فهرست مطالب

فصل اول: تعریف اصطلاحات

تعریف اندازه گیری (Measurement)

تعریف آزمون (Test)

تعریف آزمودن Testing)

تعریف سنجش (Assessment)

تعریف ارزشیابی (Evaluation)

مقایسه ی اندازه گیری، آزمودن، سنجش و ارزشیابی

ارزشیابی آموزشی و پژوهش آموزشی

تست های طبقه بندی شده فصل اول

پاسخنامه تست های طبقه بندی شده فصل اول

آزمون فصل اول

فصل دوم: مقیاس های اندازه گیری

مقیاس اسمی (Nominal scale)

مقیاس ترتیبی (Ordinal scale)

مقیاس فاصله ای (Interval scale )

مقیاس نسبی یا نسبتی (Ratio scale)

مقیاس های گسسته و پیوسته

تست های طبقه بندی شده فصل دوم

پاسخنامه تست های طبقه بندی شده فصل دوم

آزمون فصل دوم

فصل سوم: دسته بندی آزمون های مورد استفاده در روانشناسی و آموزش و پرورش

دسته بندی انواع آزمون ها

دسته بندی آزمون ها با توجه به چگونگی تهیه

تست های طبقه بندی شده فصل سوم

پاسخنامه تست های طبقه بندی شده فصل سوم

آزمون فصل سوم

فصل چهارم: دسته بندی ارزشیابی های آموزشی

دسته بندی ارزشیابی های آموزشی در قالب رویکردهای مختلف

دسته بندی ارزشیابی های آموزشی با توجه به موضوع ارزشیابی رویکردهای مختلف

دسته بندی ارزشیابی های آموزشی با توجه به ملاک های مورد استفاده رویکردهای مختلف

انواع سنجش با توجه به زمان اجرا و هدف

تفاوت ارزشیابی تراکمی و تکوینی

دسته بندی ارزشیابی های آموزشی با توجه به ارزشیابان ارزشیابی از معلم

تست های طبقه بندی شده فصل چهارم

پاسخنامه تست های طبقه بندی شده فصل چهارم

آزمون فصل چهارم

فصل پنجم: مراحل و فعالیت های ارزشیابی آموزشی

مراحل ارزشیابی آموزشی

فعالیت های ارزشیابی آموزشی

تست های طبقه بندی شده فصل پنجم

پاسخنامه تست های طبقه بندی شده فصل پنجم

آزمون فصل پنجم

فصل ششم: تهیه طرح ارزشیابی پیشرفت تحصیلی

تعریف ارزشیابی پیشرفت تحصیلی

طبقه بندی هدف های آموزشی

چگونگی توزیع سؤالات به بخش های مختلف

تست های طبقه بندی شده فصل ششم

پاسخنامه تست های طبقه بندی شده فصل ششم

آزمون فصل ششم

فصل هفتم: آزمون های تشریحی

مقدمه

قواعد تهیه سؤال های تشریحی

قواعد تصحیح آزمون های تشریحی

روش های نمره گذاری آزمون های تشریحی

تست های طبقه بندی شده فصل هفتم

پاسخنامه تست های طبقه بندی شده فصل هفتم

آزمون فصل هفتم

فصل هشتم: آزمون های کوتاه پاسخ

آزمون های کوتاه پاسخ پرسشی

آزمون های کوتاه پاسخ کامل کردنی

آزمون های کوتاه پاسخ تشخیصی یا تداعی

مزایای آزمون های کوتاه پاسخ

معایب آزمون های کوتاه پاسخ

آزمون های کوتاه پاسخ متشکل از مسائل عددی

قواعد تهیه سؤال های کوتاه پاسخ

تست های طبقه بندی شده فصل هشتم

پاسخنامه تست های طبقه بندی شده فصل هشتم

آزمون فصل هشتم

فصل نهم: آزمون های صحیح - غلط و جور کردنی

تعریف آزمون عینی

تفاوت آزمون های عینی و غیر عینی

آزمون های صحیح - غلط

انواع آزمون های صحیح - غلط

محاسن آزمون های صحیح - غلط

معایب آزمون های صحیح - غلط

قواعد تهیه آزمون های صحیح - غلط

آزمون های جور کردنی

انواع آزمون های جور کردنی

محاسن آزمون های جورکردنی

معایب آزمون های جور کردنی قواعد تهیه سؤال های جور کردنی

تست های طبقه بندی شده فصل نهم

پاسخنامه تست های طبقه بندی شده فصل نهم

آزمون فصل نهم

فصل دهم: آزمون های چندگزینه ای

تعریف آزمون چندگزینه ای

انواع آزمون های چند گزینه ای محاسن آزمون های چندگزینه ای

معایب آزمون های چند گزینه ای

قواعد تهیه آزمون های چندگزینه ای

تست های طبقه بندی شده فصل دهم

پاسخنامه تست های طبقه بندی شده فصل دهم

آزمون فصل دهم

فصل یازدهم: روش های سنجش فرآیندها و فرآورده های یادگیری غیرشناختی

آزمون های عملکردی

روش های سنجش مشاهده ای

سنجش هدف های حوزه عاطفی

تست های طبقه بندی شده فصل یازدهم

پاسخنامه تست های طبقه بندی شده فصل یازدهم

آزمون فصل یازدهم

فصل دوازدهم: روش های سنجش مستقیم رفتار

ابعاد مختلف رفتار

انواع روش های سنجش مستقیم رفتار

تست های طبقه بندی شده فصل دوازدهم

پاسخنامه تست های طبقه بندی شده فصل دوازدهم

آزمون فصل دوازدهم

فصل سیزدهم: اجرا، نمره گذاری و تحلیل آزمون

ترتیب قراردادن سؤالات آزمون به دنبال یکدیگر

تحلیل سؤال های آزمون

محاسبه ضریب دشواری سؤال

محاسبه ضریب تمییز سؤال

همبستگی سؤال با کل آزمون

همبستگی بین سؤال های آزمون

رابطه ی بین ضریب دشواری و ضریب تمییز

تحلیل گزینه های انحرافی

تحلیل سؤال های آزمون های تشریحی و آزمون های عملکردی

تحلیل سؤال های آزمون های وابسته به ملاک

استفاده از منحنی ویژگی سؤال برای تحلیل سؤال آزمون

تست های طبقه بندی شده فصل سیزدهم

پاسخنامه تست های طبقه بندی شده فصل سیزدهم

آزمون فصل سیزدهم

فصل چهاردهم: استفاده از روش های آماری برای تفسیر نمره های آزمون

تعاریف

اندازه های گرایش مرکزی

اندازه های تغییرپذیری

همبستگی بین نمرات

تست های طبقه بندی شده فصل چهاردهم

پاسخنامه تست های طبقه بندی شده فصل چهاردهم

آزمون فصل چهاردهم

فصل پانزدهم: هنجارها و نیمرخ ها

تعریف هنجار

انواع هنجارها

ویژگی های منحنی بهنجار

نیمرخ

تست های طبقه بندی شده فصل پانزدهم

پاسخنامه تست های طبقه بندی شده فصل پانزدهم

آزمون فصل پانزدهم

فصل شانزدهم: روایی آزمون

تعریف روایی و انواع آن

روایی محتوایی

روایی صوری (نوعی از روایی محتوایی است)

روایی ملاکی

روایی سازه

روایی آزمون های وابسته به ملاک

روایی آموزشی

عوامل مؤثر بر روایی آزمون

تست های طبقه بندی شده فصل شانزدهم

پاسخنامه تست های طبقه بندی شده فصل شانزدهم

آزمون فصل شانزدهم

فصل هفدهم: پایایی آزمون

روش های تعیین پایایی روش تعیین پایایی آزمون های سرعت پایایی نمرات اختلاف

عوامل مؤثر در پایایی آزمون های وابسته به هنجار

برآورد نمره حقیقی با استفاده از ضریب پایایی آزمون

خطای معیار اندازه گیری

استفاده از خطای معیار اندازه گیری برای برآورد نمره حقیقی

رابطه بین روایی و پایایی

پایایی روش های درجه بندی با رتبه بندی

پایایی آزمون های وابسته به ملاک

پایایی روش های سنجش مستقیم رفتار

تست های طبقه بندی شده فصل هفدهم

پاسخنامه تست های طبقه بندی شده فصل هفدهم

آزمون فصل هفدهم

آزمون های خودسنجی

سؤالات آزمون سراسری ۹۵

پاسخنامه آزمون سراسری ۹۵

سؤالات آزمون دکتری ۹۶

پاسخنامه آزمون دکتری ۹۶

سؤالات آزمون سراسری ۹۶

پاسخنامه آزمون سراسری ۹۶

سؤالات آزمون دکتری ۹۷

پاسخنامه آزمون دکتری ۹۷

سؤالات آزمون سراسری ۹۷

پاسخنامه آزمون سراسری ۹۷

سؤالات آزمون دکتری ۹۸

پاسخنامه آزمون دکتری ۹۸

سؤالات آزمون سراسری ۹۸

پاسخنامه آزمون سراسری ۹۸

پاسخنامه آزمون ها

منابع و مراجع

«فصل اول تعریف اصطلاحات»

در این فصل ابتدا اصطلاحات و مفاهیم موردنیاز را تعریف، سپس آنها را با یکدیگر مقایسه می کنیم.

تعریف اندازه گیری (Measurement)

در اندازه گیری، صفات یا ویژگی های افراد و اشیاء تعیین و مقدار آنها به صورت عدد یا رقم گزارش می شود. اندازه گیری فرایندی است که تعیین می کند یک شخص یا یک شیء به چه میزان از یک ویژگی برخوردار است.

در اندازه گیری، ویژگی یا صفتی را که اندازه گیری می شود، به صورت عدد و رقم مشخص می کنند؛ به طور مثال، به جای اینکه بگوییم «تهران از قم , شلوغ تر است» می گوییم «جمعیت تهران ۱۲ میلیون نفر، ولی جمعیت قم ۹۵۰ هزار نفر است.»

نانالی در تعریف اندازه گیری می گوید: «اندازه گیری دربردارنده قواعدی است برای نسبت دادن اعداد به اشیاء یا افراد، به گونه ای که صفت هایی از آن ها را به در صورت کمیت نشان میدهد. بنابراین او به کاربرد قواعد در اندازه گیری تأکید کرده است. منظور از کاربرد قاعده در اندازه گیری این است که روش های نسبت دادن اعداد باید به طور روشن بیان شود؛ به طور مثال، در نمره گذاری نتایج یک آزمون حساب، وقتی که جواب های درست دانش آموزان شمارش می شوند و به به تعداد جواب های درست آن ها نمره داده می شود، شاهد نمونه ای از کاربرد قاعده در اندازه گیری هستیم؛ بنابراین قواعد اندازه گیری، خط مشی هایی برای نشان و دادن مقادیر اشیاء مورد اندازه گیری هستند.

فایده بیان قواعد اندازه گیری این است که افرادی که می خواهند از وسایل اندازه گیری استفاده کنند، آن ها را به طور یکسان به کار می برند.

بیان قواعد اندازه گیری از جنبه های مهم استاندارد کردن یا (میزان کردن) آزمون های روانی در تربیتی است از مشخصات مهم آزمون های استاندار شده یا (میزان شده) این است که افراد. مختلفی که از این وسایل استفده می کنند، نتایج یکسانی بدست می آورند.

مثال 1: کدام ویژگی آزمون های استانداردشده (میزان شده) تضمین می کند اگر چند نفر آزمون واحدی را مورد استفاده قرار دهند، نتایج یکسانی به دست خواهند آورد؟

1) قواعد اندازه گیری

2) هنجارهای آزمون

3) نمرات استاندارد

4) اجرای متحدالشکل

پاسخ: گزینه «1» قواعد اندازه گیری که از ویژگی های آزمون های استاندارد شده است، تضمین می کند. اگر چند نفر آزمون واحدی را مورد استفاده قرار دهند، نتایج یکسانی به دست خواهند آورد.

مثال۲: اندازه گیری عبارت است از نشان دادن خصوصیت مورد اندازه گیری به طور:

1) کمی

۲) کیفی

3) کمی و کیفی

4) شکلی

پاسخ: گزینه «۱» اندازه گیری همان گونه که از شکل کلمه پیداست، «مقادیر» را بیان می کند و تعیین کننده ی میزان ویژگی است؛ بنابراین زمانی که کلمه ی اندازه گیری را مشاهده می کنیم، یعنی کمیت و مقدار مورد نظر است؛ همچنین یکی از تفاوت های بین اندازه گیری و سنجش این است که در اندازه گیری خصوصیات مورد اندازه گیری به صورت کمی بیان می شود، اما در سنجش، خصوصیات مورد سنجش به صورت کمی و کیفی بیان می شوند.

مثال 3: اندازه گیری را می توان چنین تعریف کرد:

۱) تلاش برای بیان ویژگی های افراد و اشخاص

2) شمردن تعداد واحدهای طول، عرض و ارتفاع

۳) فرآیند منظم برای کمی کردن یک متغیر پیوسته

۴) مقایسه ی پیشرفت تحصیلی دانش آموزان مختلف

پاسخ**:** گزینه «3» در تعریف اندازه گیری، ویژگی ها یا صفات اشیاء و افراد تعیین و مقدار آن ویژگی ها به صورت عدد گزارش می شود؛ بنابراین اندازه گیری عبارت است از فرایندی که تعیین می کند یک شخص یا یک شی چه میزان از یک ویژگی برخوردار است. ممکن است گزینه ۲ که به کلمه ی شمردن واحدها اشاره دارد به نادرست توسط داوطلب انتخاب شود، اما باتوجه به تعریف بالا در می یابیم که گزینه ی صحیح نمی باشد.

مثال ۴: اجرای متحدالشکل یک آزمون بیانگر رعایت کدام ویژگی اندازه گیری است؟

1) قواعد اندازه گیری

۲) کمی سازی

3) دوری از قضاوت ارزشی

4) نسبت دادن اعداد به افراد یا اشیاء

پاسخ: گزینه «۱» آزمون، وسیله یا روشی نظام مند برای اندازه گیری نمونه ای از رفتار است، بنابراین اگر آزمون به طور یک شکل برگزار شود، در واقع به قواعد اندازه گیری توجه شده است. گاهی بعضی قواعد بسیار بدیهی و ساده اند، مانند استفاده از خط کش برای اندازه گیری جهت اطلاع از طول و عرض میز و گاهی این قواعد به بیان دقیق تری نیاز دارند، مانند اندازه گیری ویژگی روانی افراد مانند هوش و یادگیری. بنابراین قواعد اندازه گیری، افراد مختلف را با ابزار یکسان آزمون می کند. فایده بیان قواعد اندازه گیری این است که کمک می کند تا افراد مختلفی که میخواهند از وسیله اندازه گیری استفاده کنند، آن را یکسان به کار برند. گزینه سوم دوری از قضاوت ارزشی، می تواند یکی از مواردی باشد که داوطلب در صورت تحلیل متن سؤال، به اشتباه انتخاب کند؛ در حالی که قضاوت ارزشی یا داوری و تعیین ارزش به ارزشیابی آزمون اشاره دارد.

تعریف، آزمون (Test)

آزمون وسیله یا روشی است نظام دار برای اندازه گیری نمونه ای از رفتار؛ هر چه نمونه بیشتر معرف ویژگی مورد نظر باشد، اندازه گیری دقیق تر و قابل اعتمادتر است. ویژگی های فیزیکی یا رفتار آشکار را به صورت مستقیم اندازه گیری کنیم؛ اما ویژگی های روانی مثل هوش، انگیزه، خلاقیت، یادگیری و مواردی از این قبیل را نمی توانیم به صورت مستقیم انداره گیری کنیم، بلکه باید آنها را به صورت غیرمستقیم اندازه گیری کنیم، یعنی رفتارهایی را که معرف آن ویژگی ها هستند، اندازه گیری کنیم.

روش ها و وسایل دیگری برای اندازه گیری ویژگی های روانی وجود دارد؛ با این حال، آزمون معمول ترین وسیله اندازه گیری ویژگی ها یا صفات روانی است. . آزمون طیف وسیعی از ابزارهای اندازه گیری را مانند مجموعه سوال های عینی چند گزینه ای، جور کردنی و صحیح - غلط، سؤال های تشریحی کوتاه پاسخ و انشایی شامل می شود؛ علاوه بر اینها شیوه های اندازه گیری روانی و تربیتی نظیر پرسشنامه و حتی آزمون های شفاهی و آزمون های حرکتی را نیز در برمی گیرد.

مثال ۵: معمول ترین وسیله اندازه گیری ویژگی های روانی ................. است.

1) مشاهده

2) آزمون

۳) مصاحبه

4) بررسی بالینی

پاسخ: گزینه «۲» معمول ترین وسیله اندازه گیری ویژگی های روانی «آزمون» است.

تعریف آزمودن Testing))

وقتی برای اندازه گیری ویژگی های روانی از آزمون استفاده می کنیم، این فعالیت را آزمودن می نامیم. آزمون، یک وسیله ی اندازه گیری است، هر وقت آزمونی به کار می رود، نوعی اندازه گیری انجام میشود.

نکته: هر نوع انداد کسری، آزمودن نید به اندازه گیری از آزمودن مفهوم گسترده تری دارد، زیرا آزمودن نوع خاصی از انداره گیری است. ممکن است اندازه گیری صورت پذیرد اما آزمونی در کار نباشد؛ مانند ثبت رفتار پرخاشگرانه یک دانش آموز از طریق مشاهده.

تعریف سنجش (Assessment)

فرایند جمع آوری اطلاعات درباره ی بادگیرندگان که به منظور کمک به تصمیم گیری در مورد رشد و پیشرفت یا شیوه های مورد استفاده برای جمع آوری اطلاعات در مورد عملکرد دانش آموزان و دانشجویان، انجام میشود.

سنجش از اندازه گیری و آزمودن گسترده تر است. اندازه گیری و آزمون فقط به کمیّت اشاره دارد یعنی زمانی که اندازه گیری یا آزمودن انجام داده می شود نتیجه به صورت اعداد (کمیّت) نشان داده می شود؛ ولی سنجش هم «کمیّ» و هم «کیفی» در سنجش، از فنون مختلفی برای اندازه گیری استفاده می شود، همچنین سنجش بر مشاهده ی عملکرد تاکید زیادی دارد. سنجش، نوعی تحلیل بالینی و بیش بینی عملکرد است، در حالی که آزمودن یک ابزار است، به عبارت دیگر، آزمودن یک وسیله اندازه گیری است. سنجش هم می تواند بصورت رسمی انجام شود (مانند اجرای آزمون ها) و هم به صورت غیررسمی (مانند مشاهده رفتار دانش آموزان در بین هم کلاسی هایشان) سالورا و بسل دانک می گویند: « سنجش تنها در جمع آوری اطلاعات خلاصه نمی شود، بلکه به جمع آوری داده ها همراه با «قصد و منظور» گفته می شود». این متخصصان عقیده دارند، سنجش در آموزش و پرورش به دو منظور به کار می رود:

1) تعیین و تشخیص مشکلات دانش آموزان

۲) تصمیم گیری در مورد آنان

اصطلاح سنجش بیشتر در «حوزه ی روانشناسی بالینی» و «آموزش و پرورش» به کار می رود. زیرا این دو حوزه از وسایل مختلف برای تعیین رفتار ها و ویژگی های روانی و تربیتی استفاده می کنند.

مثال ۶: کدام یک از موارد زیر بر مشاهده عملکرد تأکید زیادی دارد؟

۱) آزمون

۲) اندازه گیری

۳) سنجش

4) تحلیل بالینی

پاسخ: گزینه «3» سنجش بر مشاهده عملکرد تأکید زیادی دارد.

تعریف ارزشیابی (Evaluation)

ارزشیابی، فرآیند نظام دار جمع آوری، تحلیل و تفسیر اطلاعات است. ارزشیابی تعیین می کند که آیا اهداف مورد نظر تحقق یافته اند یا در حال تحقق یافتن هستند و به چه میزانی. به عبارت دیگر، ارزشیابی عبارت است از داوری ارزشی درباره کیفیت موضوع مورد ارزشیابی براساس جمع آوری، تفسیر و تحلیل اطلاعات. یکی از ویژگی های مهم ارزشیابی تلین کیفیت است؛ یعنی داوری ارزشی با توجه به «کیفیت» انجام می شود. پس از اینکه داوری ارزش انجام شد، این ارزشیابی بر مبنای تصمیم گیری برای فعالیت های عملی و اجرایی می شود.

مثال ۷: یکی از مهمترین ویژگی های ارزشیابی تعیین .................. می باشد.

1)کمیت

۲) کیفیت

۳) کمیت و کیفیت

4) اعتبار

**پاسخ:** گزینه «۲» تعیین کیفیت از مهم ترین ویژگی های ارزشیابی است.

مثال ۸: مفهوم اصطلاح ارزشیابی (Evaluation) کدام است؟

1) انتساب اعداد به ویژگی های افراد یا اشیا

2) تصحیح و نمره گذاری ورقه امتحانی دانش آموز

۳) مشخص کردن میزان همکاری دانش آموز به شیوه های مختلف

4) مقایسه نمرات دانش آموز با معیارهای از پیش تعیین شده

پاسخ: گزینه «۴» ارزشیابی با ارزیابی در تعریف به تعیین ارزش برای هر چیزی گفته میشود. (گی ۱۹۹۱) تعریف جامع تری درباره ی ارزشیابی ارائه می دهد. او می گوید: «ارزشیابی فرآیندی نظام دار برای جمع آوری، تحلیل و تفسیر اطلاعات گفته می شود، به این منظور که آیا هدف های مورد نظر تحقق یافته اند یا در حال تحقیق یافتن هستند و به چه میزانی؟» یکی از ویژگی های مهم ارزشیابی تعیین کیفیت است.

کیفیت در ارزشیابی پیشرفت تحصیلی، کیفیت به دانش، مهارت ها و توانایی هایی گفته میشود که از دانش آموزان مورد انتظار است. ارزشیابی آموزشی علاوه بر پیشرفت تحصیلی دانش آموزان، موارد دیگری را هم شامل می شود؛ مثل عملکرد معلمان، مدیران، روش های آموزشی و مواردی از این قبیل. ارزشیابی های - آموزشی، نسبت به ارزشیابی پیشرفت تحصیلی - که در آن فقط موقعیت ها و دستاوردهای یادگیری دانش آموزان و دانشجویان مورد نظر است – معنای گسترده تری دارند. مفهوم ارزشیابی پیشرفت تحصیلی به مفهوم سنجش یادگیری بسیاری نزدیک است، اما ارزشیابی آموزشی در مفهوم کلی آن معنای بسیار وسیعی دارد.

ورتن و سندرز دو تن از صاحب نظران در حوزه ی ارزشیابی آموزشی گفته اند: «در آموزش و پرورش، ارزشیابی به فعالیتی رسمی گفته می شود که برای تعیین کیفیت، اثربخشی یا ارزش یک برنامه، فرآورده یا پروژه، فرآیند، هدف یا برنامه ی درسی اجرا می شود. منظور از کیفیت در ارزشیابی پیشرفت تحصیلی، دانش ها، مهارت ها و توانایی هایی است که از دانش آموز مورد انتظار است. منظور از کیفیت در ارزشیابی آموزشی، تناسب موضوع مورد ارزشیابی با اهداف است.

مقایسه ی اندازه گیری، آزمودن، سنجش و ارزشیابی

همان طور که گفتیم ارزشیابی داوری ارزشی در مورد کیفیت موضوع مورد ارزشیابی بر مبنای تحلیل و تفسیر اطلاعات جمع آوری شده است که این اطلاعات از طریق اندازه گیری (عموما آزمودن) و سنجش جمع آوری می شوند.

سر سنجش از فنون مختلفی برای جمع آوری اطلاعات استفاده می کند. همان طور که گفتیم اندازه گیری و آزمودن هر دو به کمیت توجه دارند ولی سنجش هم کمی است و هم کیفی.

اندازه گیری از آزمودن وسیع تر است و سنجش از آزمودن و اندازه گیری وسیع تر است اما ارزشیابی هر سه اندازه گیری، آزمودن و سنجش) را در برمی گیرد. برای روشن شدن تمایز بین سنجش و ارزشیابی می توان گفت در حالی که سنجش به ما می گوید شیء مورد نظر کدام ویژگی ها را دارد، ارزشیابی قدر یا ارزش آن ویژگی ها را مشخص می کند؛ به عبارت دیگر، ارزشیابی یعنی داوری درباره ی کیفیت با ارزش اطلاعات به دست آمده از سنجش های مختلف.

ارزشیابی سنجش+ اندازه گیری + آزمودن + قضاوت

به عنوان مثال، یک معلم علوم درباره اثربخشی روش خود بر موفقیت دانش آموزان، آزمونی را تهیه، اجرا و نمره گذاری می کند و در مرحله بعد اطلاعات به دست آمده را مورد تحلیل و تفسیر قرار می دهد:

تهیه ی آزمون و نمره گذاری آن: اندازه گیری - مجموعه سؤالات تهیه شده: آزمون - اجرای سؤالات آزمون - تحلیل، تفسیر و داوری: (ارزشیابی).

چون فقط از آزمون استفاده کرده است، ما سنجش نداریم؛ زیرا سنجش از روشهای مختلفی استفاده می کند، مثلا اگر علاوه بر آزمون از راه های دیگری نیز به جمع آوری اطلاعات می پرداخت؛ به عنوان مثال از دانش آموزان می خواست به صورت فردی یک پروژه تحقیقی انجام دهند و نتایج کارشان را به طور کتبی به معلم گزارش دهند. مجموعه این اقدامات به اضافه اجرای آزمون سنجش به حساب می آید.

مثال ۹: از اصطلاحات زیر کدام یک معنای محدودتری دارد؟

1) اندازه گیری

۲) سنجش

3) آزمودن

4) ارزشیابی

پاسخ: گزینه «3» آزمودن از اندازه گیری، سنجش و ارزشیابی معنای محدودتری دارد.

ارزشیابی آموزشی و پژوهش آموزشی

شباهت: ارزشیابی آموزشی و پژوهش، آموزشی هر دو، بر روش های بررسی تجربی متکی هستند.

تفاوت ها:

1\_ هدف از ارزشیابی آموزشی تصمیم گیری است در حالی که هدف پژوهش آموزشی استنتاج است.

2\_ در پژوهش آموزشی مثل هر پژوهش دیگری تعمیم نتایج بسیار مهم است ولی ارزشیابی آموزشی بر یک پدیده آموزشی ویژه متمرکز است و تعمیم در آن نقش ندارد.

3\_ ارزشیابی آموزشی با «ارزش» سر و کار دارد ولی پژوهش «خالی از ارزش» است. پژوهشگر به دنبال حقایق علمی است و هیچ گونه قدر و ارزش مثبت یا منفی برای حقایق به دست آورده خود قائل نمی شود.

۴- پژوهش به نظریه وابسته است، ولی ارزشیابی معمولا مبنای نظری ندارد.

5\_ هدف پژوهش، تولید دانش و اضافه کردن آن به تئوری های قبلی است، در حالی که ارزشیابی به منظور حل مشکلات ویژه و کمک به تصمیم گیری انجام میشود.

تست های طبقه بندی شده فصل اول

1\_کدام یک از فعالیت های زیر موجب می شود که اعضای کلاس از همدیگر متفاوت باشند؟

(آزاد ۸۳)

1) شمارش

۲) اندازه گیری

3) دادن کمیت

4) ارزشیابی

۲- آزمون وسیله ای است عینی و محاسبه شده برای اندازه گیری:

(آزاد ۸۳)

۱) نمونه ای از حالت و رفتار معین فردا

2) ویژگی های حالت و رفتار معین فرد

۳) رفتار بارز و حالات خاص فرد

4) رفتار عمومی و حالات طبیعی فرد

۳- به مجموعه فعالیت هایی که به منظور ساخت، اجرا و نمره گذاری آزمون به کار برده می شود، چه می گویند؟

(آزاد ۸۳)

1) ارزشیابی

۲) ارزیابی

۳) اندازه گیری

4) سنجش

۴- در صورتی که گفته شود حسن به ۳۰ سؤال یک آزمون ۵۰ سؤالی پاسخ درست داده است، کدام یک از فعالیت های زیر انجام شده است؟

(آزاد ۸۳)

1) اندازه گیری

۲) سنجش

۳) ارزیابی

4) ارزشیابی

5\_ کدام یک از مفاهیم زیر دارای قدمت کمتری در متون مربوط به ارزشیابی پیشرفت تحصیلی است؟

(آزاد ۸۵)

1) اندازه گیری Measuremen

2) ارزشیابی Evaluation

3) تحقیق Research

4) سنجش Assessment

6**\_** در کدام یک از موارد زیر بیشترین اصالت را به هدف های آموزشی از پیش تعیین شده می دهند؟

۱) ارزشیابی

۲) اندازه گیری

3) ارزیابی

4) سنجش

7**\_** تعریف واژه «اندازه گیری» شامل کدام یک از مضامین زیر است؟

(آزاد ۸۶)

1) تعبیر و تفسیر کمی صفات مورد اندازه گیری

2) تعیین شاخص عددی برای صفت مورد اندازه گیری و قضاوت درباره ی آن

3) تعیین شاخص عددی برای صفت مورد اندازه گیری براساس قواعدی از پیش تعیین شده

۴) ارزشیابی و سنجش صفت ویژه ای براساس هدفهای مورد نظر

۸**-** اولین کتاب درباره ی اندازه گیری توسط چه کسی انتشار یافت؟

(سراسری ۸۷)

۱) بینه

۲) پیرسون

3) ثورندایک

4) کتل

۹**-** جمع آوری اطلاعات به منظور قضاوت در مورد پدیده ها و تصمیم گیری ................. نامیده می شود.

(سراسری ۸۸)

۱) سنجش

۲) ارزیابی

۳) اندازه گیری

4) ارزشیابی

۱۰**-** کدام عبارت و تعریف، بیانگر فرایند «سنجش» می باشد؟

(سراسری ۸۸)

۱) اختصاص کمیت به خصیصه ها یا صفات مختلف

۲) تفسیر، ارزش گذاری و داوری دربارهی عملکرد اختصاصی فرد

۳) تحلیل چندجانبه ی عملکرد اختصاصی فرد با استفاده از فنون مختلف

۴) اجرای آزمون پیشرفت تحصیلی به منظور تخمین عملکرد شغلی

11\_ ارزشیابی عبارتست از:

(آزاد ۸۹)

1) روالی که در آن اندازه گیری مجدد انجام میشود.

۲) روالی که در آن تعبیر و تفسیر اندازه گیری ها براساس هدف های مورد نظر صورت می گیرد.

۳) روالی که در آن هدفها تعبیر و تفسیر میشوند.

4) روالی که در آن اندازه ها تعبیر و تفسیر میشوند.

۱۲**-** مهمترین تفاوت اندازه گیری در علوم پایه و علوم انسانی، به متفاوت بودن ماهیت ................. ارتباط پیدا می کند.

(سراسری ۹۰)

۱) آزمودنی ها

۲) کاربردها

۳) روش های اندازه گیری

4) موضوع اندازه گیری

پاسخنامه تست های طبقه بندی شده فصل اول

1- گزینه «۴**»** در شمارش، تنها افراد یا اشیاء را از لحاظ دارا بودن یک عامل یا ویژگی بررسی می کنیم که این فعالیت نمی تواند هیچ گونه تمایزی بین اعضای کلاس قائل شود و در واقع یکی از فعالیت هایی است که در حوزهی اندازه گیری انجام می شود. دادن کمیت نیز که خود، تعریفی از اندازه گیری است، هم نمی تواند بین دانش آموزان تمایز قایل شود. تنها ارزشیابی است که بین اعضای کلاس از همدیگر براساس هدف از پیش تعیین شده و بر مبنای اندازه گیری های صورت گرفته تمایز و تفاوت قائل می شود و باعث می شود افراد از همدیگر متفاوت باشند.

٢- گزینه «۱» رفتارهای معرف ویژگی های روانی بسیار متنوع اند؛ بنابراین نمی توان همه ی آنها را اندازه گیری کرد. متخصص اندازه گیری، نمونه ای از - رفتارهای مورد نظرش را انتخاب و آزمون خود را بر پایه آن نمونه بنا می نهد، پس طبق تعریف «آزمون، وسیله یا روشی نظام دار برای اندازه گیری نمونه ای از رفتار است.»

۳- گزینه«3**»** گزینه ۱ و ۲ اشتباه می باشند؛ چون هر دو اصطلاحاتی هستند که معادل هم به کار برده می شوند و به تعیین ارزش برای هر چیزی یا داوری و ارزشی در مورد چیزی اشاره دارند؛ به عبارت دیگر ارزشیابی فرآیندی نظامدار برای جمع آوری تحلیل و تفسیر اطلاعات است، به این منظور که تعیین شود آیا هدف های مورد نظر تحقق یافته اند پا در حال تحقق یافتن هستند و به چه میزان، گزینه ۴ نیز اشتباه است؛ چون سنجش، مفهوم گسترده تری نسبت به در آزمودن و اندازه گیری دارد و به فرآیندی گفته می شود که برای گردآوری اطلاعات مورد نیاز، تصمیم گیری درباره دانش آموزان، برنامه درسی و سیاست های آموزشی استفاده می شود؛ بنابراین گزینه صحیح گزینه ۳ (اندازه گیری) است که به مجموعه فعالیت هایی که به منظور ساخت، اجرا و نمره گذاری آزمون به کار برده می شود، گفته میشود.

۴- گزینه «2» با توجه به صورت سؤال، تنها کار صورت گرفته این است که به نمره گذاری آزمونی که حسن داده است، پرداخته شده و همانطور که قبلا در سؤال 3 گفتیم اندازه گیری فعالیتی است که به منظور ساخت، اجرا و نمره گذاری آزمون به کار برده می شود؛ بنابراین گزینه 1 یعنی اندازه گیری صحیح می باشد.

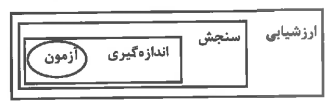
5\_ گزینه «۴» کلمه Assessment به معنای سنجش، قدمت کمتری در متون مربوط به ارزشیابی پیشرفت تحصیلی دارد.

6\_ گزینه «۱**»** اصطلاح ارزشیابی به تعیین ارزش برای همه چیز گفته می شود. ارزشیابی فرایندی نظام دار برای جمع آوری، تحلیل و تفسیر اطلاعات است و به این منظور انجام می شود که بفهمیم آیا هدف های مورد نظر تحقق یافته اند یا در حال تحقق یافتن هستند و به چه میزانی. شایان ذکر است که ارزشیابی و ارزیابی هم معنی می باشند و ارزیابی گزینه ی انحرافی است.

۷- گزینه «3» گزینه ۱ چون به تعبیر و تفسیر اشاره دارد، صحیح نمی باشد. گزینه ۲. نیز چون قضاوت را مد نظر قرار داده که در ارزشیابی مطرح است : صحیح نیست. گزینه ۴ نیز صحیح نیست، چون به طور صریح به ارزشیابی و سنجش اشاره کرده است که اصطلاحاتی متفاوت از اندازه گیری می باشند. بنابراین تنها گزینه صحیح گزینه 3 می باشد که دقیقا واژه اندازه گیری را تعریف کرده است.

8\_گزینه «3» ثورندایک اولین کتاب درباره اندازه گیری را منتشر کرد. او همچنین به نظریه شرطی سازی محرک - پاسخ، قانون اثر و قانون پاداش اشاره کرد. روش یادگیری ثورندایک، یادگیری از راه کوشش و خطا است.

9- گزینه «۴**»** سنجش مفهومی کلی دارد؛ حتی کلی تر از اندازه گیری و آزمون در سنجش به هر دو بعد کمیت و کیفیت توجه می شود. در واقع، سنجش از فنون مختلف استفاده می کند تا بتواند به یک تحلیل جامع و چندوجهی از عملکرد دست یابد. در سنجش، به نوعی تحلیل بالینی و پیش بینی عملکرد دست پیدا می کنیم، ولی ارزشیابی، قضاوت درباره کیفیت با ارزش اطلاعات به دست آمده از سنجش های مختلف است.



10\_ گزینه «3»سنجش، طبق تعریف و شکل ارائه شده در سؤال قبلی، دربرگیرنده همه راه های نمونه گیری و مشاهده است. سنجش، بیشتر در حوزه روان شناسی و آموزش و پرورش کاربرد دارد؛ ات را در این حوزه علاوه بر آزمون و پرسشنامه از وسایل و روش های مختلف دیگری هم برای تعیین رفتارها و ویژگی های روانی و ترتیبی افراد استفاده می شود؛ مثلا در سنجش به صورت رسمی از آزمون و به صورت غیررسمی از مشاهده آزمودنی استفاده می شود. علاوه بر این، گزینه ۱ به اندازه گیری اشاره دارد. گزینه ۲ که به تفسیر، ارزش گذاری و داوری اشاره دارد، مربوط به ارزشیابی است. گزینه ۴ نیز که اجرای آزمون را مد نظر دارد، فعالیتی در حوزه آزمودن می باشد؛ بنابراین تنها گزینه صحیح، گزینه 3 می باشد.

11\_ گزینه «۲» اصطلاح ارزشیابی به طور ساده به تعیین ارزش برای هر چیزی یا داوری ارزشی کردن گفته می شود. ارزشیابی را می توان به این شکل تعریف کرد: «ارزشیابی به یک فرایند نظامدار برای جمع آوری، تحلیل و تفسیر اطلاعات گفته می شود؛ به این منظور که تعیین شود آیا هدف های مورد نظر تحقق یافته اند یا در حال تحقق هستند و به چه میزانی.» با توجه به این توضیحات تنها گزینه صحیح گزینه ۲ می باشد.

۱۲\_ گزینه «۴**»** روش تحقیق در هر علمی و به تبع آن اندازه گیری در آن علم با توجه به موضوعی که آن علم به بررسی آن می پردازد، مشخص می شود. از آنجا که در علوم انسانی، انسان مورد بررسی قرار می گیرد و در علوم پایه، موضوعاتی مانند عوامل فیزیکی و شیمیایی مواد بررسی می شوند، همین امر باعث تفاوت اندازه گیری های صورت گرفته در این علوم می شود.

آزمون فصل اول

1- کدام یک از بیانات زبر صحیح است؟

1) سنجش همان اندازه گیری است و با آن تفاوتی ندارد.

2) سنجش و ارزشیابی دو اصطلاح معادل هستند.

3) ارزشیابی، سنجش، اندازه گیری و آزمودن را در بر می گیرد.

۴) سنجش همان اندازه گیری است که با وسایلی به جز آزمون انجام می شود.

۲- تفاوت سنجش اندازه گیری و آزمودن در این است که:

۱) سنجش فقط با کمیت سروکار دارد اما اندازه گیری و آزمودن با کیفیت.

2) اندازه گیری و آزمودن عموما با کمی سروکار دارند اما سنجش با کیفیت

۳) اندازه گیری و آزمودن کمی اند اما سنجش هم کمی است و هم کیفی

۴) سنجش و اندازه گیری کیفی اند اما آزمودن کمی است.

۳- کدام یک از جنبه های مهم استاندارد کردن یا میزان کردن آزمون های روانی و تربیتی است؟1

1) قواعد اندازه گیری

2) هنجارهای آزمون

3) اجرای متحدالشکل

۴) نمرات استاندارد

۴- از اصطلاحات زیر کدام یک معنای وسیع تری دارند؟

1) سنجش

۲) اندازه گیری

۳) آزمودن

۴) ارزشیابی

5\_ شباهت ارزشیابی آموزشی و پژوهشی در چیست؟

1) تعمیم نتایج

۲) داوری ارزشی

۳) روش های بررسی تجربی

۴) هر دو با ارزش سروکار دارند

6\_ هدف از ارزشیابی آموزشی چیست؟

1) حل مشکلات

۲) استنتاج

۳) تعمیم نتایج

۴) تصمیم گیری

۷- تفاوت ارزشیابی آموزشی و پژوهش آموزشی کدام گزینه است؟

۱) پژوهش با ارزش سروکار دارد ولی ارزشیابی خالی از ارزش می باشد.

2) هدف ارزشیابی آموزشی استنتاج می باشد و هدف پژوهش آموزشی تصمیم گیری است.

۳) هدف پژوهش. تمرکز بر یک پدیده می باشد ولی ارزشیابی به تعمیم نتایج می پردازد.

۴) ارزشیایی وابسته به نظریه نمی باشد ولی پژوهش مبتنی بر نظریه میباشد.

۸- اگر معلم به یکی از دانش آموزان خود بگوید که: «تو در درس ریاضی به خوبی عمل کرده ای» این کار معلم نوعی .................. است.

۱) اندازه گیری

2) سنجش

۳) آزمودن

۴) ارزشیابی

۹- کدام عبارت صحیح می باشد؟

1) گستره سنجش از آزمودن محدودتر است.

۲) سنجش یکی از ابزارهای آزمودن به حساب می آید.

۳) سنجش از آزمودن وسیع تر است

۴) آزمون مخصوص ویژگی های روانی است؛ در صورتی که سنجشی کاربرد عمومی تری دارد.

۱۰- اگر معلم ورزش بیان کند که میانگین قد دانش آموزان کلاس او ۱۵۰ می باشد، او دست به کدام فعالیت زپر زده است؟

1) آزمودن

۲) اندازه گیری

۳) سنجش

۴) ارزشیابی

فصل دوم

«مقیاس های اندازه گیری»

مقیاس مجموعه ای از اعداد (یا نمادهای دیگر است که ویژگی های آنها بر ویژگی های تجربی اشیایی که اعداد به آن ها نسبت داده می شوند، منطبق است. مقیاس های اندازه گیری به چهار دسته: اسمی، ترتیبی، فاصله ای و نسبتی تقسیم می شوند که این طبقه بندی از ساده به پیچیده و به صورت سلسله مراتبی تنظیم شده است؛ یعنی هر طبقه، خصوصیات حطبقه قبل از خود به علاوه ویژگی های مخصوص به خود را دارد.

مقیاس اسمی (Nominal scale)

این مقیاس ساده ترین و پایین ترین سطح اندازه گیری است و به دو منظور به کار میرود:

1\_ اسم گذاری: در اسم گذاری از اعداد و ارقام برای نام گذاری و تشخیص افراد و اشیاء استفاده می شود؛ اما این اعداد و ارقام هیچ گونه معنا و مفهوم ریاضی ندارند؛ مثل شماره پیراهن بازیکنان فوتبال یا شماره خیابانهای شهرها.

٢- طبقه بندی: در مقیاس اسمی از اعداد و ارقام برای طبقه بندی اشیاء و افراد نیز استفاده می شود؛ اما این اعداد و ارقام هیچ گونه معنا و مفهوم ریاضی ندارند و تنها برای مشخص کردن طبقه های مختلف در یک طبقه بندی به کار می رود: مثل طبقه بندی افراد به ۱= زن و ۲= مرد.

در این مقیاس:

عملیات مجاز آماری: فراوانی و نما است.

عملیات مجاز ریاضی: هیچ یک از چهار عمل اصلی مجاز نیست.

مثال 1: اگر افراد را از نظر دینی به 1= مسلمان، ۲= مسیحی و ۳= یهودی تقسیم کنیم، از کدام مقیاس اندازه گیری استفاده کرده ایم؟

1) ترتیبی

۲) فاصله ای

3) اسمی

۴) نسبتی

پاسخ: گزینه «3» چون فقط از اعداد و ارقام برای طبقه بندی استفاده شده است و این اعداد و ارقام هیچ گونه معنای ریاضی ندارند.

مثال ۲: در شماره گذاری پیراهن بازیکنان یک تیم فوتبال از کدام مقیاس استفاده می شود؟

1) ترتیبی

2) فاصله ای

3) اسمی

4) نسبتی

پاسخ: گزینه «3» چون این اعداد و ارقام هیچ گونه معنا و مفهوم ریاضی ندارند و فقط برای نامگذاری استفاده شده اند.

مقیاس ترتیبی (Ordinal scale)

در این مقیاس، گروهی از افراد را با توجه به یک صفت از کوچک به بزرگ یا بالعکس مرتب می کنیم، اما معلوم نیست که کدام یک از افراد به صورت مطلق آن صفت را دارند؛ همچنین مشخص نیست که این افراد از نظر آن صفت چقدر با هم فاصله دارند؛ مثلا وقتی افراد را از نظر قد به کوتاه، متوسط و بلند تقسیم می کنیم، مقیاس ترتیبی، وضعیت نسبی فرد را در یک دامنه مشخص می کند. اگر هدف از اندازه گیری، انتخاب بهترین فرد یا مناسب ترین شی باشد، این مقیاس، مقباس مناسبی است.

مشخصه مقیاس ترتیبی این است که درباره تفاوت بین نقاط در این مقیاس هیچ گونه اطلاعی موجود نیست.

در این مقیاس:

عملیات مجاز آماری: فراوانی، نما، میانه، محاسبه درصدها و ضریب همبستگی اسپیرمن است.

عملیات مجاز ریاضی: در این مقیاس هیچ یک از چهار عمل اصلی مجاز نیست.

مثال ۳: اگر افراد را از نظر سطح استرس به ۱= بالا، ۲= متوسط و ۳= پایین تقسیم کنیم، از کدام مقیاس اندازه گیری استفاده کرده ایم؟

1) اسمی

۲) ترتیبی

۳) فاصله ای

۴) نسبتی

پاسخ: گزینه «۲» در مقیاس ترتیبی، افراد را براساس یک صفت از کوچک به بزرگ مرتب می کنیم.

مثال ۴: در یک آزمون سنجش نگرش، نظر افراد در چهار گروه (کاملا مخالف ـ مخالف ـ موافق ـ کاملا موافق) سنجش شـده اسـت. مقیـاس مـورد استفاده کدام است؟

1) ترتیبی

۲) اسمی

۳) فاصله ای

4) نسبتی

پاسخ: گزینه «۱» مقیاس مورد استفاده در این طبقه بندی «مقیاس ترتیبی» است.

مثال 5: اگر گفته شود، عملکرد حسین در آزمون ریاضی بهتر از شیمی بوده است، از کدام مقیاس استفاده شده است؟

۱) اسمی

2) ترتیبی

۳) فاصله ای

4) نسبتی

پاسخ: گزینه «۲» در مقیاس ترتیبی افراد یا اشیاء باتوجه به یک صفت از بزرگ به کوچک و بالعکس مرتـب مـی شـوند؛ مثلاً بـرای مـرتـب کـردن دانش آموزان براساس ترتیب قد یا نمره درسی از مقیاس ترتیبی استفاده می شود. اگر هدف از اندازه گیری، انتخاب بهترین فرد یا مناسب ترین شیء و مواردی از این قبیل باشد، این مقیاس، مقیاسی مناسب است.

با توجه به این که بین مقیاس ترتیبی و مقیاس فاصله ای شباهت زیادی وجود دارد، داوطلب باید دقت لازم را در زمینه انتخاب گزینه درست داشته باشد.

مقیاس فاصله ای (Interval scale)

در این مقیاس رتبه اشیاء یا افراد باتوجه به یک صفت مشخص است، فاصله بین رتبه ها برابر است و این مقیاس دارای صفر قراردادی است؛ مثلاً نمرات آزمون هوش در مقیاس فاصله ای قرار دارند. وقتی می گوییم صفر موجود در مقیاس فاصله ای قراردادی است، یعنی اگر به هنگام اندازه گیری یک خصیصه در این مقیاس نمره صفر حاصل شد، به معنای نبود مطلق آن خصیصه نیست؛ به طور مثال وقتی که می گوییم نمره هوش فردی در آزمونی برابر صفر گرفته است، بدان معنا نیست که فرد اصلا توانایی هوشی ندارد یا کسی که در آزمون ریاضی نمره صفر گرفته است، بدین معنا نیست که او هیچ چیز درباره ریاضی نمی داند. از سوی دیگر چون اعمال ضرب و تقسیم در این مقیاس ممکن نیست، نمی توانیم بگوییم که پیشرفت تحصیلی کسی که در درس عربی نمره ۸ گرفته، دو برابر کسی است که در این درس نمره ۹ گرفته است.

نکته 1: تفاوت مقیاس فاصله ای و ترسی در این است که در مقیاس فاصله ای، تواصل من واحدها برابر و معلوم هستند. دردوری که در مقیاس ترتیبی، فواصل بین واحدها معلوم نیست مثلا در نمرات آزمون هوش کسی که نمره ۱۲۰ را به دست آورده، باهوش تر از کسی است که نمره ۱۱۰ را به دست آورده است؛ همچنین تفاوت بین بهره های هوشی و ۱۱۵، برابر با تفاوت بین بهره های هوشی ۱۱۰ و ۱۰۰ است یعنی فواصل مساوی اند.

در این مقیاس:

عملیات مجاز آماری: کلیه اعمال آماری مجازند.

عملیات مجاز ریاضی: جمع و تفریق مجاز است اما ضرب و تقسیم مجاز نیست.

مثال ۶: مقیاس اندازه گیری هوشبهر انحرافی کدام است؟

1) اسمى

۲) ترتیبی

۳) فاصله ای

4) نسبتی

پاسخ: گزینه «۳» هوشبهر انحرافی دارای مقیاس فاصله ای میباشد.

