تا چه حد معرف کل جامعه سوال‌های ممکن است که می‌توان از محتوای آموزش مورد نظر تهیه کرد هرچه آزمون از این لحاظ بهتر باشد دارای روایی بیشتری است. برای ساختن آزمون‌های پیشرفت تحصیلی روا د تلاش کنیم تا سوال‌های آزمون نمونه کاملی از هدف‌ها و محتوای درس باشد.

روایی صوری یکی از مشتقات روایی محتوایی است. روایی صوری به این مطلب اشاره می‌کند که سوال‌های آزمون تا چه حد در ظاهر شبیه به موضوعی هستند که برای اندازه‌گیری آن تهیه شده‌اند.

منظور از روایی ملاکی میزان ارتباط بین نمرات حاصل از یک آزمون با نمرات حاصل از یک آزمون یا وسیله اندازه‌گیری دیگر است.

روایی پیش بینی ویژه آزمون‌هایی است که برای پیش‌بینی موفقیت افراد در امور تحصیلی یا شغلی به کار می‌روند. آزمون‌هایی که برای مقاصد پیش بینی به کار می‌روند مختلف هستند. گاه از آزمون‌های پیشرفت تحصیلی فعلی یا قبلی برای پیش‌بینی پیشرفت تحصیلی آینده استفاده می‌شود.

در تعیین روایی همزمان هنگام اندازه‌گیری متغیر پیش بین علاوه بر این متغیر، متغیر ملاک نیز موجود است، در حالی که در تعیین روایی پیش بینی چنانکه دیدیم تغییر ملاک در آینده اندازه‌گیری می‌شود. نمرات آزمون واقعاً برای انجام نوعی پیش بینی به کار می‌روند اما چون در روایی همزمان داده‌های پیش بینی کننده و پیش بینی شونده همزمان جمع آوری می‌شوند هدف آزمون پیش‌بین پیش بینی ملاک نیست زیرا داده‌های ملاک در زمان اجرای آزمون پیش بین وجود دارند.

ریات داشتن داده‌های یک آزمون ملاکی با توجه به قابلیت انطباق آنها با موقعیت واقعی شخص مورد قضاوت قرار می‌گیرد. برای این کار هیچگونه شواهد تجربی وجود ندارد و باید بر اساس ملاک منطقی تصمیم گرفت

برای تعیین روایی سازه نخست باید سازه یا متغیر مورد نظر را تعریف کرد این متغیر به صورت یک سازه در نظامی از مفاهیم قرار خواهد داشت که در آن نظام این سازه به طریقی منطقی با سازه‌های دیگر در رابطه خواهد بود.

اگر آزمون جدید با آزمون موجود همبستگی زیادی داشته باشد فرض می‌شود که این دو آزمون سازه واحدی را اندازه می‌گیرند. راه دیگر تعیین روایی سازه توجه به تمایز سنی است. تحلیل عوامل یک روش آماری است که از طریق آن تعداد و ماهیت متغیرهایی را که یک آزمون اندازه می‌گیرد مشخص می‌کند

یکی از مفهوم‌های نسبتاً تازه وابسته به روایی محتوایی روایی آموزشی است که به آن روایی برنامه درسی نیز گفته می‌شود. در آزمون‌های وابسته به ملاک روایی پیش بینی به دو منظور به کار می‌رود. پیش بینی عملکرد یادگیرنده در یک هدف آموزشی بر اساس هدف آموزشی دیگر - پیش بینی یادگیری در حد تسلط

از عوامل موثر بر روایی آزمون می‌توان به موارد زیر اشاره کرد. اولین عامل تأثیر گذار دستورالعمل با راهنمای آزمون است. کیفیت سوال آزمون تأثیرگذار است. عامل دیگری که تأثیر دارد ترتیب قرار گرفتن سوال‌های آزمون به دنبال یکدیگر است.

پایایی یک وسیله اندازه‌گیری عمدتاً به ثبات و دقت نتایج حاصل از آن اشاره می‌کند. رابطه بین روایی و پایایی از این قرار است که یک آزمون باید پایا باشد تا بتواند روا باشد. یک آزمون برای اینکه روا باشد باید نخست پایا باشد یعنی پایایی شرط روایی است اما روایی برای پایایی ضروری نیست.

پایایی بین مصححان منظور از مصححان در اینجا دو یا چند نفری هستند که برگه‌های امتحانی را تصحیح می‌کنند رفتارهای افراد یا فرایندها را مورد مشاهده قرار می‌دهند یا محصولات یا فرآورده‌های رفتار را درجه‌بندی می‌کنند.

یکی از معروف‌ترین روش‌های تعیین پایایی یک آزمون روش بازآزمایی است. در این روش آزمون را در دو نوبت با گروه واحدی از آزمون شوندگان اجرا می‌کنند و نمره‌های حاصل را با هم مقایسه می‌کنند. ضریب همبستگی بین نمره‌های به دست آمده از دوباره اجرایی آزمون ضریب پایایی آزمون است. روش دیگری که برای رفع مشکلات روش بازآزمایی پیشنهاد شده است روش فرم‌های موازی یا هم ارز است در این روش دو آزمون معادل یا موازی برای یک موضوع تهیه می‌شود و آنها را در فاصله زمانی کوتاهی به یک گروه واحد از آزمون شوندگان می‌دهند. دو فرم یک آزمون در صورت موازی هستند که میانگین و واریانس آنها برابر باشد. ضریب همبستگی بین نمرات حاصل از اجرای فرم‌های موازی ضریب پایایی آن آزمون‌ها به حساب می‌آید.

در روش کودر ریچاردسون نیز آزمون تنها یک بار اجرا می‌شود اما در این روش همه ماده‌های آزمون تحلیل می‌شوند. سومین روش تعیین پایایی آزمون با تأکید بر همسانی درونی روش ضریب آلفا نام دارد که به آن ضریب آلفا یا ضریب آلفای کرائیخ یا حتی آلفا نیز گفته می‌شود. در رابطه با استفاده از ضریب پایایی آزمون‌ها را می‌توان به دو گروه آزمون قدرت و آزمون سرعت دسته‌بندی کرد. آزمون قدرت آزمونی است که در اجرای آن به آزمون شوندگان وقت کافی داده می‌شود تا به همه سوال‌ها جواب دهند. آزمون‌های سرعت آزمون‌هایی هستند که از سوال‌های بسیار آسان تشکیل شده‌اند و هدف آنها صرفاً سنجش سرعت عمل آزمون شوندگان در پاسخ دادن به سوال‌ها هست.

عوامل موثر بر افزایش پایایی آزمون به این شرح است. با افزودن بر تعداد سوالات یک آزمون پایایی آن نیز افزایش می‌یابد. متجانس و همگون کردن سوالات یک آزمون پایایی آن بیشتر می‌شود یا افزودن تعداد سوال دارای ضریب دشواری متوسط به یک آزمون پایایی آن افزایش می‌یابد.

در تفسیر نمره‌های یک آزمون همواره مفاهیم زیر مورد استفاده قرار می‌گیرد: نمره مشاهده شده که همان نمره خام است، نمره حقیقی و نمره خطا. تفاوت بین نمره مشاهده شده و نمره حقیقی نمره خطا یا خطای اندازه‌گیری نام دارد.

یکی از فواید ضریب پایایی آزمون این است که با استفاده از آن می‌توان خطای معیار اندازه‌گیری آزمون را تعیین کرد. خطای معیار اندازه‌گیری برابر است با انحراف معیار توزیع نمرات خطا. در مورد متغیرهای روانی و پرورشی معمولاً نمی‌توان آنها را اندازه گرفت اما با استفاده از ضریب پایایی آزمون می‌توان خطای معیار اندازه‌گیری آزمون را برآورد کرد.

از خطای معیار اندازه‌گیری برای برآورد نمره حقیقی استفاده می‌کنیم برای این منظور می‌توانیم فرض کنیم که اگر آزمون شونده بارها به یک آزمون پاسخ دهد نمره حاصل توزیع طبیعی خواهند داشت که میانگین آن توزیع معرف نمره حقیقی آزمون شونده است.

پایایی شرط روایی است یعنی اینکه یک آزمون باید پایا باشد تا بتواند روا باشد بنابراین هر اندازه از میزان پایایی یک آزمون کاسته شود به همان نسبت از روایی آن کم می‌شود. معروف‌ترین روش تعیین پایایی روش‌های درجه‌بندی یا رتبه‌بندی روش گیلفورد است.

فصل دوم: نکات مهم اندازه‌گیری، سنجش و ارزشیابی آموزشی تالیف ایران عرضه

۱- اندازه‌گیری عبارت است از فرایندی که تعیین می‌کند یک شخص یا یک شیء چه مقدار از یک ویژگی برخوردار است. اندازه‌گیری فرایند یا شیوه دقیق و نظام دار است. در بعضی موارد قاعده اندازه‌گیری تا حدی ساده است. فایده بیان قواعد اندازه‌گیری این هست که کمک می‌کند تا افرادی که می‌خواهند وسایل اندازه‌گیری را مورد استفاده قرار دهند آن را یکسان به کار بگیرند. چنانچه که گفتیم اندازه‌گیری یک فرایند است این فرایند به وسیله‌هایی نیاز دارد که به آن وسیله اندازه‌گیری می‌گویند.