مثال 7: در کدام یک از مقیاس های اندازه گیری، واحدها برابر است و صفر قراردادی وجود دارد؟

۱) اسمی

۲) رتبه ای

۳) فاصله ای

۴) نسبی

پاسخ: گزینه «۳» به ترتیب مقیاس ها عبارت اند از: مقیاس اسمی، مقیاس ترتیبی، مقیاس فاصله ای و مقیاس نسبی. در دو مقیاس آخـر صـفـر وجـود دارد که . در مقیاس فاصله ای صفر قراردادی و در مقیاس نسبی صفر، مطلق است. برای پاسخ به سؤال، داشتن اطلاعات ارائه شده لازم است و ممکـن اسـت داوطلب با دیدن کلمه صفر، گزینه ی ۴ را انتخاب کند که انتخاب درستی نیست.

مقیاس نسبی یا نسبتی (Ratio scale)

این مقیاس بالاترین سطح اندازه گیری است و همه ویژگی های طبقات قبلی را دارد. علاوه بر آن، دارای صفر مطلق است؛ یعنی نمره صفر در این مقیاس معنای عدم وجود کامل خصیصه ها مورد اندازه گیری است. مهمترین مقیاس نسبی، همان مقیاس اعداد از صفر تا بی نهایت است. در این مقیاس کلیه اعمال آماری و ریاضی مجاز هستند.

نکنه 2: هر مقیاسی را می توان در مقیاس های قبل از خود تبدیل نمود، اما به هیچ وجه نمی توان مقیاسی را به مقیاس بعد از خودش تبدیل کرد. مثلا مقیاس اسمی را نمی توان به ترتیبی تبدیل کرد.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ویژگی های هر مقیاس | مقیاس اسمی | مقیاس ترتیبی | مقیاس فاصله ای | مقیاس نسبتی |
| نام گذاری و طبقه بندی | نام گذاری و طبقه بندی  +  وجود ترتیب در طبقات | نام گذاری و طبقه بندی  +  وجود ترتیب در طبقات  +  فواصل مساوی در طبقات  +  وجود صفر قراردادی | نام گذاری و طبقه بندی  +  وجود ترتیب در طبقات  +  فواصل مساوی در طبقات  +  وجود صفر مطلق |

مثال ۸: کدام یک از نظام های نمره گذاری زیر اطلاعات دقیق تری را در اختیار ما می گذارد؟

۱) ۱ تا ۱۰

۲) مردودی – قبولی

۳) خوب - متوسط – ضعیف

4) E,D,C,B,A

پاسخ: گزینه «۱» نمره گذاری در یک طیف عددی، از ۱ تا ۱۰ یا ۱ تا ۲۰، در مقیاس فاصله ای و با صفر قراردادی است که اطلاعات دقیق تری را به دست می دهد و فاصله ها در این مقیاس در بین واحدها برابر هستند؛ لذا مقیاس مناسبی برای نمره گذاری است.

مثال ۹: تبدیل سطوح اندازه گیری در کدام مورد زیر امکان پذیر است؟

1) اسمی به ترتیبی

۲) ترتیبی به فاصله ای

۳) ترتیبی به نسبی

۴) نسبی به فاصله ای

پاسخ: گزینه «۴» به ترتیب مقیاس ها عبارت اند از:

۱) مقیاس اسمی ۲) مقیاس ترتیبی 3) مقیاس فاصله ای ۴) مقیاس نسبی

با توجه به این که می دانیم تبدیل مقیاس ها از آخر به اول امکان دارد و برعکس آن غیر ممکن است، بنابراین گزینه ۴ صحیح است. این تبدیل از ۴ به ۳، ۳ به ۲ و ۲ به ۱ است و برعکس آن نادرست است. گزینه ۴ تنها گزینه صحیح در میان گزینه ها است که این ترتیب را رعایت کرده است و سایر گزینه ها جهت عکس دارند و غلط هستند.

مثال ۱۰: در کدام سطح اندازه گیری تمامی عملیات ریاضی - آماری را در مورد داده ها می توان انجام دادم .

۱) ترتیبی (رتبه ای)

۲) فاصله ای

۳) نسبی

۴) طبقه بندی (اسمی)

پاسخ: گزینه «3» مقیاس نسبی با نسبتی، مقیاسی است که در آن: ۱- رتبه اشخاص با توجه به یک صفت معلوم است ۲- فاصله بین اشخاص معلوم است ۳- به علاوه فاصله از صفر منطقی نیز معلوم است. مقیاس نسبتی، مقیاس اعداد از صفر تا بی نهایت است و اندازه گیری طول، وزن، حجم و سایر صفات فیزیکی با این مقیاس میسر است.

مقیاس های گسسته و پیوسته

مقیاس گسسته: مقیاس گسسته، مقیاسی است که در آن هر یک از متغیرها فقط می توانند یک ارزش معین را به خود اختصاص دهند و ویژگی مهم آن، تساوی واحدهای شمارش در طول مقیاس است: وضعیت تأهل (مجرد یا متاهل)، جنس (مؤنث یا مذکر) و تحصیلات مثال هایی از مقیاس گسسته هستند.

مقیاس پیوسته: در مقیاس پیوسته هریک از متغیرها می توانند هر تعداد ارزش را به خود اختصاص دهند و ویژگی مهم آن تساوی واحدهای اندازه گیری است؛ بنابراین اگر ما طول یک میز را به سانتیمتر اندازه بگیریم، در سراسر طول، مقیاس یک سانتیمتر همواره یک معنا دارد.

اندازه گیری مقیاس های گسسته همیشه قطعی، اما اندازه گیری مقیاس های پیوسته تقریبی است.

حدود واقعی اعداد

حدود واقعی اعداد عبارت است از همان عدد نصف واحد اندازه گیری؛ مثلا حد واقعی عدد ۶۱،5/60 و 5/61 است.

گرد کردن

در گرد کردن، معمولا در نهایت ۲ رقم بعد از اعشار را در نظر میگیریم؛ اگر رقم سوم بعد از اعشار بزرگتر از ۵ بود، یکی به رقم قبلی اضافه می کنیم. اگر رقم سوم بعد از اعشار کوچکتر از ۵ بود، هیچ تغییری ایجاد نمی کنیم و اگر رقم سوم بعد از اعشار ۵ بود، دو حالت پیش می آید:

۱) اگر رقم سوم بعد از اعشار ۵ باشد و بعد از آن اعداد دیگری هم وجود داشته باشد، یکی به رقم دوم اضافه می کنیم.

۲) اگر رقم سوم بعد از اعشار به بود و بعد از آن عدد دیگری وجود نداشت با فقط صفر وجود داشت، دو حالت پیش می آید: اگر عدد دوم زوج بود، بدون تغییر باقی می ماند اما اگر فرد بود یکی به آن اضافه می کنیم.

رقم سوم بعد از اعشار بزرگتر از ۵ است؛ بنابراین یک واحد به عدد دوم اضافه می کنیم.

55/ 6 546/ 6

رقم سوم بعد از اعشار کوچکتر از ۵ است؛ پس رقم دوم بدون تغییر باقی می ماند.

25/ 17 2539/ 17

رقم سوم بعد از اعشاره است و بعد از آن عدد وجود دارد؛ بنابراین یک واحد به رقم دوم اضافه می کنیم.

55/ 6 545001/ 6

رقم سوم بعد از اعتبار فی است و بعد از آن فقط صفر وجود دارد و رقم دوم هم زوج است، پس رقم دوم بدون تغییر باقی می ماند.

۵۴/ 0 ۵۴۵۰۰ /۰

رقم سوم بعد از اعشار له است و بعد از آن فقط صفر وجود دارد و رقم دوم فرد است؛ پس یک واحد به رقم دوم اضافه می کنیم.

02/ 0 01500/ 0

مثال ۱۱: رقم 828001/ 13 بعد از گرد کردن چند می شود؟

1) 80/ 13

2) 82/ 13

3) 83/ 13

4) 828/ 13

پاسخ: گزینه «3» چون رقم سوم بعد از اعشار از ۵ بزرگتر است، یک واحد به عدد دوم اضافه می شود.

مثال ۱۲: رقم 7531/20 بعد از گرد کردن چند می شود؟

1) 75/ 20

2) 753/ 20

3) 8/ 20

4) 76/ 20

پاسخ: گزینه «۱» چون رقم سوم بعد از اعتبار کوچک تر از ۵ است، رقم دوم بدون تغییر باقی می ماند.

مثال ۱۳: رقم 63500/ 8 بعد از گرد کردن چند می شود؟

1) 70/ 8

2) 64/ 8

3) 63/ 8

4) 635/ 8

پاسخ: گزینه «۲» چون رقم سوم بعد از اعتبار ۵ است و بعد از آن فقط صفر وجود دارد و رقم دوم فرد است، پس یک واحد به رقم دوم اضافه می کنیم.

تست های طبقه بندی شده فصل دوم

1- مقیاسی که سلسله مراتب اشیاء یا افراد را نشان داده و واحدهای اندازه گیری آن با هم برابر باشد، کدام است؟

(سراسری ۸۱)

۱) اسمی

2) ترتیبی

3) فاصله ای

4) نسبتی

2- اگر شماره کارت ملی ایرانیان از سن آنان اطلاع دهد، این شماره در کدام مقیاس اندازه گیری بیان خواهد شد؟

(سراسری ۸۲)

1) اسمی

2) ترتیبی

۳) فاصله ای

4) نسبی

۳- مقیاس اندازه گیری مورد استفاده در آزمون های چند جوابی از کدام نوع است؟

(سراسری ۸۵)

۱) اسمی (Nominal)

۲) ترتیبی (Ordinal)

۳) فاصله ای (Interval)

۴) نسبی (Ratio)

۴- در اندازه گیری های مربوط به آموزش و پرورش، نمرات بیشتر در کدام مقیاس به دست می آید؟

(آزاد ۸۵)

۱) ترتیبی

۲) اسمی

۳) فاصله ای

۴) نسبی

5\_ اگر برای گروه بندی دانش آموزان یک ناحیه، پسران را با یک و دختران را با صفر نشان دهیم، به این سطح از اندازه گیری چه می گویند؟

(آزاد ۸۶)

۱) ترتیبی

۲) فاصله ای

3) اسمی

۴) نسبتی

6\_گروه کثیری از دانش آموزان را در نظر بگیرید که فقط می دانید آنه ایی که بهره هوشی بالاتر از ۱۲۰ دارند، مقدار (۱) و آنهایی که بهره هوشی کمتر از ۱۲۰ دارند، مقدار (0) برایشان منظور شده است. مقیاس اندازه گیری در اینجا کدام مورد است؟

(آزاد ۸۹)

1) دوارزشی - اسمی

۲) فاصله ای

۳) دوارزشی - اسمی با زیربنایی توزیع طبیعی

۴) رتبه ای

7\_ در کدام مقیاس اندازه گیری، واحدهای اندازه گیری با هم برابر نیستند؟

(سراسری ۹۰)

۱) اسمی

۲) ترتیبی

3) نسبی

۴) فاصله ای

8\_ در کدام یک از مقیاس ها مبدأ اندازه گیری، اعتباری است؟

(سراسری ۹۳)

1) نسبی

2) ترتیبی

۳) اسمی

۴) فاصله ای

پاسخنامه تست های طبقه بندی شده فصل دوم

1- گزینه «3» از مقیاس ها همواره در آزمون ها، سؤال آورده می شود و برای آنکه بتوان در سؤال، پاسخ درست را تشخیص داده باید تعریف کاملی از مقیاس ها را در ذهن داشته باشیم، مورد سؤال، اشاره به سلسله مراتب افراد و اشیاء دارد که به مقیاس فاصله ای برمی گردد. در مقیاس فاصله ای، رتبه افراد یا اشیاء باتوجه به یک صفت مشخص است و معلوم است که اشیا و یا افراد از لحاظ صفت مورد اندازه گیری چه مقدار از یکدیگر فاصله دارند. گزینه ۲ به جواب اصلی بسیار نزدیک است، زیرا مقیاس ترتیبی به مجموعه ای از رتبه ها اشاره دارد. در این مقیاس، دسته ای از افراد و اشیاء با توجه به یک صفت مرتب می شوند اما میزان دارا بودن آن صفت و فاصله آن ها از یکدیگر مشخص نیست.

۲- گزینه «۴» در مقیاس نسبی، رتبه افراد با توجه به یک صفت مشخص می شود و فاصله بین اشخاص نیز تعیین می گردد؛ بنابراین با تعریف شماره کارت ملی سر ایرانیان، باتوجه به این که کارت دارای اطلاعات فردی هر شخص می باشد و این کارت به هر فرد با یک شماره مخصوص اختصاص دارد، مترادف با تعریف مقیاس نسبی است که یک شماره (صفت) به یک فرد خاص اختصاص می یابد. گزینه ۲ می تواند داوطلب را به اشتباه بیندازد؛ زیرا مقیاس ترتیبی مقیاسی است که افراد را با توجه به صفتی خاص مرتب می کند اما این نکته را که هر فرد چه مقدار از آن صفت را دارد، مدنظر قرار نمی دهد.

۳- گزینه «3» مقیاس فاصله ای نشان میدهد که افراد (اشیا) از لحاظ صفت مورد اندازه گیری چه مقدار با هم فاصله دارند. مقیاس اندازه گیری آزمون های چندگزینه ای، فاصله ای است؛ زیرا این آزمون ها دارای صفر قراردادی هستند.

۴- گزینه «3» در اندازه گیری های مربوط به آموزش و پرورش، نمرات در مقیاس فاصله ای قرار دارند و معمولا نمره گذاری در طیف عددی ۱ تا ۲۰ قرار می گیرد که فاصله ای است.

5\_ گزینه «3» مقیاس اسمی پایین ترین سطح یا ساده ترین نوع اندازه گیری را معرفی می کند که به دو صورت (برای نام گذاری و طبقه بندی) مورد استفاده قرار می گیرد. مقیاس اسمی در جایی که برای اسم گذاری مورد استفاه قرار می گیرد از ارقام و اعداد برای نام گذاری و تشخیص اشیاء و افراد استفاده می کند، بدون اینکه این اعداد هیچ گونه مفهوم ریاضی داشته باشند؛ همچنین مقیاس اسمی زمانی که به صورت طبقه بندی مورد استفاه قرار می گیرد، از اعداد و ارقام برای طبقه بندی افراد و اشیاء مختلف استفاده می کند.

6\_ گزینه «۱» مقیاس اسمی، مقیاسی است که از اعداد برای اسم گذاری با نام گذاری و تشخیص اشیاء و افراد استفاده می کند. از آنجا که در این سؤال افراد را به دو گروه تقسیم کرده و به هر گروه یک عدد که معرف آن گروه است داده شده است، می توان گفت که از مقیاس اسمی دوارزشی استفاده شده است.

۷- گزینه «۲» مقیاس ترتیبی یا رتبه ای مجموعه ای از رتبه هاست که در آن دسته ای از افراد یا اشیاء باتوجه به یک صفت، از بزرگ به کوچک مرتب می شوند و معلوم نیست که هریک از این افراد یا اشیاء به مفهوم مطلق، چقدر از آن صفت را دارد و نیز مشخص نیست که این اشیاء یا افراد از لحاظ صفت در مورد اندازه گیری، چه مقدار از یکدیگر فاصله دارند. گزینه ای که ممکن است به اشتباه انتخاب شود، گزینه ی ۴ است؛ چون در مقیاس فاصله ای نیز رتبه اشیاء یا افراد از لحاظ صفت مورد اندازه گیری مشخص می شود اما این دو مقیاس از این لحاظ که در مقیاس فاصله ای، فاصله بین واحدها معلوم و برابر و در مقیاس ترتیبی فاصله بین واحدها معلوم نیست، از هم متمایز می شوند.

8\_ گزینه «۴» در مقیاس فاصله ای، مبدا اندازه گیری، قراردادی با اعتباری است. اعداد در مقیاس فاصله ای وسعت تفاوت میان صفات افراد یا اشیای موردنظر را نشان می دهند. این مقیاس علاوه بر اینکه همانند مقیاس رتبه ای دارای نظم و ترتیب منطقی است، میزان تقریبی تفاوت میان مقادیر مختلف صفت موردنظر را نیز نشان میدهد.

آزمون فصل دوم

۱- در کدام یک از مقیاس های اندازه گیری، تمام اعمال آماری مجاز ولى أعمال ضرب و تقسیم مجاز نیست؟

۱) اسمی

۲) ترتیبی

۳) فاصله ای

4) نسبتی

۲\_کدام یک از مقیاس های اندازه گیری صرفا به طبقه بندی و اسم گذاری افراد و اشیاء می پردازد؟

1) اسمی

۲) ترتیبی

۳) نسبتی

۴) فاصله ای

۳- اصل نمادین AxB = C مربوط به کدام مقیاس است؟

1) فاصله ای

۲) ترتیبی

3) اسمی

4) نسبتی

۴- رقم ۰۳۵۰۰/0 پس از گردکردن چند می شود؟

1) 036/ 0

2) 03/ 0

3) 04/ 0

4) 035/ 0

5\_ در فرضیه وزن با میزان بهره هوشی کودکان پایه اول دبستان رابطه دارد، مقیاس اندازه گیری ۲ متغیر وزن و بهره هوشی به ترتیب کدام است؟

۱) فاصله ای – فاصله ای

۲) نسبتی – نسبتی

۳) نسبتی – فاصله ای

۴) فاصله ای – نسبتی

۶- رقم 49400/ 9 پس از گردکردن چند می شود؟

1) 49/ 9

2) 4/ 9

3) 50/ 9

4) 494/ 9

7- تبدیل سطوح اندازه گیری در کدام مورد زیر امکان پذیر است؟

1) اسمی به ترتیبی

۲) ترتیبی به نسبی

۳) نسیبی به فاصله ای

۴) ترتیبی به فاصله ای

۸\_ وقتی افراد را از نظر طبقه اجتماعی به فقیر، متوسط و غنی تقسیم می کنیم، از کدام مقیاس اندازه گیری استفاده کرده ایم؟

۱) اسمی

۲) ترتیبی

3) فاصله ای

4) نسبی

۹- اگر گفته شود عملکرد علی در آزمون فیزیک بهتر از ریاضی بوده، از کدام مقیاس استفاده شده است؟

۱) اسمی

۲) ترتیبی

3) فاصله ای

4) نسبی

۱۰\_ مقیاس اندازه گیری در آزمون های صحیح - غلط و چند جوابی کدام است؟

۱) اسمی

۲) ترتیبی

3) فاصله ای

۴) نسبی

فصل سوم

« دسته بندی آزمون های مورد استفاده در روانشناسی و آموزش و پرورش »

دسته بندی انواع آزمون ها

آزمون ها برای دستیابی به هدف های گوناگون طرح ریزی شده اند. آن ها را می توان برای گزینش، طبقه بندی، ارزشیابی و غیره به کار برد. هر آزمون سودمند یک ویژگی یا رفتار را با دقت اندازه گیری می کند. در این فصل به دسته بندی انواع آزمون ها پرداخته می شود:

1- دسته بندی آزمون ها با توجه به ویژگی های مورد اندازه گیری

2- دسته بندی آزمون ها با توجه به چگونگی تهیه آن ها

دسته بندی آزمون ها با توجه به ویژگی های مورد اندازه گیری

الف) آزمون های سنجش توانایی

ب) آزمون های عاطفی

الف) آزمون های سنجش توانایی: آزمون های توانایی، آموخته ها، مهارت ها و استعدادهای افراد را می سنجند. آزمون های توانایی ممکن است توانایی های شناختی افراد با توانایی های روانی - حرکتی افراد را بسنجند که هر کدام کاربرد خاصی دارند. کاربرد آزمون های توانایی شناختی در مسائل تحصیلی و فعالیت های فکری و ذهنی است و کاربرد آزمون های توانایی روانی - حرکتی در زمینه های صنعتی، نظامی و تربیت بدنی است.

آنچه در آزمون های توانایی شناختی و روانی- حرکتی مهم است، این است که از فرد خواسته می شود تا حداکثر تلاش خود را بکند که بهترین نمره را کسب کند. و آزمون های سنجش توانایی خود به دو دسته آزمون های سنجش استعداد و آزمون های سنجش پیشرفت تحصیلی تقسیم می شوند.

آزمون های سنجش استعداد، توانایی و آمادگی فرد را برای انجام کاری در آینده با ظرفیت های افراد را برای یادگیری های مختلف می سنجند و به آینده توجه دارند. آزمون های سنجش پیشرفت تحصیلی، دانش و مهارتی را که فرد تا لحظه ی آزمون کسب کرده است، اندازه گیری می کنند. آن ها به گذشته توجه دارند و می خواهند آموخته های قبلی را مشخص کنند.

آزمون های سنجش استعداد؛ مانند: آزمون های هوش

آزمون های پیشرفت تحصیلی؛ مانند: امتحانات نهایی

آزمون های پیشرفت تحصیلی با محتوایی که ارزشیابی می کنند، نامگذاری می شوند؛ مثل آزمون پیشرفت ریاضی

آزمون های استعداد به تعدادی آزمون با مقاصد مختلف گفته می شود؛

مثلا آزمون هوش یا استعداد کلی که توانایی ذهنی فرد را در مورد همه فعالیت های ذهنی و فکری میسنجد و آزمون های استعداد ویژه که توانایی فرد را در زمینه های اختصاصی می سنجند. هدف آزمون های هوش کلی، پیش بینی موفقیت های همه جانبه فرد است؛ اما هدف آزمون های استعداد ویژه، پیش بینی موفقیت ها در یک زمینه خاص است. آزمون های استعداد ویژه برحسب محتوایی که می سنجند، نام گذاری می شوند؛ مثل آزمون های استعداد ریاضی یا موسیقی.

ب) آزمون های عاطفی: آزمون های شناختی و روانی - حرکتی توانایی های فکری و عملی مشخص را می سنجند؛ در صورتی که آزمون های عاطفی، ویژگی های مر عاطفی و شخصیتی فرد را می سنجند. آزمون های توانایی، حداکثر عملکرد فرد را اندازه گیری می کنند تا نشان دهند فرد در بهترین شرایط قادر به انجام چه - کاری است؛ اما آزمون های عاطفی، وضع موجود فرد را در موقعیت طبیعی زندگی او می سنجند. فرق دیگر آزمون های عاطفی و توانایی در این است که سد آزمون های توانایی، موفقیت های تحصیلی و شغلی را پیش بینی می کنند؛ اما آزمون های عاطفی به مشاوره و حل مشکلات روانی و تربیتی کمک می کنند.

آزمون های عاطفی: به آزمون های شخصیت، نگرش و علاقه تقسیم می شوند

آزمون های شخصیت: آزمون های شخصیت به اندازه گیری متغیرهای انگیزش، مزاج، سازگاری و منش اختصاص دارند.

ابزارهای اندازه گیری شخصیت به سه دسته تقدیم می شوند:

۱- پرسشنامه های گزارش شخصی: این ابزار، وسایلی هستند که شخص از طریق آن نظرش را درباره ی خود بیان می کند.

۲- پرسشنامه های جامعه سنجی: این ابزار، وسایلی هستند که از طریق آن ها افراد دیگر نظرشان را درباره شخص بیان می کنند.

٣- فنون مشاهده ای: یعنی رفتار فرد در یک موقعیت به خصوص مورد مشاهده قرار می گیرد.

یک نوع دسته بندی دیگر از آزمون های شخصیت با توجه به نحوه ی ارائه محرکه است. گاهی محرک ها به صورت سازمان یافته ارائه می شود و افراد مختلف برداشت یکسان از سؤالات دارند ولی پارهای اوقات، محرک ها به صورت سازمان نیافته ارائه می شود و افراد بر داشت یکسانی از سوالات ندارند بلکه تعابیر و تفاسیر متفاوتی دارند؛ مانند آزمون های فرافکن که شامل تصاویر مبهم و لکه های جوهر هستند که ازمون شونده باید آن ها را تعبیر و تفسیر کند.

آزمون های علاقه و نگرش: علاقه یعنی احساس فرد نسبت به یک فعالیت و نگرش یعنی احساس فرد نسبت به یک شیء، یک گروه یا نهاد اجتماعی.

مقایسه علاقه و نگرش: علاقه رجحان یک فعالیت بر سایر فعالیت ها است ولی نگرش، رجحان یک شیء، یک نهاد یا یک گروه است.

برای سنجش علاقه و نگرش، روش های مختلفی وجود دارد که معروف ترین آنها پرسشنامه علاقه و نگرش است.

دسته بندی آزمون ها با توجه به چگونگی تهیه

آزمون ها با توجه به چگونگی تهیه شدن، به دو دسته تقسیم می شوند:

الف) آزمون های میزان شده

ب) آزمون های معلم ساخته

آزمون های میزان شده یا معیارشده: این آزمون ها برای سنجش بخش بزرگی از آزمون شوندگان تهیه می شوند. این آزمون ها توسط مؤسسات یا بنگاه های آید. بانی با مراکز علمی - آموزشی برای اندازه گیری پیشرفت تحصیلی و هوش و استعداد گروه های بزرگ دانش آموزان با قصد ارزشیابی دوره های تحصیلی تهیه میشود. در این نوع آزمون ها ملاک نسبی اندازه گیری مورد استفاده قرار می گیرد.

آزمون های معلم ساخته: این آزمون های آزمون هایی هستند که معلمان برای ارزشیابی پیشرفت تحصیلی شاگردان خود می سازند. و در طول دوره پا پایان آن و اجرا می کنند. در این آزمون ها میزان توفیق یادگیرندگان در رسیدن به اهداف تعیین می شود. این آزمون ها ملاک مطلق را برای اندازه گیری به کار می برند.

نقاط ضعف آزمون های معلم ساخته:

۱- آزمون های معلم ساخته دامنه هدفهای لازم را پوشش نمی دهند؛

۲- محتوای این گونه امتحانات، جزئی و کم اهمیت است.

۳- مواد امتحانی با ارزشیابی فرایند ذهنی که امتحان باید بسنجد، کمتر مطابقت دارد.

۴- سؤالات آزمون های معلم ساخته، اغلب مبهم و غامض است.

5- مشخصات تست برای ایفای مقصود معلم، مناسب نیست.

مثال 1: کدام یک از آزمون های زیر به اندازه گیری متغیرهای انگیزش، مزاج، سازگاری و منش اختصاص دارد؟

۱) آزمون گزارش شخصی

۲) آزمون علاقه و نگرش

۳) آزمون های شخصیت

۴) آزمونهای عاطفی

پاسخ: گزینه «3» آزمون های شخصیت به اندازه گیری متغیرهای انگیزش، مزاج، سازگاری و منش اختصاص دارند.

مثال ۲: در کدام یک از انواع آزمون های زیر از ملاک نسبی اندازه گیری استفاده می شود؟

۱) آزمون های میزان شده

۲) ازمون های معلم ساخته

۳) ازمون های علاقه و نگرش

۴) آزمون های شخصیت

پاسخ: گزینه «۱» در آزمون های میزان شده، ملاک نسبی اندازه گیری مورد استفاده قرار می گیرد.

مثال 3: از مهمترین کاربردهای آزمون میزان شده (استاندارد) کدام است؟

۱) تعریف و تشریح هدف های تدریس و یادگیری

۲) مقایسه فراگیران مدارس مناطق گوناگون با یکدیگر

۳) ایجاد رغبت یادگیری و کسب عادات صحیح مطالعه

۴) اصلاح روش تدریس معلم و بهبود کمی و کیفی یادگیری فراگیران

پاسخ: گزینه «۲» یکی از کاربردهای آزمون های میزان شده، مقایسه فراگیران مدارس مناطق گوناگون با یکدیگر است. آزمون های میزان شده برای استفاده در مناطق مختلف و مقایسه ی دانش آموزان مناطق گوناگون تهیه می شوند؛ بنابراین هدف های مختلف آموزشی مناطق مختلف را باید در این آزمون ها منظور کرد؛ با این حال، این نوع آزمون ها ویژگی های خاص کلاس ها و مدارس به خصوص را نادیده می گیرند و مانند کنکور بر وجوه مشترک محتوای درسی مناطق مختلف تکیه می کنند؛ همین امر باعث اختصاص وقت و هزینه ی بیشتری می شود. هدف اصلی آزمون های استاندارد، مقایسه است: مقابله ایک فرد یا یک گروه با فرد یا گروه دیگر. گزینه ی (۱) با توجه به این که به گونه ای با متن سؤال در ارتباط است، می تواند داوطلب را در انتخاب گزینه درست، دچار تردید کند. توجه به این نکته مهم است که آزمونهای استاندارد از یک معیار و الگوی خاص پیروی می کنند؛ اما تعریف و تشریح اهداف تدریس و یادگیری، جزء کاربردهای آزمون های استاندارد نمی باشد.

مقایسه آزمون های میزان شده و معلم ساخته

1- آزمون های میزان شده از نوع هنجاری هستند و آزمون های معلم ساخته از نوع ملاکی اند؟

2- آزمون های میزان شده توسط مؤسسات بزرگ تهیه می شوند ولی آزمون های معلم ساخته از نظر طراحی، اجرا و نمره گذاری تحت نظر معلم هستند. .

۳- آزمون های میزان شده در مقایسه با آزمون های معلم ساخته به زمان و تخصص بیشتری نیاز دارند. ,

4- به دلیل آن که از آزمون های میزان شده در مناطق مختلف استفاده می شود، باید ویژگی های مشترک اجرا را در مورد آن ها در نظر گرفت.

۵- اصطلاح آزمون های میزان شده به داده های هنجاری اشاره می کند و تهیه کننده آزمون باید برای معنادار کردن نمرات آزمون، هنجارهای لازم را تهیه کند؛ اما آزمون های معلم ساخته به هنجاریابی نیازی ندارند؛ چون هدف آزمون های معلم ساخته مقایسه کردن نیست و نیازی به هنجاریابی وجود ندارد؛ بنابراین هدف اصلی آزمون های میزان شده، مقایسه یک فرد با یک گروه با فرد یا گروه دیگر است.

مثال ۴: کدام یک از موارد زیر جزء ویژگی های آزمون های میزان شده نیست؟

1) این آزمون ها توسط مؤسسات بزرگ آزمون سازی تهیه می شوند.

۲) در این آزمون ها از ملاک مطلق اندازه گیری استفاده می شود.

۳) این آزمون ها به دقت و تخصص کمتری نیاز دارند.

۴) در این آزمون ها باید ویژگی های مشترک اجرا را در نظر گرفت.

پاسخ: گزینه «۲» در این آزمون ها از ملاک نسبی اندازه گیری استفاده میشود.

مثال 5: برای بهبود جریان آموزش و بالا بردن سطح یادگیری از کدام آزمون استفاده می شود؟

1) معلم ساخته

۲) علاقه و نگرش

۳) میزان شده

4) شناختی

پاسخ: گزینه «۱» آزمون های معلم ساخته باعث بهبود جریان آموزش و بالا بردن سطح یادگیری می شوند.

چه وقت از آزمون های میزان شده استفاده می کنیم و چه وقت از آزمون های معلم ساخته؟

1\_ تصمیمات مربوط به کم و کیف جریان آموزش که هدف آن ها بهبود جریان آموزش و بالابردن سطح یادگیری است: معلم ساخته

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

۳- تصمیمات تشخیصی در مورد مشکلات یادگیرندگان: میزان شده و معلم ساخته

۴۔ تصمیمات مربوط به جایابی و پایه گزینی که نیاز به مقایسه دارد: میزان شده

۵- تصمیمات مربوط به راهنمایی و مشاوره؛ که نیاز به مقایسه هنجاری دارد: میزان شده

6\_تصمیمات مربوط به انتخاب که لازم است افراد باهم مقایسه شوند: میزان شده

۷- تصمیمات مربوط به برنامه درسی که برنامه درسی باید با سایر برنامه ها مقایسه شود: میزان شده

8\_ تصمیمات مربوط به توفیق مدارس: میزان شده

مثال ۶: از مهمترین کاربردهای آزمون های میزان شده چیست؟

۱) تشریح هدف های تدریس و یادگیری

۲) مقایسه شاگردان مدارس مختلف

۳) ایجاد، رغبت یادگیری و کسب عادات صحیح مطالعه

۴) اصلاح روش تدریس و بهبود کیفی و کمی یادگیری

پاسخ: گزینه «۲» «مقایسه شاگردان مدارس مختلف»، جز مهم ترین کاربردهای آزمون های میزان شده است.

شناختی

سنجش پیشرفت تحصیلی

توانایی روانی، حرکتی

دسته بندی آزمون ها براساس ویژگی های مورد اندازه گیری سنجش استعداد استعداد کلی

آزمون های عاطفی شخصیت استعداد ویژه

نگرش

علاقه

الف) آزمون های میزان شده (معیار شده)

دسته بندی آزمون ها باتوجه به چگونگی تهیه آن ها

ب) آزمون های معلم ساخته

مثال۷: مدیر دبیرستانی در مذاکره با پدر یکی از دانش آموزان می گوید:

وضع حسن در علوم خوب است ولی پیشرفت او در دروس اجتماعی و ادبیات در حد مطلوب نیست.» این بیان مدیر دبیرستان بر ............... دلالت دارد.

1) اندازه گیری خصوصیات

۲) ارزشیابی عملکرد

۳) نگرش وی نسبت به دانش آموز

۴) سنجش معلومات

پاسخ: گزینه «۲» این کار مدیر دبیرستان در واقع یک نوع ارائه کارنامه به صورت شفاهی است و در هر صورت، کارنامه و ارزشیابی عملکرد می تواند به صورت کتبی یا شفاهی باشد. هیچ گزینه ی دیگری نمی تواند ذهن داوطلب را به انتخاب اشتباه وادار کند و اگر به متن سؤال دقت شود، در واقع مدیر بین دروس، سطح بندی و دسته بندی از خوب تا متوسط انجام داده که این امر فقط در ارزشیابی اتفاق می افتد.

تست های طبقه بندی شده فصل سوم

۱- وجه تمایز راهبردهای فرافکن (Projective) با دیگر راهبردهای سنجش شخصیت در این است که در این راهبردها:

(سراسری ۸۵)

1) از تئوری های شخصیت استفاده می شود.

۲) صفات شخصیتی مورد تأکید واقع می شود.

۳) از محرک های سازمان نیافته استفاده میشود.

۴) درون گرایی - برون گرایی مورد تأکید واقع می شود.

۲- برای مقایسه ی کدام یک از موارد زیر بهتر است از آزمون های استاندارد شده استفاده شود؟

(سراسری ۸۵)

۱) دو نظام یا دو شیوه ی آموزشی متفاوت

۲) دانش آموزان یک کلاس با یکدیگر

۳) دانش آموزان در کلاس مختلف

۴) نتایج یک فرد در دو زمان مختلف

۳- کدام آزمون، میزان موفقیت افراد را پس از طی یک دوره آموزشی تعیین می کند؟

(سراسری ۸۶)

1) استعداد

۲) پیشرفت تحصیلی

۳) شخصیت

۴) فرافکن

۴- اگر هدف تست، تعیین این مطلب باشد که فرد در آینده در چه رشته ای موفق خواهد شد، بکارگیری چه آزمونی مناسب تر است؟

(سراسری ۸۷)

1) استعداد

۲) پیشرفت تحصیلی

۳) رغبت

۴) شخصیت

۵- اگر هدف آزمون تعیین این باشد که فرد در گذشته چه چیزهایی را یاد گرفته است، بکارگیری کدام تست مناسب تر است؟

(سراسری ۸۷)

1) استعداد

۲) رغبت

۳) شخصیت

۴) معلومات

6\_ مهم ترین ایراد واردشده به استفاده از آزمون های روانی در تصمیم گیری عبارت است از:

(سراسری ۸۸)

1) ایجاد تبعیض میان افراد جامعه

۲) نقض حریم خصوصی افراد

۳) فقدان اعتبار کافی

۴) فقدان روایی کافی

۷- ملاک یا معیار مطلق، پایه و اساس کدام یک از امتحانات زیر است؟

(آزاد ۸۸)

1) امتحانات استخدامی

۲) امتحانات کنکور دانشگاه ها

۳) امتحانات سطح دبیرستان

۴) امتحانات جهت تعیین شاگردان ممتاز

8\_استفاده از روش های سنجش و اندازه گیری در آموزش و پرورش، اطلاعات مفیدی درباره دانش آموزان به معلم می دهد که براساس آن ها:

(آزاد ۸۹)

1) معلم می تواند دانش آموزان را به دو دسته ی قوی و ضعیف تقسیم بندی کند.

۲) معلم، دانش آموزان را بهتر می شناسد و امکانات لازم را برای رشد و تکامل آنان فراهم می آورد.

۳) معلم می تواند روش تنبیه و تشویق مناسبی را در مورد دانش آموزان اعمال کند.

۴) معلم، نقاط ضعف دانش آموزان را تشخیص می دهد و به آنان گوشزد می کند.

۹- آزمون هایی که توانایی های ذهنی و حدود احتمالی موفقیت شخص را در آینده و در موقعیت های مختلف زندگی اندازه می گیرد، چه نام دارد؟

(آزاد ۸۹)

۱) آزمون هوش و استعداد

۲) آزمون پیش بین

3)آزمون پیشرفت تحصیلی

۴) آزمون رغبت

پاسخنامه تست های طبقه بندی شده فصل سوم

١- گزینه ۳» در این سؤال، به خصوصیت راهبردهای فرافکن اشاره شده است. باید به این نکته توجه کرد که استفاده از تصاویر فرافکن به شخص آزمودنی کمک می کند تا به آنچه درون او وجود دارد، اشاره کند؛ بنابراین این راهبردها سازمان یافته نیستند، بلکه به گونه ای طراحی شده اند که هر فرد به فراخور حال خود به آن اشاره کند.

۲- گزینه «1» با توجه به اینکه شیوه های آموزشی با نظام های آموزشی به صورت اساسی و بنیادین با هم تفاوت دارند، بهتر است برای مقایسه آن ها باتوجه به جدول هدف - محتوا از آزمون های استاندارد شده استفاده کرد.

۳- گزینه «۲» ارزشیابی پیشرفت تحصیلی مفهوم بسیار نزدیکی به سنجش یادگیری و اشاره به موفقیت ها و دستاوردهای یادگیری دانش آموزان و دانشجویان دارد. در این نوع ارزشیابی به ویژگی کیفیت توجه می شود. کیفیت در ارزشیابی پیشرفت تحصیلی، به دانش، مهارت ها و توانایی هایی که از دانش آموزان یا دانشجویان پس از آموزش انتظار می رود، اشاره می کند.

۴- گزینه «۱» آزمون استبداد نشان می دهد که آزمون شونده در چه زمینه ای به طور ذاتی و اکتسابی مستعد است و می تواند در آن زمینه فعالیت کند. . آزمون شونده با مهارت آزمون کننده در سؤالات روبه رو می شود و این مهارت، سبب روشن شدن نوع استعداد فرد آزمون شونده می شود؛ به هر صورت یک آزمون ساز ماهر، از سؤالات ابتکاری در جهت کشف استعدادهای افراد، بهره می برد.

5\_گزینه «۴» زمانی که یک آزمون ساز ماهر تصمیم می گیرد میزان آموخته های محفوظ در ذهن آزمون شونده را بسنجد، می تواند با مهارت و ابتکار از تست معلومات بهره بگیرد.

6\_ گزینه «1» برای انجام آزمون های روانی، آزمایش کننده پس از انجام آزمون تصمیم گیری می کند. این تصمیم گیری براساس قضاوت ذهنی فرد و آزمون کننده صورت می گیرد که منجر به ایجاد تبعیض بین افراد جامعه می شود. باید در نظر داشت که ممکن اسبت در دفعه اول، گزینه دوم نظر آزمون شونده را جلب کند؛ اما چون پس از انجام آزمون آزمون کننده تصمیم گیری می کند، بنابراین امکان تبعیض وجود دارد؛ پس گزینه (۱) صحیح می باشد.

7\_گزینه «3» ارزشیابی های آموزشی با توجه به ملاک مورد استفاده به دو دسته تقسیم می شوند: ارزشیابی وابسته به ملاک (یا ارزشیابی وابسته به ملاک و مطلق) و ارزشیابی وابسته به هنجار. در ارزشیابی وابسته به ملاک مطلق، عملکرد فرد با معمار یا ملاک از قبل تعیین شده سنجیده می شود، تا میزان دستیابی او به آن ملاک مورد قضاوت قرار گیرد؛ مانند آزمون های پیشرفت تحصیلی مقاطع مختلف. در ارزشیابی وابسته به هنجار، عملکرد فرد در مقایسه با سایر افراد سنجیده می شود؛ نه با یک ملاک از پیش تعیین شده؛ مانند امتحان کنکور دانشگاه ها.

۸- گزینه «۲» یکی از اهداف استفاده از روش های سنجش و اندازه گیری در آموزش و پرورش، فراهم کردن اطلاعات مناسب جهت هدایت یادگیرندگان به سمت هدف های آموزشی است؛ بنابراین از این روش ها می توان برای فراهم کردن اطلاعات مناسب در مورد دانش آموزان برای معلم استفاده کرد تا وسیله ای جهت شناسایی آنها و فراهم آوردن امکانات مناسب جهت رشد و تکامل آن ها باشد.

۹- گزینه «1» به طور کلی آزمون های توانایی را که آموخته ها، مهارت ها و استعدادهای افراد را می سنجد) می توان به دو دسته کلی آزمون های استعداد و آزمون های پیشرفت تحصیلی تقسیم کرد. آزمون های استعداد، ناظر بر آینده اند و برای پیش بینی موفقیت های آتی فرد به کار می روند؛ در حالی که آزمون های پیشرفت تحصیلی ناظر بر گذشته اند و برای تعیین آموخته های قبلی فرد، استفاده می شوند.

آزمون فصل سوم

۱ـ کدام عبارت صحیح است؟

۱) آزمون های میزان شده نسبت به آزمون های معلم ساخته به زمان و تخصص زیادی نیاز ندارند.

۲) آزمون های میزان شده از نوع ملاکی اند و آزمون های معلم ساخته از نوع هنجاری اند.

۳) آزمون های میزان شده و معلم ساخته هر دو توسط معلم تهیه می شوند.

۴) آزمون های میزان شده از نوع هنجاری اند و آزمون های معلم ساخته از نوع ملاکی اند.

۲- کدام آزمون، دانش و مهارتی را که فرد تا لحظه ی آزمون کسب کرده است، می سنجد؟

۱) استعداد

۲) پیشرفت تحصیلی

۳) شخصیت

4) فرافکن

۳- همه آزمون های زیر نمونه ای از آزمون های توانایی هستند به جز:

۱) استعداد ریاضی

۲) هوش

۳) فرافکن

۴) ورودی دانشگاه ها

4\_ در کدام یک از آزمون ها، محرک ها به صورت سازمان نیافته ارائه می شوند و افراد، تفاسیر و برداشت های مختلفی از متن دارند؟

1) هوش

۲) نگرش

۳) معلومات

۴) فرافکن

5\_کدام یک از بیانات زیر در رابطه با علاقه و نگرش صحیح است؟

۱) علاقه و نگرش دو جنبه از شخصیت هستند.

2) علاقه به یک نهاد و نگرش به یک فعالیت مربوط می شود.

۳) علاقه و نگرش، دو نام متفاوت برای یک موضوع واحد هستند.

4) گزینه 1 و 4

6\_ آزمون ها را براساس ویژگی های مورد اندازه گیری می توان به ........ و ........ تقسیم نمود.

۱) آزمون های میزان شده ـ آزمون های معیار شده

2) آزمون های توانایی \_ علاقه و نگرش

۳) آزمون های توانایی – آزمون های عاطفی

4) آزمون های عاطفی \_ آزمون های استعداد

۷- پرسشنامه های گزارش شخصی مربوط به کدام یک از انواع آزمون های زیر می باشند؟

1) آزمون های شخصیت

۲) آزمون های علاقه و نگرش

۳) آزمون های عاطفی

4) آزمون های معلم ساخته

8\_جهت تصمیم گیری تشخیصی در مورد مشکلات یادگیرندگان از کدام آزمون ها استفاده می شود؟

۱) آزمون های معلم ساخته و آزمون های نگرش

۲) آزمون های معلم ساخته و میزان شده

۳) آزمون های عاطفی و آزمون های توانایی

۴) آزمون های عاطفی – آزمون های استعداد

9\_ آزمون های میزان شده از نوع ........ هستند و آزمون های معلم ساخته نیز از نوع .......

1)ملاکی - هنجاری

۲) ملاکی – ملاکی

۳) هنجاری - ملاکی

۴) هنجاری – هنجاری

۱۰- زمانی که هدف ما تصمیم گیری در مورد ارتقای دانش آموزان و دانشجویان باشد، آزمون های ...... و هنگامی که هدف، تصمیم گیری در مـورد راهنمایی (مشاوره) باشد، از آزمون های استفاده می کنیم.

۱) معلم ساخته ـ معلم ساخته

۲) میزان شده - میزان شده

۳) میزان شده ـ معلم ساخته

۴) معلم ساخته .. میزان شده

فصل چهارم

«دسته بندی ارزشیابی های آموزشی»

دسته بندی ارزشیابی های آموزشی در قالب رویکردهای مختلف

اختلاف نظر بین صاحب نظران مختلف در مورد اینکه، ارزشیابی چیست و چگونه باید آن را انجام داد، به پیدایش رویکردهای ارزشیابی آموزشی مختلفی منجر شده است که به طور کلی می توان آن ها را در قالب شش رویکرد اصلی بیان کرد که هر کدام از آن ها نیز در بردارنده الگوهای ارزشیابی مختلفی هستند. این رویکردها عبارتند از:

رویکرد مبتنی بر هدف

در این رویکرد هدف های کلی و دقیق مورد تأکید هستند و ارزشیابی آموزشی به دنبال این است که تعیین کند، این هدف ها به چه میزان تحقق یافته اند. از جمله الگوهایی که در این رویکرد مورد بررسی قرار می گیرند، می توان به الگوی ارزشیابی تایلر اشاره کرد که از مراحل زیر تشکیل شده است:

1\_ تعیین هدف های کلی (غایت ها) و هدف های دقیق

۲- طبقه بندی هدف ها و غایت ها

۳- بیان هدف ها به صورت رفتاری، یعنی قابل اندازه گیری

۴- یافتن موقعیت هایی که در آن بتوان دستیابی به هدف ها را نشان داد.

5\_تولید یا تهیه روش ها و فنون اندازه گیری

6\_گردآوری داده های مربوط به عملکرد یادگیرندگان

مقایسه داده های مربوط به عملکرد هدف های رفتاری

تبا یکی از همکاران نزدیک تایلر برای برنامه ریزی درسی، یک الگوی مبتنی بر هدف پیشنهاد داد که مراحل آن عبارتند از: تعیین نیازها، بیان هدف ها، انتخاب محتوا، سازمان دادن به محتوا، انتخاب تجارب یادگیری، سازمان دادن به تجارب یادگیری و تعیین چه چیزی و چگونگی ارزشیابی.

رویکرد مبتنی بر مدیریت

هدف این رویکرد، تشخیص و تحقق بخشیدن نیازهای اطلاعاتی مدیران تصمیم گیرنده است. پدیدآورندگان الگوهای مبتنی بر این رویکرد، بر یک رویکرد سیستمی (نظام دار) تأکید دارند. مخاطب این الگوها همواره تصمیم گیرندگان آموزش و پرورش هستند. یکی از این الگوها، الگوی ارزشیابی سیپ است. این الگوی ارزشیابی دارای چهار چوبی است که مدیران و تصمیم گیرندگان را در چهار نوع تصمیم گیری زیر کمک می کند:

با ارزشیابی از بافت (Context) که هدف از این ارزشیابی فراهم کردن یک زمینه ی منطقی برای تعیین هدف های آموزشی است.

۲- ارزشیابی از درون داد (input) که در این مرحله، اطلاعات مورد نیاز درباره ی چگونگی استفاده از منابع، ماهیت توانایی های نظام آموزشی و استراتژی های بالقوه برای رسیدن به هدف ها تعیین می شود؛ همچنین در این مرحله به تصمیم گیرندگان جهت انتخاب و طراحی روش های دستیابی اهداف کمک میشود.

٣- ارزشیابی از فرایند (process) که در این مرحله سؤالاتی مانند آیا برنامه به خوبی در حال پیاده شدن است مطرح می شود که پاسخ این پرسش ها به کنترل و هدایت شیوه های اجرایی کمک می کند.

۴- ارزشیابی از فرآورده (Product). که در این مرحله معلوم می شود، نتایج به دست آمده چه هستند، بعد نتایج حاصل با هدف های برنامه مقایسه می گردند و رابطه بین انتظارات و نتایج واقعی مشخص می شود. یکی دیگر از الگوهای مربوط به این رویکرد، الگوی ارزشیابی یوسیالی یا سی اس ای است که از پنج نوع ارزشیابی تشکیل یافته است که عبارتند از: سنجش نظام، طراحی برنامه، اجرای برنامه، بهسازی برنامه و تصدیق برنامه.