- ۲- بین آزمون و پرسشنامه تفاوت‌هایی وجود دارد که به وسیله آزمون می‌توانیم توانایی‌ها را اندازه‌گیری کنیم اما پرسشنامه ویژگی‌های غیر توانایی را ارزیابی می‌کند.
- اشتراک آزمون و پرسشنامه این هست که این دو ابزار کمی‌اند، یعنی اندازه‌گیری به وسیله آنها در قالب عدد و رقم بیان می‌شود. وقتی در اندازه‌گیری یکی از ویژگی‌های روانی یا تربیتی گروه یا فردی از آزمون استفاده می‌شود، این فعالیت آزمودن نامیده می‌شود.
- ۳- تعریف آزمودن: بنا بر تعریف، وقتی که برای اندازه‌گیری یکی از ویژگی‌های روانی یا تربیتی یک فرد یا گروهی از افراد از آزمون استفاده می‌شود به این فعالیت آزمودن می‌گویند.
- تعریف سنجش: بنا بر گفته نیتکو سنجش یک اصطلاح کلی است و به صورت فرایندی تعریف می‌شود که برای گردآوری اطلاعات مورد نیاز تصمیم‌گیری درباره دانش آموزان، برنامه‌های درسی، و سیاست‌های آموزشی مورد استفاده قرار می‌گیرد.
- تعریف ارزشیابی (یا ارزیابی): اصطلاح ارزشیابی به طور ساده به تعیین ارزش برای هر چیزی یا داوری ارزش کردن گفته می‌شود. ارزشیابی به یک فرایند نظام دار برای جمع‌آوری، تحلیل، و تفسیر اطلاعات گفته می‌شود. یکی از ویژگی‌های مه ارزشیابی تعیین کیفیت است.
- ۴- ارزشیابی‌ها درست بر پایه اطلاعات دقیق استوار هستند. سنجش به جمع آوردن اطلاعات در مورد ویژگی عاطفی می‌پردازد. هدف ارزشیابی تصمیم‌گیری است ولی هدف پژوهش استخراج است.
- ۵- مقیاس به مجموعه‌ای از اعداد گفته می‌شود که ویژگی آنها به ویژگی‌های تجربی اشیایی که اعداد به آنها نسبت داده می‌شود، منطبق است. مقیاس اندازه‌گیری را به ۴ دسته تقسیم می‌کند. اسمی، ترتیبی، فاصله‌ای، نسبی. مقیاس اسمی ساده‌ترین نوع اندازه‌گیری است که به دو صورت اسم گذاری و طبقه بندی انجام می‌شود.
- مقیاس ترتیبی مجموعه‌ای از رتبه‌ها می‌باشد. دسته‌ای از اشیا یا افراد با توجه به یک صفتی که دارند از بزرگ به کوچک مرتب می‌شوند و بالعکس.
- ۶- عملیات مجاز آماری به شمارش فراوانی محاسبه میانه محاسبه ضریب همبستگی و ... می‌پردازد. عملیات مجاز ریاضی هیچ کدام از چهار عمل جمع، تفریق، ضرب و تقسیم در این مقیاس میسر نیست.
- ۷- مقیاس فاصله‌ای مقیاسی هست که رتبه اشیا با توجه به صفت مشخص است، اشیا از لحاظ صفت مورد اندازه‌گیری قرار می‌گیرند. تمایز مقیاس فاصله‌ای از ترتیبی این هست که در فاصله‌ای بین واحدها معلوم می‌باشد، اما در ترتیبی فاصله بین واحدها معلوم نیست. عملیات مجاز آماری در اینجا به محاسبه نما، میانه، میانگین، انحراف معیار و غیره می‌پردازد. عملیات مجاز ریاضی به جمع و تفریق مجاز است و ضرب و تقسیم در آن مجاز نیست.
- ۸- مقیاس نسبی مقیاسی می‌باشد که رتبه اشخاص با توجه به یک صفت معلوم است فاصله بین اشخاص معلوم است و علاوه بر این فاصله حداقل یکی از اشخاص از یک صفر مطلق نیز معلوم است. یکی از مهمترین مقیاس‌های نسبی همان مقیاس اعداد از صفر تا بی‌نهایت می‌باشد.
- ۹- مقیاس پیوسته و مقیاس گسسته.
- مقیاس گسسته متغیر مورد نظر فقط می‌تواند یکی از ارزش‌های معین را به خود اختصاص دهد. یک ویژگی مهم این نوع مقیاس تساوی واحدهای شمارش آنها در طول مقیاس است. متغیر دیگر که بیانگر مقیاس گسسته هست عبارت است از جنس، وضعیت تاهل، سطح تحصیلات و
- در مقیاس پیوسته متغیر مورد نظر می‌تواند در فاصله بین نقاط مقیاس هر تعدادی ارزش را به خود اختصاص دهد. متغیرهایی که بیانگر مقیاس پیوسته هستند شامل سرعت، زمان، طول، وزن و ... می‌باشند. اندازه‌گیری‌هایی که از متغیر گسسته به دست می‌آید دقیق و قطعی است اما اندازه‌گیری‌هایی که از متغیر پیوسته به دست می‌آید تقریبی می‌باشد. یک ویژگی مهم مقیاس پیوسته تساوی واحدهای اندازه‌گیری است
- ۱۰- حدود واقعی ارزش عدد یک متغیر پیوسته برابر می‌باشد با همان عدد به اضافه و منهای نصف واحد اندازه‌گیری.
- یک مشکلی که متخصص‌های اندازه‌گیری در کار با اعداد با آن روبرو می‌شوند تعداد زیاد ارقام اعشاری می‌باشد. برای حل این مشکل از روش گرد کردن استفاده می‌شود.
- ۱۱- این آزمون‌های استعداد و پیشرفته تحصیلی را نمی‌توان از همدیگر جدا کرد زیرا آن چیزی را که فرد در گذشته آموخته است می‌تواند برای پیش‌بینی موفقیت خود در آینده مورد استفاده قرار دهد. این آزمون‌های استعداد را برای پیش‌بینی موفقیت‌های آینده می‌توان به کار برد در حالی که پیشرفت تحصیلی برای تعیین مقدار آموخته‌ها یا توانایی‌های

موجود مورد استفاده قرار می‌دهند. معروف‌ترین آزمون استعداد آزمون هوش یا استعداد کلی هست. اهداف بیشتر آزمون‌های هوش اندازه‌گیری آن دسته از توانایی ذهنی است که در همه فعالیت‌های فکری و شناختی ضروری‌اند.

۱۲- هدف از کاربرد آزمون هوش پیش بینی موفقیت‌های همه جانبه افراد است در صورتی که آزمون‌های استعداد های ویژه برای پیش بینی موفقیت افراد در زمینه فکری و عملی مشخص به کار می‌رود. آزمون‌های شناختی و روانی توانایی فکری عملی افراد را اندازه می‌گیرند اما پرسشنامه‌های عاطفی ویژگی‌های عاطفی شخصیتی را می‌سنجند. پرسشنامه‌های سنجش عاطفی دارای تنوع زیادی هستند اما مهم‌ترین آنها پرسشنامه‌های شخصیت، نگرش علاقه است.

۱۳- پرسشنامه شخصیت در واقع هست که شخصیت را به صورت مجموعه رفتارهای فرد در موقعیت‌های اجتماعی می‌توان تعریف کرد. یعنی علاوه بر رفتارهای آشکار ویژگی‌های غیر رفتاری مثل هوش پیشرفت قد و ... به این تعریف شخصیت اضافه می‌شود. پرسشنامه شخصیت به اندازه‌گیری متغیرهای انگیزش، مزاج، سازگاری و از این قبیل اختصاص دارد. پرسشنامه‌ها و ابزارهای اندازه‌گیری شخصیت به سه دسته تقسیم می‌شود. پرسشنامه‌های گزارش شخصی که در آن فرد نظر شخصی خود را بیان می‌کند. پرسشنامه جامعه سنجی که در آن افراد دیگر نظرشان را نسبت به فرد مورد سنجش بیان می‌کنند. فنون مشاهده رفتار فرد در یک موقعیت خاص مورد مشاهده قرار می‌گیرد. ابزارهای سنجش شخصیت را می‌توان به دو دسته تقسیم کرد. محرک سازمان یافته و محرک سازمان نیافته

۱۴- آزمون‌های میزان شده برای استفاده در مناطق مختلف و مقایسه بین دانش‌آموزان مدارس تهیه می‌شود. آزمون میزان شده به وجود داده‌های هنجاری اشاره می‌کند. هدف اصلی آزمون‌های میزان شده مقایسه است مقایسه یک فرد یا یک گروه با فرد یا گروه دیگر.