رویکرد مبتنی بر مصرف کننده

در این رویکرد تأکید عمده، بر کسب اطلاعات مربوط به فرآورده های کلی آموزش است که توسط مصرف کنندگان آموزشی در انتخاب از بین برنامه های درسی مختلف، فرآورده های آموزشی و نظایر این ها مورد استفاده می شود. یکی از الگوهای مربوط به این رویکرد، الگوی ارزشیابی اسکریون است که ، عمدتأ یک ارزشیابی تراکمی است. ملاک هایی که اسکریون برای ارزشیابی از فرآورده ها به کار می گیرد، عبارتند از:

1- شواهد حاکی از تحقق یافتن هدف های مهم آموزشی

۲- شواهد حاکی از تحقق یافتن هدف های مهم غیر آموزشی

۳- نتایج حاصل از پیگیری .

4\_ آمار جانبی پیش بینی نشده

5\_ دامنه استفاده

6\_ ملاحظات اخلاقی

۷- هزینه ها

یکی دیگر از الگوهای مبتنی بر این رویکرد، الگوی ارزشیابی هدف - آزاد است که توسط اسکروین در مقابل ارزشیابی مبتنی بر هدف پیشنهاد شده است. در این الگو تمامی نتایج یک نوآوری آموزشی بدون هیچ گونه جهت گیری قبلی یا پیش داوری قبلی سنجش و ارزیابی می شود. اسکریون، این الگو را به عنوان مکمل ارزشیابی مبتنی بر هدف معرفی می کنند و معتقد است نقش عمده ی آن کاستن از سرگیری و افزایش عینیت است.

رویکرد مبتنی بر نظر متخصصان

در این رویکرد تأکید اصلی بر کاربرد مستقیم نظر متخصصان در قضاوت درباره ی کیفیت فعالیت های آموزش و پرورش است. الگوهای ارزشیابی برخاسته از نظر متخصصان را در چهار طبقه زیر میتوان دسته بندی کرد:

۱- نظام های بررسی تخصصی رسمی

۲- نظامهای بررسی تخصصی غیررسمی

۳- بررسی توسط جلسات میزگرد ویژه

۴- بررسی توسط افراد ویژه

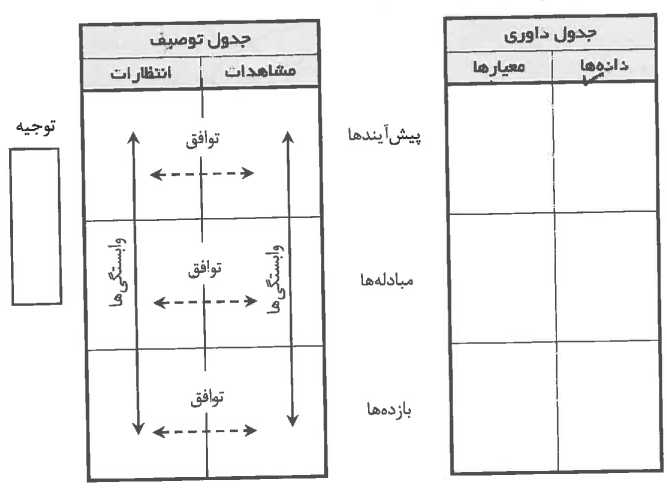
در موقعیت هایی مانند امتحانات شفاهی دوره دکتری، میزگردرهایی که برای بررسی پیشنهادهای تحقیق تشکیل می شوند، جلسات بررسی پرونده های - استخدامی و غیره از ارزشیابی مبتنی بر نظر متخصصان استفاده می شود.

رویکرد مبتنی بر اختلاف نظرهای متخصصان

در این رویکرد، اختلاف نظرهای ارزشیابان مختلف مورد تأکید قرار می گیرد. این رویکرد بر این فرض استوار است که تمام متخصصان، ارزشیابی خواسته - و ناخواسته نظرات و تعصبات خود را در جریان ارزشیابی دخالت می دهند؛ لذا برای این که نتایج درست تری از ارزشیابی به دست آوریم، باید از نظرهای موافق و مخالف چند تن از متخصصان و صاحب نظران استفاده کنیم. این رویکرد امیدوار است که با دخالت دادن نظرهای مثبت و منفی در ارزشیابی سوگیری ها را متعادل و از این طریق عدالت را برقرار کند.

رویکرد طبیعت گرایانه و مبتنی بر مشارکت کنندگان

این رویکرد بر تجارب دست اول در موقعیت ها و فعالیت های تربیتی و آموزشی تأکید دارد. از آن جهت به این رویکرد، طبیعت گرایانه می گویند که به بررسی مبتنی بر مشاهدات طبیعی تأکید می کند و از آن جهت به آن مبتنی بر مشارکت کنندگان می گویند که بر دخالت کسانی که در موضوع مورد ارزشیابی شرکت می کنند، تأکید دارد. یکی از الگوهای ارزشیابی مبتنی بر این رویکرد، الگوی ارزشیایی سیمایی است. در این الگوی ارزشیابی بر دو عمل اساسی - ارزشیابی یعنی توصیف و داوری تأکید می شود و اینها دو سیما یا دو روی ارزشیابی به حساب می آیند؛ بدین لحاظ که آن الگوی ارزشیابی سیمایی می گویند. دو عمل توصیف و داوری در سه مرحله از یک برنامه آموزشی یعنی پیش آیندها، مبادله ها و بازده ها یا پی آیندها انجام می گیرد. اینک برای کمک به ارزشیاب برای سازمان دادن به داده های جمع آوری شده و تفسیر آن ها طرح زیر ارائه شده است:



مثال ۱: در کدام یک از رویکردهای مربوط به ارزشیابی بر تجارب دست اول در موقعیت ها و فعالیت های آموزشی و تربیتی تأکید می شود؟

1) رویکرد مبتنی بر هدف

۲) رویکرد مبتنی بر مدیریت

۳) رویکرد مبتنی بر نظر متخصصان

۴) رویکرد طبیعت گرایانه و مبتنی بر مشارکت کنندگان

پاسخ: گزینه «۴» در رویکرد طبیعت گرایانه و مبتنی بر مشارکت کنندگان، بر تجارب دست اول در موقعیت ها و فعالیت های آموزشی تأکید می شود.

مثال ۲: الگوی ارزشیابی سیپ در کدام یک از رویکردهای مربوط به ارزشیابی مطرح می شود؟

1) رویکرد مبتنی بر هدف

۲) رویکرد مبتنی بر مدیریت

۳) رویکرد مبتنی بر مصرف کنندگان

۴) رویکرد مبتنی بر نظر متخصصان

پاسخ: گزینه «۲» ارزشیابی سیپ جزء الگوهایی است که در رویکرد مبتنی بر مدیریت مطرح است.

مثال 3: الگوی ارزشیابی سیمایی (که بر دو عمل ارزشیابی یعنی توصیف و داوری تأکید می کند) در کدام یک از رویکردهای مربوط به ارزشیابی مطرح می شود؟

1) رویکرد مبتنی بر نظر متخصصان

۲) رویکرد مبتنی بر مدیریت

۳) رویکرد طبیعت گرایانه مبتنی بر مشارکت کنندگان

۴) رویکرد مبتنی بر هدف

پاسخ: گزینه «3» الگوی ارزشیابی سیمایی در رویکرد طبیعت گرایانه مبتنی بر مشارکت کنندگان مطرح است.

مثال ۴: الگوی ارزشیابی اسکریون و الگوی ارزشیابی هدف - آزاد را می توان از الگوهای مربوط به رویکرد ............................................. به حساب آورد.

۱) رویکرد مبتنی بر نظر متخصصان

۲) رویکرد مبتنی بر مصرف کنندگان

۳) رویکرد مبتنی بر هدف

۴) رویکرد مبتنی بر مدیریت

پاسخ: گزینه «۲» رویکرد مبتنی بر مصرف کنندگان

دسته بندی ارزشیابی های آموزشی با توجه به موضوع ارزشیابی رویکردهای مختلف

ارزشیابی از دانشجویان و دانش آموزان: که عمدتا ارزشیابی از پیشرفت تحصیلی با ارزشیابی از میزان یادگیری آن هاست. از جمله وسایلی که به منظور ارزشیابی از می دانشجویان و دانش آموزان بکار برده می شود، می توان از آزمون های معلم ساخته و میزان شده، گزارشهای شفاهی و مشاهدات رسمی و غیر رسمی نام برد.

۲- ارزشیابی از برنامه های درسی و مواد آموزشی ارزشیابی گاهی به منظور تعیین اثربخشی برنامه درسی یا مواد آموزشی انجام می گیرد، مثل - ارزشیابی از مواد آموزشی، روش های آموزشی، کتاب های درسی.

عمده ترین ملاک در ارزشیابی از برنامه های درسی یا مواد آموزشی، پیشرفت تحصیلی یا میزان یادگیری دانش آموزان است و ملاک دیگر هزینه آن هاست؛ به منظور این است که از لحاظ هزینه در یکی از جهات زیر قابل دفاع باشد:

ما از لحاظ هزینه با سایر برنامه ها شبیه باشد، اما به پیشرفت بیشتری بینجامد. ۲- از لحاظ هزینه کمتر از سایر برنامه ها خرج بردارد، اما به پیشرفت بیشتری بینجامد. ۳- اگر پرهزینه باشد در عوض باعث پیشرفت بیشتری شود.

٣- ارزشیابی از پروژه ها و برنامه های آموزشی

۴- ارزشیابی از آموزشگاه: منظور از ارزشیابی از آموزشگاه این است که تعیین شود، هدف های پیش بینی شده برای آموزشگاه به چه میزانی تحقق یافته اند و نقاط قوت و ضعف کدام اند؟ یکی از جنبه های مهم ارزشیابی از آموزشگاه برنامه آزمون آموزشگاه می باشد؛ هرچه این برنامه جامع تر و کامل تر باشد اطلاعات به دست آمده قابل اعتماد و با ارزش تر است.

۵- ارزشیابی از کارکنان آموزشگاه: منظور از ارزشیابی از کارکنان آموزشگاه، ارزشیابی از همه کسانی است که به طور مستقیم یا غیرمستقیم، مسئول بازده های آموزشی هستند. پیچیده ترین نوع ارزشیابی، ارزشیابی از کارکنان بویژه ارزشیابی از کار معلمان است که به دلیل کم اعتباری و بی دقتی وسایل و روش های اندازه گیری، اطلاعات بی غرضانه ای را که ضروری است، به دست نمی دهد.

دسته بندی ارزشیابی های آموزشی باتوجه به ملاک های مورداستفاده رویکردهای مختلف

الف) ارزشیابی وابسته به ملاک (املاکی)

ب) ارزشیابی وابسته به هنجار (هنجاری)

الف) ارزشیابی وابسته به ملاک: وقتی که آزمودنی ها یا برنامه ها بر اساس ملاک یا معیار از پیش تعیین شده ای سنجیده شوند، به آن ارزشیابی ملاکی یا سنجش از هر نوع ملاک مطلق می گوییم که این ملاک یا معیارهای معین، همان اهداف هستند. این نوع سنجش، بیشتر در مورد ارزشیابی پیشرفت تحصیلی به کار میرود. در استفاده از ملاک مطلق در ارزشیابی پیشرفت تحصیلی، پیشرفت دانش آموزان با توجه به آنچه آموزش داده میشود، سنجیده می شود؛ یعنی میزان دسترسی به اهداف و موفقیت در آزمون های مبتنی بر ملاک مطلق، مستلزم دستیابی به اهداف است.

ارزشیابی وابسته به هنجار: این نوع ارزشیابی به یک ملاک نسبی وابسته است و اگر بخواهیم پیشرفت تحصیلی دانش آموز را ارزشیابی کنیم، باید عملکرد او را با عملکرد سایر دانش آموزان مقایسه کنیم؛ یعنی زمانی به کار میرود که ما قصد مقایسه آزمودنی ها و برنامه را داشته باشیم؛ بنابراین در این نوع ارزشیابی، می توان تعیین کرد که پیشرفت یک دانش آموز نسبت به دانش آموزان دیگر چه وضعیتی دارد، اما نمی توان مشخص کرد که پیشرفت یک دانش آموز نسبت به هدف های آموزشی با آنچه آموزش داده شده، چگونه است.

مقایسه ارزشیابی ملاکی و هنجاری

1- هدف ارزشیابی وابسته به هنجار مقایسه دانش آموزان با یکدیگر است؛ یعنی می خواهیم بفهمیم یک دانش آموز نسبت به دانش آموزان دیگر در چه وضعیتی قرار دارد، ولی در ارزشیابی ملاکی ما می خواهیم بفهمیم که هرکس تاچه اندازه، آن چیزی را که باید یاد گرفته باشد، یاد گرفته است؛ یعنی تا چه حد به اهداف رسیده است. ۲\_ ارزشیابی از نوع ملاک نسبی یا هنجاری در مورد استعداد، هوش، شخصیت و سایر جنبه های عاطفی به کار می رود، ولی ارزشیابی ملاکی در ارزشیابی از پیشرفت تحصیلی که هدف آن پی بردن به میزان دسترسی به اهداف است، به کار میرود.

3- تفاوت مهم دیگر بین آزمون های هنجاری و ملاکی در توزیع فراوانی آن هاست. در آزمون های هنجاری، یک منحنی بهنجار یا طبیعی داریم؛ یعنی بیشتر نمرات در حد میانگین هستند، نه خیلی بالاتر و نه خیلی پایین تر، اما در ارزشیابی ملاکی یک منحنی با کجی منفی داریم؛ یعنی بیشتر نمرات، بالا هستند و بیشتر افراد به اهداف دست پیدا کرده اند.

مثال ۵: اگر در یک آزمون، هدف، مقایسه فراگیران با یکدیگر باشد از کدام نوع ارزشیابی باید استفاده کنیم؟

1) تکوینی

2) ملاکی

۳) تشخیصی

۴) هنجاری

پاسخ: گزینه «۴» از اهداف ارزشیابی هنجاری «مقایسه فراگیران با یکدیگر می باشد.

مثال ۶: اگر هدف، ارزشیابی از استعداد، هوش، شخصیت و سایر جنبه های عاطفی دانش آموزان باشد، از کدام نوع ارزشیابی باید استفاده کنیم؟

1) تکوینی

۲) تراکمی

3) هنجاری

۴) تشخیصی

پاسخ: گزینه «3» از آزمون های هنجاری برای ارزشیابی استعداد، هوش، شخصیت و سایر جنبه های عاطفی دانش آموزان استفاده می شود.

مثال ۷: ضابطه مورداستفاده برای نمره گذاری در مدارس ایران عمدتا چگونه است؟

١) ملاک مدار

۲) هنجارمدار

3) غیرقابل مقایسه

۴) دلخواه معلم

پاسخ: گزینه «۱» همانگونه که همه ی ما می دانیم ملاک قبولی در مدارس ایران نمره ۱۰ و حداقل معدل ۱۰ میب اشد؛ بنابراین داوطلبان می توانند به راحتی حتی اگر با تعاریف ملاک مرجع هم آشنا نباشند. به طور حدس و با اطمینان ۱۰۰ درصد گزینه ی صحیح ملاک مدار را انتخاب نمایند.

مثال ۸: اگر در یک امتحان ورودی، کسانی را قبول کنند که به 40% سؤالات جواب داده اند، کدام نوع ارزشیابی به کار رفته است؟

1) تشخیصی

۲) تکوینی

۳) ملاکی

۴) هنجاری

پاسخ: گزینه «3» در ارزشیابی وابسته به ملاک با ارزشیابی ملاکی، ملاک یا معیار ارزشیابی، از پیش تعیین شده و عملکرد یادگیرنده باتوجه به آن سر ملاک یا معیار میشود. در این ارزشیابی، توجه به پیشرفت تحصیلی دانش آموزان و دانشجویان است. در این نوع ارزشیابی، از آزمون شوندگان انتظار می رود می که به هدف های آموزشی معیتی دست یابند و یا به سطح بخصوصی از عملکرد برسند. با توجه به متن سؤال که از آزمون شوندگان انتظار دارد که به 80 درصد سؤالات، جواب درست دهند، گزینه ی ۳ را به عنوان گزینه ی صحیح انتخاب می کنیم.

مثال ۹: در کدام یک از آزمون های زیر به داوطلبان با یکدیگر مقایسه می شوند؟

۱) دروس دانشگاهی

2) گواهینامه رانندگی

۳) نهایی پیش دانشگاهی

۴) ورودی دانشگاه ها

پاسخ: گزینه «۴» در ارزشیابی وابسته به ملاک یا ارزشیابی ملاکی، معیار ارزشیابی از پیش تعیین می شود و عملکرد یادگیرنده با توجه به آن ملاک مورد قضاوت قرار می گیرد؛ در واقع از نوعی ملاک مطلق که همان هدف های آموزشی از پیش تعیین شده هستند، استفاده می شود، اما در آزمون های وابسته به هنجار به جای یک ملاک از قبل تعیین شده، نوعی ملاک نسبی موردنظر است. در این نوع ارزشیابی، عملکرد دانش آموزان با یکدیگر مقایسه می شود. علت نام گذاری این نوع ارزشیابی، هنجار یا ماکی است که براساس آن می توان تعیین کرد پیشرفت یک دانش آموز نسبت به دانش آموز دیگر چه وضعی دارد؛ بنابراین در کنکور ورودی دانشگاه ها، داوطلبان با یکدیگر مقایسه می شوند.

انواع سنجش با توجه به زمان اجرا و هدف

الف) سنجش آغازین

ب) سنجش تکوینی

ج) سنجش تشخیصی

د) سنجش تراکمی

الف) سنجش آغازین

سنجش آغازین به دو منظور انجام می شود:

1- رفتار ورودی را تعیین می کند؛ یعنی آیا دانش آموزان اطلاعات لازم را برای یادگیری واحد جدید کسب کرده اند؟

۲- پیش آزمون را شامل می شود؛ یعنی یادگیرندگان چه مقدار از هدف ها و محتوای درس تازه را قبلا یاد گرفته اند؟ و دانش آموزان، چه اطلاعاتی درباره درس جدید دارند؟

دو کمک عمده سنجش آغازین

۱) نقطه شروع آموزش را مشخص می کند.

۲) برای جایابی و گروه بندی دانش آموزان به کار می رود (پیش آزمون به منظور جایابی یا پایه گزینی است).

اگر نتایج آزمون ورودی نشان دهد که دانش آموزان برای یادگیری درس جدید، آمادگی کافی ندارند، باید به بچه ها آموزش جبراتی داده شود. - اگر نتایج پیش آزمون نشان بدهد که همه ی دانش آموزان یا بیشتر آن ها از درس جدید، اطلاع کافی دارند، برنامه باید عوض شود.

ارزشیابی آغازین، زمانی ضروری و حتمی است که معلم از دانشها و مهارتهای دانش آموزان اطلاع کافی نداشته باشد؛ یعنی با گروه جدیدی از دانش آموزان روبه رو باشد. متخصصان آموزشی پیشنهاد کرده اند که معلم در ارزشیابی آغازین خود، ترکیبی از دو آزمون یعنی ۱- رفتار ورودی و ۲- آزمون نهایی (پیش آزمون را با هم اجرا کند.

ب) سنجش تکوینی

سنجش تکوینی در جریان آموزش اجرا می شود؛ یعنی در طول دوره آموزشی، هنگامی که فعالیت معلم ادامه دارد و یادگیری در حال تکوین است. هدف از و ارزشیابی تکوینی، پی بردن به میزان دسترسی به اهداف یا میزان یادگیری و پیدا کردن نقایص یا اشکالات روش معلم است. سنجش تکوینی یکی از مهمترین سنجش هاست؛ زیرا معلم، فرصت اصلاح دارد.

در آزمون های مورد استفاده در سنجش تکوینی، آزمون های ملاکی هستند؛ زیرا ما می خواهیم بفهمیم تا چه حد به اهداف دست یافته ایم.

برای بهره گیری کامل از نتایج ارزشیابی تکوینی، بیان هدف های آموزشی به صورت دقیق و رفتاری ضروری می باشد.

نکته دیگر این که هدف ارزشیابی تکوینی، تعیین میزان توفیق یادگیرنده در رسیدن به تک تک هدف های واحد درسی است و هیچ گونه مقایسه میان افراد مختلف صورت نمی گیرد.

مثال ۱۰: آزمون مورد استفاده معلم در آغاز یک جلسه درسی که برای سنجش آموخته های یادگیرندگان از هدف های درس جلسه قبل به کار می رود، چه نام دارد؟

۱) تکوینی

۲) تراکمی

۳) تشخیصی

۴) هنجاری

پاسخ: گزینه «۱» هدف از اجرای ارزشیابی تکوینی، اطلاع از میزان و نحوه ی یادگیری افراد از نقاط ضعف یا قوتشان در طول دوره ی آموزشی است؛ چون معلم و شاگرد هنوز در جریان آموزش هستند و امکان رفع مشکل یادگیری و برطرف کردن نواقص روش آموزشی میسر است. گزینه ی ۳ به گزینه ی صحیح بسیار نزدیک است. چون این ارزشیابی هم در طول دوره ی آموزشی انجام می شود، اما تفاوت ارزشیابی تکوینی و ارزشیابی تشخیصی این است که اولی به مشکلات کوچک یادگیری و دومی به مشکلات بزرگتر و کلی تر می پردازد و متن سؤال به گزینه ی یک اشاره کرد.

مثال ۱۱: معلمی هر دو هفته، یکبار، آزمونی بر مبنای مطالب تدریس شده در کلاس به دانش آموزان می دهد و با استفاده از نمرات حاصل، در تدریس خود به نحوی عمل می کند که نقاط ضعف دانش آموزان در یادگیری مطالب برطرف شود. این ارزشیابی از کدام نوع است؟

۱) تکوینی

۲) تراکمی

۳) تشخیصی

۴) تشویقی

پاسخ: گزینه «۱» ارزشیابی تکوینی در طول دوره به مرور صورت می گیرد. این ارزشیابی، معلم و برنامه ریزان درسی را از نواقص احتمالی آگاه می کند. گزینه ی سه هم در طول دوره انجام می شود، اما ارزشیابی تشخیص با هدف مشکلات خاص صورت می گیرد و در مقایسه، ارزشیابی تکوینی، کلی تر و وسیع تر از ارزشیابی تشخیص است.

مثال ۱۲: به منظور تشخیص نقاط ضعف دانش آموزان در یادگیری مطالب ارائه شده و کمک به رفع نارسایی های یادگیری آنان، بهتر است از کدام یک از ارزشیابی ها استفاده شود؟

۱) آغازین

2) تشخیصی

3) تکوینی

4) تراکمی

پاسخ: گزینه «3» آنچه عمدتا جهت اصلاح موضوع مورد ارزشیابی، یعنی برنامه با روش آموزشی، استفاده می شود ارزشیابی تکوینی نام دارد. این ارزشیابی با هدف اطلاع و آگاهی تولیدکنندگان برنامه ها و اصلاح های لازم صورت می گیرد. معلم با استفاده از ارزشیابی تکوینی می تواند در زمانی که هنوز یادگیری در جریان است و امکان رفع مشکلات یادگیری دانش آموزان و برطرف کردن نواقص روش آموزشی خود او میسر است، چنین کاری انجام دهد.

برای بهره گیری کامل از نتایج ارزشیابی تکوینی، بیان هدفهای آموزشی به صورت دقیق و رفتاری ضروری است. گزینه شماره دو ارزشیابی تشخیصی، سر می تواند به گزینه ی صحیح نزدیک باشد. ارزشیابی تشخیصی به یافتن ریشه های یک مشکل می پردازد. «لین و گرانداند» (۲۰۰) بار در مقایسه ارزشیابی تکوینی با ارزشیابی تشخیصی گفته اند: «مقایسه این دو با مثالی در علم پزشکی واضح می شود. برای مشکلات سادهی یادگیری از

جنبه ی ارزشیابی تکوینی استفاده میشود که به صورت کمکهای اولیه می باشد، ولی ارزشیابی تشخیصی به دنبال کشف علتهای زیربنایی مشکلات میباشد که با کمک های اولیه قابل درمان نیست.». بنابراین داوطلب باید در انتخاب خود دقت لازم را داشته باشد.

ج) سنجش تشخیصی

این سنجش هم، در جریان آموزش انجام میشود و تفاوت آن با سنجش تکوینی در این است که اگر دانش آموز با دانش آموزانی در یک زمینه خاص دارای به مشکل باشند و مشکل، از طریق ارزشیابی تکوینی شناسایی و جل نشود، معلم از همه روش های ممکن برای شناسایی و حل مشکل استفاده می کند؟

بدین جهت به آن ارزشیابی تشخیصی می گویند؛ چون از روش های مختلف برای حل مشکل استفاده میشود. - تفاوت آزمون های تشخیصی با آزمونهای پیشرفت تحصیلی این است که آزمون های پیشرفت تحصیلی، توفیق دانش آموزان در رسیدن به اهداف را می سنجند، اما آزمون های تشخیصی، نقاط قوت و ضعف را توصیف می کنند و علاوه بر شناسایی مشکل، باید راه حل مناسب هم ارائه شود.

در ارزشیابی تشخیصی، هم از آزمون های ملاکی و هم از هنجاری استفاده میشود. آزمون های استاندارد شده تشخیصی در سطح دیستان هستند.

مثال ۱۳: « ارزشیابی تشخیصی») (Diagnostic Evaluation در کدام مرحله ی آموزش انجام می شود؟ 1) در آغاز تدریس

۲) در پایان تدریس

۳) در پایان و آغاز تدریس

۴) در پایان و ضمن تدریس

پاسخ: گزینه صحیح وجود ندارد. ارزشیابی تشخیصی Diagnostic evaluation، ارزشیابی ای است که در جریان آموزش صورت می گیرد و هدف آن تشخیص مشکلات دانش آموزان در زمینهی یادگیری است؛ به همین علت به آن ارزشیابی تشخیصی می گویند. اگر گزینه ی چهار به تنهایی به «ضمن تدریس» اشاره می کرد، آن گاه می توانستیم این گزینه را به عنوان گزینه ی صحیح انتخاب کنیم.

د) سنجش تراکمی

در ارزشیابی تراکمی، تمام آموخته های دانش آموزان در طول یک دوره، در پایان آن دوره ارزشیابی می شود و هدف از آن، نمره دادن به دانش آموز و قضاوت درباره اثربخشی کلر معلم و برنامه درسی پا مقایسه برنامه های درسی با هم است، بدین جهت به آن تراکمی می گویند؛ زیرا مجموع یادگیری های دانش آموزان با یادگیری های متراکم آنها را می سنجد. هدف ارزشیابی تراکمی، تعیین صلاحیت یا عدم صلاحیت افراد برای ارتقاء به سطح بالاتر یا دوره بعدی است.

در ارزشیابی مواد و برنامه ها هم ارزشیابی تراکمی در پایان برنامه، یعنی وقتی برنامه به طور کامل اجرا شد، انجام میشود و هدف از آن جمع آوری اطلاعات بر مبتنی بر داوری های ارزشیابانه درباره برنامه برای مصرف کنندگان بالقوه برنامه است.

آزمون های مورد استفاده در ارزشیابی تراکمی، هم آزمون های هنجاری یا وابسته به ملاک نسبی هستند و هم آزمون های ملاکی یا وابسته به ملاک مطلق.

مثال ۱۴: کدام یک از عبارات زیر جزو هدف های ارزشیابی تراکمی است؟

1) گزینش فراگیران

2) بهبود فرایند های تدریس و یادگیری

۳) تعیین نتیجه کارفراگیران در یک دوره آموزشی

۴) تعیین وضع فراگیران هنگام شروع یک برنامه آموزشی

پاسخ: گزینه «3» هدف ارزشیابی تراکمی این است که تمام آموخته های دانش آموزان در طول یک دوره ی آموزشی تعیین و در نهایت، نمره به دانش آموزان ارائه شود؛ در واقع، قضاوت درباره ی اثربخشی کار معلم و برنامه ی درسی او صورت گیرد. این ارزشیابی، معمولا در پایان دوره آموزشی انجام ۱ می شود. گزینه دو می تواند برای داوطلبانی که با حدس و گمان، انتخاب می کنند، مشکل ساز شود.

تفاوت ارزشیابی تراکمی و تکوینی

۱- ارزشیابی تکوینی در جریان آموزش انجام میشود اما ارزشیابی تراکمی در پایان آموزش انجام می شود.

۲- مخاطبان با استفاده کنندگان نتایج ارزشیابی تکوینی، طراحان و تولید کنندگان برنامه هستند، ولی مخاطبان ارزشیابی تراکمی، افرادی هستند که از برنامه استفاده می کنند یا استفاده خواهند کرد.

٣- ارزشیابی تکوینی به تولید و توسعه برنامه کمک می کند، اما ارزشیابی تراکمی به ادامه، قطع و یا گسترش برنامه کمک می کند.

۴- در ارزشیابی تکوینی، ما امکان اصلاح داریم، اما در ارزشیابی تراکمی این امکان وجود ندارد.

مثال ۱۵: برای تصمیم گیری در مورد ارتقاء افراد به سطح بالاتر از کدام ارزشیابی باید بهره گرفت؟

۱) تکوینی

۲) تراکمی

۳) هنجاری

۴) تشخیصی

پاسخ: گزینه «۲» برای تصمیم گیری در مورد ارتقاء افراد به سطح بالاتر، از ارزشیابی «تراکمی» باید بهره بگیریم.

مثال ۱۶: در کدام نوع ارزشیابی، هم نقاط قوت و ضعف مشخص می شود و هم باید راهکار مناسب برای حل مشکل ارائه شود؟

۱) تشخیصی

۲) تکوینی

۳) تراکمی

۴) آغازین

پاسخ: گزینه «۱» در ارزشیابی تشخیصی نقاط قوت و ضعف دانش آموز مشخص می شود و راهکارهای متناسب با آن ارائه می شود.

دسته بندی ارزشیابی های آموزشی با توجه به ارزشیابان

الف) ارزشیابی درونی

ب) ارزشیابی بیرونی

ارزشیابی درونی: در ارزشیایی درونی، ارزشیابی توسط کارکنان و دست اندرکاران برنامه، یعنی افراد درون برنامه انجام می شود.

ارزشیابی بیرونی: ارزشیابی بیرونی توسط افرادی که خارج از برنامه هستند، صورت می گیرد.

ارزشیابی درونی، متناسب با ارزشیابی تکوینی است و ارزشیابی بیرونی متناسب با ارزشیابی تراکمی است.

ارزشیابی درونی یا تکوینی به منظور داوری درباره کیفیت مواد، روش ها و برنامه در ضمن طراحی است. هدف این ارزشیابی کمک به شکل گیری، تجدیدنظر، اصلاح و بهبود محصول آموزشی است.

اصطلاح ارزشیابی درونی به جای ارزشیابی تکوینی در مورد مواد و برنامه ها به کار می رود، ولی در مورد فعالیت های آموزشی معلمان، از اصطلاح ارزشیابی تکوینی استفاده می شود.

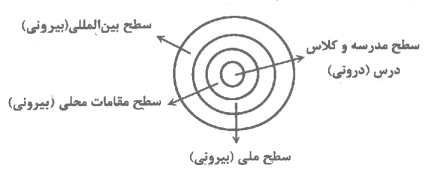
اصطلاح ارزشیابی درونی به جای ارزشیابی تکوینی غالبأ در رابطه با مواد و برنامه های آموزشی به کار می رود.اما در ارتباط با فعالیت های آموزشی معلمان، عموما از اصطلاح تکوینی استفاده م یشود.

ارزشیابی تراکمی یا بیرونی، هم نقاط قوت و ضعف را نشان می دهد و هم برای تعیین میزان توفیق در رسیدن به اهداف به کار می رود.

ارزشیابی درونی را با ارزشیابی از خود با خودارزشیابی معادل می گیرند.

به ارزشیابی درونی، بیشتر برای این منظور به کار می رود که ببینیم به هدف هایمان رسیده ایم یا نه. در صورتی که، هدف ارزشیابی بیرونی این است که مشخص شود، دانش آموزانی که با روش خاصی آموزش دیده اند، نسبت به دانش آموزانی که برنامه دیگری را گذرانده اند، چه وضعی دارند.

مک بیت و مک گلین (۲۰۰۲) می گویند ارزشیابی درونی به منظور وارسی و بازبینی از جنبه های عملکرد به وسیله صاحبان اصلی آن ها (کارکنان، دانش آموزان ) به کار میرود؛ سطح مدرسه و کلاس ولی ارزشیابی بیرونی، یعنی بازبینی و وارسی فعالیت های آموزشگاه و گزارش نتایج آن، و به وسیله کسانی که اعضای سازمان آموزشگاه به حساب نمی آیند. این دو محقق، انجام هر دو نوع ارزشیابی ضروری می دانند و رابطه بین آن ها را مانند شکل روبه رو نشان داده اند.



مثال 1۷: ارزشیابی درونی را که توسط کارکنان و دست اندرکاران برنامه اجرا می شود، می توان متناسب با ارزشیابی .................... دانست و ارزشیابی بیرونی را که توسط افرادی که خارج از برنامه قرار دارند، صورت می گیرد، می توان متناسب با ارزشیابی ................... دانست.

۱) تکوینی - تکوینی

۲) تراکمی – تراکمی

۳) تکوینی - تراکمی

۴) تراکمی - تکوینی

پاسخ: گزینه «٣» ارزشیابی درونی را می توان با ارزشیابی تکوینی و ارزشیابی بیرونی را میتوان با ارزشیابی تراکمی متناسب دانست.

ارزشیابی از معلم

تعریف: ارزشیابی از معلم عبارت است از میزان موفقیت معلمان در رسیدن به هدف های آموزشی خود

اهداف ارزشیابی از معلم

١- کمک به معلمان در جهت اصلاح و بهبود روش ها و فعالیت های آموزشی خودشان

۲- کمک به مدیران و مسئولان در جهت تصمیم گیری معقولانه تر درباره استخدام و ترفیع معلمان.

انواع روش های ارزشیابی از معلم :

چون ملاک های موفقیت معلمان متنوع اند، روش های ارزشیابی از معلمان هم متنوع هستند:

١- روش مصاحبه که به دو منظور به کار می رود: ۱- استخدام معلمان ۲- نگهداری معلمان و دادن امتیاز به آنهاء کمتر

2\_ مشاهده از کلاس درس: در این روش، اطلاعات مورد نیاز به طور مستقیم از اعمال و رفتار معلم در کلاس درس جمع آوری می شود؛

۳-استفاده از نظر دانش آموزان و دانشجویان در این روش به طور رسمی یا غیررسمی نظر دانشجویان و دانش آموزان درباره ویژگی های شخصی و حرفه ای معلم پرسیده می شود؛

۴- استفاده از یادگیری دانش آموزان و دانشجویان در این روش، میزان یادگیری با پیشرفت تحصیلی دانش آموزان با دانشجویان به عنوان ملاکی برای شایستگی معلم به کار میرود.

۵- خودسنجی یا ارزشیابی از خودت خودسنجی معلم را قضاوت درباره آموزش خود تعریف کرده اند که مناسب ترین روش برای مطالعه و بهسازی آموزش معلمان است و به روش های مختلفی انجام می شود، مثل درجه بندی شخصی، گزارش های شخصی، مواد مطالعه شخصی، مشاهده آموزش همکاران ؛ضبط صوت یا ضبط ویدیویی از تدریس خود و غیره؛

۶- ارزشیابی به وسیله همکاران در این روش، از نظر همکاران معلم برای قضاوت درباره اثربخشی کار او استفاده می شود. یکی از مشکلات این روش، پایین بودن سطح پایایی آن است؛ یعنی بین نمراتی که همکاران مختلف به یک معلم می دهند، همخوانی زیادی وجود ندارد؛

۷- بررسی آثار علمی معلم: این روش، ویژه اعضای هیأت های علمی دانشگاه ها و مؤسسات آموزشی عالی است. در این روش، آثار علمی معلم، مورد ارزشیابی قرار می گیرد. ملاک این ارزشیابی، کمیت و کیفیت آثار علمی معلم است.

نکته: هریک از روش های ارزشیابی معلم تنها بخشی از فعالیت های علمی- آموزشی معلم را می سنجد. برای به دست آوردن تصویر کامل تری از آموزش و فعالیت های علمی معلم لازم است، که از روش های مختلف ارزشیابی استفاده کنیم.

مثال ۱۸: بهترین روش برای اصلاح و بهسازی آموزش معلمان کدام است؟

۱) مصاحبه

2) مشاهده از کلاس درس

3) خودسنجی

۴) استفاده از نظر فراگیران

پاسخ: گزینه «۲» خودسنجی بهترین روش برای اصلاح بهسازی آموزش معلمان است.

تست های طبقه بندی شده فصل چهارم

۱- مهمترین تفاوت آزمون های «ملاک - مرجع» با آزمون های «هنجار - مرجع» چیست؟

(سراسری ۸۲)

۱) تعداد شرکت کنندگان در آزمون

۲) روش اندازه گیری و نوع سؤال های آزمون

3) ضریب دشواری و تشخیص سؤال های آزمون

۴) معیار تصمیم گیری درباره ی سؤال های آزمون

۲- امتحانات کوتاه (کوئیز)، که معلمان در طول یک دوره ی تحصیلی به عمل می آورند، چه نوع ارزیابی است؟

(سراسری ۸۳)

1) تشخیصی

2) تکوینی

۳) تراکمی

۴) پایانی

۳- در کدام حالت، بین ارزشیابی مستمر و ارزشیابی مقطعی، آمیختگی به وجود می آید؟

(سراسری ۸۳)

۱) هدف از امتحان مقطعی، تهیه ترازنامه پایایی باشد.

۲) حصول اهداف معینی مورد بررسی کامل قرار گیرد.

۳) نمرات طول سال در ارزیابی نهایی دخالت داده شود.

۴) نمرات پایان سال از روی نمرات کلاسی تعیین شود.

۴- ارزشیابی تشخیصی زمانی وارد عمل می شود که بخواهند:

(سراسری ۸۳)

1) تحول هر یک از دانش آموزان را در طول سال دنبال کنند.

۲) پیشرفت و نزدیکی آزمودنی به هدف تعیین شده را بررسی کنند.

3) از روند کار، ترازنامه ای فراهم آورند و تصمیم خاصی اتخاذ کنند.

۴) از توانایی آزمودنی برای شروع نوع خاصی از یادگیری مطلع شوند.

5\_ ما در کدام یک از موارد زیر، معمولا آزمون تسلط مورد استفاده قرار می گیرد؟

(سراسری ۸۳)

۱) اخذ گواهینامه رانندگی

۲) مقایسه داوطلبان کنکور با یکدیگر

۳) ارتقاء به تناسب فعالیت

۴) طبقه بندی افراد از روی معدل دیپلم

6\_کدام یک از انواع ارزشیابی ها تقریبا در حاشیه قرار دارد و مکمل عمل آموزش نیست؟

(سراسری 85)

۱) تشخیصی

۲) تکوینی

۳) مستمر

4) مجموعی

7\_ امتحانات کوتاه (کوئیز) که معلمان در طول یک دوره درسی به عمل می آورند، در زمره کدام نوع ارزشیابی قرار می گیرد؟

(سراسری ۸۶)

۱) تکوینی

۲) تراکمی

۳) تشخیصی

۴) پایانی

۸- وقتی نمره گذاری را براساس پیکه ملاک مطلق انجام می دهیم و تسلط دانش آموزان را در نظر میگیریم، اولین سؤالی که مطرح می شود، کدام است؟

(سراسری ۸۶)

1) رابطه ی تسلط با ملاک چگونه است؟

۲) منظور از تسلط چیست؟

3) ملاک مطلق چه محاسنی دارد؟

۴) ملاک رایج کدام است؟

۹- تفاوت بین «ارزشیابی تکوینی و ارزشیابی پایانی» اساسا در چیست؟

(سراسری ۸۸)

۱) ارزشیابی تکوینی در طول سال و ارزشیابی پایانی در آخر سال انجام می شود.

2) ارزشیابی تکوینی به تدریج و ارزشیابی پایانی یکباره انجام میشود.

۳) ارزشیابی تکوینی برای پاسخگویی و ارزشیابی پایانی برای ارائه بازخورد انجام می شود.

4) ارزشیابی تکوینی برای ارائه بازخورد و ارزشیابی پایانی برای دادن گواهی انجام میشود.

۱۰- انتظار می رود که نوع توزیع نمرات آزمون پیشرفت تحصیلی، که در پایان ترم از دانشجویان به عمل می آید، چگونه باشد؟

(سراسری ۸۸)

۱) کجی منفی

۲) نرمال

3) کجی مثبت

۴) اطلاعات مسئله کافی نیست.

۱۱- منظور از ملاک نسبی کدام است؟

آزاد ۸۸)

۱) ارزشیابی افراد براساس ملاک متغیر

۲) ارزشیابی افراد براساس میزان عملکرد

۳) تعیین موقعیت فراگیر براساس سطح مطلوب پیشرفت

۴) تعیین موقعیت فراگیر با توجه به وضع او در بین افراد گروه خود

۱۲- در کدام سنجش، ملاک قضاوت در مورد عملکرد فرد، مقایسه او با همگنانش است؟

(سراسری ۸۹)

۱) هدف نگر

۲) حوزه نگر

3) ملاک - مرجع

۴) نرم - مرجع

۱۳- مهم ترین هدف ارزشیابی تکوینی کدام است؟

(سراسری ۸۹)

۱) دادن گواهی قبولی یا مردودی

۲) انتخاب روش تدریس مناسب

۳) ارائه بازخورد به معلم و دانش آموزان

۴) تعیین میزان آموخته های قبلی در مورد موضوع مورد تدریس معلم به دانش آموزان

۱۴- مهم ترین تفاوت آزمون های ملاکی و هنجاری کدام است؟

(سراسری ۹۱)

1) تعداد شرکت کنندگان

۲) نحوه قضاوت در مورد نتایج

۳) نحوه اندازه گیری و شیوه مورد استفاده در آنها

۴) ضریب تشخیص و دشواری آزمون

۱۵\_ کدام نوع ارزشیابی برای بازخورد جهت اصلاح امر آموزش نقش بیشتری دارد؟

(سراسری 92)

1) ملاکی

2) تشخیصی

۳) پایانی

4) تکوینی

۱۶- کدام یک از موارد، در مورد هدف های امتحان صحیح نیست؟

(سراسری 92)

1) اصلاح روش تدریس معلم

۲) کمک به یادگیری فراگیران

3) ایجاد رقابت میان فراگیران

۴) ایجاد رغبت و کسب عادت صحیح مطالعه

۱۷- ارزشیابی تکوینی با چه هدفی انجام می شود؟

(سراسری ۹۳)

1) به دست آوردن نتیجه عمل آموزشی

۲) میزان دستیابی به هدف

۳) اصلاح جریان آموزش

۴) شناخت فرایند آموزش

۱۸- یک معلم در جلسه ای از تدریس علوم، متوجه برداشت نادرست و عدم یادگیری مورد انتظار دانش آموزان می شود. او تصمیم می گیرد تا با طرح سؤالاتی که پیش نیاز مفاهیم تدریس شده در آن جلسه کلاسی هستند، به شناخت علل فهم نادرست یادگیرنده ها پی ببرد. این ارزشیابی با چه هدفی انجام می شود؟

(سراسری ۹۴)

1) آغازین

۲) پایانی

3) تشخیصی

۴) تکوینی

۱۹- کدام مورد، بیانگر انواع اندازه گیری و ارزشیابی از نظر هدف اجرا است؟

(سراسری ۹۴)

1) شفاهی - کتبی - عملی

۲) معلم ساخته - استانداردشده

۳) تشخیصی - تراکمی - تکوینی

۴) هنجاری (گروه مرجع) - ملاک مرجع

۲۰- در الگوی سیپ، ارزشیابی از کدام مورد، زمینهای منطقی برای تعیین هدفهای آموزشی فراهم می کند؟

(سراسری ۹۴)

۱) فرآیند

۲) فرآورده

۳) بافت

۴) درون داد

پاسخنامه تست های طبقه بندی شده فصل چهارم

1-گزینه «۴» در ارزشیابی وابسته به ملاک یا ملاک - مرجع، معیار یا ملاک ارزشیابی از پیش تعیین می شود و عملکرد یادگیرنده با توجه به آن ملاک قضاوت می شود. در این ارزشیابی، نوعی ملاکت مطلق، که شامل هدف های آموزشی از پیش تعیین شده است، استفاده می شود. در ارزشیابی وابسته به هنجار و یا هنجار - مرجع به جای یک ملاک از قبل تعیین شده، تمرکز روی نوعی ملاک نسبی است. در این ارزشیابی، نحوه ی عملکرد یا گروه سنجیده می شود.

نه با ملاک از قبل تعیین شده؛ بنابراین هدف ارزشیابی هنجاری مقایسه وضعیت یک دانشجو بر دانش آموز با سایر دانشجویان دانش آموزان است، اما هدف و ارزشیابی ملاک - مرجع آن است که مشخص شود، دانشجو / دانش آموز چه مقدار از آنچه قرار بوده است بیاموزد، فرا گرفته است. گزینهی ۲ گزینه ای است

که با توجه به بدنهی سؤال که به دو آزمون ملاک - مرجع و هنجار- مرجع اشاره می نماید، داوطلبان را دچار خطا می کند و ممکن است آن ها را با توجه به پاسخ مورد نظر که در گزینه ۲ ارائه شده است، به اشتباه بیاندازد.

۲- گزینه «۲» هدف ارزشیابی تکوینی، آگاهی یافتن از میزان و نحوه یادگیری فراگیران برای تعیین نقاط قوت و ضعف یادگیری آنان می باشد. کوئیز، به به معنای امتحانات دوره ای می باشد و زمانی اجرا می شود که هنوز دوره تحصیلی به پایان نرسیده است؛ بنابراین ارزشیابی تکوینی باعث می شود، زمانی که و هنوز یادگیری در جریان است و امکان رفع مشکلات یادگیری دانش آموزان و برطرف کردن نواقص روش آموزشی مسیر است، موجب رشد و بهبود زمینه های یادگیری دانش آموزان شود.

۳- گزینه «3» ارزشیابی مستمر (ارزشیابی تکوینی) در جریان آموزش انجام می شود. ارزشیابی مقطعی (ارزشیابی تراکمی) تمام آموخته های دانش آموزان در طول یک دوره آموزشی، با هدف قضاوت درباره میزان یادگیری دانش آموزان همچنین میزان اثربخشی تدریس معلمان انجام می شود؛ اگر در طول یک دوره آموزشی، نمرات مستمر در ارزیابی نهایی دخالت داده شود، آمیختگی به وجود می آید.

۴- گزینه «۴» ارزشیابی تشخیصی با هدف تشخیص مشکلات یادگیری دانش آموزان صورت می گیرد؛ در واقع معلم. فعالیتی را در پیش می گیرد که مناسب وضع دانش آموزان است و به هر کدام از آنها کمک می کند، تا به هدف های مورد نظر معلم برسند؛ در واقع یک آزمون تشخیصی خوب، نه تنها معلم را از و نقاط قوت و ضعف دانش آموزان آگاه می کند، بلکه زمینه های ضعف و قوت او را نیز آشکار می کند؛ همچنیر از این ارزشیابی، برای اطلاع از توانایی آزمودنی برای شروع نوع خاصی از یادگیری استفاده میشود.

۵- گزینه «1» یادگیری در حد تسلط Mastery learning عبارت است از: یادگیری مطالب آموزش داده شده تا حد معین؛ مثلا این که معلم (استاد) از قبل تعیین کند که یادگیرنده دست کم ۸۵% الی ۹۰% مطالب مورد آموزش را یاد بگیرد. اخذ گواهینامه رانندگی جزء این گروه قرار می گیرد.

6\_گزینه «۱» ارزشیابی ها به این ترتیب تقسیم بندی می شوند - ارزشیابی پیش از آموزش: آمادگی، جایابی ۲- ارزشیابی ضمن آموزش تکوینی، تشخیصی ۳- ارزشیابی پس از آموزش: تراکمی یا مجموعی و ارزشیابی تشخیصی برای تعیین غلط های مصطلح و در واقع در حاشیه قرار دارد و هر وقت لازم باشد، اجرا می شود.

۷- گزینه «1» هدف ارزشیابی تکوینی، آگاهی یافتن از میزان و نحوه یادگیری فراگیران برای تعیین نقاط قوت و ضعف یادگیری آنان می باشد. کوئیز به معنی امتحانات دوره ای است و زمانی اجرا می شود که هنوز دوره تحصیلی به پایان نرسیده است؛ بنابراین ارزشیابی تکوینی باعث می شود، زمانی که هنوز یادگیری در جریان است و امکان رفع مشکلات یادگیری دانش آموزان و برطرف کردن نواقص روش آموزشی میسر است، موجبات رشد و بهبود زمینه های یادگیری دانش آموزان فراهم شود.

۸- گزینه «۲» در ارزشیابی وابسته به ملاک یا ارزشیابی ملاکی، معیار یا ملاک ارزشیابی از پیش تعیین می شود و عملکرد یادگیرنده با توجه به آن ملاک قضاوت می شود. در این نوع ارزشیابی، از نوعی ملاک مطلق که همان هدف های آموزشی از پیش تعیین شده هستند، استفاده میشود. در این نوع ارزشیابی از آزمون شوندگان انتظار می رود که به اهداف معینی دست یابند یا به سطح خاصی از عملکرد برسند. با این تعاریف، اولین سؤال این است که منظور از تسلط چیست و تسلط در چه سطح و معیاری مورد نظر است.

۹- گزینه «۴» هدف ارزشیابی تکوینی، آگاه ساختن فراگیران از نواقص کارشان است؛ همچنین در رابطه با پیشرفت تحصیلی شاگردان، آن ها برای تعیین نقاط قوت و ضعف یادگیری شان از نتایج ارزشیابی تکوینی بهره می برند؛ مثلا در پایان هر واحد درسی، یک آزمون دقیق و مختصر که حاوی اهداف آموزشی آن واحد است، اجرا می شود و براساس نتایج، یادگیرندگان متوجه می شوند، به کدام یک از اهداف دست یافته اند (یا دست نیافته اند). نکته دیگر اینکه در ارزشیابی تکوینی، هدف، بررسی میزان توفیق یادگیرنده در رسیدن به اهداف درسی است و هیچ گونه مقایسه ای میان افراد مختلف و مدنظر نمی باشد. هدف از ارزشیابی پایانی، بررسی تمام آموخته های دانش آموزان در یک دوره آموزشی می باشد؛ ضمن آنکه به هدف نمره دادن به دانش آموزان و قضاوت درباره اثربخشی کار معلم نیز توجه می شود. به دلیل آنکه این ارزشیابی در پایان دوره صورت می گیرد، به آن ارزشیابی پایانی نیز می گویند (نام دیگر آن ارزشیابی تراکمی نیز هست).

هدف کلی ارزشیابی تکوینی، ارائه بازخورد و در ارزشیابی پایانی با تراکمی، ارائه نمره و نیز تعیین اثربخشی طرح آموزشی و میزان توفیق معلم می باشد.

۱۰- گزینه «1» آزمون پیشرفت تحصیلی که در پایان ترم به عمل می آید، نوعی آزمون پایانی یا تراکمی است که در آن نمره ارائه می شود. در اینجا انتظار می رود که اکثر دانشجویان، نمرات قبولی و مثبت را کسب کنند و تعداد اندکی نمره پایین بگیرند؛ بنابراین این آزمون دارای کنجی منفی است.

در آزمون پیشرفت تحصیلی، تمام آموخته های دانشجویان، اندازه گیری و نمره گذاری می شود.



۱۱- گزینه «۴» ارزشیابی هنجاری و آزمون های مبتنی بر ملاک نسبی در سنجش هوش و استعداد و نیز در سنجش شخصیت و سایر جنبه های عاطفی به کاره می روند. در این نوع ارزشیابی، عملکرد هر فرد با توجه به عملکرد سایر افراد و رتبه ای که هر فرد در مقایسه با سایر افراد کسب می کند، مورد قضاوت قرار می گیرد.

۱۲- گزینه «۴» آزمون های نرم - مرجع (Norm Referenced) آزمون هایی هستند که ملاک موفقیت در آزمون پس از انجام آزمون مشخص می شود. در این آزمون ها، نحوه ی عملکرد با گروه سنجیده می شود و نه با ملاکی که از قبل تعیین شده است. این ملاک موفقیت، از عملکرد کل افراد آزمودنی حاصل می شود؛ به طور مثال، میانگین و انحراف معیار آزمون و موقعیت آن ها در آزمون مشخص می شود. کنکور، نمونه روشنی از این نوع آزمون است. مسعود گزینه سه (ملاک - مرجع) می تواند به جواب اصلی سؤال نزدیک باشد، اما اگر به متن سؤال دقت کنیم، متوجه این تفاوت می شویم که در آزمون های ملاک - مرجع، ملاک ارزشیابی از قبل تعیین شده است و آزمودنی با توجه به آن ملاک مورد قضاوت قرار می گیرد؛ بنابراین تضاد میان این دو گزینه مشخص

میشود. در آزمون نرم - مرجع، ارزشیابی نسبی است، در حالی که در آزمون ملاک - مرجع، ارزشیابی به نوعی ملاک مطلق را در نظر می گیرد.

۱۳- گزینه «3» آنچه به منظور کمک به اصلاح موضوع مورد ارزیابی، یعنی برنامه با روش آموزشی استفاده می شود، ارزشیابی تکوینی نام دارد. هدف از کاربرد ارزشیابی تکوینی در رابطه با پیشرفت تحصیلی دانش آموزان و دانشجویان، آگاهی یافتن از میزان و در نحوه یادگیری آنان برای تعیین نقاط قوت و ضعف یادگیری و نیز تشخیص مشکلات روش آموزشی معلم در رابطه با هدف های آموزشی است. اسم این قدرت تشخیص به معلم امکان می دهد که سطح بندی خاص خود را در کلاس اجرا کند. ضمن آنکه شاگردان از نقاط قوت و ضعف هر یک از دروس(Formative evaluation) خود آگاه می شوند و در یک مقایسه با گروه همسالان، جایگاه فعلی خود را درک و برای رسیدن به جایگاه بالاتر تلاش بیشتری می کنند. گزینه (۴) به تعریف ارزشیابی تشخیصی اشاره دارد که می تواند داوطلب را در انتخاب گزینه صحیح دچار تردید کند.

۱۴- گزینه «۲» مهم ترین تفاوت آزمون های ملاکی و هنجاری در نحوه قضاوت در مورد نتایج است؛ زیرا در آزمون های ملاکی، دانش آموزان را با توجه به د ملاک خاص (نمره تعیین شده از قبل) ارزیابی می کنند، ولی در آزمون های هنجاری، دانش آموزان در مقایسه با یکدیگر ارزشیابی می شوند؛ در نتیجه مهم ترین تفاوت آزمون ملاکی و هنجاری مربوط به نحوه قضاوت در مورد نتایج آنها است.

۱۵- گزینه «۴» ارزشیابی های آموزشی با توجه به زمان و هدف استفاده از آن ها به چهار دسته ی ارزشیابی آغازین، ارزشیابی تکوینی، ارزشیابی تشخیصی و ارزشیابی تراکمی تقسیم می شوند. آنچه عمدتا برای کمک به اصلاح موضوع مورد ارزشیابی، یعنی برنامه یا روش آموزشی مورد استفاده قرار بر می گیرد، ارزشیابی تکوینی نام دارد. هدف از اجرای ارزشیابی تکوینی، آگاه ساختن تولید کنندگان برنامه از نواقص برنامه خود و کمک به اصلاح

آنهاست. هدف از کاربرد ارزشیابی تکوینی در رابطه با پیشرفت تحصیلی دانش آموزان و دانشجویان، آگاهی یافتن از میزان و نحوه یادگیری آنان برای تعیین نقاط قوت و ضعف یادگیری و نیز تشخیص مشکلات روش آموزشی خود توسط معلم است.

16\_گزینه «4» امتحان به مفهوم علمی، نوعی ارزشیابی است که براساس اصول معینی انجام می سود و در تعریف آن می توان گفت که امتحان عبارت است از تعیین میزان تغییرات حاصل در جنبه های خاصی از رفتار فراگیر، در فواصل زمانی معین و در جهت نیل به هدف های آموزشی و پرورشی. هدف های اساسی امتحان عبارت اند از:

۱) کمک به یادگیری فراگیر

۲) کمک به فراگیران برای آشنایی با هدف های آموزشی درس

۳) برانگیختن رغبت فراگیران به یادگیری

۴) تصمیم گیری برای شروع مراحل بعدی تدریس

۵) اصلاح روش تدریس

۶) ملاک ارتقای فراگیران و طبقه بندی آنان.

نکته: ایجاد رقابت میان فراگیران از اهداف امتحان محسوب نمی شود.

۱۷- گزینه «3» امتحان مرحله ای یا تکوینی در خاتمه تدریس هر فصل با بخش از مطالب، براساس محتوای تدریس شده در آن فصل، به منظور سنجش میزان پیشرفت تدریجی یادگیری دانش آموزان و اصلاح روش تدریس و آموزش به عمل می آید.

18\_گزینه «٣» امتحان ورودی یا تشخیصی در ابتدای هر دوره آموزشی، به منظور تعیین معلومات و رفتار ورودی و سنجش آمادگی دانش آموزان برای یادگیری درس جدید، براساس محتوای آموزش پیش نیاز آن درس به عمل می آید.

به طور کلی ۴ نوع امتحان از نظر هدف داریم:

1\_ امتحان ورودی یا تشخیصی: به منظور تعیین معلومات و رفتارهای ورودی، سنجش دانش آموزان برای یادگیری درس جدید؛

۲- امتحان مرحله ای یا تکوینی: در خاتمه درسی تدریس شده؛ به منظور سنجش پیشرفت تدریجی دانش آموزان و اصلاح روش تدریس معلم؛

۳- امتحان پایانی با تراکمی: در پایان دوره آموزشی از کل مطالب تدریس شده؛

۴- امتحان تحقیقی: مقایسه روش های تدریس، ارزشیابی از روش تدریس و...

۱۹\_ گزینه«3» أنواع امتحان از نظر هدف:

1\_ امتحان ورودی پا تشخیصی: به منظور تعیین معلومات و رفتارهای ورودی و سنجش دانش آموزان برای یادگیری درس جدید؛

۲۔ امتحان مرحله با تکوینی: در خاتمه درس تدریس شده، به منظور سنجش پیشرفت تدریجی دانش آموزان و اصلاح روش تدریس معلم؛

۳- امتحان پایانی با تراکمی: در پایان دوره آموزشی از کل مطالب تدریس شده؛

۴- امتحان تحقیقی: مقایسه روش های تدریس، ارزشیابی از روش تدریس و...

۲۰- گزینه «3» در الگوی سیپ، ارزشیابی از بافت یا شرایط، زمینه ای برای تعیین هدفهای آموزشی را فراهم می آورد. استافل بیم، اهداف الگوی ارزشیابی سیپ (Cipp) را موارد زیر می داند:

1\_ هدف الگوی ارزشیابی سیپ، اثبات کردن نیست؛ بلکه بیشتر برای رشد و پیشرفت برنامه ها است.

2\_ ایجاد و تهیه اطلاعات مفید برای شقوق مختلف تصمیمات.

٣- کمک به افراد برای قضاوت و بهبود ارزش چندگانه برنامه آموزشی با موضوع آموزشی

۴- کمک به رشد و بهبود سیاست ها و خط مشی برنامه ها الگوی سیپ از جمله الگوهای مبتنی بر مدیریت است که به منظور تسهیل در تصمیم گیری مدیران، ارائه و تدوین شده و الگویی کل نگر و جامع است که می تواند یک برنامه را به صورت سیستماتیک و همه جانبه بررسی کند. مدل Ci6p از چهار نوع ارزشیابی (زمینه، درون داد، فراگرد و برون داد)، دو نقش (تصمیم گیری و پاسخگویی) و سه مرحله (تعیین کردن، به دست آوردن و فراهم ساختن) تشکیل شده است.

آزمون فصل چهارم

1- اگر در یک آزمون ورودی، کسانی پذیرفته شوند که به ۷۵% سؤالات پاسخ درست بدهند، کدام نوع ارزشیابی به کار رفته است؟

1) تکوینی

۲) هنجاری

۳) تشخیصی

۴) ملاکی

۲- کدام یک از عبارات زیر جزو هدف های ارزشیابی تراکمی است؟

1) اصلاح فرایند تدریس و یادگیری

۲) تعیین نتیجه کار فراگیران در یک دوره آموزشی

3) گزینش فراگیران

۴) تعیین وضع فراگیران هنگام شروع یک برنامه

۳- معلمی هر هفته یک آزمون از مطالبی که تدریس کرده است، تهیه و اجرا می کند تا نقاط ضعف دانش آموزان را برطرف کنند. این ارزشیابی از کدام نوع است؟

1) تکوینی

۲) تراکمی

۳) تشخیصی

۴) تشویقی

۴- ارزشیابی تراکمی در کدام مرحله از آموزش انجام می شود؟

1) در آغاز تدریس

۲) در حین تدریس

۳) در آغاز و حین تدریس

۴) در پایان تدریس

5\_خود ارزشیابی معادل کدام اصطلاح زیر است؟

۱) ارزشیابی ملاکی

۲) ارزشیابی درونی

۳) ارزشیابی هنجاری

۴) ارزشیابی از فرایند

6\_ یک تفاوت مهم بین آزمون های هنجاری و ملاکی این است که:

1) در آزمون های ملاکی، یک منحنی با کجی مثبت وجود دارد، اما در آزمون های هنجاری یک منحنی با کجی منفی وجود دارد.

۲) در آزمون های ملاکی، یک منحنی بهنجار وجود دارد، اما در آزمون های هنجاری یک منحنی با کجی منفی وجود دارد.

ه) در آزمون های ملاکی، یک منحنی با کجی مثبت وجود دارد، اما در آزمون های هنجاری یک منحنی بهنجار وجود دارد.

۴) در آزمون های هنجاری، یک منحنی بهنجار وجود دارد، اما در آزمون های ملاکی یک منحنی با کجی منفی وجود دارد.

7- آزمون های مورد استفاده در سنجش تکوینی، ................................. هستند.

1) هنجاری

۲) ملاکی

۳) هنجاری و ملاکی

۴) هیچ کدام

۸- در ارزشیابی تشخیصی از کدام آزمون ها استفاده می شود؟

1) ملاکی

۲) هنجاری

۳) هنجاری و ملاکی

۴) هیچ کدام

9\_ ارزشیابی تشخیصی در چه زمانی اجرا می شود؟

1) در ابتدای آموزشی

۲) در ابتدا و انتهای آموزش

3) در انتهای آموزش

۴) در جریان آموزش

۱۰- ارزشیابی آغازین در چه زمانی ضرورت پیدا می کند؟

۱) زمانی که معلم از دانش ها و مهارت های قبلی دانش آموزان خود، اطلاع کافی نداشته باشد.

۲) زمانی که معلم با دانش آموزان آشنایی کامل دارد.

۳) ارزشیابی آغازین، همیشه ضرورت دارد.

۴) هیچ یک از موارد

فصل پنجم

«مراحل و فعالیت های ارزشیابی آموزشی»

مراحل ارزشیابی آموزشی

ارزشیابی آموزشی به تصمیم گیری برای منظور های مختلف کمک می کند. این تصمیم گیری ها در رابطه با هدف ها، راهبردها، روشها و ابزارهای اندازه گیری و مواردی از این قبیل هستند.گی (۱۹۹۱) مراحل ارزشیابی را به سه دسته تقسیم کرده است که هر مرحله با یک سؤال ویژه مرتبط می شود. این مراحل عبارتند از:

ا- مرحله ی طراحی: چه کاری انجام خواهیم داد؟ (زمان آینده)

۲- مرحله ی فرایندی: چگونه آن را انجام میدهیم؟ (زمان حال)

۳- مرحله ی فراوردهای: چگونه آن را انجام داددا یم؟ (زمان گذشته)

مرحله ی طراحی

این مرحله شامل تصمیم درباره فعالیت هایی است که قرار است انجام شود. در این مرحله اقدامات زیر صورت می گیرد:

تحلیل موقعیت: در این مرحله موقعیت موجود را تحلیل می کنیم. در این مرحله، تشخیص موانع و محدودیت ها و یک برآورد واقع بینانه از امکانات برای تعیین هدف های ارزشیابی، ضروری است. برای تحلیل موقعیت، معلم باید از آموخته های قبلی دانش آموزان، اطلاعات کافی کسب کند و امکانات موجود را به دقت برآورد کند. از جمله کارهایی که برای تحلیل موقعیت باید انجام دهیم، جمع آوری سوابق اطلاعاتی و تعیین موانع موجود است. از فواید بازنگری سوابق پژوهشی، این است که نشان خواهد داد که کدام رویکردهای ارزشیابی مفید و کدام غیر مفید می باشد.

تعیین و توصیف هدف ها: هدف های ارزشیابی به دو دسته کلی و جزبی تقسیم می شوند. به هدف های کلی معمولا غایت می گویند و به هدفهای جزیی هدف می گویند. غایت ها یا هدف های کلی، مبهم اند و به سادگی قابل اندازه گیری نیستند، ولی هدف های جزیی، دقیق و قابل تبدیل به کمیت و اندازه گیری هستند. باید غایت ها تحلیل شوند و به صورت هدف های دقیق درآیند.

هدف های کلی، خود به دو دسته فرایندی و فراورده ای با تولیدی تقسیم می شوند.

هدف های فرایندی به کوشش هایی گفته میشود که در ضمن اجرای برنامه انجام خواهند گرفت.

در مقابل، هدف های فرآورده ای به بازده و نتایج کوشش ها و اقدامات اشاره می کند؛ بنابراین چنین فرض می شود که تحقق یافتن هدف های فرایندی، یعنی تحقق یافتن هدف های فراورده ای.

باید توجه داشت که هدف ها به فعالیت های بعدی ارزشیابی، جهت می دهند و نهایتا در ارزشیابی میزان تحقق این هدف ها سنجش می شوند؛ پس چون هدف ها اساس فعالیت های بعدی ارزشیابی را تشکیل می دهند، باید آن ها را برحسب ربط داشتن، اندازه پذیر بودن و... مورد ارزشیابی قرار داد.

در ارزشیابی از برنامه ها و پروژه های آموزشی، هر دو نوع هدف به کار می روند، اما در ارزشیابی از پیشرفت تحصیلی، فقط هدف های فراوردهای مورد استفاده قرار می گیرند.

توصیف پیش نیازها یا رفتارهای ورودی: رفتار ورودی به آمادگی فرد برای یادگیری هدف های آموزشی اشاره می کند. رفتارهای ورودی همان رفتارها یا پیش نیازهایی هستند که یادگیرنده، پیش از شروع به یادگیری هدف های تازه، باید آنها را کسب کرده باشد. آموزش و ارزشیابی درست، مستلزم توصیف و اندازه گیری این پیش نیازهاست. برای تهیه و توصیف رفتارهای ورودی، معلم باید ابتدا به تهیه فهرست هدف های آموزشی اقدام کند، بعد معلوم نماید که یادگیری هر یک از هدف ها چه پیش نیازهایی لازم دارد. معمولا انجام چنین کاری به روش تحلیل تکلیف نیاز دارد.

انتخاب و تولید وسایل اندازه گیری: برای جمع آوری اطلاعات ضروری در ارزشیابی آموزشی، وسایل یا ابزارهایی مورد نیازند؛ مثل آزمون های پیشرفت تحصیلی، آزمون های نگرش یا فرم های ثبت مشاهده ای. اگر این وسایل موجود باشند باید به بررسی و انتخاب مناسب ترین آن ها اقدام کرد و اگر موجود نباشد، باید آنها را تهیه کرد. توصیف راهبردها یا استراتژی ها: منظور از استراتژی یا راهبرد، فعالیت با تدبیری است که به منظور کمک به تحقق یافتن هدف های برنامه به کار می رود. در مورد انواع آن، هم می توان از استراتژی های آموزشی نام برد و هم از استراتژی های برنامه درسی و غیره.

انتخاب طرح پژوهشی: طرح های پژوهشی بسیار متنوع اند و غالبا" به دو دسته طرح های آزمایشی و طرح های شبه آزمایشی تقسیم می شوند.

تدارک یک برنامه زمان بندی: تهیه و تدارک یک برنامه زمان بندی برای هر فعالیت جدی آدمی، از جمله ارزشیابی آموزشی، ضروری است. اساسا یک برنامه زمان بندی، در برگیرنده فهرستی از فعالیت های اصلی مربوط به اقدامات پیشنهادی و زمان شروع و خاتمه هر یک از فعالیت هاست.

مثال ۱: کدام مرحله از ارزشیابی، به آینده توجه دارد؟

۱) فرایندی

۲) طراحی

۳) فراوردهای

۴) گزینه های ۲ و ۳

پاسخ: گزینه «۲» مرحله طراحی از ارزشیابی، به آینده توجه دارد.

مثال ۲: در ارزشیابی از پیشرفت تحصیلی، کدام هدف ها مورد استفاده قرار می گیرند؟

۱) جزیی

۲) فرایندی

۳) فراورده ای

۴) همه موارد

پاسخ: گزینه «3» در ارزشیابی از پیشرفت تحصیلی، هدف های فراورده ای مورد استفاده قرار می گیرند.

مرحله فرایندی یا اجرایی

در این مرحله، تصمیمات مربوط به اقدامات عملی یا اجرایی گرفته می شوند. از نخستین اقدامات این مرحله می توان اجرای پیش آزمون و آزمون رفتار در ورودی را نام برد. براساس نتایج این آزمون ها مشخص می شود، هدف هایی که در مرحله قبل تعیین شده اند، مناسب بوده اند یا نیاز به تجدی نظر دارند. در هدف اساسی در این مرحله از ارزشیابی این است که معلوم شود آیا کوششها آن گونه که منظور بوده، به کار گرفته شده اند، تا میزان توفیق در هدر هدفهای فرایندی تعیین شود. گی معتقد است که مرحله فرایندی ارزشیابی همان چیزی است که ارزشیابی تکوینی نامیده می شود و هدف اصلی ارزشیابی تکوینی، کمک به اصلاح برنامه است.

مثال 3: اجرای پیش آزمون و آزمون رفتار ورودی در کدام مرحله ارزشیابی صورت می گیرد؟

۱) فرایندی

۲) طراحی

۳) فراوردهای

۴) گزینه های او ۲

پاسخ: گزینه «۱» در مرحله فرایندی، اجرای پیش آزمون و آزمون رفتار ورودی صورت می گیرد.

مرحله فرآورده ای

در مرحله فراورد های تصمیمات در پایان برنامه یا پروژه و در پایان دوره یا مقطعی از آن، یعنی زمانی که فعالیت ها کامل می شوند، اتخاذ می گردند. نام دیگر این مرحله از ارزشیابی، ارزشیابی تراکمی است. به طور کلی در مرحله فرآوردهای ارزشیابی، تصمیمات مربوط به اثربخشی موضوع مورد ارزشیابی، اتخاذ می شود؛ یعنی میزان اثربخشی برنامه تعیین شود. در این مرحله تعیین می شود که هدفهای از پیش تعیین شده تا چه حد تحققه یافته اند. نتایج این مرحله از ارزشیایی برای اشخاص داخل و خارج سیستم، بازخورد لازم را فراهم می آورند و مبنای تصمیمات مربوط به فعالیتها را در آینده تشکیل می دهند. پوفام (۱۹۷۵) می گوید: «نتایج این مرحله از ارزشیابی به دیگران کمک می کنند که تصمیم بگیرند، آیا برنامه آموزشی مورد ارزشیابی را ادامه دهند، تغییر دهند یا مورد تأکید قرار دهند.»

خلاصه تصمیم ها و فعالیت های سه مرحله اصلی ارزشیابی (اقتباس از گی1991)

|  |  |  |  |
| --- | --- | --- | --- |
| مرحله ارزشیابی | زمان اجرا | نوع تصمیم ها | فعالیت های ویژه |
| مرحله  طرح ریزی | پیش از اجرای  آموزش،  برنامه، پروژه و غیره | 1\_ نتایجی که قرار است به دست آیند.  ۲- اعمالی که باید انجام شوند. | 1\_تحلیل داده ها  2\_ تعیین و توصیغ هدف ها  3\_توصیف پیش نیازها  4\_انتخاب یا تهیه وسایل اندازه گیری  5\_توصیف استراتژی ها  6\_ انتخاب طرح پژوهشی  7\_ تدارک برنامه زمان بندی |
| مرحله فرایندی | ضمن اجرای آموزش،  برنامه ،پروژه و غیره | 1-به چه میزان مراحل طرح ریزی شده اجرا می شوند؟  ۲- چه تغییراتی باید داده شود تا پیشرفت لازم به دست آید؟ | 1\_ اجرای پیش از آزمون  ۲- اجرای آزمون رفتار ورودی  ٣- سنجش مناسب بودن هدفها  ۴۔ جمع آوری داده ها به طور دوره ای، مثل اجرای آزمون  ۵- تحلیل اثربخشی استراتژی ها |
| مرحله  فراورده ای یا تولیدی | پس از اجرا آموزش، برنامه، پروژه و غیره | 1\_ اثر بخشی کلی آموزش، پروژه ، برنامه و غیره  2\_ اقدامات آتی | 1\_ جمع آوری داده های مربوط به هدف ها، مثل اجرای پس از آزمون  2\_ جمع آوری داده های مربوط به بازده های پیش بینی نشده  3\_ تحلیل و تفسیر داده ها  4\_تهیه گزارش |

فعالیت های ارزشیابی آموزشی

۱- تعیین، انتخاب، پالایش یا تغییر غایت های برنامه و هدف های آموزشی؛

۲- مشخص کردن استانداردها یا هدف ها

٣- تهیه طرح ارزشیابی مناسب؛

۴- انتخاب یا تولید روش های جمع آوری داده ها؛

5\_ جمع آوری داده های مناسب؛

6\_ پردازش، خلاصه کردن و تحلیل داده ها؛

۷- مقایسه داده ها با استانداردها؛

8- گزارش و بازخورد دادن از نتایج؛

۹- هزینه، سود یا اثربخشی.

در ارزشیابی تراکمی، فعالیت های بالا به طور کامل و یکباره انجام می شود، اما در ارزشیابی تکوینی، این فعالیت ها به طور دوره ای و تکراری صورت می گیرد.

تست های طبقه بندی شده فصل پنجم

1\_ مهم ترین اقدام در ارزشیابی عبارت است از:

(آزاد 84)

1) راهنمایی صحیح فرد

۲) شناسایی کامل فرد

۳) اخذ تصمیم مناسب برای فرد

۴) مقایسه نمره های فرد

2- یک فهرست وارسی باید به عوض مقیاس درجه بندی استفاده شود. وقتی:

(آزاد 84)

۱) داوری ها بلی- خیر انجام می شود.

۲) یک فرآورده، ارزشیابی می شود.

۳) فقط مستلزم برداشت های کلی می باشد.

۴) فقط یک دانش آموز در یک زمان مشاهده می شود.

۳- برای تعیین و اندازه گیری حدود احتمالی موفقیت فرد در آینده، از کدام نوع آزمون استفاده می شود؟

(آزاد ۸۸)

۱) آزمون های هوش و استعداد

۲) آزمون رغبت

۳) آزمون پیشرفت تحصیلی

۴) آزمون غیر کلامی

پاسخنامه تست های طبقه بندی شده فصل پنجم

1- گزینه «3» همان طور که از اسم ارزشیابی پیداست، این اصطلاح به صورت ساده، به تعیین ارزش یا داوری ارزشی گفته میشود. تعریف جامع تری از ارزشیابی عبارت است از یک فرایند نظام دار برای جمع آوری، تحلیل و تفسیر اطلاعات به منظور تعیین این که آیا هدف های موردنظر تحقق یافته اند یا در حال تحقق اند و به چه میزانی. در این تعریف از ارزشیابی، ابتدا به جمع آوری اطلاعات از منابع گوناگون پرداخته می شود، سپس براساس آنها در مورد فرد مورد ارزیابی تصمیم گیری می شود.

٢- گزینه «۱» در روش فهرست وارسی، تنها تعیین می شود که یک ویژگی حضور دارید یانه و در روش مقیاس درجه بندی، میزان حضور یک رفتار، درجه بندی می شود؛ بنابراین تنها زمانی که داوری ها به صورت بلی - خیر صورت می گیرد، باید از روش فهرست وارسی استفاده کرد.

3\_گزینه «1» به طور کلی، آزمون های توانایی را (که آموخته ها، مهارت ها و استعدادهای افراد را می سنجند) می توان به دو دسته کلی آزمون های استعداد و آزمون های پیشرفت تحصیلی تقسیم کرد. آزمون های استعداد ناظر بر آینده اند و برای پیش بینی موفقیت های آتی فرد به کار می روند، در حالی که آزمون های پیشرفت تحصیلی ناظر بر گذشته اند و برای تعیین آموخته های قبلی فرد استفاده می شوند.

آزمون فصل پنجم

۱- تحلیل موقعیت از اقدامات کدام مرحله از مراحل ارزشیابی است؟

۱) فرآورده ای

۳) فرایندی

۳) طراحی

۴) گزینه های او ۳

۲- در ارزشیابی پیشرفت تحصیلی یادگیرندگان، غالبأ چه نوع هدف هایی استفاده می شود؟ .

۱) فرایندی

۲) فراورده ای

۳) هر دو

۴) هیج کدام

۳- در کدام مرحله ارزشیابی، تصمیمات مربوط به اقدامات عملی یا اجرایی گرفته می شوند؟

1) مرحله فرایندی

۲) مرحله طراحی

۳) مرحله فراورده ای

۴) هیچ کدام

۴- نام دیگر مرحله فراورده ای چیست؟

۱) ارزشیابی تکوینی

۲) ارزشیابی تشخیصی

۳) ارزشیابی تراکمی

۴) سنجش آغازین

5\_ متخصص ارزشیابی، زمانی به انتخاب یک طرح پژوهشی اقدام می کند که بخواهد:

1) رابطه بین دو یا چند متغیر را بررسی کند.

۲) تعداد کسانی را که به حد تسلط رسیده اند، مشخص کند.

۳) تعیین کند آیا برنامه مورد ارزشیابی موفق بوده است یا نه.

۴) گزینه های ۲ و ۳

6\_ در ارزشیابی از برنامه ها و پروژه های آموزشی، غالبا از چه هدف هایی استفاده می شود؟

۱) مرحله فرایندی

۲) فراورده ای

۳) گزینه ه ای ۱ و ۲

۴) هیچ کدام

۷- فعالیت یا تدبیری را که «به منظور کمک به تحقق یافتن هدف های برنامه به کار می رود، در اصطلاح ................................................ گویند.»

1) طراحی

۲) تحلیل موقعیت

۳)، تحلیل راهبردی

۴) تعیین هدف

8\_ هدف اصلی ارزشیابی تکوینی چیست؟

۱) تعیین هدف

۲) اصلاح برنامه

۳) متوقف کردن برنامه

۴) گزینه های ۱ و ۲

۹- نام دیگر مرحله فرایندی چیست؟

۱) ارزشیابی تکوینی

۲) ارزشیابی تشخیصی

۳) ارزشیابی تراکمی

۴) سنجش آغازین

10\_انجام تحلیل و تفسیر داده ها مربوط به کدام مرحله ارزشیابی می باشد؟

1) مرحله طرح ریزی

۲) مرحله فرآورده ای یا تولیدی

۳) مرحله فرایندی

۴) گزینه های او ۳

فصل ششم

«تهیه طرح ارزشیابی پیشرفت تحصیلی»

تعریف ارزشیابی پیشرفت تحصیلی

ارزشیابی پیشرفت تحصیلی به معنای سنجش عملکرد یادگیرندگان و مقایسه نتایج حاصل بـا هـدف های آموزشی از پیش تعیــن شـده است. تصمیم گیری درباره این که آیا فعالیت های آموزشی معلم و کوشش های یادگیری دانش آموزان به نتایج مطلوب رسیده است و به چه میزانی؟

برای این مسئله باید دو فعالیت صورت گیرد:

۱ـ مشخص کردن اهداف

۲ـ اندازه گیری عملکرد دانش آموزان براساس اهداف.

هدف های کلی که غایت نامیده می شوند، بیاناتی آرمانی هستند که منظور از ایجاد دوره های آموزشی را نشان می دهند. این هدف های کلی، معمولاً به وسیله برنامه ریزان در سطح وزارت تهیه و تدوین می شوند و به صورت مقاصد آرمانی دوره های آموزشی، در اختیار معلمان و سایر مربیان قرار می گیرند.

هدف های عینی و دقیق با استفاده از محتوای درس های مختلف توسط معلمان و متخصصان تهیه میشود. هدف های دقیق آموزشی، برخاسته از غایت های کلی هستند. در ارزشیابی آموزشی، عملکرد دانش آموزان با توجه به اهداف دقیق سنجیده میشود. به طور کلی هدف های آموزشی یا هدف های یادگیری، اهدافی هستند که معلم از آموزش خود و یادگیرنده از یادگیری خود دارد.

اهداف دقیق آموزشی یا اهداف رفتاری برحسب عملکرد دانش آموزان و قابل اندازه گیری اند. امتیاز مهم این گونه هدف های آموزشی آن است که به سهولت قابل شناسایی و اندازه گیری است. همچنین معلم به کمک هدف های رفتاری، منظور و مقصود خود را از آموزش موضوع های درسی به طور دقیق مشخص می کند و به دانش آموزان و دانشجویان تفهیم می کند که در پایان درس چه انتظاراتی از آنان خواهند داشت.

به عنوان مثال، دانش آموزان پس از مطالعه این فصل قادر خواهند بود:

۱ـ انواع باکتری ها را از نظر شکل ظاهری نام ببرند.

۲- راه های انتقال باکتری ها را به انسان توضیح دهند.

چنان که میبینیم، فعل های اهداف رفتاری بر بر حسب رفتار یادگیرنده بیان شده اند؛ اگر به جای این افعال از افعالی مانند فهمیدن، دانستن و درک کردن استفاده میشد، هدف های حاصل، هدف رفتاری نبودند؛ زیرا این گونه افعال، مبهم هستند و برای افراد مختلف، معانی متفاوتی دارند. متخصصـان آمـوزش و ارزشیابی توصیه می کنند که بهتر است هنگام نوشتن هدف های آموزشی رفتاری، علاوه بر کاربرد فعل هـا و عبارت های مشخص کننده رفتار، شرایط یا وضعیتی را که رفتار باید در آن انجام پذیرد و نیز ملاک یا معیار توفیق رفتار نیز ذکر شود. به این نوع هدف های رفتاری هدف های کامل رفتاری می گویند.

هدف های کامل رفتاری سه ویژگی دارند:

۱) برحسب رفتار اندازه پذیر دانش آموز هستند

۲) دارا بودن شرایط یا وضعیت عملکرد

3) دارا بودن ملاک عملکرد

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

هدف بالا هر سه ویژگی هدف رفتاری را دارد:

۱ـ انتخاب کند: این هدف براساس رفتار اندازه پذیر دانش آموز نوشته شده است.

۲- از میان فهرستی از ماشین ها و وسایلی که معلم در اختیار او می گذارد: این هدف دارای شرایط عملکرد است.

۳ـ حداقل پنج ماشین را: این هدف دارای ملاک عملکرد است. نوشتن هدف های رفتاری کامل، کار تهیه سؤال های امتحانی را به غایت آسان می کند؛ زیرا آن ها به معلم می گویند که در نوشتن هر سؤال چه نوع عملکردی را از دانش آموزان در نظر داشته باشد، چه شرایطی برای پاسخ دهی دانش آموزان فراهم آورد و با چه معیارها یا ملاک هایی، کار آنها را ارزیابی کند.

مثال ۱: ویژگی هدف های رفتاری عبارتند از:

۱) اندازه پذیر بودن

۲) دارا بودن شرایط عملکرد

۳) دارا بودن ملاک عملکرد

۴) هر سه مورد

پاسخ: گزینه «۴» اندازه پذیر بودن، دارا بودن شرایط عملکرد و دارا بودن ملاک عملکرد از ویژگی های هدف های رفتاری هستند.

طبقه بندی هدف های آموزشی

بلوم، هدف های آموزشی را به سه حیطه تقسیم کرد:

۱) حیطه شناختی

۲) حیطه روانی - حرکتی

3) حیطه عاطفی

حیطه شناختی

هدف های حوزه ی شناختی به جریان هایی که با فعالیت های ذهنی و فکری آدمی در ارتباط است، مربوط می شوند. این حیطه، مهـم تـریـن حـوزه یادگیری زیرا اکثر فعالیت ها و هدف های آموزشی به این حیطه مربوط میشوند. آزمون های مورد نیاز برای سنجش هـدف های حـوزه شناختی، آزمـون هـای توانایی های شناختی هستند که غالباً به صورت کتبی، تهیه و اجرا میشوند.

بلوم و همکاران، هدفهای حوزهی شناختی را به شش طبقه کلی تقسیم کرده اند. هر یک از این طبقه ها، یک نوع فرایند شناختی را توصیف می کنند. این شش طبقه به شکل سلسله مراتبی، از ساده به پیچیده و از عینی به انتزاعی مرتب شده اند این سطوح عبارت اند از:

۱ـ دانش: دانش شامل یادآوری (بازخوانی و بازشناسی) امور است. در واقع، دانش عبارت است از تکرار پاسخ هایی که قبلا در موقعیت یادگیری، تمرین و آموخته شده اند.

هدف های نمونه:

یادگیرنده بتواند انبساط و انقباض را تعریف کند.

یادگیرنده بتواند انواع سبک های ادبی را نام ببرد.

یادگیرنده بتواند سه ویژگی مهم میانگین را ذکر کند.

۲ـ فهمیدن: توانایی درک منظور و مقصود یک مطلب، مثلاً خلاصه کردن، درک شباهتها و تفاوتها، ترجمه، مثال گفتن، تفسیر و تعبیر کردن.

هدف های نمونه:

یادگیرنده بتواند حوزه شناختی را توضیح دهد.

یادگیرنده بتواند یک متن عربی را به فارسی جمه کند.

۳ـ کاربستن: توانایی استفاده از امور انتزاعی، قواعد و قوانین اصول اندیشه ها و روش ها در موقعیت های عینی و عملی.

هدف های نمونه:

یادگیرنده بتواند مسائل ریاضی را حل کند.

یادگیرنده بتواند برای خودش یک رشته ورزشی مناسب تعیین کند.

۴ـ تجزیه و تحلیل: توانایی شکستن یک مطلب به اجزای تشکیل دهنده آن.

هدف های نمونه:

یادگیرنده بتواند در یک مقاله، عقاید را از واقعیت ها متمایز کند.

یادگیرنده بتواند اجزای مختلف یک طرح تحقیق را مشخص کند.

5\_ ترکیب: توانایی کنار هم گذاشتن اجزاء و عناصر برای ایجاد یک الگو یا ساختار نو. ترکیب، نامی است که بلوم و همکاران او (۱۹۵۶) برای آفرینندگی و خلاقیت به کار برده اند.

هدف های نمونه:

یادگیرنده بتواند یک قطعه جدید صنعتی طراحی کند.

یادگیرنده بتواند یک فیلمنامه آموزشی بنویسد.

۶ـ ارزشیابی: توانایی قضاوت یا داوری کمی و کیفی درباره امور با توجه به ملاک های معین.

هدف های نمونه:

یادگیرنده بتواند یک مقاله فلسفی را نقد و بررسی کند.

یادگیرنده بتواند با استفاده از ملاک های معین، یک پایان نامه دانشجویی را داوری کند.

مثال ۲: اگر دانش آموزان بتوانند منظور گوینده یک داستان یا یک شعر را بیان کنند، به کدام یک از سطوح یادگیری در حیطه شناختی دست یافته اند؟

1) کاربرد

۲) ارزشیابی

۳) ترکیب

4) درک و فهم

پاسخ: گزینه «۴» بلوم و همکارانش معتقدند که هدف های حوزه ی شناختی به جریان هایی که با فعالیت های ذهنی و فکری آدمی سیروکار دارند، مربوط می شوند. این طبقه بندی شامل: ۱- دانش ۲- درک و فهم ۳- کاربستن ۴- تحلیل 5- ترکیبی 6-ارزشیابی است. اگر دانش آموزی بتواند منظور گوینده یک داستان یا یک شعر را بیان کند، در واقع او به توانایی تفسیر، ترجمه و خلاصه کردن دست یافته است و توانسته موضوع را درک. و آن را تفسیر و تعبیر کند. با توجه به صورت سؤال، گزینه شماره (۳) می تواند در ذهن داوطلب تردید ایجاد کند، اما باید توجه داشت که ترکیب، یعنی توانایی کنار هم گذاشتن عناصر و اجزاء برای ایجاد یک الگو و ساختار جدید؛ در حالی که برای عنوان کردن منظور نویسنده داستان یا شاعر، کودک باید مطلب را فهمیده باشد و لازم نیست که آن اجزاء را کنار هم قرار دهد و مطلب جدیدی بسازد.

مثال 3: کشف روابط بین اجزاء و طریقه سازمان بندی، جزء کدام یک از طبقات حیطه شناختی است؟

۱) دانش

۲) فهمیدن

۳) کاربستن

۴) تجزیه و تحلیل

پاسخ: گزینه «۴» کشف روابط بین اجزا و طریقه سازمان بندی جزء تجزیه و تحلیل حیطه شناختی است.

مثال 4- برای اندازه گیری برآیندهای تشخیص (مفروضه های بیان نشده، رابطه بین افکار و ... ) از کدام مقوله باید استفاده کرد؟

۱) ارزشیابی

۲) ادراک

۳) تجزیه و تحلیل

۴) کاربرد

پاسخ: گزینه «3» تمام گزینه ها به سطوح شناختی اشاره دارد. این طبقه بندی حیطه شناختی عبارت است از: ا- دانش ۲- درک و فهم و کاربستن ۴- تحلیل ها ترکیب ۶ ارزشیابی.

زمانی که به تشخیص موضوعی اشاره می شود، در حیطه ی شناختی بلوم به تجزیه و تحلیل مربوط میشود.

مثال ۵: اگر یک دانشجوی رشته علوم تربیتی پس از پایان تحصیلات، به عنوان یک معلم، امتحانات را به خوبی برگزار کند و نتایج آن را به نحو مطلوب ارزشیابی کنید، به کدام طبقه از حیطه شناختی دست یافته است؟

1) ارزشیابی

۲). تجزیه و تحلیل

۳) ترکیب

۴) کاربستن (کاربرد)

پاسخ: گزینه «۴» با توجه به سؤال، می توان به راحتی به این نکته پی برد که این دانشجو پس از فارغ التحصیلی، تمام مطالبی را که یاد گرفته است به کار می بندد؛ بنابراین حتی اگر طبقه بندی شناختی را نیز در حافظه نیاوریم، می توانیم گزینه ی صحیح را انتخاب کنیم. طبقه بندی عبارت است از:

۱) کسب دانش (در دوره ی دانشجویی)

۲) درک و فهم (در پایان تحصیلات)

3) کاربستن (برگزاری درست امتحان و ارزشیابی صحیح) باتوجه به این که در گزینه های سؤال به کلمه ارزشیابی هم، اشاره شده است و نیز در طبقه بندی داریم

۴) تجزیه و تحلیل

5) ترکیب

۶) ارزشیابی

بنابراین داوطلب باید ترتیب را در نظر بگیرد و به همسانی سؤالی و گزینه ی یک توجه کند.

مثال ۶: می خواهیم در مورد یک بیت شعر قضاوت کنیم. در کدام مورد، ملاک قضاوت ما خارجی خواهد بود؟

1) قابل فهم بودن برای اکثریت

۲) کوتاه و روان بودن

۳) دارا بودن مفاهیم عمیق

۴) تطابق شعر با اوزان شعری

پاسخ: گزینه «۴» ارزشیابی، بالاترین مورد در حیطه شناختی است که درباره تطابق این موضوع با متن سؤال، زمانی که تطابق شعر با اوزان شعری ملاک قضاوت خارجی قرار می گیرد، مصداق پیدا می کند.

مثاال 7- قضاوت هوشیارانه درباره کارها و روش ها به کدام یک از حیطه های زیر تعلق دارد؟

۱) انفعالی

۲) روانی - حرکتی

۳) شناختی

۴) عاطفی تا

پاسخ: گزینه ۳» هدف های آموزشی در سه حوزه قرار می گیرند که عبارتند از: حوزه شناختی، حوزه عاطفی و حوزه روانی - حرکتی این حوزهها در گروه های خود، زیرمجموعه های طبقه بندی شده دارند. در حوزه شناختی، اهداف به جریان هایی که با فعالیتهای ذهنی و فکری آدمی سروکار دارند، مربوط می شوند. داوطلب باید به این حوزه ها تسلط داشته باشد و طبقه بندی حوزه شناختی را که شامل دانش، فهمیدن، کاربستن، تحلیل، ترکیب و ارزشیابی است، بشناسد. بنابر متن سؤال، نیاز به قضاوت هشیارانه در این طبقه بندی به قسمت ارزشیابی اشاره دارد که به داوری کمی و کیفی درباره امور می پردازد.

مثال 8- از داوری آگاهانه راجع به اندیشه ها، کارها، راه حل ها و روش ها جزء کدام یک از طبقات حیطه شناختی است؟

۱) ترکیب

۲) ارزشیابی

۳) فهمیدن

۴) کاربستن

پاسخ: گزینه «۲» ارزشیابی، داوری آگاهانه راجع به اندیشه ها، کارها، راه حل ها و روش ها است.

مثال ۹- اگر دانش آموزی بتواند منظور یک گوینده داستان را بیان کند، به کدام یک از سطوح یادگیری در حیطه شناختی دست یافته است؟

1) فهمیدن

۲) کاربرد

۳) ترکیب

۴) ارزشیابی

پاسخ: گزینه «۱» چون دانش آموز توانسته است منظور گوینده را بفهمد و درک کند، در مرحله فهمیدن از حیطه شناختی قرار دارد.

مثال ۱۰: فهم معنای ظاهری یا تحت اللفظی پیام ها در کدام یک از طبقات حیطه شناختی قرار دارد؟

1) فهمیدن

۲) تجزیه و تحلیل

۳) دانش

4) کاربستن

پاسخ: گزینه «3» فهم معنای ظاهری یا تحت اللفظی پیام ها، جز «دانش» از حیطه شناختی است.

مثال ۱۱: اگر یادگیرنده بتواند به تکرار پاسخ هایی که قبلا در موقعیت یادگیری و تمرین آموخته است، بپردازد مربوط به کدام مرحله حیطه شناختی است؟

1) فهمیدن

۲) ترکیب

۳) دانش

۴) ارزشیابی

پاسخ: گزینه «۲» تکرار پاسخ هایی که یادگیرنده قبلأ در موقعیت های یادگیری و تمرین آموخته است، جز «دانش» از حیطه شناختی است.

مثال ۱۲: اگر یادگیرنده بتواند به خلق یک داستان بپردازد، مربوط به کدام حیطه شناختی می باشد؟

۱) دانش

۲) ترکیب

۳) فهمیدن

۴) ارزشیابی

پاسخ: گزینه «۲» خلاقیت یا خلق یک چیز، جز «ترکیب» از حیطه شناختی می باشد.

مثال ۱۳: در طبقه بندی اهداف آموزش «بلوم»، کدام حیطه شناختی از سطح عالی تری برخوردار است؟

1) تجزیه و تحلیل

۲) دانش

۳) درک و فهم

۴) کاربرد

پاسخ: گزینه «۱» مراحل طبقه بندی حوزه شناختی از نظر بلوم به ترتیب عبارتند از:

۱) دانش ۲) درک و فهم ۳) کاربستن ۴) تجزیه و تحلیل ۵) ترکیب 6) ارزشیابی

بنابراین تجزیه و تحلیل که جزء گزینه های سؤال می باشد صحیح است.

اخیرا اندرسون کراتول و همکاران، در طبقه بندی حوزه شناختی بلوم تجدیدنظر کرده اند و طبقه بندی جدیدی را با نام یک طبقه بندی برای یادگیری آموزش و سنجش، به وجود آوردند. در طبقه بندی تازه حوزه شناختی، یک بعد دانش و یک بعد فرایند شناختی وجود دارد. بعد دانش شامل دانش امور واقعی، دانش مفهومی، دانش روندی و دانش فراشناختی است.|

بعد فرایند شناختی در برگیرنده به یاد آوردن، فهمیدن، کار بستن، تحلیل کردن، ارزشیابی کردن و آفریدن است. هر دو بعد، به صورت سلسله مراتبی، یعنی م از عینی به انتزاعی و از ساده به پیچیده تنظیم یافته اند. طبقه بندی تجدیدنظر شده ی حوزه شناختی را می توان برای مقاصد مختلفی استفاده کرد؛ از جمله آنها می توان این موارد را نام برد: کمک به معلمان در انتخاب فعالیت های آموزشی، تدوین برنامه های درسی و تحلیل هدف های یادگیری.

حوزه عاطفی

حوزه عاطفی شامل هدفهایی است که تغییرات حاصل در علایق، نگرش ها، ارزش ها و سازگاری را نشان میدهد. هرچند که معلمان در ضمن آموزش هدف های حوزه شناختی چه مستقیم و چه غیرمستقیم در آموزش هدف های حوزه عاطفی نیز می کوشند، اما کاربرد حوزه عاطفی در سطح آموزشگاه ها به اندازه کاربرد حوزه شناختی شهرت نیافته است.

کراتول، بلوم و ماسیا هدف های حوزه عاطفی را به ۵ طبقه تقسیم کرده اند. طبقه بندی حوزه عاطفی نیز به صورت سلسله مراتبی و از ساده به پیچیده است. این طبقه بندی شامل پنج طبقه اصلی به شرح زیر است:

۱) دریافت کردن (توجه کردن): حساسیت نسبت به محرکهای محیطی و میل به دریافت کردن با توجه کردن به آنها.

هدف نمونه: یادگیرنده به سبک های مختلف معماری توجه می کند.

۲) پاسخ دادن، توجه فعالانه به پدیده ها و انجام کارهایی درباره آنها یا فعالیت حاکی از پذیرش یک عقیده یا خط مشی.

هدف نمونه: یادگیرنده به طور داوطلبانه، مجلات مربوط به معماری را مطالعه می کند.