۱۵- روش‌های مختلف اندازه‌گیری و سنجش در آزمون‌ها را می‌توان به نوع سوال یا ماده دسته‌بندی کرد. در اندازه‌گیری به سوال آزمون یا هر تکلیف یا تمرینی که به آزمون شونده داده می‌شود، ماده می‌گویند. وسایل مختلف اندازه‌گیری با توجه به نوع سوال یا ماده به بسته پاسخ و باز پاسخ تقسیم می‌شوند. در بسته پاسخ هم سوال و هم پاسخ تهیه می‌شود و در اختیار و از آن می‌خواهد تا در مورد آن تصمیم بگیرد. در باز پاسخ دانش آموز خود پاسخ سوال را تولید می‌کند. آزمون‌های بسته پاسخ و باز پاسخ به ترتیب عینی و غیرعینی نامیده می‌شود. به جای اصطلاحات بسته پاسخ و باز پاسخ از آزمون‌های گزینه ثابت و سنجش‌های عملکرد پیچیده استفاده کرده‌اند.

۱۶- صاحب نظران ورتن و سندرز ارزشیابی تمام رویکردهای ارزشیابی آموزشی را به ۶ دسته تقسیم می‌کنند. رویکرد مبتنی بر هدف - رویکرد مبتنی بر مدیریت- رویکرد مبتنی بر مصرف کننده- رویکرد مبتنی بر نظر متخصصان- رویکرد طبیعت گرایانه و مبتنی بر مشارکت کنندگان. این تقسیم بندی شش گانه بر تقسیم بندی دوگانه هاوس از ارزشیابی‌های آموزشی استوار است. این تقسیم بندی بیشتر مبنای فلسفی و ایدئولوژی دارد. رویکرد ارزشیابی پیرو فلسفه عینیت گرایی است

۱۷- رویکرد شهود گرایی - کثرت گرایی بر این اندیشه استوار است که ارزش به تاثیر برنامه بر فرد فرد افراد وابسته است نه بر اکثریت آنها. در رویکرد کثرت گرایی ارزشیابی آموزشی احساسات دریافت‌های فردی ملاک خوب و درست به حساب می‌آید. متعادل کردن و وزن دادن داورهای و ملاک‌های مختلف در این رویکرد عمدتاً جنبه شهودی دارد که به آن عنوان شهودگرایانه دادند. به لحاظ اجرایی در ارزشیابی‌های شهودگرا و کثرت گرا تمام افراد و گروه‌هایی که تحت تاثیر برنامه آموزش بودند به عنوان داوران استفاده می‌شوند.

در رویکرد ارزشیابی مبتنی بر هدف ابتدا مقاصد یک فعالیت آموزشی تعیین می‌شود و بعد از آن در ارزشیابی کوشش انجام می‌گیرد تا نشان داده شود که آن مقاصد تا چه اندازه تحقق یافته‌اند. الگوی ارزشیابی تایلر اولین الگوی ارزشیابی آموزشی در تاریخ پیدایش آن به دهه ۱۹۳۰ میلادی برمی‌گردد. مفهوم تایلر از ارزشیابی آموزشی این هست که تعیین کند هدف آموزشی برنامه آموزشگاه با برنامه درسی تا چه میزان تحقق یافته است.

۱۸- الگوی ارزشیابی آموزشی تایلر حوزه برنامه‌ریزی درسی را تحت تاثیر قرار داده است.

الگوهای ارزیابی که از رویکرد مبتنی بر مدیریت آمدند برای کمک به مدیران و تصمیم گیرندگان به عمل آمده‌اند. پدید آورندگان این الگوهای ارزشیابی بر یک رویکرد سیستمی تاکید دارند.

۱۹- منظور از ارزشیابی از دانش آموزان و دانشجویان عمدتاً این است که پیشرفت تحصیلی و میزان یادگیری آنها مورد ارزیابی واقع شود. عمده‌ترین ملاک در ارزشیابی آموزشی، هدف‌های آموزشی از پیش تعیین شده، هستند. ارزشیابی پیشرفت تحصیلی مهمترین نوع ارزشیابی آموزشی است که تعیین می‌کند دانش‌آموزان یا دانشجویان تا چه میزان به هدف‌های آموزشی خود که از پیش تعیین شده بودند، رسیدند. ارزشیابی آموزشی بعضی مواقع برای تعیین اثربخشی برنامه درسی یا مواد آموزشی انجام می‌شود.

۲۰- ارزشیابی از برنامه‌های درسی و مواد آموزشی عمده‌ترین ملاک پیشرفته تحصیلی میزان یادگیری دانش‌آموزان و دانشجویان می‌باشد. ملاک دیگر ارزشیابی از برنامه درسی و مواد آموزشی هزینه آن‌ها

۲۱- منظور از ارزشیابی از آموزشگاه این هست که تعیین شود که اهدافی که پیش بینی شده برای آموزشگاه به چه میزان تحقق پیدا کرده و نقاط قوت و ضعف آن چیست. مانند دیگر ارزشیابی‌ها در ارزشیابی از آموزشگاه پیشرفت تحصیلی دانش‌آموزان و دانشجویان مهم‌ترین بخش برنامه آزمون و منبع کسب اطلاعات است.

ارزیابی از کارکنان آموزشگاه یک ارزشیابی است که از همه ساکنانی که به طور مستقیم و غیر مستقیم مسئول بازدهی‌های آموزشی موسسه آموزشی هستند، انجام می‌شود. مهمترین نوع ارزشیابی از کارکنان آموزشگاه ارزشیابی معلم است. پیچیده‌ترین نوع ارزشیابی، ارزشیابی از کارکنان به ویژه از کارمندان است.

۲۲- آن چیزی که برای کمک به اصطلاح ارزشیابی یعنی برنامه یا روش آموزش مورد استفاده قرار می‌گیرد ارزشیابی تکوینی می‌باشد. ارزشیابی آموزشی به تصمیم‌گیری برای مسائل مختلف کمک می‌کند این تصمیم‌گیری‌ها در رابطه با هدف هستند. این تصمیم‌گیری‌ها به سه دسته تقسیم می‌شوند. مرحله طراحی- مرحله فرایندی - مرحله فرآورده‌ای. مرحله طراحی شامل یک سری از تصمیم‌هایی است درباره فعالیت‌هایی که قرار است انجام بگیرد. در این مرحله اقداماتی از قبیل تحلیل موقعیت تعیین هدفها توصیف پیش نیازها انتخاب یا تولید وسایل و ابزار مورد نیاز انتخاب طرح و ... انجام می‌شود. در اولین گام مرحله طرح‌ریزی موقعیت موجود را

تحلیل می‌کنیم که از جمله کارهایی که برای این باید انجام دهیم جمع آوری سوابق اطلاعاتی و برآورد و تعیین موانع موجود است. ۲۳- برای جمع آوری اطلاعات مهم و لازم در ارزشیابی آموزشی به ابزارهایی نیاز هست از جمله آنها می‌توان به آزمون‌های پیشرفته تحصیلی، آزمون‌های نگرش‌سنجش، یا فرم‌های ثبت مشاهده‌ای را نام برد. تهیه وسایل اندازه‌گیری مانند آزمون‌ها و پرسشنامه‌ها کاری هست که به دقت و صلاحیت نیاز استراتژی یا راهبرد فعالیت است که برای کمک به تحقق یافتن هدف‌های برنامه به کار برده می‌شود. اولین تصمیم استراتژی این هست که معین کنیم توالی ارائه این موضوع‌ها چگونه باشد.

۲۴- دومین مرحله ارزشیابی مرحله فرایندی است. در این مرحله تصمیم‌هایی که مرتبط با اقدامات عملی و اجرایی گرفته می‌شود. هدف مهم و اصلی ارزشیابی در مرحله فرایندی هدایت فعالیت‌های آتی برای بهبود فعالیت‌های آموزش و یادگیری است. گی گفته است که مرحله فرایندی ارزشیابی همان چیزی است که ارزشیابی تکوینی نامیده می‌شود و هدف ارزشیابی تکوینی را کمک به اصلاح برنامه آموزشی می‌داند.

در مرحله فرآورده‌ای یا تولیدی برعکس مرحله فرایندی تصمیمات در پایان برنامه اتخاذ می‌شود. هدف این مرحله جمع آوری داده‌هایی است که متخصص ارزشیابی را در تصمیم‌گیری‌هایش درباره اثربخشی کلیه دانش آموزشی برنامه درسی و ... کمک نماید. در این مرحله مشخص می‌شود که هدف‌هایی که از پیش تعیین شده‌اند تا چه مقدار تحقق یافته‌اند. به طور کلی در مرحله فرآورده‌ای ارزشیابی تصمیمات مربوط به اثربخشی موضوع مورد ارزشیابی اتخاذ می‌گردند.

۲۵- ارزشیابی به یک فرایند نظام‌دار برای جمع آوری و تحلیل و تفسیر اطلاعات گفته می‌شود برای اینکه معین شود آیا هدف‌های مورد نظر تحقق پیدا کرده‌اند یا در حال تحقق یافتن هستند. غایت‌ها له بیاناتی هستند که به صورت کلی و آرمانی می‌باشند که قصد و منظور از ایجاد و اداره دوره‌های آموزشی یک جامعه را نشان می‌دهند. در مقابل هدف کلی نوع دیگری هدف وجود دارد که دقیق‌تر است و معلمان و متخصصان آموزشی با توجه به هدف‌های کلی و با استفاده از محتوای درس‌های مختلف آنها را تهیه می‌کنند. هدف آموزشی یا هدف یادگیری یک درس مقصودی است که معلم از آموزش خود و یادگیرنده از یادگیری خود دارد. برای آموزش و هم برای ارزشیابی نیاز است هدف‌های کلی به هدف‌های دقیق تجزیه شوند.

۲۶- هدف‌های آموزشی ابتدا به سه طبقه کلی با نام‌های حوزه شناختی، حوزه عاطفی و حوزه روانی-حرکتی تقسیم شده‌اند. هدف‌های حوزه شناختی به جریان‌هایی که با فعالیت ذهنی و فکر آدم سر و کار دارد، مربوط می‌شود. حوزه عاطفی شامل هدف‌هایی است که تغییرات حاصل در علاقه‌ها نگرش‌ها و ارزش‌ها را نشان می‌دهد. حوزه روانی-حرکتی به مهارت‌های حرکتی یا حرکات بدنی مربوط است. هدف‌های حوزه عاطفی و حوزه روانی-حرکتی با آزمون کتبی به خوبی قابل سنجش نیستند.

۲۷- بلوم و همکارانش هدف‌های حوزه شناختی را به ۶ طبقه کلی طبقه‌بندی کردند. دانش- فهمیدن- کار بستن- تحلیل- ترکیب- ارزشیابی. هر کدام از این طبقه‌ها یک نوع فرایند شناختی را توصیف می‌کند.

۲۸- برای حوزه روانی حرکتی تعداد طبقه‌بندی وجود دارد که در اینجا طبقه‌بندی سیمپسون مورد نظر می‌باشد که شامل ۷ طبقه اصلی و تعدادی خورده طبقه است. ادراک- آمادگی- پاسخ هدایت شده- مکانیسم- پاسخ پیچیده آشکار- انطباق- ابتکار.

۲۸- برای تهیه سوال‌های صحیح - غلط از سوی متخصص‌ها قواعدی پیشنهاد شده است.

موضوع‌های مهم در ارتباط با هدف‌های آموزشی را در سوال‌های خود بیاوریم.

سوال‌های طرح شود که درک و فهم را بسنجد نه یادآوری کلمه به کلمه مطالب را.

سوالی طرح شود که درست یا غلط بودن آن قابل دفاع باشد.

از جمله‌های پیچیده استفاده نشود.

در هر یک سوال بیشتر از یک موضوع قرار داده نشود.

تا جایی که امکان دارد از زبان کمی و دقیق استفاده شود نه از زبان کیفی و غیر دقیق.