۳) ارزش گذاردن: احساس یا باور نسبت به ارزشمند بودن یک چیز یا ابراز عقیده یا نگرش درباره ی ارزش یک چیز.

هدف نمونه: شرکت فعال در به نمایش گذاشتن آثار هنری معاصر.

۴) سازمان دادن به ارزش ها: دسته بنبری و نظام دار کردن ارزش ها به گونه ای که بعضی از آنها مهم تر از بعضی دیگر جلوه کنند.

هدف نمونه: یادگیرنده به سیاست ها و رفتارهای اجتماعی مختلف در قبال رفاه عمومی، بیشتر از منافع محدود گروه ها بها می دهد.

۵) شخصیت پذیرفتن: تبدیل نظام ارزشی به صورت یک سبک زندگی یا یک دیدگاه فلسفی، این کار از درهم آمیزی ارزش ها و تعیین روابط بیان آن ها فراتر می رود و نوعی فلسفه پایدار زندگی را شامل می شود.

هدف نمونه: یادگیرنده در موقعیت های مختلف، متناسب با نظام ارزشی خود رفتار می کند.

مثال ۱۴: تغییرات ایجاد شده در نگرش و علاقه فرد نسبت به یک رشته مربوط به کدام حوزه رفتاری می باشد؟ ۱) شناختی

۲) عاطفی

۳) روانی - حرکتی

۴ گزینه های ۱ و ۳

پاسخ: گزینه «۲» تغییرات ایجاد شده در نگرش و علاقه فرد نسبت به یک رشته، مربوط به حوزه عاطفی می باشد.

مثال 15: اگر فرد ارزش های کسب شده را در رفتار خود نشان دهد، مربوط به کدام حیطه حوزه عاطفی می باشد؟

1)شخصیت پذیرفتن

۲) ارزش گذاری

3) دریافت

۴) سازمان دادن

پاسخ: گزینه «۱» در شخصیت پذیرفتن» از حیطه حوزه عاطفی، فرد، ارزش های کسب شده را در رفتار خود نشان می دهد.

حوزه روانی - حرکتی

حوزه روانی- حرکتی به زمینه مهارت های حرکتی با حرکات بدنی مربوط است؛ به عبارت دیگر، هر فعالیتی که علاوه بر جنبه روانی، جنبه جسمانی هم داشته باشد، در این حوزه جای می گیرد. تفاوت عمده هدف های این حوزه با هدف های حوزه شناختی در این است که هدف های حوزه شناختی صرفا به فعالیت های ذهنی ختم می شوند، اما هدف های حوزه روانی - حرکتی، هم به فعالیت های ذهنی و عاطفی و هم به فعالیت های جسمی نیاز دارند. برخلاف حوزه شناختی، هدف های حوزه عاطفی و حوزه روانی - حرکتی با آزمون های متداول به خوبی قابل سنجش نیستند. سنجش هدف های حوزه عاطفی و حوزه روانی - حرکتی با وسایلی غیر از آزمون های مورد استفاده در حوزه شناختی امکان پذیر است.

سیمپسون حوزه روانی - حرکتی را به هفت طبقه از ساده به پیچیده، تقسیم کرده است که عبارتند از:

1- ادراک حسی استفاده از حواس برای هدایت کنش های حرکتی

هدف نمونه: یادگیرنده با چشیدن غذاء مزه آن را تشخیص می دهد.

۲- آمادگی یا آمایه: آماده بودن برای انجام یک عمل یا یک رشته اعمال معین

هدف نمونه: یادگیرنده آماده است به توپی که به سوی او پرتاب می شود، ضربه بزند.

۳- پاسخ هدایت شده: عمل کردن به کمک هدایت یک الگو یا اعمالی که در مراحل اولیه یادگیری انجام می شوند.

هدف نمونه: یادگیرنده، قلم خطاطی را مطابق با مربی حرکت می دهد.

۴- مکانیسم یا عادت: انجام اعمالی که نسبتا مشخص، عادتی و خودکار هستند.

هدف نمونه: یادگیرنده بتواند اولین سرو را در منطقه سرو، در ۷۰ درصد از موارد بگیرد.

5- پاسخ پیچیده آشکار: انجام رشته اعمالی که تا حدودی به صورت خودکار یا عادتی در آمده اند یا ترکیب تعدادی از عادت ها.

هدف نمونه: یادگیرنده بتواند در طول خیابان رانندگی کند.

۶- انطباق: استفاده از مهارت های قبلا آموخته شده برای انجام تکالیف تازه، اما مربوط به گذشته.

هدف نمونه؛ یادگیرنده نوشته های دارای درجات مختلف دشواری را با سرعت های متفاوت تایپ می کند.

۷- ابتکار: خلق الگوهای تازه حرکتی برای حل مسائل غیر معمول با برخورد با موقعیت های جدید.

هدف نمونه: یادگیرنده بتواند یک قطعه آهنگ و موسیقی بسازد و اجرا کند.

جدول مشخصات

مهمترین گام در تهیه آزمون های پیشرفت تحصیلی، تهیه جدول مشخصات است. جدول مشخصات، یک جدول دو بعدی است که در ردیف بالای آن، بعد محتوا و در ستون کنار آن بعد هدف قرار دارد.

آزمونی روایی دارد که سؤال های آن حاوی نمونه درستی از همه شما و محتوای درس باشد جدول مشخصات به معلم کمک کند با از مجموع بزرگی از سؤالات ممکن، نمونه ای را که به خوبی معرف محتوا و هدف های درس باشد، یک آزمون خوب پیشرفت تحصیلی، آزمونی است که به بهترین شکل، منعکس کننده تمامی هدف های آموزشی و محتوای برنامه درسی باشد. از لحاظ نظری بهترین آزمون برای یک درس یا یک موضوع، آزمونی است که تمام مطالب و کلیه هدف های آموزشی آن درس را شامل می شود.

مثال ۱۶: اولین گام در تدوین آزمون های پیشرفت تحصیلی چیست؟

۱) تعیین هدف های آموزشی

2) انتخاب نوع سوال آزمون

۳) طرح سؤالات اولیه و اجرای آزمایشی

4) بارم بندی سوالات در جهت عینیت بخشی

پاسخ: گزینه «۱» ارزشیابی، یک فرایند نظامدار برای جمع آوری، تحلیل و تفسیر اطلاعات است؛ به این منظور که تعیین می شود، آیا هدف هـای مـورد نظر محقق شده اند و به چه میزانی؟ ارزشیابی پیشرفت تحصیلی را می توان به صورت سنجش عملکرد یادگیرندگان و مقایسه ی نتایج حاصـل بـا هـدف هـای آموزشی از قبل تعیین شده تعریف کرد که این کار به منظور تصمیم گیری درباره نتایج فعالیت های آموزشی معلم و کوشش های یادگیرندگان، صـورت ر می گیرد؛ بنابراین در ارزشیابی تحصیلی به این نکات توجه ویژه ای میشود: ۱- تعیین هدفهای آموزشی ۲- سنجش یا اندازه گیری عملکرد یادگیرندگان.

مثال ۱۷: آزمونی که سؤالات آن حاوی نمونه درستی از هدف ها و محتوای درس باشد، در اصطلاح می گویند دارای .......... میباشد.

۱) پایایی

۲) روایی

۳) گزینه های ۱ و ۲

۴) نمونه درست

پاسخ: گزینه «۲» آزمونی که سؤالات آن حاوی نمونه درستی از هدف ها و محتوای درس باشد، دارای روایی است؛ زیرا روایی، مربوط بودن سؤالات آزمون با هدف آزمون را بررسی می کند.

مثال ۱۸: هنگامی که در تهیه سؤالات آزمون، اهمیت نسبی هر یک از هدف های رفتاری را تعیین می کنیم (جدول مشخصات را تهیه می کنیم) ایـن کار موجب میشود تا ........

۱) روایی آزمون افزایش یابد.

۲) پایایی آزمون افزایش یابد.

۳) محاسبه ضریب دشواری امکان پذیر شود.

۴) محاسبه ضریب تمییز، به سادگی میسر شود.

پاسخ: گزینه «۱» مهم ترین گام در تهیه ی آزمون های پیشرفت تحصیلی، تهیه ی جدول مشخصات برای موضوعی است که قرار است از آن آزمون تهیه شود. در این جدول، محتوای یک درس و همه ی هدف های یادگیری، لحاظ می شود؛ در صورتی همه ی محتوا و اهداف مورد نظر در آزمـون مـدنظر قـرار گیرد ، این آزمون روایی بالا خواهد داشت.

گزینه ی دو می تواند ذهن داوطلب را در انتخاب درست، دچار تشویش و انتخاب صحیح را به انتخاب درست تبدیل کند، امـا اگـر تفـاوت روایی و پایایی را بدانند، اشتباه رخ نمی دهد.

مثال ۱۹: یکی از فواید استفاده از جدول مشخصات در تهیه آزمون ها آن است که:

۱) عینیت آزمون را افزایش میده

۲) تهیه آزمون را تسهیل و تسریع می کند.

۳) روش تصحیح، نمره گذاری و تفسیر نتایج آزمون را مشخص می کند.

۴) نمونه بهتری از محتوا و هدف های آموزشی به دست می دهد.

پاسخ: گزینه «۴» برای افزایش روایی در آزمون ها از جدول مشخصات استفاده . کنند. جدول مشخصات، هم در تهیه آزمون های وابسته به ملاک و هم در تهیه آزمون های وابسته به هنجار استفاده میشود. در این جدول اهداف آموزشی معلم و تمام بازده های یادگیری افراد لحاظ می شود .

گزینه ۴، دربرگیرنده تمامی گزینه های دیگر است. داوطلب باید دقت کند که گزینه سه را به جای گزینه ی صحیح انتخاب نکند.

چگونگی توزیع سؤالات به بخش های مختلف

برای اینکه بفهمیم و در بخش چند سؤال باید اختصاص یابد، به ۳ نکته باید توجه کنیم:

۱ـ به هربخش باتوجه به حجم آن چه مقدار ساعت اختصاص داده شده است.

۲- کدام یک از اهداف، از نظر نگهداری و کاربرد، در آینده ارزش بیشتری دارند.

۳- اهمیت کدام بخش بیشتر است.

برای محاسبه نسبتا ساعاتی که صرف تدریس هر بخش شده است و تعداد سؤالاتی که از هر بخش باید تهیه شود، از فرمول های زیر استفاده می کنیم.

تعداد ساعت هایی که صرف تدریس بخش شده

تعداد کل ساعت های تدریس

تعداد کل سؤالات نسبت ساعاتی که صرف تدریس هر بخش شده = تعداد سؤالات هر بخش

مثال۲۰: اگر کتاب علوم پنجم دبستان ۱۰ بخش داشته باشد و ۷۷ ساعت تعداد کل ساعت های تدریس باشد و اگر معلم ۱۱ ساعت صرف تدریس فصل اول کرده باشد (با فرض اینکه همه فصول اهمیت یکسان دارند)، اگر معلم بخواهد آزمونی تهیه کند که دارای ۵ سوال است، چند سؤال باید از فصل اول طرح گردد؟

1) 12

2) 9

3) 7

4) 11

پاسخ: گزینه «3»

7 14/7 50 142/0 تعداد سؤال هر بخش

142/ 0 نسبت ساعاتی که صرف هر بخش شد

مثال ۲۱: بر چه اساس، تعداد کل پرسش های یک آزمون را تعیین می کنند؟

1) اهمیت نسبی مطالب

۲) تعداد صفحات با حجم کتاب ام

۳) زمان مصرف شده جهت آموزش مطالب

۴) عینی یا انشایی بودن پرسش ها

پاسخ گزینه «۴» به راحتی از متن سؤال می توان فهمید که تعداد صفحات یا حجم کتابها در تعیین مقدار پرسش، دخالت مستقیم دارد. هر چند سایر عوامل نیز در گزینه های دیگر آورده شده اند، اما برای انتخاب درست باید به نکته اصلی اشاره کرد؛ بنابراین تعداد انتخاب سؤال، به تعداد صفحات و حجم کتاب، اهمیت نسبی مطالب، تعداد صفحات و زمان صرف شده جهت آموزش بستگی دارد؛ باید به این نکته توجه کرد که در آزمون های انشایی، سؤالات، تعداد کمتری دارند و در آزمون های چند گزینه ای تا صحیح - غلط و جور کردنی (آزمون های عینی) تعداد سؤالات بیشتر هستند.

جیمز پوفام برای تهیه طرح آزمون های وابسته به ملاک چنین نظر میدهد که بهتر است برای هر دسته از سؤال ها، که طبقه واحدی. از رفتارها یا عملکرد یادگیرندگان را می سنجد، یک مشخصات آزمون پونام جداگانه تهیه شود که دارای بخش های زیر است:

1- توصیف کلی: توصیف اعمالی که دانش آموز باید انجام دهد تا کاری را که از او خواسته شده است، کامل کند.

۲- سؤال نمونه: یک سؤال نوعی که ویژگی های سؤال های آزمون را که تشکیل دهنده حوزه با حیطه مورد اندازه گیری خواهند بود، مشخص می سازد.

۳- ویژگی محرک: قاعده یا منشور تهیه سؤال های آزمون از طریق توصیف ویژگی سؤال ها به گونه ای که سؤال ها موازی یکدیگر باشند.

4- ویژگی های پاسخ: الف) مشخص کردن طبقه پاسخ هایی که آزمون شونده از میان آن دست به انتخاب خواهد زد

ب) مشخص کردن معیارهایی که پاسخ های تهیه شده توسط آزمون شونده به وسیله آن ها مورد قضاوت قرار خواهند گرفت.

5- توضیحات اضافی: این بخش پیوست، معمولا شامل فهرست با توضیحات مربوط به محتوای مورد سنجش است.

گرانلاند (۱۹۸۸) درباره طرح پوقام می گوید: اگر این طرح به خوبی تهیه شود، می تواند ماهیت عملکردهای مورد اندازه گیری دانش آموزان و ویژگی های دقیق سؤال های آزمونی را که قرار است تهیه شود، به طور کامل مشخص کند.»

آزمون های پیشرفت تحصیلی متداول، که به آزمون های کتبی با آزمون های مداد و کاغذی معروف اند، به طور عمده با حوزه شناختی سروکار دارند،

اما هدف های آموزشی مربوط به حوزه عاطفی و روانی - حرکتی به کمک وسایل دیگر سنجش از جمله مشاهده، مقیاس درجه بندی، فهرست وارسی، واقعه نگاری، ثبت حضور و غیاب، گزارش های کار عملی و مانند این ها سنجش می شوند.

نکته: هم در تهیه آزمون های وابسته به ما و هم در تهیه آزمون های وابسته به هنجار تهیه جدول مشخصات ضروری است. با این حال، تفاوت عمده بین این دو نوع جدول، مشخصات آن است که در بعد هدف جدول مشخصات آزمون های وابسته به هنجار، معمولا به جای هدف های دقیق آموزشی، اصطلاحات و عبارت های کلی تری به کار می روند که معرف طبقات هدف های آموزشی یا انواع یادگیری هستند.

مثال ۲۲: کدام نوع آزمون برای سنجش حیطه شناختی بکار برده می شود؟

1) مشاهده

۲) مقیاس درجه بندی

۳) واقع نگاری

۴) کتبی (امداد-کاغذی)

پاسخ: گزینه «۴» از آزمون های کتبی (مداد - کاغذی) برای سنجش حیطه شناختی استفاده می کنند.

تست های طبقه بندی شده فصل ششم

1- وقتی دانشجویی با شرکت در چند جلسه آموزش کامپیوتر، به این باور می رسد که کار با کامپیوتر ارزشمند است، طبق طبقه بندی هدف های آموزشی بلوم، این باور در کدام حوزه قرار می گیرد؟

(سراسری ۸۱)

۱) عاطفی

۲) شناختی

۳) روانی - حرکتی

4) عاطفی و شناختی

۲- کدام اقدام، قبل از طراحی سؤال های آزمون پیشرفت، باید مورد توجه آزمون ساز قرار گیرد؟

(سراسری ۸۱)

۱) طراحی جدول مشخصات

۲) تصمیم درباره نوع ارزشیابی

3) مشخص نمودن تعداد سؤالها

۴) تعیین نوع سؤال های آزمون

۳- توانایی در ارائه یک طرح تحقیقی، جزو کدام یک از طبقات حوزه شناخت است؟

(سراسری ۸۲)

۱) کار بستن

۲) تحلیل

۳) ترکیب

۴) ارزشیابی

۴- ترجمه یک متن از زبان عربی به فارسی، مربوط به کدام یک از هدف های آموزشی زیر است؟

(آزاد ۸۲)

1) دانش

۲) درک و فهم

۳) تجزیه و تحلیل

۴) ترکیب

5- گام اول در تهیه یک آزمون پیشرفت تحصیلی کدام است؟

۱) نوشتن سؤالات آزمون

۲) تهیه جدول مشخصات

3) تعریف هدف های درس

۴) تعیین تعداد سؤالات

۶- هریک از سؤالات آزمون معلم ساخته باید

(آزاد ۸۲)

۱)دانستنی های مهم را ارزشیابی کند.

۲) مشکل ترین بخش درس را ارزشیابی کند.. .

۳) جزئیات را اندازه گیری کند.

۴) هوش را اندازه گیری کند.

7- شعر زیر را به نثر ساده برگردانید.

« خاطری چند اگر از تو شود شاد بس است زندگانی به مراد همه کس نتوان کرد» .

سؤال بالا در چه سطحی قرار دارد؟

(سراسری ۸۳)

1)فهمیدن

2) کاربستن

۳) تحلیل

4) ترکیب

۸- وجود جدول طرح آزمون موجب می شود که .............

(سراسری ۸۲)

۱) سؤالات، تنها سطوح بالای هدف ها را اندازه بگیرند.

2) تعداد سؤالات با اهمیت هدفها و محتوا متناسب باشد.

۳) به هدف های کلی، سؤالات تشریحی اختصاص داده شود.

۴) از فصول با حجم بیشتر، سؤالات بیشتری طرح شود.

۹- آزمونی که برای اندازه گیری پیشرفت تحصیلی تهیه می شود، کدام دو عنصر را در نظر می گیرد؟

(سراسری ۸۳)

۱) هدف های پرورشی و محتوای درس

۲) روایی و پایایی آزمون

3) دشواری و قدرت تمییز سؤالات

۴) گذشته و حال دانش آموزان

۱۰- جدول بودجه بندی سؤالی به خاطر کدام یک از اهداف زیر در ساختن آزمون های پیشرفت تحصیلی تدوین می شود؟

آزاد ۸۳)

۱) اطمینان از دستیابی به هدف های آموزشی

۲) تخصیص سؤال های آزمون براساس فصول با مواد آموزشی

3) تنظیم سؤال ها بر اساس اصول آزمون سازی

,۴) سهولت در تعیین نوع سؤال ها و گزینه های آن

۱۱- کدام یک از فرایندهای ذهنی در قلمرو « درک و فهم» دانش آموز قرار می گیرد؟

(سراسری ۸۴)

1) تسلط بر دانش اصول و قواعد کلی

۲) توانایی در تجزیه و تحلیل یک مطلب

۳) قدرت ابتکار و نوآوری مطلب جدید

۴) توانایی در تفسیر نکات مهم یک مطلب

۱۲- به منظور کار عملی، از دانشجویان خواسته شده است که یکی از مسائل آموزش و پرورش را بررسی کنند و راه حل های پیشنهادی خود را ارائه دهند. این تکلیف، کدام سطح یادگیری را ارزیابی می کند؟

(سراسری ۸۴)

۱) کاربرد

2) تجزیه

۳) ترکیب

۴) ارزشیابی

۱۳- در هنگام آموزش یک هدف عاطفی، معلم از دانش آموزان می خواهد که توضیحات بیشتری ارائه دهند. این دانش آموزان، در کدام سطح از یادگیری عاطفی قرار دارند؟

(سراسری ۸۵)

1) واکنش

۲) ارزش گذاری

۳) سازماندهی ارزش ها

۴) توجه (دریافت)

۱۴- بر اساس روش رفتارگرایان، آزمون های پیشرفت تحصیلی بر مبنای کدام یک از موارد زیر تنظیم می شوند؟

(آزاد 85)

۲) تلاش برای یادگیری مطلب

۲) هدف های مربوط به مهارت های پایان دوره ,

۳) میزان مهارت دانش آموز نسبت به دانش آموزان دیگر

۴) توانایی انطباق مطالب درسی با دنیای واقعی

۱۵- کدام یک از موارد زیر در تعیین تعداد کل سؤالات امتحانی مهم است؟

(آزاد 86)

1) زمان آزمون، سن و تعداد شاگردان

۲) نوع سؤالات و مدت زمان تدریس

۳) نوع سؤالات، سن و پایه تحصیلی

۴) سن، پایه تحصیلی و جنسیت دانش آموز

۱۶- درصد سؤالات هر فصل از مواد تدریس شده، نسبت به کل سؤال های یک آزمون، تابع کدام عامل زیر است؟

(آزاد 86)

1) میزان علاقه دانش آموزان

۲) هدف های آموزشی هر فصل

۳) مدت زمان صرف شده برای تدریس فصل

۴) حجم فصل و مواد تدریس شده

۱۷- « دانش آموز، فاعل و مفعول یک جمله را مشخص می کند. این فعالیت، جزء کدام سطح از سطوح هدف های بلوم است؟

(آزاد 86)

1) ارزشیابی کردن

۲) فهمیدن

۳) تحلیل کردن

۴) ترکیب کردن

۱۸- نوشتن طرح درس برای یک واحد درسی، در کدام سطح از سطوح هدف های بلوم قرار می گیرد؟

(آزاد 86)

۱) ترکیب

۲) تحلیل

۳) قضاوت

۴) به کار بستن

۱۹- محاسبه مساحت متوازی الاضلاع براساس اعداد داده شده، در کدام سطح از سطوح هدف های بلوم قرار می گیرد؟

(آزاد 86)

1) فهمیدن

۲) دانش

۳) ترکیب

۴) به کار بستن

۲۰- در ساخت یک آزمون، کدام مرحله بر سایر مراحل ارجحیت دارد؟

(سراسری 88)

۱) تعیین اهداف قابل قبول

۲) اهداف آموزشی معلم

۳) تعریف و تعیین صفت مورد اندازه گیری

۴) تهیه جدول بودجه بندی سؤال

۲۱- نگارش انشاء توسط کودکان مقطع ابتدایی، در کدام طبقه از طبقات حیطه شناختی بلوم قرار می گیرد؟

(سراسری 88)

1) کاربرد

۲) ترکیب

۳) تحلیل

۴) ادراک

۲۲- « توانایی تهیه و طرح تدریس یک مطلب ویژه» بر کدام سطح از حیطه شناختی بلوم تأکید دارد؟

(آزاد88)

1) به کار بستن

۲) ترکیب

۳) تجزیه و تحلیل

۴) ارزشیابی

۲۳- « توانایی تفسیر انواع گوناگون اطلاعات» در کدام طبقه از حیطه شناختی قرار می گیرد؟

(آزاد 88)

1) ترکیب

۲) تجزیه و تحلیل

۳) به کار بستن

۴) فهمیدن و درک

۲۴- اولین مرحله در تهیه طرح مقدماتی آزمون یا امتحان کدام است؟

(آزاد 88)

۱) تعیین روش و دستورالعمل اجرای آزمون با امتحان

۲) تعیین موضوع آزمون یا امتحان

۳) تعیین و انتخاب نمونه آزمون با امتحان

۴) مشخص کردن موارد استعمال آزمون با امتحان

۲۵- کدام یک از موارد زیر، ملاک مناسبی برای تهیه و تنظیم جدول تقسیم بندی محتوای کتاب به منظور تهیه سؤال نیست؟

(آزاد 88)

1) اهمیت نسبی هریک از هدف های آموزشی

۲) تعداد صفحات

۳) میزان تسلط معلم بر مطالب

۴) میزان وقتی که صرف تدریس هر مبحث شده است.

۲۶- در تهیه ی جدول بودجه بندی سؤالات، توجه به کدام مورد از اهمیت بیشتری برخوردار است؟

(سراسری 93)

۱) دشواری محتوا

۲) سن دانش آموزان

۳) اهمیت محتوا

۴) میزان علاقه دانش آموزان به هر بخش از مطالب

پاسخنامه تست های طبقه بندی شده فصل ششم

1- گزینه «۱» بلوم و همکارانش، طبقه بندی هایی را ارائه داده اند که یکی از این طبقات به حوزه ی عاطفی اشاره دارد و ارزش گذاری، یکی از سطوح این طبقه است. ارزش گذاری طبق تعریف، احساس یا باوری پایدار است که به ارزشمند بودن یک اندیشه یا ابراز عقیده و نگرش در مورد ارزش یک موضوع اشاره دارد؛ به عنوان مثال دفاع کردن از نظر شخصی یا گروند، می تواند نوعی ارزش گذاری تلقی شود. گزینه ی دو گزینه ای است که می تواند داوطلب را به اشتباه بیندازد؛ زیرا بلوم و همکارانش درباره حیطه شناختی نیز طبقه بندی ای ارائه داده اند؛ بنابراین ارزشیابی کردن به معنای داوری کردن با ملاک قراردادن چیزی در طبقه بندی حیطه شناختی وجود دارد و می تواند منجر به انتخاب گزینه ی نادرست شود.

۲- گزینه «۱» ارزشیابی پیشرفت تحصیلی، مستلزم انجام دو فعالیت است:

۱- تعیین و توصیف اهداف آموزشی با هدف های یادگیری ۲- سنجش یا اندازه گیری عملکرد یادگیرندگان آزمونهای پیشرفت تحصیلی، یکی از وسایل مهم و اندازه گیری پیشرفت تحصیلی هستند و برای تهیه آنها، نیاز به تهیه جدول مشخصات داریم. جدول مشخصات، منطبق بر موضوعی است که قرار است آزمون از آن تهیه شود جدول مشخصات شامل دو بعد است: ردیف افقی، بعد محتوا است که مطالب درسی در آن نوشته می شود و بعد هدف در ردیف عمودی قرار اس می گیرد که فهرست هدف های آموزشی را شامل می شود. گزینه ی دو، تصمیم گیری درباره ی نوع ارزشیایی، گزینه ی نزدیک به گزینه ی صحیح است. با توجه به به این موضوع که سؤال به آزمون پیشرفت تحصیلی اشاره دارد، ذهن داوطلبی را که آمادگی کامل ندارد به انتخاب این گزینه نادرست هدایت می کند.

۳- گزینه «٣» تمام گزینه ها به حیطه شناختی و طبقه بندی بلوم اشاره دارد. با توجه به این که ارائه طرح پژوهش، انجام کاری جدید و پدید آوردن حرکتی نو است، گزینه دربست، ترکیب است. توانایی کنار هم گذاشتن عناصر و اجزا برای ایجاد یک الگو یا ساختار نو، ترکیب نام دارد. با توجه به مفهوم تحلیل در گزینه ی ۲، این گزینه می تواند باعث تردید در افراد داوطلب شود. تجلیل، یعنی شکستن یک مطلب به اجزا و عناصر تشکیل دهنده. توجه به این نکته اهمیت دارد که در پژوهش و ارائه ی جدید به دنبال تجزیه و تحلیل موضوع نیستیم، بلکه هدف، ایجاد یک الگوی جدید است؛ بنابراین، تحلیل در پی یافتن ماهیت مسائل است که با معنی ترکیب در تضاد می باشد.

۴- گزینه «۲» فهمیدن به توانایی درک منظور یا مقصود یک مطلب اشاره دارد؛ همچنین فهمیدن، به پاسخ های بیشتر از آنچه فرد، قبلا تمرین و آموخته است، نیاز دارد. این پاسخ های اضافی عبارتند از: ترجمه یا برگردان، تفسیر، تعبیر، خلاصه کردن و کشف شباهت ها و تفاوت ها. از آنجایی که در ترجمه یک متن، فرد علاوه بر دانستن مینای لغات متن و قواعد زبان، باید متن مورد نظر را به شکلی که در زبان مقصد، کاملا روش و گویا باشد، برگرداند، نیاز است که متن مورد نظر را ابتدا فهمیده باشد.

۵- گزینه «3» در ارزیابی پیشرفت تحصیلی، دو اقدام اساسی ضروری است: ۱) تعیین هدف های آموزشی ۲) سنجش یا اندازه گیری عملکرد یادگیرندگان. گزینه ای که ممکن است در این سؤال به اشتباه انتخاب شود، گزینه دو می باشد که مهمترین گام در تهیه آزمون پیشرفت تحصیلی به شمار می آید.

6- گزینه «۱» آزمون های معلم ساخته، آزمون هایی هستند که معلمان برای سنجش پیشرفت تحصیلی دانش آموزان دوره اجرا می کنند. این نوع آزمون ها غالب برای تعیین میزان توفیق یادگیرندگان در رسیدن به هدف های دقیق آموزشی طرح ریزی می شوند. با دقت در گزینه ها، درمی یابیم که دانستنی های مهم به هدف های آموزشی مربوط است که در نظر معلم می باشد؛ بنابراین، آزمون های معلم ساخته به ارزشیابی آنها می پردازند.

۷- گزینه «۱» طبق تعریف، در طبقه بندی حوزه شناختی بلوم، فهمیدن، توانایی درک منظور یا مقصود یک مطلب است. فهمیدن به پاسخهایی بیشتر از آنچه قبلا تمرین و آموخته شده است، نیاز دارد. این پاسخ ها شامل ترجمه، تفسیر و تعبیر، خلاصه کردن و کشف شباهت ها و تفاوت ها است.

8- گزینه «۲» این جدول شامل دو بعد محتوا - اهداف آموزشی است که با توجه به محتوای درس و هدف های یادگیری در سه حوزه شناختی، عاطفی و روانی حرکتی؛ تکمیل می شود. آزمون مناسب، آزمونی است که به بهترین شکل، تمام اهداف آموزشی و برنامه درسی را منعکس می کند؛ همچنین تعد سؤالات باید با اهمیت اهداف و محتوای آموزشی متناسب باشد.

9- گزینه «۱» همان طور که در سؤال قبل نیز اشاره شد، یک آزمون خوب پیشرفت تحصیلی، آزمونی است که به بهترین شکل منعکس کننده ی تمام 1 - اهداف آموزشی و ۲- تمام محتوای برنامه درسی باشد. از لحاظ نظری، بهترین آزمون برای یک درس یا یک موضوع، آزمونی است که تمام مطالب و کلیه هدف های آموزشی آن درس (موضوع) را شامل شود.

۱۰- گزینه «۱» یک آزمون مناسب پیشرفت تحصیلی، آزمونی است که به بهترین شکل منعکس کننده تمام هدف های آموزشی و همه ی محتوای برنامه درسی باشد. با تهیه جدول بودجه بندی، اطمینان می یابیم که آزمون، تمام هدف های درس را مد نظر قرار داده و به ارزشیابی کل هدف های آموزشی درس پرداخته است.

۱۱- گزینه «۴» با توجه به سه حیطه عاطفی، روانی - حرکتی و شناختی، باید توجه داشته باشیم که حیطه شناختی بلوم شامل ( 1- دانش ۲- درک 3- کاربرد ۴- تحلیل و ترکیب ۵- قضاوت است. بنابراین توانایی درک و فهم، به توانایی در تفسیر نکات مهم یک مطلب اشاره دارد.

۱۲- گزینه «3» زمانی که دانشجویان مسئله ای را بررسی می کنند؛ در واقع در مرحله چهارم حیطه شناختی بلوم قرار می گیرند که شامل تحلیل و ترکیب مسائل است. تحلیل، یعنی شکستن یک مطلب به عناصر و اجزای تشکیل دهنده و ترکیب، یعنی کنار هم گذاشتن این عناصر و اجزا و ایجاد الگویی با ساختار جدید.

۱۳- گزینه «۱» در آموزش و پرورش، دو نوع هدف وجود دارد: ۱- کلی (اهداف آرمانی) ۲- دقیق (اهداف رفتاری) اهداف آموزشی به سه حوزه کلی تقسیم میشوند: 1- حوزه شناختی ۲- حوزه عاطفی 3-حوزه روانی - حرکتی

در حوزه عاطفی، شاهد پنج طبقه هستیم. دریافت کردن (توجه کردن)، پاسخ دادن، ارزش گذاردن، سازمان دادن به ارزش ها، شخصیت پذیرفتن وقتی معلم از شاگرد بخواهد، توضیح بیشتری ارائه دهد، او را در مرحله پاسخ دادن و واکنش قرار میدهد.

۱۴- گزینه «۲» در روش رفتارگرایان، عمل آموزشی به اجزای کوچکتری تقسیم می شود که به صورت دقیق و در غالب اهداف آموزشی بیان می شوند. این اهداف، مبنای آموزش و ارزشیابی از فراگیران است.

۱۵- گزینه «۲» اگر سؤالات به کار گرفته شده در آزمون، از نوع سؤالات عینی باشند، می توان سؤالات زیادی را در آزمون به کار برد؛ اما اگر بخواهیم از سؤالات انشایی استفاده کنیم، باید سؤالات کمتری طراحی کنیم؛ همچنین هر چه زمان تدریس بیشتر باشد، باید سؤالات بیشتری و هر چه زمان تدریس کمتر باشد باید سؤالات کمتری را در آزمون گنجاند؛ مانند سؤالات ارزشیابی تکوینی؛ چون زمان کمتری صرف تدریس آنها شده است. بنابراین نوع سؤالات و زمان تدریس، دو مورد مهم در تعیین کل سؤالات امتحانی به شمار می آیند.

۱۶- گزینه «۳» برای به دست آوردن درصد سؤالات هر فصل از موارد تدریس شده، باید نسبت ساعت های تدریس شده برای آن فصل را به دست آوریم؟ اگر بخواهیم تعداد سؤالات را به دست بیاوریم، باید نتیجه را در تعداد کل سؤال های مدنظر برای آزمون ضربه می کنیم تا تعداد سؤالات هر بخش به دست آید.

ساعت های صرف تدریس شده برای هر بخش

= نسبت ساعتهای تدریس هر بخش

کل ساعات تدریس

۱۷- گزینه «3» در تحلیل، بادگیرنده باید یک مطلب را به اجزاء یا عناصر تشکیل دهنده آن تجزیه کند که در این سؤال، با توجه به این که دانش آموز باید فاعل و مفعول جمله را مشخص کند، فعالیت در سطح تحلیل می باشد.

۱۸- گزینه «1» در ترکیب، یادگیرنده باید بتواند عناصر و اجزای مختلف را برای ایجاد یک الگو یا ساختار نو کنار هم قرار دهد. با توجه به این که برای نوشتن طرح درس، یادگیرنده باید عناصر و اجزای مختلف مانند اهداف، روشهای تدریس، مواد و وسایل و... را با هم برای ایجاد یک طرح جدید به کار گیرد، می توان گفت که این فعالیت در سطح ترکیب می باشد.

۱۹-گزینه «۴» در سطح کار بستن، فرد باید توانایی استفاده از امور انتزاعی، قواعد و قوانین، اصول، اندیشه ها و روش ها را در موقعیت های عینی و عملی داشته باشد. با توجه به این که برای محاسبه مساحت متوازی الاضلاع، فرد باید قاعده محاسبه آن را به کار گیرد، می توان گفت که این فعالیت در سطح به کار بستن است.

۲۰- گزینه «3» آزمون چیست؟ وسیله ای نظام دار و روشی هدفمند است که نمونه ای از رفتار را اندازه می گیرد. با این تعریف، برای ساختن یک آزمون باید به اهداف آزمون توجه کرد که واقعا قرار است چه چیزی را اندازه گیری کند. وقتی صفت ها و اهداف مورد نظر تعریف شوند، اندازه گیری در آن آزمون آسان تر می شود؛ به عنوان مثال، یک متخصص اندازه گیری، نمونه ای از رفتار را انتخاب و آزمون را براساس آن پایه ریزی می کند.

۲۱- گزینه «۲» حوزه شناختی به طور سلسله مراتبی از عینی و ساده به انتزاعی و پیچیده می رود.

۱- دانش: یادآوری و تکرار

2ـ فهمیدن: درک مطلب

۳ـ کاربستن: توانایی کاربرد قواعد و اصول

۴- تحلیل: توانایی شکستن مطلب به اجزاء و عناصر آن

5- ترکیب: توانایی کنار هم گذاشتن عناصر و اجزاء برای خلق ساختار نو

۶- ارزشیابی: توانایی قضاوت با مدرک های معین

همان گونه که در مورد 5 توضیح داده شد، هدف، آفرینندگی و خلاقیت است.

وقتی کودک بتوانـد بـا کنـار هـم گذاشتن عناصر و اجزای در دسترس، بـه خـلـق یک الگوی جدید و نو بپردازد، در مرحله ترکیـب قـرار دارد.

نوشتن انشاء در مورد تعطیلات تابستانی و یا تدریس یک موضوع ابتکاری، نمونه روشنی در این زمینه می باشد.

۲۲- گزینه «۲» در ترکیب، یادگیرنده باید بتواند عناصر و اجزای مختلف را برای ایجاد یک الگو یا ساختار نو کناره هم قرار دهد. با توجـه بـه ایـن کـه بـرای نوشتن طرح درس، یادگیرنده باید عناصر و اجزای مختلف را (اهداف، روش های تدریس، مواد و وسایل و ...) با هم برای ایجاد یک طرح جدید به کار گیرد، می توان گفت این فعالیت در سطح ترکیب می باشد.

۲۳- گزینه «۴» در اهداف سطح فهمیدن، فرد به پاسخهایی بیشتر از آنچه قبلا آموخته است، نیاز دارد. این پاسخ های اضافی عبارتند از: ترجمه یا برگردان، تفسیر و تعبیر، خلاصه کردن و کشف شباهت ها و تفاوت ها.

۲۴- گزینه «۲» مهم ترین گام در تهیه آزمون های پیشرفت تحصیلی، تهیه جدول مشخصات برای موضوعی است که قرار است آزمـون از آن تهیـه شـود. برای تهیه این جدول مشخصات، باید ابتدا موضوع آزمون و اهداف آن تعیین شود.

۲۵- گزینه «۳» به طور کلی برای تعیین درصد سؤال هایی که به بخش های مختلف محتوا و هدفها تعلق می گیرند، جواب سؤال های زیر را باید تعیین کرد:

۱ـ کدام بخش از محتوا و کدام دسته از هدف ها حجم بیشتری دارد و وقت بیشتری صرف آموزش آنها شده است؟

2ـ کدام هدف ها از لحاظ نگهداری در حافظه و کاربرد در آینده ارزش بیشتری دارند؟

۳- کدام قسمت یا قسمت های درس یا موضوع درس از اهمیت بیشتری برخوردارند؟

۲۶- گزینه «۳» در جدول بودجه بندی سؤالات، بر توجه به اهمیت محتوا تأکید بیشتری می شود؛ زیرا هدف، بودجه بندی محتوای یک مبحث است تا خصوصیات دیگر آن.

آزمون فصل ششم

۱- در طبقه بندی اهداف آموزشی بلوم کدام طبقه از حیطه شناختی از سطح عالی تری برخوردار است؟

۱) درک و فهم

۲) کاربرد

۳) تجزیه و تحلیل

۴) دانش

۲- وقتی دانش آموزی از طریق قواعدی که آموخته است، بتواند مسائل فیزیک را حل کند، در چه سطحی از حیطه شناختی قرار داد؟

۱) دانش

۲) درک و فهم

۳) تجزیه و تحلیل

4) کاربرد

3- اگر قسمتی از کتاب تاریخ بیهقی به دانش آموزان داده شود تا به زبان خود آن را بیان کنند، این کار در چه سطحی از حیطه شناختی قرار داد؟

1) دانش

۲) ترکیب

۳) ارزشیابی

۴) درک و فهم

۴- قضاوت هوشیارانه درباره کارها و روش ها، به کدام یک از حیطه های زیر تعلق دارد؟

1) انفعالی

۲) روانی - حرکتی

۳) شناختی

۴) عاطفی

5- وقتی یادگیرنده در موقعیت های مختلف، متناسب با نظام ارزشی خود رفتار می کند، در چه سطحی از حیطه عاطفی قرار دارد؟

۱) ارزش گذاری

۲) پاسخ دادن

۳) سازمان دهی به ارزش ها

4) شخصیت پذیرفتن

6- کدام یک از فعل های زیر را نمی توان در بیان هدف های رفتاری به کار بست؟

1) توضیح دهد

۲) بفهمد

۳) طبقه بندی کند

4) خلاصه کند

۷- کدام یک از اهداف زیر یک هدف رفتاری نیست؟

۱) دانش آموز بتواند مسائل ریاضی را حل کند.

2) دانش آموز بتواند طول کلاس را با متر اندازه بگیرد.

3) دانش آموزی تعریف سرمایه گذاری را درک کند.

۴) داستان را به زبان خود بیان کند.

8 - اولین گام در تهیه آزمون های پیشرفت تحصیلی چیست؟

۱) تعیین نوع سؤالات آزمون

۲) تعیین نوع مقیاس اندازه گیری

۳) تهیه جدول مشخصات آزمون

۴) توزیع سؤالات در میان بخش های مختلف درس

۹- اگر دانش آموز با استفاده از فرمول های ریاضی، بتواند مساحت کلاس را به درستی حساب کند، این هدف در چه سطحی از حیطه شناختی قرار دارد؟

۱) دانش

۲) فهمیدن

3) کار بستن

۴) تحلیل

10- نوشتن یک انشاء در کدام یک از سطوح حیطه شناختی قرار می گیرد؟

۱) دانش

2) فهمیدن

3) تحلیل

۴) ترکیب

فصل هفتم

«آزمون های تشریحی»

آزمون های کتبی یا مداد و کاغذی

پیشرفت تحصیلی را میتوان به دو دسته عینی و غیر عینی (ذهنی) تقسیم کرد.

آزمون های عینی، آزمون هایی هستند که در آنها هم سؤال و هم پاسخ سوال در اختیار آزمون شوندگان قرار می گیرد و آزمون شوندگان در مورد پاسخ ها می کنند؛ بدین دلیل به این آزمون ها بسته پاسخ هم می گویند. آزمون های عینی به سه دسته تقسیم می شوند:

۱) صحیح - غلق

۲) جور کردنی

3) چند گزینه ای

آزمونه ای عینی نسبت به غیرعینی آسان تر است و نظر شخصی مصحح،هیچ گونه دخالتی در تصیح نداد.نام دیگر آزمون های عینی به روش های انتخاب پاسخ است؛ زیرا پاسخ دهنده فقط به انتخاب یا گزینش پاسخ مبادرت می کند.

آزمون های غیرعینی، پاسخ دهنده، پاسخ را می سازد؛ از جمله آزمون های غیر عینی می توان آزمون های تشریحی، کوته پاسخ و امتحانات شفاهی، پروژه های عینی و غیره را نام برد که به آن ها آزمون های باز پاسخ هم گفته می شود. آزمون های کوته پاسخ حد واسط آزمونهای عینی و ذهنی هستند.

نکته: آزمون های عینی، توانایی باز شناسی یا تشخیص را می سنجند.اما آزمون های تشریحی یا کوته پاسخ، عموما توانایی باز خوانی آزمودنی ها را اندازه می گیرد.

امثال ۱: کدام عبارت صحیح می باشد

۱) آزمون های عینی، توانایی بازخوانی آزمون شوندگان را می سنجند.

۲) آزمون های تشریحی یا کوته پاسخ، توانایی بازخوانی آزمون شوند

3) آزمون های تشریحی، توانایی بازشناسی آزمون شوندگان را بسنجند.

4) هم آزمون های عینی و هم آزمون های تشریحی می توانند توانایی بازخوانی آزمون شوندگان را می سنجند.

پاسخ: گزینه «۲» آزمون های تشریحی یا کوته پاسخ، توانایی بازخوانی توانایی بازخوانی آزمون شوندگان را می سنجند.

مثال ۲: اگر معلم بخواهد توانایی بازشناسی را در آزمون شوندگان اندازه گیری کند،کدام یک از انواع آزمون های زیر را پیشنهاد می کنید؟

1) کوته پاسخ

۲) آزمون شفاهی

3) آزمون های عینی

4) آزمون های انشایی

پاسخ: گزینه «3» آزمون های عینی می توانند توانای بازشناسی را در آزمون شوندگان اندازه گیری کنند.

مثال 3: به نظر عده زیادی از متخصصان سنجش و اندازه گیری، سوال نویسی .....

۱) آسان است.

2) هنر است.

3) اطلاعات نظری نیاز دارد.

4) فقط به تجربه نیاز دارد

پاسخ: گزینه «۲» با توجه به این که در یک مجموعه سوال، باید کل آزمون شوندگان را در نظر گرفت، در یک آزمون، سؤالات آسان، متوسط و مشکل بند گنجانده می شوند تا بتوانند افراد را دسته بندی کنند و اهداف آموزشی نیز مورد سنجش قرار گیرد؛ بنابراین به نظر می رسد، طراح سؤال علاوه بر داشتن علم در زمینه مربوط و نیز تجربه ی کافی برای طرح سؤال، باید هنر نیز داشته باشد.گزینه های ۳ و ۴، زیرمجموعه ی گزینه ی ۲ می باشند که ممکن است انتخاب درست داوطلب را تحت تأثیر قرار دهند.

مثال ۴: فرق اساسی سوال های عینی و تشریحی با انشایی این است که:

۱) تعداد سؤال ها در امتحان انشایی کمتر است.

2) محتوای سؤال های عینی عملی تر است.

۳) نوشتن سؤال های عینی آسان تر است.

۴) سؤال های انشایی، هدف های مهم تر تعلیم را اندازه می گیرند.

پاسخ: گزینه «۴» فرق اساسی سوال های عینی و تشریحی این است که سؤال های انشایی، هدف های مهم تر تعلیم را اندازه می گیرند. آزمون های تشریحی را با توجه به آزادی عمل آزمون شونده در پاسخ دادن به سؤال های آزمون، به دو دسته گسترده پاسخ و محدود پاسخ تقسیم می کنند:

آزمون های گسترده پاسخ

آزمون هایی هستند که آزمون شوندگان هیچ محدودیتی در پاسخگویی به آنها ندارند و آزادی کامل دارند تا پاسخ را هر طور که می خواهند، بپرورانند و بنابراین، محدودیتی از نظر مقدار پاسخ و زمان پاسخگویی ندارند. همچنین، آزمون شونده باید اطلاعات و دانش های خود را ارزیابی کند و اندیشه های منطقی منسجم و بدیع روی کاغذ آورد؛ بنابراین، این آزمون ها برای سطوح بالای حیطه شناختی یعنی تحلیل، ترتیب و ارزشیابی مناسب اند.

مثال ۵: از جمله امتیازهای « سؤال های باز پاسخ»، کدام است؟

1) تقلب را کاهش میدهد. |

۲) نمره گذاری آن کاملا عینی و پایا است

۳) پوشش وسیعی از محتوا و هد فها را ممکن می سازد.

۴) عامل حدس و گمان را کاهش می دهد.

پاسخ: گزینه «۴» در آزمون های باز پاسخ، هیچ گونه محدودیتی برای آزمون شونده منظور نمی شود و او عملا آزاد است، هر طور که مایل باشد، پاسخ خود را بپروراند و سازماندهی کنند. آزمون شونده در این نوع آزمون، هم از لحاظ زمان پاسخ دهی و هم از جهت مقدار پاسخ، آزادی کامل دارد. آزمون های باز پاسخ مولفه ای نیز دارند که این مؤلفه ها عبارت اند از: آسان بودن تهیه سؤالات، سنجش سطوح بالای حیطه شناختی (در قسمت های تحلیل و ترکیب و ارزشیابی)، وجود قدرت پرورش اطلاعات و ارائه موقعیت های نزدیک به واقعیت. با توجه به سطوح بالای حیطه شناختی، احتمال حدس و گمان کاهش می یابد به همین دلیل، در صورتی که با توجه به متن سؤال، تحلیل صورت گیرد، می تواند باعث انتخاب نادرست توسط داوطلب شود. این بدان معنی است که با دیدن عنوان باز پاسخ در متن سؤال، ذهن آزمون شونده به سمت محتوا و هدف می رود که نادرست می باشد.

مثال۶: اگر مایل به اندازه گیری توانایی شاگرد از نظر انتخاب و سازماندهی معلومات باشیم، بهتر است از سوالات نوع ................. استفاده کنیم.

1) معدود پاسخ

۲) کوته پاسخ

۳) عینی

۴) انشایی

پاسخ گزینه «۴» در پاسخ به سؤالات انشایی توسط آزمون شونده، معلومات در سطوح تحلیل و سازماندهی سنجیده می شود و هدف این آزمون توانایی آزمون شونده در پروراندن جواب سؤال ها و بیان آن هاست؛ بنابراین یک آزمون تشریحی، توانایی افراد را بهتر از سایر آزمون ها نمایش میدهد.

گزینه ها با متن سؤال ارتباطی ندارند.