تا جایی که ممکن است از طرح سوال‌های منفی پرهیز شود. از کلمات راهنما یا اشاره کننده به جواب درست استفاده در طرح سوال سعی شود تعداد سوال‌های غلط بیشتر از تعداد سوال‌های درست باشد. طول سوال‌های صحیح و سوال‌های غلط تقریباً باید با هم، هم اندازه انتخاب شود.

۳۰- سوال‌های جور کردنی بندی از نوع بالا با سوال‌های جور کردنی معمولی از لحاظ تعداد ماده‌های فهرست پرسش و پاسخ فرق دارند. یکی از امتیاز سوال‌های جور کردنی از نوع رده‌بندی یا ماده فهرست این است که در تصحیح جواب‌ها می‌توان از پاسخنامه‌های جداگانه استفاده کرد. نکته‌ای که در رابطه با سوال‌های جور کردنی م این هست که فهرست پرسش‌ها و فهرست پاسخ‌ها را هم می‌توان در ستون موازی در کنار هم قرار داد و هم می‌توان آنها را پشت سر هم نوشت.

به طور کل در راهنمای سوال سه دسته اطلاع را باید قرار داد. اساس رابطه بین پرسش‌ها و پاسخ‌ها یا مبنای جور کردن آنها- روش مشخص کردن پاسخ‌ها یا جور کردن پرسش‌ها و پاسخ‌ها- و اینکه از یک پاسخ تنها یک بار استفاده شود یا بیشتر از یک بار.

۳۱- یکی از کاربردهای ابتکاری سوال‌های چند گزینه‌ای استفاده از رایانه است. این آزمون‌ها تاکنون در حوزه آموزش پزشکی بیشترین مورد استفاده را داشتند و معلوم شده که وقتی با آزمون‌های سنتی به کار می‌رود روش بسیار مفید و معتبری است. به طور کل برای امتیازهای آزمون‌های چند گزینه‌ای می‌توان به مواردی اشاره کرد. این آزمون‌ها از بقیه آزمون‌های عینی انعطاف پذیرتر هستند- آزمون‌های چند گزینه‌ای نسبت به آزمون‌های صحیح- غلط کمتر امکان حدس زدن کورکورانه را می‌دهند- پاسخ این سوال‌ها به سادگی و با عینیت کامل قابل تصحیح هستند- معلمان و دانش آموزان این نوع آزمون را به سایر آزمون‌های عینی ترجیح می‌دهند.

۳۲- محدودیت‌های آزمون‌های چند گزینه‌ای شامل: ساختن این آزمون‌ها بسیار دشوار است- در مقایسه با بعضی آزمون‌های دیگر خواندن این آزمون و پیدا کردن گزینه درست وقت زیادی را می‌گیرد- معلمان اغلب سوال‌های چند گزینه‌ای را در حد سنجش اطلاعات جزئی می‌نویسند.

۳۳- سوال‌های آزمون‌های تشریحی گسترده پاسخ برای سنجش هدف‌های تحلیل و آفرینندگی مناسب‌ترین سوال است. در آزمون تشریحی محدود پاسخ آزمون شونده در دادن پاسخ به سوال آزادی کامل ندارد و باید پاسخ خود را در چهارچوب شرایط خاصی محدود کند. بنابراین سوال‌های آزمون محدود پاسخ برای اندازه‌گیری توانایی یادگیرندگان در سطوح فهمیدن، کار بستن و تحلیل مناسب اند، ولی برای سنجش توانایی ترکیب و ارزشیابی مفید نیستند.

۳۴- در روش تحلیلی نمره گذاری که به آن روش امتیازبندی گفته می‌شود پاسخ نمونه خود را به اجزای کوچک‌تری تقسیم کنید و برای هر جز مشخص نمره یا امتیاز جداگانه در نظر بگیرید. در روش کلی نمره گذاری که به آن روش درجه‌بندی گفته می‌شود پاسخ نمونه به اجزا و قسمت‌های کوچک تقسیم نمی‌شود بلکه تنها به صورت یک معیار به کار می‌رود. ایبل درباره نمره گذاری به روش کلی گفته است که در مجموع این روش نمره گذاری آسان‌تر و سریع‌تر از روش تحلیلی است.

روش دیگر نمره گذاری آزمون تشریحی روش ویژگی‌های اصلی نام دارد در این روش معلم ویژگی اصلی پاسخ دانش آموزان و هر سوال را سنجش می‌کند و برای هر یک از آنها نمره‌های ۰ تا ۴ که معرف عالی تا غیر قابل قبول اند را منظور می‌کند. این نوع آزمون‌ها امتیازهای آنها به این شرح است: تهیه آزمون‌های تشریحی آسان است. این آزمون توانایی تولید پاسخ را می‌سنجد نه توانایی انتخاب پاسخ را. این آزمون‌ها موقعیت‌های واقعی‌تری را در اختیار آزمون‌شوندگان می‌گذارد. این آزمون‌ها بر روش مطالعه یادگیرندگان تاثیر مثبت به جا می‌گذارد.

در نتیجه آزمون‌های تشریحی نمونه ضعیفی از محتوا و هدف‌های درس را اندازه می‌گیرند. مشکل دیگر آزمون‌های تشریحی ذهنی بودن نمره گذاری مصححان و یکنواخت نبودن ارزشیابی آنان است.

۳۵- آزمون کوتاه پاسخ در بین آزمون‌های عینی و تشریحی قرار دارد. آزمون کوتاه پاسخ تعدادی سؤال مختصر را در بر می‌گیرد که پاسخهای آنها به یک عدد، علامت، کلمه، عبارت یا جمله محدود می‌شود. آزمون‌های کوتاه پاسخ را به سه دسته پرسشی، کامل کردنی و تشخیصی طبقه بندی می‌کنند.

در آزمون کوتاه پاسخ نوع پرسشی سؤال به صورت یک جمله پرسشی طرح میشود و آزمون‌شونده پاسخ کوتاهی به آن میدهد. در آزمون کوتاه پاسخ نوع کامل کردنی سؤال به صورت یک جمله ناقص نوشته میشود که در آن یک جای خالی وجود دارد و آزمون‌شونده با پر کردن جای خالی به سؤال جواب میدهد و در آزمون کوتاه پاسخ نوع تشخیصی یا تداعی صورت سؤال ارائه می‌شود و آزمون‌شونده باید جواب آن را تداعی کند.

۳۶- یک دسته از روش‌های سنجش واقعی یا اصیل سنجش عملکردی نام دارد.

آزمون‌های عملکردی یا فرایند یا شیوه اجرا، فرآورده با محصول کار و یا ترکیبی از این دو سر و کار دارند. ماهیت عملکرد مورد سنجش مشخص می‌کند که تاکید بر چه قسمتی باید باشد. فرآورده چیزی است که یادگیرنده تولید می‌کند و فرایند جریانی است که او مورد استفاده قرار می‌دهد یا به فرآورده دست یابد.

ویژگی آنها به چهار دسته تقسیم می‌شود. تاکید بر کاربرد، تاکید بر سنجش مستقیم، استفاده از مسائل واقعی، ترغیب و تشویق تفکر باز.

۳۷- در روش نمونه کار از یادگیرنده درخواست می‌شود تا کاری را انجام دهد که معرف عملکرد واقعی مورد سنجش باشد. در سنجش ساختارمند معلم هم بر تکالیف سنجش و هم بر شرایطی که در آن دانش آموز باید پاسخ دهد کنترل دارد.

۳۸- کار پوشه به یک مجموعه از کارهای یادگیرنده گفته می‌شود که نشان دهنده کوشش و موقعیت و پیشرفت او در یک زمینه خاص می‌باشد.

کار پوشه حاوی کارهایی است که در یک مدت معین انتخاب شده است نه تمام کارها.

طبق گفته سانتروک به طور کل چهار دسته مدرک را می‌توان در کار پوشه دانش آموز قرار داد. کارهای انجام شده توسط دانش آموز، کارهای تهیه شده توسط دانش آموز، اظهار نظر دیگران درباره دانش آموز، تولیدات دانش آموز.

برای تهیه کار پوشه به چهار نکته باید توجه داشت. استفاده‌هایی که از کار پوشه به عمل خواهد آمد، کسانی که به آن دسترسی دارند، تناسب داشتن هدف تهیه کار پوشه با کارهایی که قرار است در کار پوشه قرار بگیرد، ملاک‌هایی که در ارزشیابی از کار پوشه به کار خواهد رفت.

از امتیازهای خاص این روش می‌توان به مواردی اشاره کرد. با این روش هم می‌توان سنجش کرد و هم آموزش داد. به کمک این روش می‌توان موفقیت یادگیرندگان را سنجش کرد و هم جریان رشد موفقیت آنها را به یادگیرندگان کمک می‌کند تا مسئولیت انتخاب هدف خود را بپذیرند و بر جریان پیشرفت نظارت داشته باشند.

۳۹- ساده‌ترین نوع مقیاس درجه‌بندی مقیاس درجه‌بندی عددی است. متخصص‌ها برای هر سه روش مشاهده‌ای راهنماهایی ارائه داده‌اند. روش‌های سنجش مشاهده‌ای دارای امتیازها و محدودیت‌هایی هستند که از جمله امتیازها می‌توان به این موارد اشاره کرد. آسان بودن استفاده با تکالیف و موقعیت‌های گوناگون، بالا بردن قدرت پیش بینی نتایج سنجش، به دست آوردن داده‌های کمی و اطلاعات کیفی و ... از جمله محدودیت‌هایی که برای روش سنجش مشاهده‌ای وجود دارد شامل این موارد می‌باشد. امکان نتیجه‌گیری‌های غلط، نشان ندادن علت‌های رفتار، دشواری ثبت مشاهدات همزمان، قطعیت غیر موجه.

۴۰- سنجش رفتاری یک سری امتیازها و محدودیت‌هایی دارد. با استفاده از سنجش رفتاری می‌توان ویژگی‌های رفتار مورد نظر را شناسایی و سنجش کرد یا استفاده از روش‌های سنجش رفتاری می‌توان داده‌های مورد نیاز پژوهش‌های آموزشی را گردآوری کرد. از این روش برای سنجش ویژگی‌های دانش آموزان می‌توان بهره برد. محدودیت‌های روش

سنجش رفتاری عبارتند از استفاده از این روش وقتگیر و لازمه صرف نیروی زیاد است. در پاره مواقع به علت وجود پدیده واکنش پذیری این روش‌ها نیز از روایی کامل برخوردار نیستند و ...

ویژگی عاطفی باید ملاک‌هایی را دارا باشد که از جمله این ملاک‌ها، دربرگیرنده احساس یا هیجان باشد. ویژگی معمولی یا نوع عاطفی فرد باشد. هر ویژگی عاطفی باید دارای سه ملاک خاص شدت، جهت و هدف باشد. شدت درجه یا نیروی احساس یا عاطفه است. جهت به جنبه مثبت یا منفی احساس گفته می‌شود. هدف به موضوعی که عاطفه یا احساس متوجه آن است، اشاره می‌کند.

۴۱- یکی از روش‌های تحلیل سوال‌های آزمون‌های تشریحی و عملکردی روش ویتنی و سبیرز است.

نتیجه آرماتی آزمون‌های وابسته به ملاک این است که در آنها اکثر آزمون شوندگان اکثر سوال‌ها را درست جواب دهند که در چنین حالتی دشواری سال‌ها بسیار بزرگ و ضریب تمیز بسیار کوچک خواهد بود.

در این فصل نظریه صفت مکنون و الگوی سوال پاسخ را توضیح دادیم در این نظریه‌ها با استفاده از منحنی ویژگی سوال ضرایب دشواری و تمیز سوال تعیین می‌شود.

۴۲- نمره معیار دیگری که در آزمون‌های میزان شده توانایی کاربرد زیادی دارد نمره سنی معیار است در این مقیاس استاندارد میانگین ۱۰۰ و انحراف معیار ۱۶ است.