آزمون های مسدود پاسخ

در این آزمون ها، آزمون شونده ها آزادی کامل دارند، بلکه صورت سوال ها طوری طرح شد، که آزمون شونده را مجبور می کندن در چهار چوب مناسبی پاسخ دهد و همچنین آزمون شونده از نظر مقدار پاسخ و زمان پاسخ محدودیت دارد، در آزمون های محدود پاسخ، با محدود کردن پاسخ آزمون شوندگان، به یک موضوع مشخص، تصحیح پاسخ آنان را هماهنگ تر و آسان تر می کنیم، اما این محدودیت سبب می شود که آزمون شونده نتواند توانایی خود را در ترکیب کردن اندیشه ها و بیان آن ها به صورتی منسجم و منطقی ابراز کنند. بنابراین این آزمون ها، برای به سطوح پایین حیطه شناختی یعنی فهمیدن، کاربرد و تحلیل مناسب اند.

مثال 7- برای اندازه گیری هدف های سطح ترکیب از آزمون های ................... می توان استفاده کرد.

1) کوته پاسخ

2) آزمون های شفاهی

۳) آزمون های عینی

4) آزمون های تشریحی گسترده پاسخ

پاسخ گزینه «۴» برای اندازه گیری هدف های سطح ترکیب از آزمون های تشریحی گسترده پاسخ استفاده می شود.

مثال ۸: اگر معلم بخواهد توانایی فهمیدن و کاربرد را در آزمون شوندگان اندازه گیری کند، کدام یک از انواع آزمون های زیر را پیشنهاد می کنید؟

1) آزمون های تشریحی گسترده

2) آزمون های تشریحی محدود پاسخ

3) آزمون های شفاهی

۴) آزمون های کوته پاسخ

پاسخ گزینه «۲» با استفاده از آزمون های تشریحی محدود پاسخ، می توان توانایی فهمیدن و کاربرد را در آزمون شوندگان اندازه گیری کرد.

مثال9 : تفاوت بین دیدگاه رفتارگرایی و شناختی را از لحاظ توجه آن ها به یادگیرنده توضیح دهید.

(این سؤال چه نوعی سؤالی است؟)

1) تشریحی گسترده پاسخ

۲) تشریح محدود پاسخ

۳) کوته پاسخ

۴) هیچ کدام

پاسخ: گزینه «۱» تشریحی از نوع گسترده پاسخ می باشد.

محاسن آزمون های تشریحی

۱) تهیه آزمون های تشریحی از آزمون های عینی آسان تر است؛

۲) این نوع آزمون ها، تنها وسیله ی موجود برای سنجش توانایی آزمون شونده در پروراندن جواب سؤال ها و بیان آنهاست.

۳) این آزمون ها توانایی پاسخ دادن به سؤال ها را می سنجند، نه توانایی انتخاب پاسخ ها را؛

۴) این آزمون ها نسبت به آزمون های عینی موقعیت واقعی تری را به آزمون شوندگان عرضه می کنند؟

۵) این ها بر روش مطالعه یادگیرندگان تأثیر مثبت می گذارند و آن ها را وادار می کنند تا روابط میان مطالب را جستجو و اطلاعات را به راه های معنی دار سازماندهی کنند.

مثال ۱۰: کدام گزینه صحیح می باشد؟

1) تهیه آزمون های تشریحی از آزمون های عینی دشوارتر است.

۲) آزمون های عینی، موقعیت واقعی تری را نسبت به آزمون های تشریحی فراهم می کنند.

۳) آزمون های عینی بر روش مطالعه دانش آموزان تأثیر مثبت می گذارند.

۴) آزمون های تشریحی، توانایی پاسخ دادن را می سنجند، نه توانایی انتخاب پاسخ ها را

پاسخ: گزینه ۴» آزمون های تشریحی، توانایی پاسخ دادن را می سنجند، نه توانایی انتخاب پاسخ ها را

معایب آزمون های تشریحی

1) این آزمون ها نمونه کوچکی از محتوای درس و هدف های آموزش را اندازه می گیرند.

2) تصحیح برگه های امتحانی این آزمون ها، نمی تواند با دقت و به طور عینی انجام گیرد. نمره گذاری این آزمون ها ذهنی است.

۳) تصحیح برگه های امتحانی این آزمون ها بسیار وقت گیر است.

از جمله معایب دیگر آزمون های تشریحی این است که در یک زمان محدود امتحانی، نمی توان تعداد زیادی سوال تشریحی به آزمون شوندگاری ارائه داد، لذا بر همیشه تعداد سؤال های یک آزمون تشریحی، بویژه آزمون های گسترده پاسخ، بسیار محدود است. مشکل دیگر آزمون های تشریحی، ذهنی بودن نمره گذاری یک مصحح و یکنواخت نبودن ارزشیابی آنان است. سومین عیب آزمون های تشریحی، وقت گیر بودن تصحیح پاسخ سؤالها است. برخلاف پاسخ آزمونهای عینی که بوسیله ی ماشین یا افراد غیر متخصص قابل تصحیح است، تصحیح اوراق امتحانی آزمون های تشریحی به افراد با تخصص بالا نیاز دارد.

مثال ۱۱: کدام گزینه از معایب آزمون های تشریحی به شمار نمی آید؟

۱) تصحیح برگه های این آزمون وقت گیر میباشد.

۲) تصحیح برگه های این آزمون نمی تواند به دقت و به طور عین انجام شود.

3) این آزمون ها توانایی پاسخ دادن را می سنجند.

۴) این آزمون های نمونه کوچکی از محتوای درس و هدف ها را اندازه گیری می کنند.

پاسخ: گزینه «3» معایب آزمون های تشریحی: ۱ - تصحیح برگه های این آزمون وقت گیر میباشد ۲- تصحیح برگه های این آزمون نمی تواند به دقت و به طور عینی انجام شود ۳- این آزمون های نمونه کوچکی از محتوای درس و هدف ها را اندازه گیری می کنند.

مثال ۱۲: کدام یک از موارد زیر، از مزیت های عمده سؤال های انشایی به شمار می رود؟

۱) دانش آموزان در پاسخ دادن به آن ها آزادی عمل بیشتری دارند.

۲) نوشتن آن ها آسان تر است و وقت کمتری می گیرد.

۳) قضاوت ذهنی تصحیح کننده در نمره گذاری آن ها مؤثر است.

۴) سطوح عالی تر توانایی های ذهنی را اندازه می گیرند.

پاسخ: گزینه «۴» آزمون های گسترده پاسخ یا انشایی (تشریحی) محاسنی دارند که یکی از آنها به سطوح عالی تر توانایی های ذهنی اشاره دارد. در این روش، آزمون شوندگان میدانند که یادگیریشان مورد سنجش قرار می گیرد؛ بنابراین با احتمال بیشتری به دنبال روابط میان مطالب به جستجو خواهند پرداخت و اطلاعات را به طور معنی دارتری سازماندهی خواهند کرد. هر چند که در حین پاسخ دهی به سؤالات، دانش آموزان از آزادی عمل برخوردارند، اما باید به این نکته توجه داشـت کـه این آزادی عمل در یک چهارچوب معین (متن سؤال و حیطه پاسخدهی) قرار دارد؛ بنابراین انتخاب گزینه ی صحیح، منوط به دانستن محاسن سؤالات تشریحی است. محاسن آزمون های تشریحی گسترده پاسخ یا انشایی عبارتند از: ۱- سنجش سطوح بالای شناختی ۲ـ آسان بودن تهیـه سـؤال هـا ۳ـ نشان دادن قدرت پروراندن اطلاعات ۴- ارائه موقعیتی واقعی تر 5- کاهش قدرت حدس و گمان.

مثال ۱۳: یکی از معایب سؤالات تشریحی این است که:

۱) دانش آموز می تواند حدس بزند.

۲) در مورد کودکان قابل اجرا نیست.

۳) نمره گذاری آنها اعتبار زیادی ندارد.

۴) هدف های مهم را اندازه نمی گیرد.

پاسخ: گزینه «۳» از معایب آزمون های تشریحی (انشایی) می توان به مشکل بودن تصحیح و نمره گذاری اشاره کرد. به طور کلی، معایب این آزمون ها از نظر صاحب نظران عبارتند از: ۱) این آزمون ها نمونه ی کوچکی از محتوای درسی و اهداف آموزشی را اندازه گیری می کنند؛ ۲) تصحیح برگه های امتحانی نمی تواند به طور عینی و با دقت صورت گیرد؛ ۳) تصحیح اوراق بسیار وقت گیر است.

مثال ۱۴: رعایت کدام شرط در بهسازی سؤال های انشایی مؤثر است؟

۱) مشخص بودن موضوع سؤال

۲) انتخاب سؤال عیناً از متن کتاب

۳) دادن حق انتخاب به امتحان شوندگان

۴) طرح سؤالاتی که تعبیر و تفسیر آن نزد همه یکسان نباشد.

پاسخ: گزینه «۱» برای تهیه سؤال های انشایی (تشریحی) قواعدی وجود دارد که یکی از آنها روشن بودن صورت سؤال ها، بدون ابهام و کلی گویی می باشد که با گزینه ی ۱ همخوانی دارد. در گزینه ۱ به موضوع مشخص بودن سؤال اشاره می شود؛ در صورتی که هیچ کدام از گزینـه هـای دیگر، بـه ایـن م .. قواعد 10 گانه اشاره نمی کنند و داوطلب، در صورت اطلاع کافی از این قواعد، به راحتی می تواند پاسخگوی این سؤال باشد. وقتی موضوع سؤال با کلماتی مانند، « طبقه بندی کنید» یا « بررسی کنید»، « تعریف کنید» مشخص شوند، آزمون شونده به راحتی به سؤال پاسخ می دهد.

قواعد تهیه سؤال های تشریحی

تفاوت عمده آزمون های تشریحی و آزمون های عینی، که بزرگ ترین امتیاز آزمون های تشریحی میباشد، این است که آزمون شوندگان را وامی دارد تا اندیشه های خود را به طور منطقی، منسجم و سازمان یافته بیان کنند و با این کار، خلاقیت خود را نشان دهند.

1) در نوشتن سؤال ها باید دقت کرد که سؤال ها به طور مستقیم به هدف های آموزشی مربوط شوند. به دلیل کم بودن تعداد سؤال ها در آزمون های تشریحی ر نسبت به آزمون های عینی، دقت عمل در نوشتن این نوع سؤال ها ضرورت بیشتری دارد و هر سؤال را باید با توجه به یک هدف دقیق آموزشی نوشت.

۲) سؤال ها را تنها به اندازه گیری هدف هایی محدود کنید که با سایر آزمون ها به خوبی قابل اندازه گیری نیستند. سؤال های تشریحی گسترده پاسخ، بهترین سؤال برای طبقه تحلیل، ترکیب و ارزشیابی هستند و از سؤال های محدود پاسخ برای طبقات پایین می توان استفاده کرد.

۳) صورت سؤالهای تشریحی را باید با عبارات و کلمات واضح و روشن نوشت.

۴) از کاربرد کاربرد کلمات چه کسی، چه وقت و کجا بپرهیزید؛ چون این کلمات، تنها معلومات را اندازه می گیرند. به جای این کلمات، از «چرا»، «چگونه» و «به چه دلیل» استفاده کنید.

۵) تا حد امکان از سؤال های تازه و موقعیت های جدید استفاده کنید. استفاده از موقعیت های جدید در طرح سؤال ها، قدرت کاربرد و درک و فهم را اندازه می گیرد.

۶) سؤالات مربوط به موضوعات و عقاید بحث انگیز باید طوری مطرح شوند که آزمون شونده شواهد لازم را برای مستند کردن عقاید انتخابی بیان کند، نه این که از او بخواهند تا صرفاً عقاید شخصی خود را شرح دهد.

۷) به آزمون شوندگان، حق انتخاب چند سؤال از میان تعدادی سؤال را ندهید؛ زیرا:

۱- تهیه سؤال هایی که از نظر درجه دشواری همسنگ باشند، کار آسانی نیست.

۲- آزمون شوندگان، توانایی انتخاب سؤال هایی که جواب آنها را بهتر میدانند، ندارند.

۳ـ دادن حق انتخاب، مشکل نمونه گیری را در آزمون های تشریحی دشوارتر می کند.

۴۔ چون آزمون شوندگان قوی تر، بیشتر به سراغ سؤال های دشوار می روند، ممکن بودن حق انتخاب، در نهایت به ضرر این گونه افراد تمام شود. با این حال در مواردی که هدف سؤال ها اندازه گیری میزان یادگیری نیست، بلکه صرفا اندازه گیری مهارت پاسخ دهندگان در نوشتن و پروراندن مطالب است، دادن حق انتخاب از میان تعدادی سؤال قابل دفاع می باشد

8) برای پاسخ دادن به سؤال ها زمان کافی در نظر بگیرید. داشتن وقت، هم فرصتی برای جواب دادن است و هم فرصتی را برای فکر کردن آزمون شونده در نظر میگیرد.

9) با نوشتن سوال هایی که جواب کوتاه دارند، تعداد سؤالات را افزایش دهید و مشکل ضعف نمونه گیری آزمونهای تشریحی را کاهش دهید.

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

مثال ۱۵: کدام گزینه در مورد قواعد تهیه سؤال های تشریحی صحیح نمی باشد؟

1) سؤال ها را تنها باید به اندازه گیری هدف هایی محدود کنید که با سایر آزمون ها به خوبی قابل اندازه گیری نیستند.

2) امکان حق انتخاب پاسخ به سؤال را برای دانش آموزان فراهم آورید.

۳) از موقعیت های جدید برای طرح سؤال استفاده کنید.

۴) از کاربرد کلمات چه کسی، چه وقت و کجا بپرهیزید.

پاسخ: گزینه «۲» در سؤال های تشریحی فرد حق انتخاب پاسخ را ندارد، بلکه آن را می نویسد.

مثال ۱۶: رعایت کدام شرط در بهسازی سؤال های انشایی مؤثر است؟

۱) مشخص بودن موضوع سؤال

۲) انتخاب سؤال عینا از متن کتاب

۳) دادن حق انتخاب به آزمودنی

۴) طرح سؤال های آسان

پاسخ: گزینه «1» مشخص بودن موضوع سؤال در بهسازی سؤال های انشایی مؤثر است.

قواعد تصحیح آزمون های تشریحی

1) پاسخ سؤال ها را تنها براساس هدفی که در سوال گنجانیده شده است، تصحیح کنید. از دخالت دادن نکات دستوری، املایی و انشایی و جزء این ها خودداری کنید.

۲) با نوشتن یک پاسخ نمونه با الگو برای هر سؤال به عنوان کلید، از دخالت عوامل نامربوط جلوگیری کنید.

۳) پاسخ ها را سؤال به سؤال تصحیح کنید، نه ورقه به ورقه؛ یعنی ابتدا پاسخ سؤال اول همه ی اوارق را تصحیح کنید، بعد پاسخ های سؤال دوم اوراق و و همینطور تا آخر ادامه دهید. این روش، امکان مقایسه پاسخ های آزمون شوندگان مختلف را به یک سؤال فراهم و کار نمره گذاری را آسان می کند.

۴) هنگام تصحیح ورقه های امتحانی از شناسایی نام صاحبان آن خودداری کنید.

۵) از یکی، دو نفر بخواهید تا پاسخ سؤال هایی را که شما تصحیح کرده اید، تصحیح کنند. این روش، تنها راه بررسی عینیت و دقت ارزشیابی شما از پاسخ های دانش آموزان است.

۶) تمام پاسخ های آزمون شوندگان به یک سؤال را در یک جلسه و بدون وقفه زمانی تصحیح کند.

۷) به نمرات سؤال هایی که قبلا تصحیح کرده اید، نگاه نکنید.

۸) روی برگه های آزمون، اشتباهات دانش آموزان را تصحیح کنید و اظهارنظرهای خود را بنویسید.

9) از روش های متنوع نمره گذاری استفاده کنید.|

روش های نمره گذاری آزمون های تشریحی

1- روش تحلیلی (امتیازبندی): در این روش، معلم، پاسخ را به اجزاء کوچکتری تقسیم می کند و بعد برای هر قسمت، نمره ای در نظر می گیرد؛ همچنین در این روش، عواملی چون «قدرت بیان»، «ساختمان منطقی پاسخ» و «ذکر دلایل» امتیازهای جداگانه ای دریافت می کنند.

٢- روش برداشت کلی (درجه بندی): در روش برداشت کلی، معلم یکبار پاسخ را می خواند، یک برداشت کلی می کند و به آن نمره میدهد. در این روش سر به هیچ عامل واحدی، امتیاز خاصی داده نمی شود؛ بلکه همه عوامل مورد نظر قرار می گیرند و کل پاسخ یکباره قضاوت می شود.

گی (۱۹۹۱) معتقد است که روش تحلیلی برای آزمون های وابسته به ملاک و روش کلی برای آزمون های وابسته به هنجار مناسب تر است.

۳- روش ویژگی اصلی: در روش ویژگی اصلی، معلم ویژگی اصلی پاسخ را می سنجد و نسبت به آن از صفر تا ۴ نمره می دهد.

مثال ۱۷: کدام روش نمره گذاری در تصحیح سؤال های انشایی یا تشریحی دقیق تر و عینی تر است؟

۱) کلی یا ترکیبی

۲) تطبیقی

۳) درجه بندی

۴) تحلیلی

گزینه «۴» در تصحیح اوراق انشایی (تشریحی) می توان از روش های متنوع نمره گذاری استفاده نمود که از جمله آنها می توان بـه ایـن مـوارد اشاره کرد.

۱- روش تحلیلی (نمره گذاری)

۲- روش کلی (درجه بندی)

۳- ویژگی های اصلی

در روش تحلیلی( analytical)، پاسخ نمونه که توسط مصحح تهیه می شود، به اجزای کوچک تری تقسیم میشود و برای هر جزء مشخص، نمره یا امتیاز جداگانه ای در نظر گرفته میشود. این موضوع، باعث رعایت جانب اعتدال در تصحیح اوراق می شود؛ در نتیجه نظر مصحح، کمتر در نتیجه تصحیح اثرگذار خواهد بود. گزینه های ۱ و ۳ به یک عنوان اشاره دارد و گزینه ی شماره ۲ کاملا انحرافی استر در روش کلی (درجه بندی) تنها معیارگذاری میشود و معلم به نسبت برداشت خود از برگه ها نمره گذاری می کند، بنابراین، به دقت روش تحلیلی (نمره گذاری) عمل نمی نماید.

پرسش های شفاهی

پرسش های شفاهی را می توان نوعی سؤال تشریحی دانست که به صورت غیررسمی طرح میشود و جواب آن هم به صورت شفاهی در کلاس ارائه می شود. پرسش شفاهی را نمی توان نوعی آزمون به حساب آورد و از نتایج آن برای نمره گذاری میزان یادگیری دانش آموزان استفاده کرد، اما از این نوع پرسش می توان برای بهبود روش های یادگیری دانش آموزان و شیوه های آموزشی معلم استفاده کرد؛ زیرا امکان بازخوردهای فردی برای معلم و دانش آموزان را فراهم می کند. وقتی سن یادگیرنده ها کم است، زمانی که معلولیت دارند یا وقتی می خواهیم قدرت بیان و اعتماد به نفس آن ها را اندازه یه، آزمون های شفاهی مناسب هستند. یکی از مزایای پرسش های شفاهی این است که دانش آموز از معلم می خواهد نکات مبهم سؤال را برای او مشخص کند. یکی دیگر از موارد استفاده از پرسش شفاهی، کاربرد آن برای دانش آموزان معلول جسمی است که قادر به جواب های کتبی نیستند؛ لذا در چنین شرایطی بهترین روش، آزمون شفاهی میباشد.

مثال ۱۸: زمانی که سن یادگیرندگان پایین است، کدام یک از آزمون های زیر مناسب میباشد.

۱) کوته پاسخ

۲) آزمون شفاهی

۳) آزمون گسترده پاسخ

۴) آزمون عینی

پاسخ: گزینه «۲» آزمون های شفاهی برای یادگیرندگان، که سن پایینی دارند، مناسب است.

مثال ۱۹: کدام مورد زیر از مزایای آزمون های شفاهی به شمار می رود؟

۱) دانش آموز راحت تر می تواند پاسخ سؤالات را بیان کند.

۲) از طریق آنها بهتر میتوان توانایی تحلیل، ترکیب و ارزشیابی آزمون شوندگان را سنجید.

۳) طرح آنها برای معلم راحت تر است و از نظر اجرا از بقیه آزمون ها آسان تر هستند.

۴) دانش آموز میتواند از معلم بخواهد تا نکات مبهم سؤال را برای او مشخص کند.

پاسخ: گزینه «۴» از مزایای آزمون های شفاهی این است که دانش آموز می تواند از معلم بخواهد که نکات مبهم سؤال را برای او مشخص کند.

تست های طبقه بندی شده فصل هفتم

۱- در تصحیح اوراق انشایی، برای خنثی کردن اثر نحوه ی پاسخگویی به اولین سؤال بر نمره گذاری سؤالات بعدی، باید ..............

(سراسری ۸۲)

۱) اوراق را بدون اسم تصحیح کرد.

۲) الگوی دقیق پاسخ ها را فراهم آورد.

۳) اوراق را سؤال به سؤال تصحیح کرد.

۴) کل پاسخ را یکجا در نظر گرفت.

۲- در سؤال های نوع یادآوری، آزمودنی باید:

(آزاد ۸۲)

۱) پاسخ را مقایسه کند.

۲) پاسخ را بازشناسد.

۳) پاسخ را تشخیص دهد.

۴) پاسخ را خود، تولید کند.

۳- در ارتباط با تهیه و اجرای آزمون های تشریحی، کدام روش توصیه نشده است؟

(سراسری ۸۳)

۱) از کلماتی مثل «چه کسی»، « چه وقت» و «هرگز» استفاده کردن

۲) زمان لازم برای پاسخگویی به هر سؤال را جداگانه در نظر گرفتن

۳) حق انتخاب چند سؤال از بین تعدادی سؤال را به آزمودنی ها دادن

۴) به جای ورقه به ورقه تصحیح کردن، سؤال به سؤال تصحیح کردن

۴- پاسخ سؤالات انشایی، اصولا چگونه باید تصحیح شود؟

(سراسری ۸۳)

1) اوراق مختلف در زمان های مختلف و به دفعات، تصحیح شود.

۲) هر ورقه را حداقل سه ارزیاب تصحیح کنند و جداگانه نمره بدهند.

3) کل پاسخ یکجا خوانده شود و یک نمره کلی در نظر گرفته شود.

۴) پاسخ کلی به اجزای کوچکتر تقسیم و به هر یک نمره داده شود.

5- عینیت نمره گذاری سؤال های انشایی زمانی افزایش می یابد که:|

(سراسری ۸۳)

١) اوراق امتحانی را دیگران تصحیح کنند.

۲) حجم پاسخ های آزمودنی خیلی محدود شود.

۳) کل پاسخ ها یکجا در نظر گرفته شود.

۴) پاسخ ها با کلید از پیش تهیه شده، تصحیح شود.

6-کی و پژوهش ها نشان می دهد که وقتی معلمان تعداد زیادی از اوراق را تصحیح می کنند، معمولا به .................. اوراق، نمره ی بیشتری می دهند.

(سراسری ۸۴)

1) ثلث اول

۲) ثلث دوم

۳) ثلث آخر

۴) ثلث اول و آخر

7- ورقه ی تشریحی زیست شناسی و ۱۰۰ ورقه ی تشریحی شیمی را در اختیار تعدادی معلم گذاشته ایم و از آن ها خواسته ایم کیه اوراق را جداگانه تصحیح کنند. نتیجهی به دست آمده مطابق جدول زیر است. درباره ی اوراق زیست شناسی، « درصد عدم توافق» بین معلمان چقدر است؟

(سراسری ۸۴)

|  |  |  |  |
| --- | --- | --- | --- |
| نظر معلمان  اوراق | مردود توسط  همه معلمان | قبول توسط  همه معلمان | قبول توسط یک عده و مردود توسط عده ی دیگر |
| زیست شناسی | 25 | 35 | 40 |
| شیمی | 18 | 27 | 55 |

1) 25

2) 35

3) 40

4) 60

۸۔ چنانچه در آزمون های انشایی به دانش آموزان، شانس انتخاب داده شود تا از بین سؤالات به چند سؤال به دلخواه پاسخ دهند، این اقدام چگونه بر تفسیر نمره های آزمون تأثیر می گذارد؟

(سراسری ۸۵)

۱) تفسیر ارزشیابی ملاکی را دقیق تر می سازد.

۲) بر تفسیر ارزشیابی ملاکی نمره ها اثر می گذارد.

۳) بر تفسیر ارزشیابی هنجاری نمره ها اثر می گذارد.

۴) بر تفسیر ارزشیابی های هنجاری و ملاکی اثر می گذارد.

۹- در کدام نوع از سوالات زیر بهتر می توان عمق مطلب را تعیین کرد؟

(آزاد ۸۵)

۱) انشایی

۲) چند گزینه ای

3) شفاهی

۴) جورکردنی

۱۰- کدام گزینه جزء قواعد تهیه ی آزمون های انشایی به شمار نمی رود؟

(سراسری ۸۶)

۱) از کاربرد کلمات کجا، چه کسی و چه وقت در طرح سؤال ها بپرهیزید.

۲) از سؤالات بدیع و موقعیت های جدید استفاده کنید.

۳) به آزمون شوندگان، چند سؤال اختیاری بدهید.

۴) سؤال های آزمون، مستقیما به هدف های آموزشی مربوط شوند.

11- عینی بودن یک آزمون، بیشتر بر کدام مورد تأکید دارد؟

( سراسری 86)

1) نمره گذاری

2) تهیه آزمون

۳) هدف گذاری

4) اجرای آزمون

۱۲- برای اندازه گیری اهداف سطح ترکیب در حوزه شناختی، کدام یک از آزمون های زیر مناسب تر است؟

(آزاد 86)

۱) جورکردنی

۲) تشریحی

۳) چهارگزینه ای

4) عملکردی

۱۳- مهمترین ویژگی سؤالات عینی آن است که:

(سراسری 87)

۱) حالات روحی آزمودنی در پاسخ گویی اثری ندارد.

۲) نمره گذاری آن مستقل از نظر مصحح باشد.

۳) پاسخ های پیشنهادی در درجات مختلفی از صحت هستند.

۴) سؤالات آن معرف دقیقی از حیطه مورد سنجش باشد.

۱۴- آزمونی عینی نامیده میشود که .........

(سراسری 89)

۱) نتایج رفتار پاسخگو قابل مشاهده باشد.

2) نمره گذاری آن مستقل از نظر مصحح باشد.

۳) اطمینان از عملی بودن آنها وجود داشته باشد.

4) سوالات آن معرف دقیقی از حیطه مورد سنجش باشد.

۱۵- دادن حق انتخاب به آزمودنی ها (مثلاً پاسخ به ۵ سؤال از ۱۰ سؤال) چه مشکلی را ایجاد می کند؟

(سراسری 89)

۱) ارزیاب براساس نگرشی که به یک سؤال دارد، میزان سخت گیری متفاوتی خواهد داشت.

۲) داوطلبان به سؤالات متفاوتی پاسخ خواهند داد و امکان مقایسه آنها مشکل خواهد بود.

۳) سلیقه شخصی دانش آموزان وارد عمل خواهد شد و ارزیاب را تحت تأثیر قرار خواهد داد.

۴) داوطلبان فرصت طلب، سؤالات آسان را انتخاب خواهند کرد و موفق تر خواهند شد.

۱۶- کدام ویژگی در مورد امتحانات کتبی مصداق ندارد؟

(آزاد 89)

۱) انعطاف پذیری شرایط امتحان

2) صرفه جویی اقتصادی در جریان امتحان

۳) کمی تأثیر حالات روانی و آزمودنی بر نتایج امتحان

4) وسعت و گستردگی اندازه گیری توانمندی موردنظر آزمودنی

۱۷ـ اثر خطای هاله ای در تصحیح کدام سؤالات بیشتر است؟

(سراسری 93)

1) تشریحی

۲) کوتاه پاسخ

۳) چند گزینه ای

4) جور کردنی

۱۸ـ در خصوص سؤالات تشریحی، کدام مورد به عنوان عامل خطا محسوب می شود؟

(سراسری 94)

۱) تقسیم بارم هر سؤال به اجزای پاسخ

۲) تصحیح سؤال به سؤال برگه های امتحانی

3) پرهیز از نمره دهی براساس شناخت قبلی از آزمودنی ها

۴) اعطای حق انتخاب به آزمودنی ها برای پاسخ دهی به تعدادی سؤال

19-در صورتی که سؤالات آزمون ورودی کارشناسی ارشد رشته روانشناسی به صورت تشریحی طراحی شود، کدام مورد درست است؟

(سراسری 94)

1) نمره گذاری آن عینی خواهد بود.

2) روایی محتوایی آن پایین خواهد بود.

3) هدف های آموزشی بیشتر و دقیق تری را اندازه خواهد گرفت.

4) اجرای آزمون راحت تر خواهد بود.

پاسخ تست های طبقه بندی شده فصل هفتم

1- گزینه «۲» در این تست به قواعد تصحیح أوراق انشایی (تشریحی) اشاره شده است، اما باید به این نکته توجه نمود که سؤال به دنبال گزینه ای است که بر اثر جواب یک سؤال بر نمره گذاری سایر سؤالات را خنثی می کند. با توجه به این که این اثر بر کل آزمون تأثیر می گذارد، بسیار مهم تلقی می گردد.

گزینه ۳، صحیح است؛ زیرا زمانی که اوراق، سؤال به سؤال تصحیح می شوند، امکان مقایسه پاسخ دهندگان مختلف به یک سؤال فراهم شده است و کاربرد یک ملاک واحد و هماهنگ برای همه آزمون شوندگان فراهم می شود و امکان دخالت نظر مصحح در نتیجه کل آزمون داوطلبان کاهش می یاید. همانطور که اشاره شد، تمام گزینه ها به قواعد تصحیح اوراق تشریحی اشاره دارند که این موضوع، کار داوطلب را اندکی مشکل می کند و او باید با دقت به بدنهی سؤال و متن سؤال، به دنبال تشخیص جواب صحیح باشد.

2- گزینه «۴» در این گونه سؤالات آزمون شونده باید به بازخوانی آنچه قبل فرا گرفته است بپردازد و خود، به تولید پاسخ بپردازد.

۳- گزینه «3» اگر به آزمودنی در پاسخ به سؤالات تشریحی حق انتخاب سؤال داده شود، با ۳ مشکل روبه رو خواهیم شد:

الف) تهیه سؤال هایی که از نظر درجه دشواری هم سنگ باشند، کار آسانی نیست.

ب) آزمون شوندگان، توانایی انتخاب سؤال هایی که جواب آنها را بهتر می دانند، ندارند.

ج) از آنجا که آزمون شوندگان قوی تر، بیشتر به سراغ سؤال های دشوار می روند، ممکن است داشتن حق انتخاب، در نهایت به ضرر این گونه افراد تمام شود. توجه: با ایجاد امکان انتخاب از میان سؤالات، در واقع امکان مقایسه بین افراد را کاهش داده ایم.

۴- گزینه «۴» برای آنکه بتوان سؤالات انشایی را درست تصحیح نمود؛ قواعد نه گانه ای وجود دارد که گزینه ۴ یکی از این قواعد می باشد؛ پس پاسخ درست را از قبل به اجزا و آیتم های کوچک تقسیم کنید و از قبل آنها را نمره گذاری کنید تا هم یک الگو و نمونه به نام کلید داشته باشید و هم از دخالت عوامل نامربوط در تصحیح سؤال جلوگیری کنید.

5-گزینه «۴» در زمینه ی تصحیح سوال های تشریحی و انشایی، بهترین روش برای جلوگیری از دخالت نظر مصحح در نمره گذاری آن است که به سؤالات، از پاسخ داده، یک کلید تهیه نموده، به جزء جزء قسمت ها توجه کرده و نمره گذاری انجام شود تا دخالت نظر مصحح در نمره گذاری به حداقل ممکن برسد. |

6- گزینه «1» برای اینکه این مشکل را کاهش دهیم، باید به ۹ قاعده ی تصحیح اوراق تشریحی توجه داشته باشیم تا میزان دخالت نظر مصحح را در تصحیح اوراق به کمترین حد ممکن برسانیم؛ از جمله این قواعد می توان به تصحیح سؤال به سؤال هر کدام از اوراق اشاره کرد که باعث می شود مصحح نسبت به همه برگه ها یکسان عمل کند. در هر صورت در ابتدای تصحیح اوراق، ممکن است شرایط روحی و جسمی مصحح، بسیار زیاد در نمره گذاری دخیل شود.|

۷- گزینه «3»

درصد عدم توافق ۴۰ = ۶۰ – ۱۰۰ درصد توافق ۶۰ = ۳۵ + ۲۵

۸- گزینه «3» در ارزشیابی وابسته به ملاک یا ارزشیابی ملاکی، ملاک ارزشیابی از پیش تعیین می شود و عملکرد یادگیرنده با توجه به آن ملاک مورد قضاوت قرار می گیرد.

ارزشیابی وابسته به هنجار با ارزشیابی هنجاری، نوعی ارزشیابی است که به جای یک ملاک مطلق از پیش تعیین شده، به نوعی ملاک نسبی وابسته است.

در آزمون های انشایی، نباید به آزمودنی ها حق انتخاب چند سوال به دلخواه را داد؛ زیرا این کار، استفاده از ملاکی یکنواخت در تصحیح را از بین می برد و امکان مقایسه پاسخها با یکدیگر را سلب می کند (ارزشیابی هنجاری نمرات را مخدوش می کند).

9- گزینه «3» می توان سؤال شفاهی را نوعی سؤال تشریحی دانست که به صورت غیررسمی طرح می شود و جواب آن هم به صورت شفاهی در کلاس ارائه می شود. این نوع سؤالات از آنجا که همان سؤالات تشریحی هستند، می توانند توانایی ترکیب تحلیل و ارزشیابی را در آزمون شونده بسنجند؛ علاوه بر این، یک پرسشگر ماهر می تواند پاسخهایی را در دانش آموزان برانگیزاند که نشان دهنده فرایندهای فکری او باشند؛ بنابراین با این نوع سؤالات بهتر می توان تا عمق مطلب را تعقیب کرد.

۱۰- گزینه «۳» در آزمون های انشایی نباید به آزمودنی ها حق انتخاب چند سؤال به دلخواه را داد؛ زیرا این کار، استفاده از ملاکی یکنواخت در تصحیح را از بین می برد و امکان مقایسه پاسخ ها با یکدیگر را سلب می کند (ارزشیابی هنجاری نمرات را مخدوش می کند).

۱۱- گزینه «1» آزمون تستی - چند گزینه ای، متداول ترین آزمون عینی مورد استفاده در روانشناسی و آموزش و پرورش است. این نوع آزمون از لحاظ یکنواختی سؤال ها و همچنین سهولت تصحیح پاسخنامه، ارجح می باشد؛ بنابراین، تأکید عمده بر نمره گذاری آزمون است که نظر مصحح در آن دخالتی ندارد.

۱۲- گزینه «۲» در سؤال های جور کردنی و چهار گزینه ای، یادگیرنده هم سؤال و هم جواب را در اختیار دارد و فقط باید به بازشناسی آنچه قبلا آموخته است، بپردازد. آزمون های عملکردی نیز با مهارت سروکار دارند (مهارت در استفاده از فرایندها و شیوه های اجرایی و نیز مهارت در تولید فرآورده ها). سر آزمون شونده در پاسخ به سؤال های تشریحی آزاد است مطالبی را ارائه دهد که به نظر او منطقی و مفید هستند و مطالب خود را به گونه ای که می خواهد، است. سازمان دهد؛ همچنین، آزمون شونده باید اطلاعات و دانش های خود را ارزیابی کند و اندیشه های خود را به نحوی منطقی، منسجم و بدیع روی کاغذ بیاورد؛

بنابراین سؤال های آزمون های تشریحی گسترده پاسخ، برای سنجش هدفهای تحلیل، ترکیب و ارزشیابی مناسب ترین سؤال ها هستند.

۱۳- گزینه «۲» آزمون های عینی، از آن دسته آزمون هایی هستند که هم سؤال و هم جواب در اختیار آزمون شوندگان قرار می گیرد. این آزمون شوندگان هستند که بر حسب مهارت، معلومات، سرعت و تشخیص خود درباره پاسخ تصمیم می گیرند.

1- صحیح - غلط

این آزمون ها از سه دسته اصلی تشکیل شده اند 2- جور کردنی

3- چند گزینه ای

ویژگی مهم این سؤالات در این نکته نهفته است که در تصحیح جواب ها، نظر مصحح هیچ گونه دخالتی ندارد و جواب سؤال هنگام تهیه سؤال مشخص میشود.

۱۴- گزینه «۲» آزمون های پیشرفت تحصیلی کتبی را می توان به دو دسته تقسیم کرد:

الف) عینی (بسته پاسخ): ۱- چند گزینه ای ۲- جور کردنی ۳- صحیح - غلط ۴- کامل کردنی

ب) ذهنی (غیرعیتی)

آزمون های عینی (objective) آزمون هایی هستند که در آنها هم سؤالات و هم جواب ها در اختیار آزمون شوندگان قرار داده میشود و آزمون شوندگان درباره پاسخ های داده شده، تصمیم هایی می گیرند یا اعمالی را انجام می دهند.

ویژگی مهم این آزمون ها در این است که تصحیح جواب ها به طور کاملا دقیق و عینی انجام می گیرد و در آن نظر شخصی مصحح هیچ گونه دخالتی ندارد. آزمون چند گزینه ای، متداول ترین آزمون عینی است که به لحاظ یکنواختی سؤال ها و سهولت تصحیح، بهترین نوع آزمون های عینی تلقی میشود.

۱۵- گزینه «3» اگر به آزمودنی ها حق انتخاب چند سؤال را از مجموعه سؤالات بدهیم، داوطلبان به سؤالات مختلفی جواب خواهند داد و امکان مقایسه را از بین خواهند برد.

۱۶- گزینه ۳۰» امتحانات کتبی که شامل طیف وسیعی از انواع آزمون ها می شود، هم در شرایط اجرای آزمون انعطاف دارد و هم می توان با استفاده از آن ها طیف گسترده ای از توانمندی های آزمودنی را اندازه گیری کرد؛ علاوه بر این، آزمون های کتبی در مقایسه با آزمون های عملکردی، که عملکرد آزمون شونده را اندازه گیری می کنند، از نظر اقتصادی به صرفه تر هستند. یکی از ضعف های این گونه آزمون ها این است که حالات روانی آزمودنی می تواند بر نتایج آزمون تأثیر زیادی داشته باشد؛ به همین دلیل به معلمان توصیه می شود که هنگام برگزاری آزمون، شرایط روانی مناسبی را برای آزمون شوندگان فراهم آورند تا از اضطراب آنها جلوگیری کنند.

۱۲- گزینه «1» اثر خطای هاله ای، تحت تأثیر قرار گرفتن مصحح براساس ویژگی های آزمودنی است. در آزمون های تشریحی که پاسخ ها و معانی مختلفی ارائه می شود، در صورت وجود خطای هاله ای، یک سوگیری آشکار در تصحیح اتفاق می افتد؛ اما در آزمون های جور کردنی، چند گزینه ای و کوتاه پاسخ، به دلیل وجود کلید مشخص، سوگیری اتفاق نمی افتد.

۱۸- گزینه «۴» عامل خطا در سؤالات تشریحی:

1- اعطای حق انتخاب به آزمودنی ها برای پاسخ به تعدادی سؤال؛

۲- تصحیح براساس خصوصیات آزمودنی؛

۳- وابستگی به توانایی نویسندگی، فصاحت بیان، زیبانویسی و.. .

۱۹- گزینه «۲» در صورت تشریحی بودن سؤالات کنکور، روایی محتوایی سؤالات پایین خواهد آمد.

آزمون فصل هفتم

1- مناسب ترین آزمون برای سنجش بالاترین طبقات هدف های حوزه شناختی کدام است؟

1) تشریحی گسترده پاسخ

۲) کوته پاسخ

۳) عینی

۴) تشریحی محدود پاسخ

۲- کدامیک از ویژگی های آزمون های تشریحی نیست؟

1) تهیه آزمون های تشریحی از عینی آسان تر است.

2) این آزمون ها توانایی پاسخ دادن به سؤال ها را می سنجند، نه توانایی انتخاب پاسخ ها را

۳) این آزمون ها بر روش مطالعه یادگیرندگان تاثیر مثبت بر جای می گذارند.

۴) تصحیح جواب آن ها به طور کاملا دقیق و عینی انجام می گیرد.

3- بزرگترین امتیاز آزمون های تشریحی این است که:

۱) تهیه آنها آسان است.

۲) آنها موقعیت های عینی تری را نسبت به آزمون های عینی، به آزمون شوندگان عرضه می کنند.

۳) آزمون شوندگان را وادار می کنند تا اندیشه های خود را به طور منطقی و منسجم بیان کنند.

۴) تقلب را کاهش می دهند.

۴- در یک پاراگراف، تفاوت میان آزمون سنجش و اندازه گیری را توضیح دهید. این سؤال چه نوع سؤالی است؟ 1) تشریحی گستر پاسخ

۲) تشریحی محدود پاسخ

3) کوته پاسخ

۴) هیچ کدام

5- فقط در صورتی می توان به آزمون شونده جق، انتخاب تعدادی سؤال از میان سؤال های بیشتری را داد که هدف از آزمون، سنجش ......... باشد.

۱) توانایی یادآوری مطالب و سازمان دادن به آن ها

۲) مقایسه عملکرد آزمون شوندگان با یکدیگر

۳) سرعت جواب دادن به سؤالها

۴) مهارت در نوشتن و پروراندن مطالب

6- آزمون های گسترده پاسخ برای سنجش کدام مورد مناسب می باشند؟

۱) تحلیل و ترکیب

۲) کاربرد و تحلیل

۳) تحلیل، ترکیب و ارزشیابی

۴) کاربرد، تحلیل و ترکیب

۷- چرا در طرح سؤال های تشریحی، نباید به آزمون شوندگان حق انتخاب چند سؤال را از میان تعدادی سؤال داد؟

1) طرح سوال های هم سنگ از لحاظ دشواری، کار سختی است.

۲) آزمون شوندگان، توانایی انتخاب سؤال هایی که جواب آنها را بهتر می دانند، ندارند.

۳) چون حق انتخاب در نهایت به ضرر دانش آموزان ضعیف می باشد.

۴) گزینه های او ۲

۸- کدام عبارت در مورد قواعد تصحیح سؤال های انشایی صحیح نمی باشد؟

1) برای هر سؤال با نوشتن کلید، از دخالت عوامل نامربوط جلوگیری کنید.

۲) فقط خود معلم باید به تصحیح جواب های دانش آموزان بپردازد.

3) هنگام تصحیح از شناسایی نام صاحبان آن ها خودداری کنید.

۴) از روش های متنوع نمره گذاری استفاده کنید.

۹- عدم دقت در تصحیح سؤالات کدام نوع آزمون وجود دارد؟

1) کوته پاسخ

۲) شفاهی

3) تشریحی

4) عینی

10- کدام عبارت زیر صحیح می باشد؟

1) در تصحیح سؤالات آزمون های تشریحی باید از یک روش خاص استفاده کرد.

۲) همواره به نمرات سؤال هایی که قبلا تصحیح کرده اید، نگاه کنید.

۳) پاسخهای سؤالات را ورقه به ورقه، نه سؤال به سؤال، تصحیح کنید.

۴) از یکی دو نفر بخواهید تا پاسخ سؤال هایی را که تصحیح کرده اید، تصحیح کنند.

فصل هشتم

«آزمون های کوتاه پاسخ»

آزمون های کوتاه پاسخ، جد واسط بین آزمون های عینی و تشریحی هستند. این آزمون ها از این جهت به آزمون های تشریحی شبیه هستند که صورت سؤال را معلم تهیه می کند ولی پاسخ را دانش آموز تهیه می کند و از این جهت شبیه به آزمون های عینی هستند که جواب، مختصر و پاسخ، روشن است؛ اسم همچنین، نظر معلم در تصحیح دخالت ندارد و تصحیح آن با دقت زیادی انجام می گیرد.

سؤال های آزمون کوتاه پاسخ، که برای سطوح پایین حیطه شناختی یعنی دانش، فهمیدن و کار بستن مناسب است، از آزمون شونده می خواهند تا کلمه، عبارت، جمله، عدد یا علامتی را در پاسخ به یک سؤال یا تکمیل یک جمله بنویسد. آزمون های کوتاه پاسخ به سه دسته تقسیم می شوند:

1) پرسشی

۲) کامل کردنی

3) تشخیصی یا تداعی

آزمون های کوتاه پاسخ پرسشی

در شکل پرسشی، تنه ی سؤال به صورت پرسشی مطرح می شود و بهترین نوع آزمون های کوتاه پاسخ است؛ زیرا هم سؤال به راحتی تهیه میشود و هم این سؤال ها برای آزمون شوندگان راحت ترند و ابهام کمتری دارند. به عنوان نمونه: عمل غذاسازی در گیاهان سبز چه نام دارد؟

آزمون های کوتاه پاسخ کامل کردنی

در آزمون های کامل کردنی، سؤال به صورت یک جمله ناقص نوشته می شود که در آن یک جای خالی وجود دارد و آزمون شونده با پر کردن جای خالی به سؤال جواب میدهد.

به عنوان نمونه: عمل غذاسازی در گیاهان سبز را .............. می نامند.

آزمون های کوتاه پاسخ تشخیصی یا تداعی

در این شکل، صورت سؤال به صورت یک مطلب یا موضوع است و آزمون شونده باید پاسخ را تداعی کند. به عنوان نمونه: بعد از نام هر عنصر، علامت اختصاری آن را بنویسید.

الف) هیدروژن

ب) منیزیم

ج) کربن

همانطور که قبلا گفته شده از این نوع آزمون ها بیشتر برای اندازه گیری اهداف سطح پایین استفاده می شود؛ البته از این آزمون ها در مواردی مانند موارد زیر نیز می توان استفاده کرد:

الف) توانایی تفسیر ساده اطلاعات و کاربرد قواعد، مانند شمارش تعداد هجاهای یک کلمه؛

ب) توانایی حل مسائل عددی در ریاضیات و علوم؛

ج) توانایی دست کاری نمادهای ریاضی و حل کردن معادله های ریاضی و شیمیایی

مزایای آزمون های کوتاه پاسخ

1) تهیه این سؤال ها و تصحیح آنها آسان است؛

۲) اجرای آنها ساده است؛

۳) امکان تقلب را کاهش می دهد؛

۴) نسبت به آزمون های عینی جور کردنی و صحیح غلط، اطلاعات تشخیصی بیشتری را در اختیار معلم قرار می دهد؛

۵) در مقایسه با آزمون های عینی صحیح - غلط و چند گزینهای احتمال حدس کورکورانه را کاهش می دهد؛ در آزمون های صحیح - غلط، احتمال شانسی جواب دادن ۵۰ درصد و در چهار گزینه ای ۲۵ درصد است. معایب آزمون های کوتاه پاسخ

۱) چون پاسخ ها خیلی مختصرند، نمی توان آن ها را برای سطوح بالای حیطه شناختی استفاده کرد و کاربرد آنها به اندازه گیری هدف های سطح دانش، سمی و حداکثر فهمیدن و به کار بستن در موارد بسیار نادر محدود می شود.

۲) کاربرد این آزمون ها ممکن است باعث حفظ مطالب شود.

3) تصحیح این آزمون ها ممکن است به دقت و سرعت آزمون های عینی نباشد و دانش آموزان ممکن است به سؤال به صورت های مختلفی پاسخ دهند؛

در نتیجه مصحح باید تصمیم بگیرد که جوابها درست اند یا غلط یا نیمه درست. این امر ممکن است باعث عدم دقت در تصحیح اوراق امتحانی و کاستن از میزان پایایی آزمون شود.

آزمون های کوتاه پاسخ متشکل از مسائل عددی

این آزمون ها، از سؤال های مربوط به محاسبات عددی تشکیل می شوند و نوع ویژه ای از آزمون های کوتاه پاسخ هستند. مهمترین مزیت این آزمونها، آن است که می توان آن ها را در موقعیت های تازه به کار برد و فهم و کار بستن دانش آموزان را سنجش کرد.

مثال: در تساوی ۳۰ = ۶+ ۴x، مقدار X چقدر است؟

قواعد تهیه سؤال های آزمون های کوتاه پاسخ از نوع عددی عبارتند از :

1- تا حد امکان از اعداد ساده استفاده کنید؛ چون هدف این گونه آزمون های آزمون کردن فهم دانش آموز از جریان یا فرایند محاسبات است نه دقت در محاسبات عددی؛

٢- تا حد امکان مسائلی را مطرح کنید که جواب صحیح دارند؛

۳- میزان دقت مورد انتظار را تعیین کنید؛

۴- اگر لازم است آزمون شوندگان در پاسخ خود واحد اندازه گیری را نشان دهند، این مطلب را به آنان اطلاع دهید؛

5- در صورت امکان، مسائل پیچیده و چند مرحله ای را به تعدادی مسأله یک مرحله ای ساده تقسیم کنید؛

6- تا آنجا که امکان دارد میسائل عددی را به زبان ساده و مختصر بیان کنید.

قواعد تهیه سؤال های کوتاه پاسخ

۱) برای آزمون، جدول مشخصات تهیه کنید؛

۲) هر سؤال باید موضوع مهمی را شامل شود؛

۳) صورت سؤال را کامل و روشن بنویسید به طوری که به پاسخ مشخص و معینی نیاز داشته باشد؛

۴)صورت سؤال را از روی مطالب کتاب عینا نقل نکنید؛ زیرا ممکن است باعث تشویق دانش آموزان به حفظ کردن مطالب شود؛ -

۵) در سؤال هایی که پاسخهای آنها اعداد هستند، واحد مقیاس و میزان دقتی را که در محاسبات باید رعایت شود، مشخص کنید. این کار، تکلیف آزمون شوندگان را مشخص می کند و موجب صرفه جویی در دقت دانش آموزان می شود؟

۶) در سؤال های کوتاه پاسخ کامل کردنی، تنها کلمات و عبارات مهم را حذف کنید

۷) در سؤال های کامل کردنی که برای ارزشیابی پیشرفت تحصیلی به کار می روند، تعداد زیادی جای خالی قرار ندهید؛ زیرا ممکن است، سؤال مبهم و ناخوانا شود. گاهی می توان در سؤال های کوتاه پاسخ از نوع کامل کردنی، از دو جای خالی استفاده کرد، (آن هم در شرایطی که صورت سؤال طوری طرح شود که . هیچ گونه ابهامی در آن نباشد).

۸) جای خالی سؤال های کامل کردنی را تا آنجا که ممکن است در قسمت پایانی سؤال قرار دهید تا آزمون شونده پیش از رسیدن به جای خالی از موضوع مورد سؤال اطلاع کامل حاصل کند.

۹) جای خالی سؤال های کامل کردنی را همواره یک اندازه تعیین کنید تا باعث حدس جواب نشوند.

۱۰) تا حد امکان به جای سؤال های کامل کردنی از سؤالات پرسشی استفاده کنید؛ زیرا پاسخ دادن به آنها برای دانش آموزان آسان تر است، همچنین تهیه سؤالات پرسشی آسان تر است و ابهام کمتری دارند؟

۱۱) در سؤال های کامل کردنی از کاربرد اشارات دستوری و مواردی که جواب سؤال را مشخص می کنند، خودداری کنید؛ اشاراتی از قبیل زمان افعال و مفرد و جمع بودن آنها.

تستهای طبقه بندی شده فصل هشتم

۱- اگر سؤال کامل کردنی به صورت جمله پرسشی مطرح شود، آن را چه می نامند؟

(آزاد ۸۹)

۱) سؤال پرسشی

۲) سؤال انشایی

۳) سؤال غیر عینی

۴) سؤال کوته جواب

پاسخنامه تست های طبقه بندی شده فصل

۱- گزینه «۴» در آزمون های کوته پاسخ از نوع کامل کردنی، سؤال به صورت یک جمله ناقص نوشته میشود که در آن یک جای خالی وجود دارد و آزمـون شونده با پر کردن جای خالی به سؤال جواب می دهد؛ اگر همین جمله را به صورت استفهامی مطرح کنیم، به یک سؤال کوته پاسخ از نوع پرسشی تبدیل می شود بنابراین گزینه (۴) صحیح میباشد.

آزمون فصل هشتم

1- تفاوت آزمون های تشریحی با کوته پاسخ در این است که:

۱) تهیه آزمون های کوته پاسخ آسان تر از تشریحی است.

۲) در آزمون های کوته پاسخ، تصحیح پاسخ ها با دقت و عینیت بیشتری انجام می گیرد.

3) نسبت به تشریحی، امکان تقلب را کاهش می دهد.

۴) همه ی موارد

۲- « یک متر، چند سانتی متر است؟»، این سؤال چه نوع سؤالی است؟

1) تشریحی محدود پاسخ

۲) کوته پاسخ از نوع پرسشی

3) کوته پاسخ از نوع تشخیصی

۴) هیچکدام

۳- در آزمون های کوته پاسخ از نوع کامل کردنی، جای خالی باید در کدام قسمت سؤال قرار گیرد؟

1) تفاوتی نمی کند

۲) در وسط سؤال

۳) در اول سؤال

۴) در انتهای سؤال

۴- سؤال های کوتاه پاسخ مربوط به مسائل عددی، تا کدام سطح یادگیری را می توانند بسنجند؟

۱) به یاد آوردن

۲) ارزشیابی

۳) کار بستن

۴) فهمیدن

۵- امتیاز مهم آزمون های مسائل عددی این است که:

۱) نیاز به وقت زیادی ندارند.

۲) به سادگی تهیه می شوند.

3) هدف های طبقه های بالای حیطه شناختی را می سنجند.

۴) می توانند هدف های فهم و کاربرد را بسنجند.

6- شباهت آزمون های کوته پاسخ با آزمون های عینی در این است که:

1) جواب، مختصر و روشن است

۲) نظر معلم در تصحیح دخالت ندارد.

۳) دقت زیادی در تصحیح می باشد.

۴) هر سه مورد هر

۷- بهترین نوع آزمون های کوتاه پاسخ کدام است؟

۱) کامل کردنی

۲) تشخیصی

۴) گزینه های ۱ و ۳

۸- «بعد از نام هر استان نام مرکز آن استان را بنویسید». این سؤال در کدام دسته از آزمون های کوتاه پاسخ قرار می گیرد؟

1) پرسشی

۲) کامل کردنی ۳

) گزینه های ۱ و ۲

۴) تشخیصی یا تداعی

۹- کدام مورد از مزایای آزمون های کوتاه پاسخ نمی باشد؟

1) اجرای آنها آسان است.

۲) اطلاعات تشخیصی بیشتری را در اختیار معلم قرار می دهند.

٣) از آنها می توان برای سطوح بالای حیطه شناختی استفاده کرد.

۴) امکان تقلب را کاهش می دهد.

۱۰- پاسخ دادن به کدام نوع سؤال در آزمون کوتاه پاسخ، برای دانش آموز آسان تر است؟

۱) کامل کردنی

۲) پرسشی

۳) پرسشی

۴) هیچ کدام

فصل نهم «آزمون های صحیح - غلط و جور کردنی»

تعریف آزمون عینی

آزمون های عینی، آزمون هایی هستند که در آنها هم سؤال ها و هم جواب سؤال ها در اختیار آزمون شوندگان قرار می گیرند و دانش آموزان در مورد جواب های داده شده، تصمیم گیری می کنند. ویژگی مهم این آزمون ها این است که تصحیح جواب آنها به طور کاملا دقیق و عینی انجام می گیرد و نظر شخصی مصحح هیچ گونه دخالتی در آن ندارد؛ همچنین این آزمون ها، معمولا توانایی بازشناسی یا تشخیص را می سنجند؛ برخلاف آزمون های تشریحی که عموما توانایی بازخوانی را می سنجند. آزمون های عینی به سه دسته تقسیم می شوند: ۱) صحیح - غلط ۲) جور کردنی ۳) چند گزینه ای

تفاوت آزمون های عینی و غیرعینی

1- آزمون های عینی، آزمون هایی هستند که هم سؤال و هم جواب در اختیار آزمون شونده قرار می گیرد و آزمون شوینده باید دست به انتخاب بزند.

۲- تصحیح آزمون های عینی آسان تر و دقیق تر از غیر عینی است.

۳- آزمون های عینی، توانایی بازشناسی را می سنجند و آزمون های تشریحی و کوته پاسخ، توانایی بازخوانی را می سنجند.

مثال 1: خاصیت مشترک انواع سوال های عینی این است که:

۱) تصمیم اصلی در مورد جواب سؤال، مستلزم قضاوت ذهنی است.

۲) مرور زمان در سؤال تأثیر ندارد.

۳) جواب هر سؤال در موقع تهیه آن مشخص می شوند

۴) صجت جواب درجات مختلفی دارد.

پاسخ: گزینه «3» خاصیت مشترک انواع سؤال های عینی این است که جواب هر سؤال در موقع تهیه آن مشخص می شود.

مثال ۲: کدام مورد جزء آزمون های عینی نمی باشد؟

۱) جور کردنی

۲) تشخیصی

۳) چند گزینه ای

۴) صحیح - غلط

پاسخ: گزینه «۲» آزمون های عینی شامل ۱- جور کردنی ۲- صحیح - غلط ۳۔ چند گزینه ای است.

مثال 3: یکی از نقاط ضعف آزمون های عینی این است که:

1) عامل حدس در آن نقش تعیین کننده ای دارد

۲) عامل عینیت در آن به سادگی قابل حصول نیست.

3) عدم تناسب برای ارزیابی قدرت استدلال و سازماندهی مطالب

۴) عدم تناسب برای سنجش فهم و درک مطالب ها

پاسخ: گزینه «3» هر چند آزمون های عینی یکی از رایج ترین آزمون ها هستند، اما معایبی نیز دارند که یکی از آنها گزینه ی ۳ است. سایر معایب عبارتند از: ۱- سخت بودن مراحل تهیه سؤال ۲- سطوح بالای شناختی را اندازه گیری نمی کند (فقط در سطح دانش است) 3- استدلال و پروراندن مطالب را اندازه گیری نمی کند.

مثال ۴: یکی از مزایای آزمونهای عینی این است که:

1) برای ارزشیابی همه طبقات هدف های آموزشی، مفید است.

۲) بیشتر از آزمون های دیگر، مورد علاقه دانشجویان است.

۳) تهیه سؤالات این نوع آزمون ها از تهیه بقیه سوالات آسان تر است.

۴) نمونه بزرگتری از محتوای درس و هدف های آموزشی را ارائه میدهند.

پاسخ: گزینه «۴» آزمون های عینی هم سؤال و هم جواب را در اختیار آزمون شونده قرار میدهند؛ بنابراین برای سنجش توانایی بازشناسی یادگیرنده در سطح وسیع کاربرد دارند و به اهداف آموزشی وسیع تری می پردازند که یادگیرنده آن را فرا گرفته است و در این آزمون ها محک زده میشود.

آزمون های صحیح - غلط

این آزمون های ساده ترین نوع آزمونهای عینی هستند که در آنها تعدادی جمله داده می شود و شخص باید درست یا غلط بودن آنها را مشخص کند. سؤالهای صحیح - غلط، متداول ترین نوع سؤالهای عینی هستند و پرورش کاران از زمان سقراط تاکنون از آنها استفاده کرده اند.

انواع آزمون های صحیح - غلط

1- صحیح - غلط معمولی: یک جمله خبری، یک محاسبه عدی، یک معادله یا غیره به دانش آموز می دهیم و از او می خواهیم درستی یا نادرستی آن را تشخیص دهد.

به عنوان نمونه: جامعه عبارت است از گروهی از افراد، اشیاء یا حوادث که حداقل یک صفت یا ویژگی مشترک دارند. صحیح- غلط

۲- نوع بله - نه: در این حالت، یک جمله پرسشی به شخص داده می شود و او باید جواب بله یا نه را مشخص کند.

به عنوان نمونه: آیا آب در ۱۰۰ درجه سانتیگراد به جوش می آید؟ بله - نه

٣- نوع اصلاحی: یک جمله به دانش آموز داده می شود که در آن غلطی وجود دارد. او باید غلط را تشخیص دهد و درستش را بنویسد.

به عنوان نمونه: در مواد اهمی، رابطه بین جریان و ولتاژ، غیر خطی است.

در مثال بالا دانش آموز باید کلمه غیر خطی را خط بزند و به جای آن کلمه خطی را اضافه کند.

نکته: در نوع اصلاحی علاوه بر توانایی باز شناسی، توانایی باز خوانی را هم می سنجیم؛ در واقع این نوع سؤال، ترکیبی از نوع صحیح - غلط و نوع کوته پاسخ است. در نوع اصلاحی سوال صحیح – غلط بهتر است. در زیر بخشی از جمله که قرار است مورد قضاوت آزمون شونده قرار گیرد خط کشیده شود تا او بخش دیری از سوال را که مدنظر معلم نبوده است، مورد قضاوت قرار ندهد.

مثال ۵: این نوع سؤال صحیح - غلط، در واقع ترکیبی از سؤالات نوع صحیح - غلط و کوته پاسخ می باشد؟

1) بله - نه

۲) صحیح - غلط معمولی

۳) نوع اصلاحی

۴) موافق - مخالف

پاسخ: گزینه «3» آزمون نوع اصلاحی» از آزمون های صحیح - غلط، ترکیبی از سؤالات نوع صحیح - غلط و کوته پاسخ می باشد.

۴- نوع موافق و مخالف: جمله ای داده میشود و شخص باید موافقت یا مخالفت خود را با آن بیان کند.

به عنوان نمونه: دانش توسط خود شخص ساخته می شود. موافق / مخالف

5- نوع خوشه ای یا چند صحیح- غلط: در این حالت، تنه سؤال به صورت یک جمله ناتمام نوشته می شود و به دنبال آن چند جواب می آید و دانش آموز باید صحیح یا غلط بودن هر یک از پاسخ ها را مشخص کند. این نوع سؤال ها به معلم اجازه می دهند که در هر سؤال به جای یک مطلب، چندین مطلب بگنجاند و از این طریق، در جای آزمون و زمان خواندن آزمون به وسیله آزمون شوندگان صرفه جویی کند.

به عنوان نمونه

میانه:

الف) یکی از اندازه های گرایش مرکزی است. صحیح - غلظ

ب) نقطه ۵۰ درصدی را نشان میدهد. صحیح - غلظ

ج) میانه بیشتر از میانگین تحت تأثیر نمرات خیلی بزرگ یا خیلی کوچک قرار می گیرد. صحیح - غلط

نکته ۲: بهترین مورد استفاده از آزمون های صحیح – غلط ، سنجش دانش ها، وقایع و رویدادها است.

محاسن آزمون های صحیح - غلط

ا- سهولت و سادگی تهیه

۲- عینی بودن تصحیح جواب ها

۳- بیان موضوعات درسی در قالب عبارت های ساده

۴- فراوانی تعداد آن ها در هر جلسه امتحان

معایب آزمون های صحیح - غلط

1- اطلاعات جزیی و غیر مهم را می سنجند؛

۲- حدس زدن کورکورانه در آن ها دخالت زیادی دارد؛

۳- استفاده بیش از حد آنها باعث حفظ مطلب می شود

۴- سطوح پایین حیطه شناختی را می سنجند و توانایی های پیچیده سطح بالا را اندازه نمی گیرند.

ایل (۱۹۷۹) گفته است که ایرادهای بالا به طراحان سؤال های صحیح - غلط وارد است، نه به خود این سؤال ها. این سؤال ها، الزاما اطلاعات حفظی را مورد تأکید قرار نمی دهند، بلکه آن ها را می توان به گونه ای نوشت که توانایی های سطح بالای یادگیری را نیز بسنجند.

قواعد تهیه آزمون های صحیح - غلط

١- سؤالات باید تا حد امکان مختصر، ساده و روشن نوشته شوند؛ برای این منظور، در هر سؤال تنها یک مطلب مهم را سنجش کنید از کاربرد اصلاحات فنی و پیچیده پرهیز نمایید و ساختمان جمله را تا حد امکان ساده بنویسید.

۲- سؤال باید طوری نوشته شود که برای دانش آموزان کاملا آگاه، موضوع سؤال کاملا درست یا کاملا غلط آشکار باشد.

۳- از کاربرد کلماتی چون همه، بعضی وقتها، غالبأ، هرگز و از این قبیل باید پرهیز کرد. اگر بدون استفاده از این کلمات امکان نوشتن یک جمله صحیح یا غلط میسر نیست، بهتر است از نوع دیگری از سؤال ها استفاده شود. . ۴- باید سعی شود طول سؤالات صحیح و سؤالات غلط هم اندازه باشد.

5- تعداد سؤالات صحیح و سولات غلط تقریبا به یک اندازه باشند.

6- سؤال های صحیح - غلط را نباید از روی جمله های کتاب نوشت؛ چون این امر باعث می شود که دانش آموزان به حفظ کردن مطالب تشویق شوند؛ پس باید سعی کرد مطالب را با عبارات تازه و شیوه های نو نوشت؛ ۷- هر سؤال باید به یک هدف آموزشی مهم مربوط باشد.

آزمون های جور کردنی

این آزمون ها از سه قسمت تشکیل می شوند: 1) راهنما 2) فهرست پرسش ها 3) فهرست پاسخ ها

این آزمونها برای مقایسه مناسبند. در این آزمون ها باید سؤال ها و پاسخ ها متجانس و همگور باشند.

انواع آزمون های جور کردنی

جور کردنی معمولی

که در آنها معمولا تعداد پاسخ ها باید بیشتر از پرسش ها باشد.

به عنوان نمونه: راهنمایی: در جای خالی کنار ستون سمت راست، حرف اول نام کاشفان و مخترعان ستون سمت چپ را بنویسید. از هر نام می توانید بیشتر از یکبار استفاده کنید.

() ۱- پروتون الف) هاروی

() ۲- نوترون ب) فلمینگ

() ۲- پرتو ایکس ج) ماری کوری

() ۴- پنی سیلین د) رادفورد

() ۵- رادیوم ه) چادویک

ی) سالک

ز) دی باکی

ر) رونتگن

جورکردنی از نوع رده بندی این نوع از آزمون های

این نوع از آزمون جور کردنی، به رده خاصی از علوم اختصاص دارد. در جور کردنی از نوع رده بندی، تعداد پرسش ها می تواند بیشتر از پاسخ ها یـا مسـاوی آنها باشد. این نوع سؤال ها بیشتر برای سنجش توانایی انتقال یادگیری و توضیح و درک مطلب مناسب اند. یک تفاوت سؤال های جور کردنی رده بندی با سایر سؤالهای جور کردنی این است که در سؤالهای رده بندی، یک پاسخ را می توان بیش از یک در به کار برد، بنابراین تعداد پاسخها می تواند هـم انـدازه یا بیشتر از تعداد پرسشها باشد، ولی در سؤال های جور کردنی معمولی، تعداد پاسخ ها غالبا بیشتر از تعداد پرسش ها است.

به عنوان نمونه: راهنمایی: در جای خالی ستون سمت راست تعیین کنید که کدام فلز، کدام نافلز و کدام شبه فلز است.

( ) ۱- قلع الف) فلز

( ) ۲- سدیم ب) نافلز

( ) ۳- فلوئور ج) شبه فلز

( ) ۴- سیلسیم

( ) ۵ - اکسیژن

( ) ۶- ید

مثال ۶: این نوع سؤال های جورکردنی، بیشتر برای سنجش توانایی انتقال یادگیری، توضیح و درک مطلب مناسبند؟

۱) جورکردنی معمولی

2) جورکردنی از نوع رده بندی

۳) موافق ـ مخالف

۴) اصلاحی

پاسخ: گزینه «۲» سؤال های جور کردنی از نوع رده بندی، بیشتر برای سنجش توانایی انتقال یادگیری، توضیح و درک مطلب مناسبند.

مثال 7 : بهترین مورد استفاده از آزمون های جور کردنی چیست؟

۱) سنجش درک و فهم

۲) تجزیه و تحلیل

۳) به کار بستن تواناییها

۴) آزمون دانش، اصطلاحات تاریخ و رویدادها

پاسخ: گزینه «۴» بهترین مورد استفاده از آزمونهای جور کردنی، آزمون دانش، اصطلاحات تاریح و رویدادها می باشد.

نکته3:آزمون های جور کردنی برای آزمون دانش، اصطلاحات، تعاریف، تاریخ ها بر رویدادها بسیار مناسبند گوانلاند (۱۹۷۱) گفته است: « آنجا که بازده بادگیری بر توانایی تشخیص روابط بین امور تأکید دارد و می توان تعدادی پرسش و پاسخ متجانس یافت، «سوال های جور کردنی مهم ترین سوال ها هستند»

محاسن آزمون های جورکردنی

۱- تعداد زیادی سؤال را می توان طراحی کرد؛ زیرا پاسخ دادن به سؤالات وقت گیر نیست؛

۲، نسبت به صحیح ـ غلط، حدس زدن را کاهش می دهد؛

۳- برای سطوح بالاتر حیطه شناختی هم قابل استفاده است؛

۴- تصحیح آن بسیار آسان است.

معایب آزمون های جورکردنی

۱- اگر زیاد مورد استفاده قرار گیرند، یادگیری به سمت یادگیری تداعی می رود و سطحی می شود؛

۲- ممکن است در راهنمایی، ابهام وجود داشته باشد؛

۳- گاهی اوقات، پیداکردن مجموعه ای از پرسش ها، که بتوان برای آن ها مجموعه ای از پاسخ های همگون پیدا کرد، کار دشواری است؛

۴- طولانی بودن پرسش ها و پاسخ ها؛

مثال ۸: در یک سؤال جور کردنی، ۶ پرسش وجود دارد و باید با ۷ پاسخ که درباره ی دانشمندان، هنرمندان و مخترعان است، جور شود. عیب این سؤال کدام است؟ 1

) تعداد کم پرسش ها

۲) تعداد زیاد پرسش ها

۲) تجانس کم پرسش ها

۴) ناهمسانی تعداد پرسش ها و پاسخ ها

پاسخ: گزینه «3» سؤالات جور کردنی قواعدی دارند که یکی از آن ها همگون بودن پرسش ها و پاسخ ها و بالا بردن میزان تجانس آنهاست؟ هر بنابراین سؤال ها و جواب ها باید همه به یک مفهوم یا موضوع اشاره داشته باشند؛ در نتیجه به همان نسبتی که بخواهیم قدرت تمییز یا تشخیص یک سؤال را بالا ببریم، باید بر تجانس و همگونی عناصر پرسش و پاسخ آن بیفزاییم. گزینه ی ناهمسانی تعداد پرسش ها و پاسخ ها، می تواند داوطلب را دچار اشتباه کند و با توجه به ناهماهنگی تعداد پرسش برابر باشد؛ در صورت ناهماهنگی پرسش ها تعداد سن هاء برابر نیز می تواند باشد و حتی می تواند چند برابر پرسش، گزینه پاسخ در متن سؤال گنجانده شود؛ بنابراین، دقت داوطلب به انتخاب درست او کمک شایانی می کند.

قواعد تهیه سؤال های جورکردنی

.1- پرسش و پاسخ های متجانس یا همگون انتخاب کنید؛ یعنی این که مربوط به یک مفهوم، طبقه، یا زمینه باشند؛

۲- طول پرسش ها و پاسخ ها را کوتاه انتخاب کنید. فهرست کوتاه مطالب هم برای معلم بهتر است، هم برای دانش آموزان. فهرست کوتاه به معلم در تهیه مطالب متجانس کمک می کند و باعث افزایش دقت و تمرکز حواس دانش آموزان می شود؛ در

۳- در راهنمایی، اطلاعات لازم را برای جور کردن بنویسید؛

۴- همه پرسش ها و پاسخ های یک سؤال را در یک صفحه قرار دهید؛

۵- در صورت امکان، فهرست پاسخ ها را به طور منطقی مرتب کنید. برای مرتب کردن مطالب، نام ها را به ترتیب حروف الفبا، تاریخ ها را به ترتیب زمان و اعداد را از بزرگ به کوچک و بالعکس مرتب کنید؛

6- در فهرست پرسش های سؤال های جور کردنی، از کاربرد جملات ناتمام خودداری کنید؛ .

۷- پرسش ها را با شماره و پاسخ ها را با حروف نشان دهید؛

8- هر یک از پاسخ ها باید برای تمام پرسش های یک سؤال، درست جلوه کنند؛

9- از نوشتن سؤال هایی که پرسش ها و پاسخ های آنها کاملا با هم جور می شوند، بپرهیزید.

تست های طبقه بندی شده فصل نهم

۱۔ اگر طول پرسش ها و پاسخ های سؤالات جور کردنی کوتاه باشد، انجام دادن کدام یک از موارد زیر راحت تر خواهد شد؟

(سراسری 83)

1) تضمین روایی محتوا

۲) ازدیاد اعتبار آزمون

۳) رعایت تجانس دو ستون

۴) اندازه گیری اهداف سطح بالا

۲ـ در طرح سؤالات جور کردنی، تنظیم فهرست پاسخ ها به طور منطقی چه امتیازی دارد؟

(سراسری 83)

1) در وقت دانش آموز صرفه جویی می شود.

2) انسجام بین پاسخ های سؤالات تضمین میشود.

۳) پاسخ ها و پرسش ها با یکدیگر جور می شوند.

۴) یک پاسخ معین با چند پرسش جور می شود.

۳ـ از جمله قواعد تهیه سؤال های جور کردنی (Matching items) این است که:

(سراسری 84)

۱) در هر دو ستون، تنوع سؤالات در حد مطلوب باشد.

۲) تعداد پرسش ها و پاسخ های دو ستون برابر و یکسان باشد.

3) هر دو ستون از کلمات مترادف یا متضاد تشکیل شود.

4) پرسش ها یا پاسخ های یک سؤال، به یک موضوع مربوط باشند.

۴- یک آزمون از نوع صحیح ـ غلط، شامل دو سؤال است. احتمال اینکه در این آزمون، شانسی و تصادفی بتوان نمره کامل گرفت، چقدر است؟

(سراسری 84)

1)

2)

3)

4)

5- کدام نوع از سؤال های عینی برای اندازه گیری توانمندی یادآوری معانی لغات و نام ها مناسب است؟

(آزاد 88)

۱) سؤال های کامل کردنی

۳) سؤال های جورکردنی

۲) سؤال های صحیح ـ غلط

۴) سؤال های چندگزینه ای

۶- « سؤال ها باید با یکدیگر متجانس و تعداد پاسخ ها بیش از تعداد سؤال ها باشد.» مورد یاد شده به قواعد کدام سؤال اشاره دارد؟

(آزاد ۸۸)

۱) سؤال های کامل کردنی

2) سوال های صحیح - غلط

۳) سؤال های انشایی

۴) سؤال های چندگزینه ای

۷ـ در کدام یک از انواع سؤال ها، فراگیر می تواند پاسخ صحیح را با احتمال ۵۰ درصد و صرفاً براساس حدس انتخاب کند؟

(آزاد 89)

۱) صحیح – غلط

۲) چهار گزینه ای

۳) کامل کردنی

۴) جور کردنی

۸- در طرح سؤال های کامل کردنی .......

(آزاد 89)

۱) در هر سؤال می توان بیش از یک جای خالی در نظر گرفت.

۲)کلمه یا عبارت حذف شده باید از مفاهیم مهم و با ارزش جمله باشد.

۳) جای خالی را می توان در ابتدای جمله در نظر گرفت.

۴) پاسخ ها می تواند متعدد و یا دارای درجات مختلفی از صحت باشد.

۹- در کدام یک از سؤال های عینی تعداد پاسخ ها باید از تعداد سؤال ها بیشتر باشد؟

(آزاد ۸۹)

۱) کامل کردنی

۲) چند گزینه ای

۳) کوته جواب

۴) جور کردنی

پاسخنامه تست های طبقه بندی شده فصل نهم

1- گزینه «3» برای تهیه سؤال های جور کردنی، قواعد نه گانه ای وجود دارد که رعایت تجانس و همگون بودن پرسش ها و پاسخ های مربوط، یکی از آن ها می باشد؛ این امر یعنی همه پرسش ها و پاسخ ها به یک مفهوم، طبقه یا زمینه خاص مربوط می باشند.

۲- گزینه «1» همان طور که اشاره شد، سؤال های جور کردنی قواعد نه گانه ای دارد که یکی از آنها مرتب کردن و فهرست کردن پاسخ ها به طور منطقی می باشد؛ اگر مطالب به طور منطقی مرتب شده باشد، دانش آموزان بدون صرف وقت زیاد آن را پیدا خواهند کرد، در غیر این صورت ممکن است وقت زیادی برای یافتن هر پرش، به هدر برود.

۳- گزینه «۴» برای سؤال های جور کردنی، ۹ قانون وجود دارد که گزینه چهار بیان کننده اولین قاعده برای تهیه ی سوال های جور کردنی است. این قاعده به این موضوع اشاره دارد که باید پرسش ها و پاسخ های متجانس و همگون مطرح شوند؛ یعنی سعی شود پرسش ها و پاسخ های یک سؤال به یک مفهوم یا یک زمینه مربوط شود و مطالب پراکنده و غیرمرتبط در یک سؤال قرار نگیرد.

۴- گزینه «3» با توجه به اینکه آزمون 2 سؤال دارد، شانس انتخاب درست برای هر سؤال است؛ بنابراین شانس انتخاب درست و کامل می شود.

۵- گزینه «۱» آزمون های کوتاه پاسخ از مجموعه ای سؤال مختصر، که غالبا برای سنجش هدف های آموزشی سطح پایین طرح می شوند، تشکیل می یابد. سؤال های آزمون های کوتاه پاسخ از آزمون شونده می خواهند تا جمله، عبارت، کلمه، عدد یا علامتی را در پاسخ به یک سؤال یا تکمیل یک جمله بنویسد. آزمون های کامل کردنی را، که از نوع آزمون های کوتاه پاسخ هستند، می توان برای یادآوری معانی لغات و نام ها به کار برد؛ بنابر این گزینه ۱ از بقیه گزینه ها صحیح تر می باشد.

6- گزینه «۲» در سؤال های جور کردنی باید پرسش ها یا پاسخ های یک سؤال، همگی به یک مفهوم، طبقه، با زمینه مربوط باشند. فهرست های نامتجانس ا پرسش ها و پاسخ ها، تداعی های کلامی را اندازه می گیرند؛ در نتیجه، اطلاعات سطحی آزمون شوندگان را می سنجنده همچنین در این نوع سؤال ها، نباید هر , سؤال با یک پاسخ جور شود، بلکه باید تعداد پاسخ ها را بیشتر از تعداد سؤالات نوشت؛ اگر هر پرسش با یک پاسخ جور شود و آزمون شونده، پاسخ همه پرسش ها جز یکی از آنها را بداند، آخری را نیز خود به خود جواب خواهد داد؛ همچنین، اگر در یکی از پرسشها اشتباهی صورت بگیرد، در آخرین پاسخ هم این اشتباه صورت خواهد گرفت.

۷- گزینه «1» از آنجا که در سؤال های صحیح - غلط، آزمودنی باید از بین دو گزینه دست به انتخاب بزند، با احتمال ۵۰ درصد می تواند گزینه صحیح را انتخاب کند.

۸- گزینه «۲» در طرح سؤال های کامل کردنی باید در هر سؤال، تنها یک جای خالی در نظر گرفت، بهتر است جای خالی را در انتهای جمله در نظر گرفت، پاسخ ها نیز نباید متعدد و یا دارای درجات مختلفی از صحت باشند؛ همچنین عبارت حذف شده، باید از مفاهیم مهم و با ارزش جمله باشد.

۹- گزینه «۴» در سؤال های جور کردنی نباید هر سؤال با یک پاسخ جور شود بلکه باید تعداد پاسخها را بیشتر از تعداد سؤالات نوشت؛ چون در سؤال هایی که هر سؤال با یک پاسخ جور می شود، اگر آزمون شونده پاسخ همه پرسش ها به جز یکی از آنها را بداند، آخری را نیز خود به خود جواب خواهد داد؛ همچنین اگر در یکی از پرسشها اشتباهی صورت گیرد، در آخرین پاسخ هم این اشتباه صورت خواهد گرفت.

آزمون فصل نهم

۱- یکی از تفاوت های آزمون های تشریحی و عینی این است که:

1) آزمون های تشریحی، توانایی بازشناسی را می سنجند و آزمون های عینی، توانایی بازخوانی را

۲) آزمون های تشریحی، توانایی بازخوانی را می سنجند و آزمون های عینی، توانایی بازشناسی را

۳) آزمون های عینی، هم توانایی بازخوانی را می سنجند و هم توانایی بازشناسی را، أما آزمون های تشریحی فقط توانایی بازشناسی را می سنجند.

۴) هیچ یک از موارد.

2- ساده ترین شکل سؤالات عینی کدام است؟

1) جور کردنی

۲) چند گزینه ای

۳) صحیح - غلط

۴) گزینه های 2و 3

3- در تهیه سؤال های جورکردنی، کدام یک از موارد زیر پیشنهاد شده است؟

1) تعداد ماده های ستون پرسش از تعداد ماده های ستون پاسخ بیشتر باشد.

۲) تعداد ماده های ستون پاسخ و پرسش برابر باشند.

۳) تعداد ماده های ستون پرسش از تعداد ماده های ستون پاسخ کمتر باشد.

۴) تعداد ماده های ستون پاسخ و پرسش بنا بر ماهیت سؤال تغییر کند. |

۴- بهترین مورد استفاده از آزمون های صحیح - غلط کدام است؟

1) سنجش دانش وقایع و رویدادها

۲) سنجش درک و فهم

۳) قدرت تجزیه و تحلیل مسایل

۴) اندازه گیری توانایی کاربرد

5- ما در یک سؤال جورکردنی، هفت پرسش آمده است که باید با ۱۰ پاسخ درباره تاریخ ها، شهرها و کشورها جور شوند. اشکال این سؤال در چیست؟

1) تجانس زیاد پاسخ ها

۳) تجانس کنم پاسخ ها

۳) تعداد زیاد پرسش ها

۴) تعداد زیاد پاسخ ها

6-کدام گزینه جزء محاسن آزمون های صحیح - غلط نمی باشد؟

1) سهولت و سادگی تهیه

۲) عینی بودن تصحیح جواب ها

۳) بیان موضوعات در قالب عبارت های ساده

۴) کمی تعداد آنها در هر جلسه امتحان

7- در یکی از معایب اصلی آزمون های صحیح و غلط، گزینه ..................... می باشد.

1) سهولت تهیه

۲) عدم سنجش سطوح بالای حیطه شناختی

۳) زیادی تعداد در هر جلسه امتحان

۴) عینی بودن جواب ها

۸- در این نوع آزمون از کاربرد کلماتی مانند غالبأ، هرگز، بعضی وقت ها و .................... باید پرهیز شود.

1) جور کردنی

۲) تشریحی

۳) کوته پاسخ

۴) صحیح - غلط

۹- کدامیک از آزمون های زیر برای مقایسه مناسب می باشند؟

۱) صحیح - غلط

۲) جور کردنی

۳) کوته پاسخ

۴) گزینه های ۱ و 2

10-کدام مورد، جزء محاسن آزمون های جور کردنی محسوب نمی گردد؟

۱) تعداد زیادی سؤال را می توان طرح کرد.

۲) نسبت به سؤال های صحیح - غلط باعث افزایش حدس زدن می شود.

3) تصحیح آن ها آسان است.

۴) برای سطوح بالای شناختی قابل استفاده می باشند.

فصل دهم

«آزمون های چند گزینه ای»

آزمون های چندگزینه ای متداول ترین و بهترین نوع آزمون های عینی هستند.

مثال۱: کدام یک، بهترین و متداول ترین نوع آزمون های عینی می باشد؟

1) صحیح – غلط

۲) جور کردنی

۳) چند گزینه ای

4) کامل کردنی

پاسخ: گزینه «۳» آزمون های چند گزینه ای، بهترین و متداول ترین نوع آزمون های عینی هستند.

تعریف آزمون چند گزینه ای

آزمون چند گزینه ای، تعدادی سؤال دارد که هر یک از آن ها، از یک قسمت اصلی و تعدادی گزینه (پاسخ) تشکیل میشود و آزمون شونده از میان گزینه های پیشنهادی، گزینه صحیح را انتخاب می کند. قسمت های مختلف یک سؤال چند گزینه ای عبارتند از:

۱) قسمت اصلی یا تنه سؤال: در برگیرنده ی مسأله یا موضوعی است که سؤال، باید آن را اندازه بگیرد.

۲) قسمت درست یا پاسخ سؤال: یکی از گزینه های پیشنهادی که پاسخ صحیح سؤال است.

۳) گزینه های انحرافی: سایر گزینه ها به غیر از گزینه درست.

نکته: نقش گزینه های انحرافی، منحرف کردن پاسخ دهندگانی است که پاسخ درست سؤال را نمی دانند.

تنه سؤال ممکن است به صورت یک پرسش یا یک جمله ناتمام باشـد که پرسش، بهتر است و گزینه های پیشنهادی، جواب جمله انستفهامی یا تکمیل کننده ی جمله ناقص هستند.

تعداد گزینه های سؤال های چندگزینه ای بین ۲ تا ۵ متغیر است؛ هرچه تعداد گزینه ها بیشتر باشد، احتمال حدس زدن کمتر میشود، اما چون پیدا کردن، بیش از سه گزینه ی انحرافی کار دشواری است، آزمون های چهارگزینه ای، که در آنها احتمال حدس زدن جواب درست است، یعنی ۲۵ درصد از همه متداول ترند.

مثال ۲: در یک آزمون پنج گزینه ای، دانش آموزی به ۲۰ سوال کاملا تصادفی پاسخ داده است. احتمالاً چند سؤال را وی صحیح پاسخ داده است؟

1) 4

2) 5

3) 10

4) 15

پاسخ: گزینه «۱» باتوجه به تعداد گزینه های موجود، احتمال پاسخ در است. بنابراین تعداد سوالی که داوطلب می تواند به طور احتمالی پاسخ درست را انتخاب کند، 4 سوال است.

مثال ۳: احتمال این که بتوان با روش شانس و حدس کورکورانه به تمام سؤال های یک آزمون ۵ سؤالی دو گزینه ای، پاسخ درست داد، چقدر است؟

1)

2)

3)

4)

پاسخ: گزینه «۴» احتمال پاسخ در یک آزمون، که سؤال آن دو گزینه ای است، است؛ با این تعریف که سؤال ها دارای ۲ گزینه هستند. تعداد 5 سؤال مورد نظر است؛ بنابراین:

مثال ۴: در یک سؤال چهار گزینه ای، چند درصد احتمال دارد، جواب صحیح انتخاب نشود؟

1) ۵۰ درصد

2) ۲۵ درصد

۳) ۷۵ درصد

۴) ۴۵ درصد

پاسخ: گزینه «۲» احتمال عدم انتخاب پاسخ صحیح است، یعنی ۷۵ درصد.

مثال ۵: سؤال های چند گزینه ای دارای کدام یک از ویژگی های زیر هستند؟

1) ساده

۲) صرفه جویی در وقت و هزینه

۳) به سهولت به آن ها پاسخ داده میشود.

۴) از دقت خیلی زیادی برخوردار نیستند.

پاسخ: گزینه «٣» متداول ترین آزمون عینی، آزمون چند گزینه ای است که مورد استفاده در روانشناسی و آموزش و پرورش است. این آزمون می تواند در یک زمان محدود، با هزینه ای اندک، تعداد زیادی سؤال را در اختیار تعداد زیادی داوطلب قرار دهد و آن ها را مورد آزمون قرار دهد؛ بنابراین گزینه ی ۳ صحیح است. گزینه ی شماره ی2 یکی از ویژگی های گزینه ی 3 می باشد که نباید به اشتباه، توسط داوطلب انتخاب شود.

انواع آزمون های چندگزینه ای

1- نوع تنها گزینه درست: ساده ترین نوع سؤال های چندگزینه ای است که تنها یکی از گزینه ها صحیح است و بقیه گزینه ها کاملا غلط هستند.

2- نوع بهترین گزینه: همه گزینه ها درست اند اما یکی از همه درست تر یا بهتر است

۳- نوع منفی: تنه سوال به صورت منفی است و همه گزینه ها به جز یکی از آنها درست است.

سال های چند گزینه ای به این نمونه ها محدود نمی شوند و معلمان و طراحان می توانند با استفاده از خلاقیت خود، گونه های دیگری از این سؤالات را طراحی کنند.

نکته 2: وقتی پیدا کردن گزینه های انحرافی دشوار است، بهتر است از سؤالان معنی استفاده کنیم.

نکنه ۳: آزمون های چند گزینه ای قادر به اندازه گیری اکثر بازده های یادگیری از دانش آموزان تا ارزشیابی، هستند. در واقع، هر گونه دانش و فهمی را بتوان از طریق آزمون های عینی دیگر سنجید، از طریق آزمون های چند گزینه ای هم می توان سنجید.

مثال ۶: کدام مورد، ساده ترین نوع سؤال چند گزینه ای محسوب می شود؟

1) نوع بهترین گزینه

۲) نوع منفی

۳) نوع تنها گزینه درست

۴) گزینه های1و ۲

پاسخ: گزینه «3» ساده ترین نوع سؤال چند گزینه ای «نوع تنها گزینه درست» می باشد.

محاسن آزمون های چند گزینه ای

1- نسبت به سایر آزمون های عینی انعطاف پذیرترند. آنها علاوه بر دانش، توانایی، استدلال، قضاوت و بسیاری بازدههای مهم دیگر یادگیری را می سنجند.

٢- از طریق آن ها می توان تعداد زیادی سوال مطرح کرد؛ زیرا خیلی وقت گیر نیستند.

٣- نسبت به آزمون های صحیح - غلط، احتمال حدس کورکورانه را کاهش میدهند.

۴- آنها به سادگی قابل تصحیح هستند.

۵- معلمان و دانش آموزان، آن ها را به سایر آزمون ها ترجیح می دهند.

6- اگر گزینه های انحرافی، به درستی انتخاب شوند، منبع مناسبی برای تشخیص مشکلات دانش آموزان هستند.

مثال 7: برای اندازه گیری یک نمونه کامل از محتوای مطالب آموزش داده شده، کدام نوع آزمون مناسب تر خواهد بود؟

۱) جور کردنی

۲) چند گزینه ای

۳) گسترده پاسخ

۴) کوتاه جواب

پاسخ: گزینه «۲» آزمون چند گزینه ای، متداول ترین آزمون عینی مورد استفاده در روان شناسی و آموزش و پرورش است. این نوع آزمون ها از لحاظ یکنواختی سؤال ها، حساسیت کم در مقابل حدس زدن کورکورانه و تصحیح پاسخنامه ها، بهترین نوع آزمون عینی هستند که می توانند محتوای کاملی از مطالب مورد نظر را بسنجند.

معایب آزمون های چندگزینه ای

۱- ساختن این آزمون ها دشوار زیرا پیداکردن گزینه های انحرافی مشکل است.

۲ـ معلمان، معمولاً این آزمون ها را در حد سنجش اطلاعات جزیی به کار می برند.

3- در مقایسه با آزمون های صحیح ـ غلط، خواندن این آزمون ها و پیدا کردن گزینه ی درست، مستلزم صرف وقت زیادتری است؛ به ویژه اگر گزینه ها خیلی شبیه به هم ساخته شوند.

۴- زمانی که به پاسخ های غلط، نمره منفی داده می شود، افرادی که بیشتر ریسک می کنند و افرادی که در این سؤال ها خبره و ماهرند، بیشتر سود میبرند.

5- دانش آموزان قوی، که دارای دیدگاه واگرا هستند، بیشتر ضرر می کنند.

مثال 8: کدام مورد زیر، جزء معایب سؤال های چند گزینه ای است؟

۱) قدرت نگارش را مستقیما اندازه نمی گیرند.

۲) شیوه ی نگارش در نمره اثر می گذارد.

3) دانشجو می تواند پاسخ را عوض کند.

۴) به صورت کاملا ذهنی نمره گذاری می شوند.

پاسخ: گزینه «۱» آزمون های عینی هرچند از رایج ترین آزمون های روانشناسی و آموزش و پرورش است، اما دارای معایبی است که این معایب عبارتند از:

۱- آزمون سازی دشوار، که موجب می شود گنجاندن گزینه های انحرافی در متن سؤالات مشکل باشد.

۲- سؤالات چند گزینه ای در حد سنجش اطلاعات جزیی و کم اهمیت است.

۳- باتوجه به این که نمره منفی در این گونه سؤالات وجود دارد، آن دسته از امکان موفقیت دارند. .

۴ـ باعث هدر رفتن وقت دانش آموزان در مقایسه با آزمون های صحیح ـ غلط میشود.

در سؤالات تستی، قدرت تحلیل و بررسی دانش آموزان، مورد ارزیابی قرار نمی گیرد و پرورش و پختگی اطلاعات آزمـون شـونده از نظـرهـا دور می ماند؛ علاوه بر این، با دقت در گزینه ها متوجه می شویم که گزینه ۲ از ویژگی های آزمون های چند گزینه ای نیست، بلکه مربوط بـه آزمـون هـای تشریحی است. گزینه ۳ را نیز نمی توان به عنوان یک ضعف برای این گونه آزمون ها تلقی کرد، گزینه ۴ نیز اشتباه است؛ چون سؤالات چند گزینه ای، کاملاً عینی هستند نظر مصحح در نمره گذاری تأثیر نمی گذارد.

قواعد تهیه آزمون های چندگزینه ای

۱ ـ هر سؤال باید یک موضوع مهم یا یک هدف آموزشی را اندازه بگیرد.

۲- بیشتر از یک مسأله یا یک مطلب نباید در سؤال قرار گیرد؛ چون نه تنها سؤال را پیچیده می کند، بلکه اگر آزمون شونده در پاسخگویی به سؤال موفق نباشد، معلوم نیست که کدام یک از مطالب را ندانسته و علت عدم موفقیت وی در چه بوده است.

۳- سؤالات باید واضح و روشن و قابل فهم باشند؛ مگر در مواردی که سؤال برای اندازه گیری توانایی خواندن و فهمیدن با توانایی کلامی آزمون شوندگان فهمیدن با توانایی کلامی آزمون شوندگان نوشته شده باشد.

۴- از تکرار مطالب در گزینه ها خودداری شود.

5ـ مطالب اصلی سؤال باید در تنه سؤال گنجانده شود.

۶- همه گزینه های یک سؤال باید متجانس و به موضوع واحدی مربوط باشند. در غیر این صورت سؤال از حالت چند گزینه ای خارج میشود و هر یک از گزینه ها به صورت یک سؤال صحیح غلط در می آید.

۷- پاسخ درست یا باید، تنها پاسخ درست، باشد یا، درست ترین پاسخ، باشد.

8- برای اندازه گیری فرایندهای پیچیده ذهنی باید از موقعیت های تازه استفاده کرد تا فقط دانش سنجیده نشود.

۹- گزینه های انحرافی باید طوری نوشته شود که توجه آزمون شوندگان بی اطلاع را به خود جلب کند. این گزینه ها باید ظاهر منطقی داشته باشند و در یک بررسی، درست جلوه کنند. معلم می تواند از اشتباهات معمول دانش آموزان برای نوشتن گزینه های انحرافی استفاده کند.

۱۰ـ باید بین تنه سؤال و گزینه ها از لحاظ دستوری هماهنگی وجود داشته باشد؛ در غیر این صورت، آزمون شوندگان از این ناهماهنگی به عنوان یک راهنما استفاده می کنند و گزینه درست را حدس می زنند.

۱۱- از به کارگیری اشاره هایی دستوری نامربوط باید خودداری شود (اشاراتی مانند زمان افعال و یا مفرد و جمع بودن آنها).

۱۲ـ در سؤالات منفی باید کلمات منفی را برجسته کرد؛ مثلاً زیر آنرها خط کشید تا دانش آموز دچار اشتباه نشود.

۱۳- از نوشتن سوال هایی که متن سوال در آن منفی و گزینه ها هم منفی است، یعنی منفی مضاعف، باید خودداری کرد. سؤال هایی که چنین هستند، بسیار دشوارند و خواندن آنرها آزمون شونده را گیج و سردرگم می کند.

۱۴- تا حد امکان از کاربرد عبارت هایی «مثل همه آنچه در بالا گفته شد» و «تمامی موارد فوق» خودداری شود؛ زیرا اگر آزمون شونده بداند که دو گزینه صحیح است، به سمت پاسخ درست راهنمایی می شود.

15- سوالات را باید مستقل از هم نوشت. پاسخ دادن به یک سؤال نباید شرط لازم برای پاسخ دادن به سؤال دیگر باشد و جواب دادن به یک سؤال، نباید به پاسخ دادن به سؤال دیگر کمک کند.

۱۶- از طرح سؤالات گمراه کننده باید پرهیز شود؛ زیرا هدف اصلی، اندازه گیری میزان فهم و دانش آزمون شوندگان است.

۱۷- دو گزینه متضاد را، که یکی از آنها درست است، به کار نبرید؛ زیرا آزمون شونده، پاسخ خود را به یکی از این دو گزینه محدود می کند؛ بنابراین بهتر سمی است یا به طور کلی از کاربرد متضادها صرف نظر کنید، یا از دو زوج متضاد برای سؤال های چهار گزینه ای استفاده نمایید.

۱۸- طول گزینه درست در سؤال های مختلف باید متغیر باشد؛ چون اگر در چند گزینه به یک اندازه باشد، باعث راهنمایی یادگیرنده در پاسخ گویی به سؤال می شود.

۱۹- محل گزینه درست را باید به شکل تصادفی انتخاب کرد. ساده ترین راه برای انتخاب محل گزینه های درست، استفاده از صفحات کتاب است.

۲۰- برای هر سؤال، بین ۳ تا ۵ گزینه در نظر بگیرید. دانش آموزان کلاس دوم از عهده پاسخ دادن به سوال های سه گزینه ای برمی آیند. دانش آموزان کلاس سوم و چهارم می توانند سؤال های ۳ تا ۴ گزینه ای را جواب دهند و سؤالهای ۴ تا ۵ گزینه ای برای دانش آموزان کلاس ششم و بالاتر مناسب است.