توضیح بهنجار توزیعی است که بر منحنی بهنجار منطبق و دارای یک سری ویژگی‌ها می‌باشد. دو قسمت منحنی بهنجار نسبت به خط وسط آن قرینه دو امتداد منحنی بهنجار در بی‌نهایت محور افقی را قطع می‌کنند. در منحنی بهنجار میانگین، میانه و نما بر روی هم قرار دارند و ...

۴۳- روایی اصطلاحی است که به هدفی که آزمون برای تحقق بخشیدن به آن درست شده اشاره می‌کند. لین و گراثاند روایی را به صورت یک ارزشیابی از کفایت و مناسبت تفسیرها و استفاده‌های نتایج سنجش تعریف کردند. آزمون دارای روایی که برای اندازه‌گیری آنچه مورد نظر است کافی و مناسب باشد. آزمون‌های مورد استفاده در آموزش و پرورش و روانشناسی دارای سه هدف عمده هستند.

روایی محتوایی به این مطلب اشاره می‌کند که نمونه سوال‌های مورد استفاده در یک آزمون تا چه حد معرف کل جامعه سوال‌های ممکن است که می‌توان از محتوای آموزش مورد نظر تهیه کرد هرچه آزمون از این لحاظ بهتر باشد دارای روایی بیشتری است. برای ساختن آزمون‌های پیشرفت تحصیلی رواد تلاش کنیم تا سوال‌های آزمون نمونه کاملی از هدف‌ها و محتوای درس باشد.

روایی صوری یکی از مشتقات روایی محتوایی است. روایی صوری به این مطلب اشاره می‌کند که سوال‌های آزمون تا چه حد در ظاهر شبیه به موضوعی هستند که برای اندازه‌گیری آن تهیه شده‌اند.

منظور از روایی ملاکی میزان ارتباط بین نمرات حاصل از یک آزمون با نمرات حاصل از یک آزمون یا وسیله اندازه‌گیری دیگر است.

روایی پیش بینی ویژه آزمون‌هایی است که برای پیش‌بینی موفقیت افراد در امور تحصیلی یا شغلی به کار می‌روند. آزمون‌هایی که برای مقاصد پیش‌بینی به کار می‌روند مختلف هستند. گاه از آزمون‌های پیشرفت تحصیلی فعلی یا قبلی برای پیش‌بینی پیشرفت تحصیلی آینده استفاده می‌شود.

۴۴- یکی از مفهومی‌های نسبتاً تازه وابسته به روایی محتوایی روایی آموزشی است که به آن روایی برنامه درسی نیز گفته می‌شود. در آزمون‌های وابسته به ملاک روایی پیش‌بینی به دو منظور به کار می‌رود. پیش‌بینی عملکرد یادگیرنده در یک هدف آموزشی بر اساس هدف آموزشی دیگر - پیش‌بینی یادگیری در حد تسلط.

از عوامل موثر بر روایی آزمون می‌توان به موارد زیر اشاره کرد. اولین عامل تأثیر گذار دستورالعمل با راهنمای آزمون است. کیفیت سوال آزمون تأثیرگذار است. عامل دیگری که تأثیر دارد ترتیب قرار گرفتن سوال‌های آزمون به دنبال یکدیگر است.

۴۵- پایایی یک وسیله اندازه‌گیری عمدتاً به ثبات و دقت نتایج حاصل از آن اشاره می‌کند. رابطه بین روایی و پایایی از این قرار است که یک آزمون باید پایا باشد تا بتواند روا باشد. یک آزمون برای اینکه روا باشد باید نخست پایا باشد یعنی پایایی شرط روایی است اما روایی برای پایایی ضروری نیست.

پایایی بین مصححان منظور از مصححان در اینجا دو یا چند نفری هستند که برگه‌های امتحانی را تصحیح می‌کنند رفتارهای افراد یا فرایندها را مورد مشاهده قرار می‌دهند یا محصولات یا فرآورده‌های رفتار را درجه‌بندی می‌کنند.

یکی از معروف‌ترین روش‌های تعیین پایایی یک آزمون روش بازآزمایی است. در این روش آزمون را در دو نوبت با گروه واحدی از آزمون‌شوندگان اجرا می‌کنند و ثمره‌های حاصل را با هم مقایسه می‌کنند. ضریب همبستگی بین نمره‌های به دست آمده از دوباره اجرایی آزمون ضریب پایایی آزمون است. روش دیگری که برای رفع مشکلات روش بازآزمایی پیشنهاد شده است روش فرم‌های موازی یا هم‌ارز است در این روش دو آزمون معادل یا موازی برای یک موضوع تهیه می‌شود و آنها را در فاصله زمانی کوتاهی به یک گروه واحد از آزمون‌شوندگان می‌دهند. دو فرم یک آزمون در صورت موازی هستند که میانگین و واریانس آنها برابر باشد. ضریب همبستگی بین نمرات حاصل از اجرای فرم‌های موازی ضریب پایایی آن آزمون‌ها به حساب می‌آید.

در روش کودر ریچاردسون نیز آزمون تنها یک بار اجرا می‌شود اما در این روش همه ماده‌های آزمون تحلیل می‌شوند.

۴۶- پایایی شرط روایی است یعنی اینکه یک آزمون باید پایا باشد تا بتواند روا باشد بنابراین هر اندازه از میزان پایایی یک آزمون کاسته شود به همان نسبت از روایی آن کم می‌شود.

معروف‌ترین روش تعیین پایایی روش‌های درجه‌بندی یا رتبه‌بندی روش گیلفورد است.



ایران عرضه

مرجع نمونه سوالات

آزمون های استخدامی

به همراه پاسخنامه تشریحی

خدمات ایران عرضه:

- ارائه اصل سوالات آزمون های استخدامی
- پاسخنامه های تشریحی سوالات
- جزوات و درسنامه های آموزشی

برای دانلود رایگان جدیدترین سوالات استخدامی بخش های زیر، روی آن کلیک نمایید:

سنجش و اندازه گیری پیشرفت تحصیلی - راهنمای معلم در ارزشیابی توصیفی - سوالات آموزش و پرورش

« انتشار یا استفاده غیر تجاری از این فایل، بدون حذف لوگوی ایران عرضه، مجاز می باشد »