۲۱- عبارت هایی نظیر «هیچ یک از موارد» یا «هیچ کدام از آنها» را زیاد به کار نبرید. استفاده از این گزینه ها مشکل حدس زدن را از بین نمی برد، بلکه تنها اثر آن را کاهش می دهد؛ ضمنأ فرایند رسیدن به جواب درست نیز شامل بازشناسی است، نه بازخوانی.

۲۲- در آزمون های تکوینی برای شناسایی مشکلات دانش آموزان از گزینه ی «نمی دانم» در گزینه ها استفاده کنید.

۲۳- در سؤال هایی که تنه آنها جمله ی ناتمام است، جای خالی باید در قسمت آخر جمله قرار گیرد.

۲۴ - تنها در صورت لزوم، از آزمون های چند گزینه ای استفاده کنید. مواردی یافت می شود که بهتر است از آزمون های دیگر به جای آزمون های چند گزینه ای استفاده کرد. برای مسایل محاسباتی ریاضی و علوم، آزمون های کوته پاسخ و برای سنجش عقاید و واقعیت ها، آزمون های صحیح - غلط، توصیه میشود. به طور کلی آن نوع سؤالی را به کار برید که بهتر از بقیه، هدف شما را به طور مستقیم و مؤثر سنجش می کند.

مثال ۹: در سؤالات چند گزینه ای:

۱) تعداد گزینه ها کم باشد.

۲) سؤالات باید راهنمای یکدیگر باشند.

۳) گزینه ها نباید شبیه به هم باشد.

۴) موضوع باید روشن و بدون ابهام باشد.

پاسخ: گزینه «۴» آزمون های عینی، از جمله آزمون های چندگزینه ای، قواعدی دارند که عبارتند از:

۱) هر سؤال باید موضوع مهم یا هدف آموزشی را اندازه بگیرد.

۲) بیشتر از یک مسأله یا یک مطلب در هر سؤال قرار نگیرد.

۳) سؤال ها کاملا روشن و قابل فهم باشد.

۴) از تکرار مطالب خودداری شود.

مثال ۱۰: قابلیت انعطاف کدام یک از انواع سؤال های عینی بیشتر است؟

۱) جور کردنی

۲) صحیح- غلط

۳) چند گزینه ای

۴) کامل کردنی

پاسخ: گزینه «3» سؤالات چند گزینه ای قابلیت انعطاف بیشتری نسبت به سؤالات عینی دیگر دارند.

مثال ۱۱: در یک سؤال چند گزینه ای باید:

1) قسمت عمده سؤال در گزینه ها باشد.

۲) یکی از دو واژه متضاد، گزینه درست باشد.

۳) قسمت عمده سؤال در متن آن عرضه شود.

۴) بین متن سؤال و گزینه هم تشابه ظاهری وجود داشته باشد.

پاسخ: گزینه «3» در یک سؤال چند گزینه ای باید قسمت عمده سؤال در متن آن عرضه شود.

نکته ۴: چون هیچ نوع آزمونی وجود ندارد که به تنهایی از عهده سنجش همه هدف های سطوح مختلف یادگیری برآید، بهتر است از آزمون های ترکیبی، شامل سؤال های مختلف تشریحی، عینی و کوته پاسخ، استفاده کنیم.

مثال ۱۲: برای کنترل عامل حدس در امتحانات تستی چه راه حلی پیشنهاد می کنید؟

1) استفاده از سؤال های آسان

۲) استفاده از سوالات دشوار

3) افزودن بر تعداد گزینه ها

۴) کم کردن تعداد سؤال ها

پاسخ: گزینه «٣» افزودن بر تعداد گزینه ها، یکی از راه هایی است که عامل حدس را در آزمون های عینی کاهش می دهد.

تست های طبقه بندی شده فصل دهم

1- در سؤال های چند گزینه ای، هر چه میزان همگنی گزینه ها زیادتر باشد، سؤال .................. است؟

(سراسری ۸۱)

۱) دشوارتر

۲) ساده تر

۳) مطلوب تر

۴) پیچیده تر

۲- کدام عامل، در افزایش حدس و گمان مؤثر است؟

(سراسری 88)

1) زیاد بودن تعداد سؤال ها

۲) کم بودن تعداد گزینه ها

3) ممیز بودن سؤال ها

۴) دادن وقت کافی برای پاسخ به سؤال ها

۳- در پاسخگویی به آزمون ها، کدام عامل در کاهش حدس و گمان تأثیر گذار است؟

(سراسری ۸۲)

1) دشوار بودن گزینه ها

۲) کاهش تعداد سؤال ها

۳) زیاد بودن تعداد گزینه ها

۴) قدرت تشخیص بالای سؤال ها

۴- سؤال های پیشنهادی در یک آزمون چند گزینه بهتر است به کدام یک از صورت های زیر باشد؟

(آزاد ۸۲)

۱) نامتجانس

۲) ناهمگن

3) موازی

۴) کاملا متجانس

5- کدام نوع سؤال عینی دارای قابلیت انعطاف بیشتری است؟

(آزاد ۸۲)

۱) چند گزینه ای

۲) کامل کردنی

۳) صحیح - غلط

۴) کوته پاسخ

۶- کدام جمله، تنه مناسب تری برای یک سؤال چند گزینه ای است؟

(سراسری ۸۳)

۱) اوراق سؤالات انشایی را باید بدون اسم تصحیح کرد.

۲) اوراق سؤالات انشایی را حتما باید ورقه به ورقه تصحیح کرد

3) در تصحیح اوراق سؤالات انشایی باید دو قاعده ................. رعایت شود.

4) در تصحیح اوراق تسؤالات انشایی کدام قاعده باید رعایت شود؟

۷- در طرح سؤالات چند گزینه ای، تنها محدودیتی که توصیه شده عبارت است از:

(سراسری ۸۳) :

۱) استفاده نکردن از ذوق و سلیقه شخصی

۲) نوشتن تنه سؤال به صورت جمله سؤالی

۳) استفاده نکردن از جملات منفی و دشوار

۴) حفظ ویژگی اصلی سؤال (تنه و تعدادی گزینه)

۸- رعایت کدام یک از موارد زیر در نوشتن سؤال های پاسخ بسته، ضروری است؟

(آزاد، ۸۳)

۱) سعی شود تا آنجا که میسر است سؤال ها مطابق با نظر یادگیرندگان باشد.

۲) در هر سؤال فقط یک مطلب مهم مورد ارزشیابی قرار گیرد.|

۳) آزمون یا سؤال ها بر اساس توانایی دانش آموزان زرنگ ساخته شود.

۴) در هر آزمون حداقل چند سؤال صحیح و غلط وجود داشته باشد.

۹- گزینه پیشنهادی نامناسب گزینه ای است که:|

(آزاد ۸۳)

۱) کلیه اعضاء گروه ضعیف آن را پاسخ درست بدانند.

۲) تمام افرادی که مطالب را یاد گرفته اند، آن را انتخاب کنند.

3) کلیه افراد ضعیف آن را به عنوان پاسخ درست انتخاب کنند.

۴) کلیه اعضای گروه ضعیف آن را به عنوان پاسخ درست انتخاب کنند.

۱۰- یک گزینه پیشنهادی درست که خوب تدوین شده است، گزینه ای است که تمام:

(آزاد ۸۳)

۱) اعضای گروه ضعیف آن را به عنوان پاسخ درست انتخاب کنند.

2) افرا گروه های قوی و ضعیف آن را به عنوان پاسخ درست انتخاب کنند.

۳) افراد ضعیف آن را به عنوان پاسخ غلط انتخاب کنند.

۴) افراد قوی آن را به عنوان پاسخ نادرست انتخاب نمایند.

۱۱- اگر دانش آموزی با حداقل شایستگی، یکی از گزینه های انحرافی یک سؤال را حذف کند، در آن صورت احتمال حدس پاسخ صحیح سؤال چقدر است؟

(آزاد ۸۴)

1)

2)

3)

4)

۱۲- کدام عبارت زیر در مورد آزمون های چند گزینه ای غلط است؟

(آزاد 84)

۱) جملات باید ادبی بوده رعایت قواعدی دستوری بشود.

۲) بین متن سؤال و گزینه ها نباید تشابه ظاهری باشد.

۳) دوای منفی به ویژه منفی مضاعف نباید بکار رود.

۴) قسمت عمده مطلب در متن سوال گنجانده شود.

۱۳- سؤالات عینی به کدام دسته از سؤالات زیر اطلاق می شود؟

(آزاد 85)

۱) سؤالات منطبق با دنیای واقعی

۲) سؤالاتی که نتایج آن را بتوان بخوبی مشاهده کرد.

۳) سؤالات شبیه رفتار عملی

4) سؤالات چند گزینه ای و جور کردنی

۱۴- در سؤالات چند گزینه ای بهتر است طول پاسخ ها به چه ترتیبی تنظیم شود؟

(آزاد 85)

۱) گزینه صحیح طول کوچکتر و گزینه های غلط طول بزرگتری داشته باشند.

۲) گزینه صحیح طول بزرگتری داشته باشد. .

3) به طور تصادفی طول گزینه ها بزرگ و کوچک شود.

۴) همه گزینه ها طول یکسانی داشته باشند.

۱۵- کدام یک از موارد زیر جزء قواعد تهیه آزمون های عینی نیست؟

۱) احتراز از منفی مضاعف

٢) مربوط بودن تمام گزینه ها به یک موضوع

3) نوشتن صورت سؤالات حتی المقدور به صورت مثبت

۴) گنجانیدن دو مطلب در هر سؤال میک

۱۶- نقش گزینه های انحرافی در سؤالات چند گزینه ای کدام است؟

(آزاد 86)

۱) جلب توجه فراگیران بی اطلاع

۲) افزایش دقت آزمون

۳) گمراه کردن آزمودنی های بی دقت

۴) افزایش اطمینان فراگیران مطلع

۱۷- کدام یک از عبارات زیر در مورد مقایسه سؤالات عینی و انشایی صحیح است؟

(آزاد 86)

1) سؤالات عینی برای اندازه گیری قوه استدلال فراگیران مناسب است.

۲) پایایی و دقت سؤالات عینی بیشتر است.

3) سوالات انشایی بخش وسیع تری از آموخته های فراگیران را اندازه گیری می کند.

۴) نتایج اندازه گیری آزمون های انشایی دقیق تر است.

۱۸- گزینه های گمراه ساز در آزمون چند گزینه ای باید دارای کدام خواص باشد؟

(سراسری 87)

۱) دانش آموزان ضعیف را جذب کند.

۲) دانش آموزان قوی را جذب کند.

۳) دانش آموزان متوسط بیشتر جذب شوند.

۴) همه دانش آموزان را جذب کند.

۱۹- مهمترین مزیت سؤالات چندگزینه ای نسبت به سؤالات تشریحی عبارت است از:

(سراسری 88)

۱) تعداد زیاد سؤالات

۲) عینی بودن نمره گذاری

۳) اقتصادی بودن

۴) سهولت اجرا

۲۰- در یک آزمون دوسؤالی چهار گزینه ای، احتمال این که بتوان با توسل به حدس، به هر دو سؤال پاسخ صحیح داد، کدام است؟

(سراسری 88)

1)

2)

3)

4)

21- سؤال عینی یعنی این که:

(آزاد88)

۱) ارزش گذاری یا تصحیح آن بر طبق نظر معلم انجام می گیرد.

۲) ارزش گذاری یا تصحیح آن براساس کیفیت پاسخ است.

۳) ارزش گذاری یا تصحیح آن در حضور دانش آموز انجام می گیرد.

۴) ارزش گذاری یا تصحیح آن به دور از قضاوت شخصی انجام می گیرد.

22- رعایت کدام یک از نکات زیر در تهیه و نوشتن سؤال های عینی لازم است؟

(آزاد88)

1) بکار بردن نشانه های هدایت کننده نظیر « همیشه»، «هرگز» و ...

۲) هر سوال دربرگیرنده یکی از هدف های مهم تدریس باشد.

۳) طرح سؤال های گمراه کننده و گول زننده

۴) هر سؤال به نحوی با سؤالی های دیگر ارتباط داشته باشد.

۲۳- سؤال های چندگزینه ای، کدام یک از توانمندی های فراگیران را ارزیابی می کند؟

(آزاد 88)

۱) بازشناسی

۲) یادآوری

۳) حافظه

4) مقایسه

۲۴- امتحان شونده در پاسخ به سؤال های عینی:

(آزاد ۸۹)

۱) تا حدی آزاد است.

۲) بطور کلی آزاد است.

۳) تا حدی مقید است.

۴) مقید است.

۲۵- اگر فراگیری، در یک امتحان ۸۰ سؤالی از نوع چهار گزینه ای، ۵۰ سؤال را درست و ۳۰ سؤال را غلط جواب داده باشد، نمره خام او در این امتحان با استفاده از فرمول تصحیح امتحانات عینی عبارتست از:

(آزاد ۸۹)

1) 50

2) 20

3) 40

4) 30

۲۶- از بکار بردن افعال و کلمات منفی در متن سؤال ها باید پرهیز شود زیرا:

(آزاد89)

1) ادراک سؤال های منفی بویژه مضاعف بسیار آسان است.

۲) سؤال های منفی معمولأ مطالب و مفاهیمی را که فراگیر میداند، اندازه گیری می کند.

۳) انسان های عادی و طبیعی به لحاظ روانی مثبت بین و مثبت خوان هستند.

۴) هدف معلمان از اجرای امتحان بیشتر اندازه گیری آنچه که فراگیر نمیداند، است.

۲۷- آزمونی را عینی می نامند که:

(آزاد 89)

1) سؤال های آن عینا از مواد درسی انتخاب شده است.

2) سؤال های آن با خواست دانش آموزان مطابقت داشته باشد.

3) معلم در انجام آزمون دخالت نداشته باشد.

۴) تصحیح سؤال های آن مستقل از نظر تصحیح کننده است.

۲۸- کدام عبارت، تندی مناسب تری برای یک سؤال چند گزینه ای است؟

(سراسری 90)

1) اوراق سؤالات انشایی را باید ورقه به ورقه تصحیح کرد.

۲) در تصحیح سؤالات انشایی کدام قاعده باید رعایت شود؟

3)در تصحیح اوراق سؤالات انشایی باید دو قاعدهی .................. و .................. رعایت شود.

۴) ارواق سؤالات انشایی را باید بدون توجه به نام صاحب برگه تصحیح کرد، زیرا ..................

۲۹- در کدام حالت، گزینه های انحرافی سؤال برای افراد بی اطلاع از جذابیت بیشتر برخوردارند؟

(سراسری 92)

۱) مرتب کردن آن ها به صورت تصادفی

۲) مرتب کردن آن ها با نظم و الگویی خاص

۳) نوشتن آن ها براساس اشتباه رایج دانش آموزان

۴) استفاده از قیود مطلق در تهیه آن ها

پاسخنامه تست های طبقه بندی شده فصل دهم

۱- گزینه «۱» تهیه سؤالات چند گزینه ای دارای قواعد ۲۴ گانه ای است که هشتمین قاعدهی آن به این موضوع اشاره دارد که همه گزینه های یک سؤال باید متجانس و مربوط به موضوع واحدی باشند. همه ی گزینه های یک سؤال باید از لحاظ محتوا، همگون یا متجانس باشند، نه این که هر یک از آنها به مطلب ر جداگانه ای اشاره کنند. اگر در زمان طرح سؤال، طراح به جای گزینه های انحرافی، از گزینه های کاملا شبیه و همگون با گزینه ی اصلی استفاده کند، آزمون بسیار دشوارتر خواهد بود. گزینه ۴ با توجه به ماهیت جواب درست، گزینه ای است که داوطلبان را دچار نوعی عدم تصمیم گیری جهت انتخاب گزینه صحیح می کند، اما با توجه به این که پیچیده بودن به دلیل دشواری موضوع رخ می دهد، می توان نتیجه گرفت که گزینه ی ۱ صحیح است.

۲- گزینه «۲» تعداد گزینه های سوال های چند گزینه ای (گزینه درست و گزینه های انحرافی) از ۲ تا ۵ گزینه متغیر است. از لحاظ نظری هر چه تعداد گزینه ها بیشتر باشد، امکان حدس زدن کم است، اما به سبب این که پیدا کردن بیش از سه پاسخ انحرافی کار دشواری است، آزمون های چهار گزینه ای که امکان حدس زدن جواب درست، یک در چهار، است، شهرت بیشتری کسب کرده اند. گزینه ی شماره ۴ یکی از نزدیک ترین جواب ها به گزینه ی صحیح مندی است؛ زیرا با توجه به بدنه ى سؤال و نوع سؤال درباره ی افزایش حدس و گمان، داوطلب به تحلیل سؤال دست می زند و در صورتی که از پاسخ درست مطمئن نباشد، به اشتباه، گزینه ی چهارم را برمی گزیند.

۳- گزینه «٣» قواعد تهیه سؤال های چند گزینه ای به ۲۴ قاعده میرسد. یکی از این دسته بندی ها به این نکته اشاره دارد که هر چه تعداد گزینه ها از لحاظ نظری بیشتر باشد، احتمال حدس زدن پاسخ سؤال کمتر خواهد بود و گزینه سوم به صراحت به این مطلب اشاره می کند. گزینه اول، که به دشوار بودن گزینه ها اشاره دارد، می تواند یکی از نزدیک ترین پاسخ ها به گزینه صحیح باشد، اما به دشواری گزینه ها در قواعد تهیه سؤال های چند گزینه ای، اشاره ای نکرده است و تنها، شکل گزینه ۱. به همراه بدنه سؤال، می تواند علت تشخیص نادرست باشد.

۴- گزینه «۴» از آنجایی که سؤالات متجانس در یک آزمون باعث افزایش پایایی می شود و پایایی نیز شرط روایی است، این گزینه صحیح می باشد. علاوه بر این، گزینه های هر سؤال نیز باید با یکدیگر متجانس باشند و یک چیز را بیان کنند، نه این که هر گزینه موضوعی را مد نظر قرار دهد.

5- گزینه «۱» سؤال های چندگزینه ای از آن جایی که شامل انواع مختلفی همچون، نوع تنها گزینه درست، نوع بهترین گزینه و نوع منفی هستند، انعطاف بیشتری نسبت به بقیه گزینه ها دارند. سؤال های کوتاه پاسخ، که سؤال های کامل کردنی نیز جزیی از آن ها به شمار می روند، از آزمون شونده میخواهند تا کلمه، عبارت، جمله، عدد یا علامتی را در پاسخ به یک سؤال یا تکمیل یک جمله بنویسند. در آزمون های صحیح - غلط نیز آزمون شونده باید صحیح یا غلط بودن یک جمله یا سؤال را مشخص کند؛ بنابراین سؤال های چندگزینه ای نسبت به بقیه گزینه ها از انعطاف بیشتری برای طرح سؤال برخوردارند. در این نوع از سؤال ها تنها کافی است که ویژگی اصلی سؤال چند گزینه ای (یعنی تنه سؤال و تعدادی گزینه) حفظ شود و بیش از این، رعایت هیچ محدودیتی ضروری نیست.

6- گزینه «۴» تنه سؤال: قسمت اصلی و متن اصلی سؤال را تشکیل می دهد و دربرگیرنده موضوع و مسأله اصلی سؤال است که مورد اندازه گیری قرار می گیرد. با توجه به گزینه ها، تنها گزینه چهار دارای این ویژگی می باشد گزینه های ۱ و ۲ به صورت اخباری بیان شده اند و هیچ گونه نشانه ای از این که سؤال باشند، در آن ها پیدا نمی شود. گزینه ۳ که ممکن است به اشتباه به عنوان گزینه صحیح انتخاب شود، بیشتر برای سؤال های کوتاه پاسخ از نوع کامل کردنی مناسب است.

۷- گزینه «۴» سؤال های چند گزینه ای دارای انواعی از قبیل تنها گزینه ی درست، بهترین گزینه و نوع منفی می باشند؛ علاوه بر این ها، که انواع متداول سؤال های چند گزینه ای هستند، معلمان و طراحان سؤال های چند گزینه ای می توانند گونه های مختلف دیگری را که خود مناسب می دانند، مورد استفاده قرار دهند. کافی است ویژگی اصلی سؤال چند گزینه ای (یعنی تنه سؤال و تعدادی گزینه) حفظ شود و بیش از این رعایت هیچ محدودیت دیگری ضروری نیست.

8-گزینه «۲» با دقت در گزینه ها متوجه می شویم که گزینه ۱ غلط است. از آنجا که هر کدام از یادگیرندگان ممکن است نظر خاصی در مورد نوشتن سؤال داشته سند باشند، نویسنده نمی تواند نظر تک تک آن ها را مد نظر قرار دهد. گزینه ۳ نیز به خاطر این که فقط دانش آموزان زرنگ را در نظر می گیرد، اشتباه است. گزینه ۴ نیز اشتباه است؛ چون، چه در سؤال های صحیح و غلط و چه در بقیه سؤال های بسته پاسخ، در هر سؤال فقط باید یک مطلب مهم مورد ارزشیابی قرار گیرد.

۹- گزینه «۲» این سؤال یک ایراد دارد و آن، این است که گزینه های ۱ و ۳ و ۴ همگی یک مطلب را بیان می کنند. به طور کلی گزینه پیشنهادی نامناسب، گزینه ای است که تمام افرادی که از موضوع مطلع هستند، آن را به عنوان پاسخ صحیح و افراد غیر مطلع، آن را به عنوان پاسخ غلط انتخاب کنند.

۱۰- گزینه «۳» به طور کلی یک گزینه پیشنهادی مناسب، گزینه ای است که تمام اعضای گروه قوی آن را انتخاب کنند و تمام اعضای گروه ضعیف، آن را به اشتباه، به عنوان پاسخ درست برگزینند، سؤالی که به این شکل باشد، دارای قدرت تمیز مثبت و کامل (۱) می باشد.

۱۱- گزینه «۱» از آنجا که سؤال چهار گزینه ای است، احتمال پاسخ صحیح دادن در کل است. وقتی که آزمون شونده یک گزینه را حذف می کند در واقع به یک سؤال سه گزینه ای تبدیل می شود؛ در چنین حالتی احتمال پاسخ صحیح دادن او است.

۱2- گزینه «۲» یکی از قواعد تهیه آزمون های چند گزینه ای این است که بین متن سؤال وگزینه ها باید تشابه ظاهری وجود داشته باشـد کـه ایـن مـورد درگزینه ی ۲ به صورت معکوس بیان شده است؛ بنابراین گزینه ۲ می باشد.

۱۳- گزینه «۴» از بین گزینه های موجود صحیح ترین گزینه، گزینه ۴ می باشد؛ البته سؤالات عینی، محدود به سؤالات چند گزینه ای و جورکردنی، نمی شـود و سؤالات صحیح - غلط را نیز شامل می شود.

۱۴- گزینه «۴» گزینه درست نباید درچند سؤال پشت سر هم از سایر گزینه ها بلندتر یا کوتاه تر باشد؛ زیرا ممکن است بـه آزمـون شـونده ـ صـرف نظـر از این که جواب درست سؤال را می داند یا نمی داند ـ در پیدا کردن گزینهی درست کمک کند.

۱۵- گزینه «۴» یکی از اصول مهم در تهیه آزمون های عینی (صحیح ـ غلط جور کردنی و چند گزینه ای) این است در هر سؤال باید یک مطلب اساسی گنجانده شود. با توجه به این که گزینه چهار، عکس این موضوع را مطرح کرده است، این گزینه پاسخ صحیح می باشد.

۱۶- گزینه «۱» هدف از گزینه های انحرافی در سؤال های چند گزینه ای این است که توجه آزمون شوندگان بی اطلاع از موضوع سؤال را به خود جلب کند. این گزینه ها باید ظاهری منطقی داشته باشند و در یک بررسی سطحی، درست جلوه کنند؛ یعنی نباید آشکارا غلط باشند، آن گونه که هرکسی بتواند غلط بودن آنها را تشخیص دهد.

۱۷- گزینه «۲» گزینه ۱ صحیح نمی باشد؛ چون این سؤالات انشایی هستند که برای اندازه گیری قوه استدلال فراگیران مناسب می باشند نه سؤالات عینی. گزینه ۴ صحیح نمی باشد؛ چون با توجه به این که در هر آزمون، چند سؤال انشایی به کار می رود، نمی توان سطح وسیعی از مطالب یک درس را مـورد آزمون قرار داد. گزینه ۱ نیز با توجه به این که در تصحیح سؤال های انشایی، نظر مصححان دخالت دارد و نسبت به سؤالات عینی از دقت کمتری برخوردارند، صحیح نمی باشد. تنها گزینه صحیح، گزینه ۲ می باشد؛ زیرا سؤالات عینی دارای پرسش های بیشتری می باشد و نظر مصححان در تصحیح آن ها دخالت ندارد؛ بنابراین پایایی و دقت بیشتری نسبت به سؤالات انشایی دارند.

۱۸- گزینه «۱» گزینه رافی، نقش منحرف کردن آزمون شوندگانی را برعهده دارد که پاسخ درست سؤال را نمی دانند؛ بنابراین دانش آموزی کـه بـدون آمادگی (ضعیف) در آزمون شرکت . کند، جذب گزینه انحرافی می شود.

۱۹- گزینه «۲» آزمون های تستی و چندگزینه ای متداول ترین آزمون های عینی مورد استفاده در روانشناسی و آموزش و پرورش هستند. علت اینکه آن ها بهترین آزمون های عینی هستند چیست؟ به دلیل آنکه برعکس آزمون های تشریحی، که نیاز به تحلیل و جمله بندی دارند و هر شخص پاسخ دهنده با ادبیات خاص به آنها پاسخ می دهد، ، آزمون های چندگزینه ای، تنها یک جواب مشخص دارند و نمره گذاری آنها بسیار راحت است.

20-گزینه «۴» با توجه به اینکه هر کدام از گزینه ها احتمال درست بودن دارند و ما دو سؤال چهارگزینه ای داریم، احتمال درست جواب دادن دو سؤال به طور حدسی به این ترتیب است.

۲۱- گزینه «۴» آزمون های پیشرفت تحصیلی را می توان به دو دسته عینی و ذهنی تقسیم کرد. در تصحیح سؤالات عینی نمی کند؛ به همین خاطر به آن ها عینی می گویند.

22- گزینه «2» طبق قواعد تهیه سوالات عینی، باید از طرح سوال هایی که در آن ها نشانه های هدایت کننده نظیر همیشه، هرگز و... استفاده شده است، همچنین از طرح سوال های گمراه کننده و گول زننده پرهیز کرد، سوال ها را مستقل از یکدیگر نوشت و هر سوال باید یکی از هدف های مهم تدریس را مدنظر داشته باشد.با دقت در گزینه ها در می یابیم که گزینه های 1 و 3و 4 به صورت عکس آمده اند؛ بنابراین گزینه 2 صحیح می باشد.

23-گزینه «1» با توجه به این که در سوال های چند گزینه ای، آزمون شونده فقط باید از بین گزینه های موجود یکی را انتخاب کند، تنها به بازشناسی آنچه قبلا آموخته است، می پردازد.

24- گزینه «4» آزمون های عینی، آزمون هایی هستند که در آن ها هم سوال و هم جواب در اختیار آزمون شونده قرار می گیرد، و او باید درباره جواب های ارائه شده، اعمالی را انجام دهد یا تصمیماتی را بگیرد.

25- گزینه «3» با توجه به فرمول زیر جواب صحیح به دست می آید.

26- هیچ کدام از گزینه ها صحیح نیست. به طور کلی باید از نوشتن سوال هایی که در آن ها هم متن سوال، هم گزینه ها منفی هستند، یعنی منفی مضاعف پرهیز کرد؛ چون خواندن آنها بسیار دشوار است و خواننده را گیج و سردرگم می کند.

27- گزینه «4» آزمون های پیشرفت تحصیلی را می توان به دو دسته عینی و ذهنی تقسیم کرد. در تصحیح سوالات عینی نظر شخصی مصحح دخالت ندارد؛ به همین منظور به آنها عینی می گویند.

28 گزینه «2» متخصصان آزمون سازی از جمله گرانلاند، پیشنهاد کرده اند که آزمون ساز تازه کار، بهتر است کوشش کند تا متن سوال های خود را در قالب جمله های پرسشی طرح کند؛ زیرا به این شکل، هدف سوال به طور آشکار تری بیان می شود.همچنین لین و گرانلاند گفته اند که نوشتن صورت سوال های چند گزینه ای به صورت پرسشی، آسان تر است و برای دانش آموزان کم سن طبیعی تر جلوه می کند. این گونه سوال ها به طور روشن تری مسایل را بیان می کنند. با این حال، چنان که طرح سوال در قالب جمله پرسشی، خیلی طولانی تر از طرح آن به صورت جمله نا تمام باشد، بهتر است از همان شکل جمله نا تمام استفاده کرد؛ بنابراین گزینه 2 از بقیه گزینه ها صحیح تر است.

29- گزینه «3» آزمون چند گزینه ای، شامل تعدادی سوال است که هر یک از آن ها از یک قسمت اصلی و تعداد گزینه (پاسخ) تشکیل می شود و آزمون شونده از میان گزینه های پیشنهادی، گزینه صحیح را انتخاب می کند. قسمت های مختلف سوال های چند گزینه ای عبارت اند از: قسمت اصلی یا تنه سوال، گزینه درست یا پاسخ درست و گزینه های انحرافی.

غیر از گزینه درست، تعدادی گزینه دیگر نیز برای هر سوال تهیه می شود که به آن ها گزینه های انحرافی می گویند. نقشه گزینه انحرافی، منحرف کردن آزمون شوندگانی است که پاسخ درست سوال را نمی دانند. گزینه ای انحرافی باید به گونه ای نوشته شوند که توجه آزمون شوندگان بی اطلاع از موضوع سوال را، به خود جلب کنند. این گزینه ها باید ظاهری منطقی داشته باشند و در یک بررسی سطحی، درست جلوه کنند. معلم می تواند از اشتباهات معمول دانش آموزان در کلاس برای نوشتن گزینه های انحرافی سوال های خود استفاده کند.

آزمون فصل دهم

۱ـ نقش گزینه های انحرافی این است که:

1) آزمون شوندگانی که پاسخ صحیح را میدانند گمراه کنند.

۲) هم آزمون شوندگان قوی و هم ضعیف را گمراه کنند.

3) آزمون شوندگانی که پاسخ صحیح را نمی دانند گمراه کنند.

۴) هیچ کدام از موارد

۲- در آزمون های چندگزینه ای، هرچقدر تعداد گزینه ها بیشتر باشد، امکان حدس زدن ............

۱) افزایش می یابد

۲) کاهش می یابد

۳) افزایش می یابد، در صورتی که زمان آزمون طولانی باشد.

۴) کاهش می یابد، در صورتی که زمان آزمون کوتاه باشد.

۳ـ در آزمون های چندگزینه ای چه وقت بهتر است از سؤالات منفی استفاده شود؟

۱) وقتی تهیه گزینه های انحرافی غلط که درست جلوه کند، دشوار باشد.

۲) وقتی تعداد سؤالات آزمون زیاد باشد.

۳) وقتی تهیه گزینه های انحرافی آسان باشد.

۴) گزینه های ۲ و ۳

4- کدام یک از موارد زیر از معایب آزمون های چندگزینه ای نیست؟

۱) اغلب برای سنجش اطلاعات جزیی و کم اهمیت به کار می روند.

۲) این آزمون ها نسبت به آزمونه ای صحیح - غلط، احتمال حدس کورکورانه را افزایش میدهند.

۳) ساختن این آزمون ها بسیار دشوار است.

۴) در مقایسه با صحیح ـ غلط به وقت بیشتری نیاز دارند.

5- بهترین منبع برای انتخاب گزینه های انحرافی کدام است؟

۱) تجارب معلم

۲) غلط های کتاب درسی

۳) کتاب راهنمای معلم

4) اشتباهات دانش آموزان

۶ـ وقتی پیدا کردن گزینه های انحرافی دشوار است، بهتر است از چه نوع سؤالاتی استفاده شود؟

۱) جور کردنی

2) صحیح - غلط

۳) منفی

4) اصلاحی

7- کدام نوع از آزمون های عینی از لحاظ شمول سنجش، سایر آزمون ها را در برمی گیرد؟

۱) جور کردنی

۲) چند گزینه ای

۳) صحیح ـ غلط

4) مخالف و موافق

۸- هدف اصلی از طرح سؤالات چند گزینه ای چیست؟

۱) گمراه کردن آزمون شوندگان

۲) اندازه گیری میزان فهم و دانش آزمون شوندگان

۳) شناسایی مشکلات دانش آموزان

۴) گزینه های ۱ و ۳

۹- بهترین راه شناسایی مشکلات دانش آموزان در آزمون های تکوینی استفاده از ........ می باشد.

۱) هرگز

۲) نمی دانم

3) نباید

4) اگر

۱۰- بهترین راه سنجش همه هدف های سطوح مختلف یادگیری، استفاده از آزمون های ........ می باشد.

۱) تشریحی

۲) کوته پاسخ

۳) عینی

۴) همه موارد

فصل یازدهم

«روش های سنجش فرایندها و فرآورده های یادگیری غیرشناختی»

آزمون های کتبی یا مداد - کاغذی که در فصول قبل توضیح داده شد، عمدتا برای اندازه گیری هدف های آموزشی حوزه ی شناختی مناسب اند. از آنجا که هدف های حوزه شناختی، بخش مهمی از کلیه درس ها و موضوع های درسی همه دوره های تحصیلی را شامل می شوند، آزمون های کتبی، شهرت زیادی کسب کرده اند، ولی برای حوزه های روانی - حرکتی و عاطفی به تنهایی کافی نیستند و برای سنجش هدف های آموزشی این حوزه ها باید از تدابیر و وسایل دیگری استفاده کنیم. این وسایل عبارتند از: .

۱- آزمون های عملکردی

۲- روش های سنجش مشاهده ای

۳- روش های ویژه ی سنجش هدف های حوزه ی عاطفی

آزمون های عملکردی

آزمون های عملکردی، که به آنها آزمون های واقعی یا اصیل هم گفته می شود، فرایندها و فرآورده های یادگیری دانش آموزان و دانشجویان را به طور مستقیم سنجش می کنند.

فرایندهای عالی فکری و مهارت های تفکر انتقادی را با ارزشیابی از عملکرد و محصول عملکرد، بهتر می توان سنجید؛ همچنین از طریق آزمون های سی عملکردی می توانیم، مهارت های شناختی سطح بالا و تکالیف عملکردی مهم مانند توانایی نوشتن، ایجاد ارتباط های کلامی و حل مسأله را سنجش کنیم.

آزمون های عملکردی با مهارت سر و کار دارند؛ مهارت در استفاده از فرایندها و شیوه های اجرایی و نیز مهارت در تولید فرآورده ها.

مهارت در حل مسأله ی ریاضی و مهارت در برقرار کردن ارتباط در دروس فارسی و زبان خارجی، نمونه هایی از مهارت در استفاده از فرایندها و شیوه های اجرایی هستند. دروس هنر و موسیقی، آموزش صنعتی و تجاری، نمونه هایی از مهارت در تولید فرآورده ها هستند. سر به طور کلی هدفهای آموزش و یادگیری را می توان به دو دسته تقسیم کرد:

١- هدف هایی که به دانستن درباره ی امور مربوط می شوند.

٢- هدف هایی که به دانستن درباره ی نحوه انجام دادن عمل یا کار مربوط می شوند.

- آن دسته از وسایل اندازه گیری که می کوشند. تا دامنه و صحت دانش شخص را بسنجند، آزمون های کتبی نامیده می شوند.

- آن دسته از وسایل اندازه گیری که می کوشند بسنجند، چگونه کسی می تواند کاری را انجام دهد، آزمون های عملکردی نام دارند. آزمون های عملکردی با فرایند یا شیوه اجرا و فرآورده یا ترکیبی از این دو سروکار دارند. فرآورده چیزی است که یادگیرنده تولید می کند و فرایند، جریانی است که او مورد استفاده قرار می دهد تا به فرآورده دست یابد. در بسیاری از موارد، هم فرایند و هم فرآورده، هر دو، به عنوان جنبه های مهم عملکرد سنجش می شوند. عملکرد مورد سنجش مشخص می کند که تأکید بر فرایند است یا بر فرآورده ها.

در انجام کارهایی مثل سخنرانی با نواختن آلات موسیقی، تاکید فرایند و در کارهایی مثل نقاشی و شعر، تأکید بر فرآورده است.

مثال 1: در این نوع آزمون ها فرایند و فرآورده های یادگیری دانش آموزان و دانشجویان به طور مستقیم سنجش می شود؟

۱) کتبی

۲) سنجش مشاهده ای

۳) عملکردی

۴) گزینه های1و ۳

پاسخ: گزینه «3» در آزمون های عملکردی، فرایندها و فرآورده های یادگیری دانش آموزان به طور مستقیم سنجش میشود.

ویژگی های آزمون های عملکردی

١- تأکید بر کاربست: یعنی آیا یادگیرنده علاوه بر دانستن، می تواند دانش را به کار ببندد؟

۲- تأکید بر سنجش مستقیم: یعنی سنجش هدف ها به طور مستقیم به جای غیرمستقیم.

۳- استفاده از وسایل واقعی: استفاده از مسائل و موقعیت هایی که در زندگی عادی یافت می شوند یا شبیه به آن ها هستند.

۴- ترغیب و تشویق تفکر بازی تشویق به پیدا کردن راه حل های مختلف برای مسائل و درگیر شدن با مسائل به طور گروهی.

آزمون های عملکردی به چهار دسته تقسیم می شوند:

1) آزمون های کتبی عملکردی بیشترین فاصله را با عملکرد در زندگی واقعی دارد

۲) آزمون های شناسایی

۳) انجام عملکرد در موقعیت های شبیه سازی شده

۴) نمونه کار کمترین فاصله را با عملکرد در زندگی واقعی دارد

در این، تقسیم بندی، آزمون، نوع اول، یعنی آزمون کتبی عملکردی، بیشترین فاصله را با عملکرد در زندگی واقعی دارد، ولی آخرین نوع آزموں، یعنی نمونه کار، کمترین فاصله را با عملکرد در زندگی واقعی دارد. البته، این روش ها را هم می توان جداگانه به کار بست و هم با یکدیگر مورد استفاده قرار داد.

1- آزمون کتبی عملکردی: تفاوت بین آزمون های کتبی عملکردی با آزمون های کتبی فصول قبل در این است که، آزمون های کتبی عملکردی عمدتا بر کاربست دانش و مهارت شهر موقعیت های عملی با شبیه سازی شد، تأکید می کند آزمون های عملکردی یا باز نمونه های پایانی، مراحل میانی عملکردی سیر که برای رسیدن به بازده ای مطلوب پایانی ضروری هستند. همچنین یادگیری را سنجش می کنند. آزمون های عملکردی را می توان به عنوان وسیله سنجش اطلاعات مقدماتی فرد، درباره کاری که انجام خواهد داد به کار برد؛مثلا اطلاعات فرد از ساختار ابزاری که مورد استفاده قرار خواهد ، پیش از این که عملا شروع به کار کردن با دستگاه کند. اندازه گیری شود. این نوع اطلاعات، به صورت کتبی و در صورت لزوم به صورت گروهی، قابل سنجش هستند. ۲- آزمون شناسایی آزمون سامانی: روشی است که توانایی یادگیرنده را در تشخیص ویژگی های امور و تشخیص نقاط قوت و ضعف آنها سنجش می کند. این آزمون ها، در آموزش های صنعتی، علوم زیستی و سایر دسته های علمی کاربرد دارد. مثلا از یادگیرنده می خواهیم تا با گوش کردن صدای یک دستگاه معیوب، مثل موتور اتومبیل، از روی صدا مشکل را تشخیص دهد. یا از آنها بخواهیم جانوران مختلفی را که در آزمایشگاه وجود دارند شناسایی کنند. از آزمون های شناسایی می توان به عنوان وسایل سنجش غیرمستیم عملکرد استفاده کرد، مثلا از یک مترجم ماهر انتظار می رود که اطلاعات , بیشتری در مورد ترجمه، نسبت به یک مترجم تازه کار داشته باشد؛ بنابراین از آزمون های شناسایی می توان، برای ترند کردن تعداد زیادی از داوطلبان مشاغل استفاده کرد. گرانلاند (۱۹۸۸) در رابطه با آزمون های شناسایی گفته است «معمول تر، آن است که از آزمون های شناسایی به عنوان یک تدبیر آموزشی برای آماده کردن دانش آموزان و دانشجویان در اجرای عمل در موقعیت های واقعی یا شبیه سازی شده، استفاده شود. »

۳- انجام عملکرد در موقعیت های شبیه سازی شده: در انجام عملکرد در موقعیت های شبیه سازی شده، از یادگیرنده خواسته می شود تا در یک موقعیت شبیه سازی شده یا متنوعی، همان اعمالی را انجام دهد که در موقعیت های واقعی ضروری هستند.

۴- نمونه کار: در این روش از یادگیرنده می خواهیم به انجام اعمالی بپردازد که معرف عملکرد واقعی مورد سنجش هستند، به این خاطر، این روش نزدیک ترین روش سنجش به عملکرد واقعی یادگیرنده در محیط های طبیعی است، مثلا" از یک دانشجوی تربیت معلم می خواهیم درسی را به گروهی از دانش آموزان آموزش دهند. در روش نمونه تازه نمونه اعمالی که می خواهیم، یادگیرنده انجام دهد باید شامل عناصر مهم عملکرد کلی باشد که در شرایط کنترل شده اجرا می شود. روش نمونه کاره برای سنجش توانایی افراد در انجام کارهای مختلف، از دیرباز، در صنعت مشاغل به فعالیت های عملکردی متداول نیاز داشته است و می توان از آن ها در محیط های آموزشی نیز حداکثر استفاده را نیز به عمل آورد.

مثال ۲: کدام یک از آزمون های عملکردی، بیشترین فاصله را با عملکرد در زندگی واقعی دارند؟

1) آزمون های شناسایی

۲) نمونه کار

۳) آزمون های کتبی عملکردی

۴) انجام عملکرد در موقعیت شبیه سازی شده

پاسخ: گزینه «3» آزمون های کتبی عملکرد بیشترین فاصله را با عملکرد در زندگی واقعی دارند.

مراحل تهیه آزمون های عملکردی

1- بازدهی مورد نظر عملکرد را مشخص کنید.

۲- موقعیت آزمون را واقع بینانه برگزینید. .

۳- از راهنمایی ها و دستورالعمل هایی که به روشنی موقعیت آزمون را مشخص می کنند، همچنین از روش ها و فنون گوناگون استفاده نمایید.

کار پوشه: نوع دیگری از آزمون های عملکردی، که به تازگی شهرت زیادی کسب کرده است، کارپوشه، نام دارد. کار پوشه نوعأ مجموعه ای از کارهای یک دانش آموز است که کوشش، پیشرفت و موفقیت تحصیلی او را در یک زمینه خاص نشان می دهد. در روش کار پوشه، یادگیرنده برای نشان دادن توانایی ها و پیشرفت خود، نمونه هایی از کارهایی را که در طول یک مدت معین (ثلث، نیمسال) انجام داده است با خود به کلاس می آورد و برای ارزشیابی در اختیار معلم می گذارد. از آنجا که مطالب موجود در پوشه، جنبه تراکمی دارند، به هم مربوط اند و در طول یک دوره زمانی انجام گرفته اند، می توانند یک فیلم از چگونگی پیشرفت یادگیرنده فراهم آورند.

کار پوشه دو هدف دارد:

۱- ارائه بهترین کارها

۲- نشان دادن رشد تحصیلی یادگیرنده

مزیت مهم کارپوشه آن است که یادگیرندگان با اندیشیدن درباره نمونه کارهای خود، ارزشیابی از خود را می آموزند.

روش کار پوشه، ابتدا در صنعت به کار بسته شد و با موفقیت مواجه گردید؛ پس از آن که مربیان، موفقیت این روش را در صنعت مشاهده کردند، به استفاده از آن در آموزش تشویق شدند، اما پژوهش ها درباره اثر بخشی این روش نشان داد که روش کار پوشه نتوانسته است موفقیت های حاصل در صنعت را در زمینه آموزش و پرورش به دست آورد.

روش های سنجش مشاهده ای

مهمترین فنون مبتنی بر مشاهده، که مورد استفاده معلمان هستند، عبارتند از:

۱) فهرست وارسی

۲) مقیاس درجه بندی

۳) واقعه نگاری

۱- روش فهرست وارسی

از طریق روش فهرست وارسی مشخص می شود که آیا ویژگی های مورد نظر در عملکرد شخصی مورد سنجش، موجودند یا نه. در این روش فقط بود یا نبود یک ویژگی ثبت میشود، نه درجه و فراوانی آن فهرست وارسی، هم می تواند برای سنجش فرآورده به کار رود و هم برای سنجش فرایند. هنگام تهیه فهرست وارسی به منظور ارزشیابی از فرایند عملکرد، تأکید بر رفتارهای فرد می باشد؛ ولی هنگام تهیه فهرست وارسی برای ارزشیابی از محصول یا فرآورده رفتار، بر جنبه ها یا ویژگی های قابل مشاهده محصول مورد سنجش تأکید می شود.

یک فهرست وارسی خوب دارای ویژگی های زیر است:

۱ـ فهرست نسبتا کوتاه

۲- هر ماده فهرست به طور کاملاً روشن، موضوع و مورد ارزشیابی را بیان می کند.

۳ـ هر ماده بر یک رفتار یا ویژگی قابل مشاهده می کند.

۴- تنها رفتارها یا ویژگی های مهم منظور می شوند.

5ـ ماده های فهرست به گونه ای آرایش می یابند که، کل فهرست را بسادگی می توان مورد استفاده قرار داد. ۲- روش مقیاس درجه بندی

برخلاف روش فهرست وارسی، که در آن فقط وقوع یا عدم وقوع یک رفتار ثبت میشود، در مقیاس درجه بندی، درجه یا مقدار رفتار یا ویژگی مورد نظر تعیین می گردد. مقیاس درجه بندی ، هم برای ارزشیابی از فرآورده به کار می رود و هم برای ارزشیابی از فرایند.

انواع مقیاس درجه بندی

۱ـ مقیاس درجه بندی عددی: که ساده ترین نوع مقیاس درچه بندی است. در این مقیاس، مشاهده گر با کشیدن دایره به دور یک عدد یا علامت گذاری آن، میزان یا درجه آن رفتار یا ویژگی را مشخص می کند مقیاس درجه بندی عددی، زمانی قابل استفاده است که ویژگی ها یا کیفیت های مورد درجه بندی قابلیت طبقه بندی کردن در طبقات محدودی را داشته باشند و نیز در رابطه با طبقه ای که معرف هر عدد است، توافق وجود داشته باشد.

به عنوان مثال باکشیدن دایره دور یکی از اعداد، نشان دهید که میزان شرکت دانش آموزان در فعالیت های فوق برنامه چقدر است؟

۱) خیلی کم

۲) کم

۳) زیاد

۴) خیلی زیاد

مقیاس درجه بندی نگاره ای: در این مقیاس، به جای استفاده از اعداد از خطوط افقی استفاده میشود و مشاهده گر باید داوری خود را درباره ویژگی مورد نظر، در طول یک خط مستقیم و با گذاشتن علامت نشان دهد.

به عنوان مثال با گذاشتن علامت (×) روی خط افقی، نظر خود را در مورد فعالیت های سیاسی بیان کنید.

بسیار علاقه مند نسبتاً علاقه مند دارای علاقه متوسط نسبتاً بی علاقه بسیار بی علاقه

۳ـ مقیاس درجه بندی نگارهای توصیفی: در این مقیاس، نقاط روی مقیاس توصیف میشوند. توصیف هایی که برای این منظور به کار می روند، عبارت هایی هستند که به صورت رفتاری نشان می دهند، دانش آموز یا فرد مورد مشاهده و مورد ارزشیابی در مراحل مختلف مقیاس، چگونه عمل مینماید.

مثلاً : با قراردادن علامت × روی خط مربوط به هر سؤال، ویژگی مورد نظر را ارزشیابی کنید.

۱- کیفیت کار این دانش آموزان را در انجام تکالیف درسی چگونه ارزیابی می کنید؟

کار خیلی خوب است و هیچ گونه نیازی به نظارت ندارد.

کار خوب است و به ندرت نیاز به نظارت دارد.

کارخوب است و گاه گاه نیاز به نظارت دارد.

کار قابل قبول است اما احتیاج به نظارت فراوان دارد.

کار ضعیف است و نیاز به نظارت فراوان دارد.

۳- روش واقعه نگاری

روش واقعه نگاری یا ثبت رویداد به توصیفات واقعی گفته میشود که معلم در نتیجه ی مشاهده زندگی دانش آموز به دست می آورد. این روش، بیشتر جنبه تحلیلی دارد تا ارزشیابی. این روش، بیشتر برای سنجش هدف های حوزه عاطفی و روانی- حرکتی قابل استفاده است.

سنجش هدف های حوزه عاطفی

هدف های حوزه عاطفی را نمی توان با آزمون های توانایی سنجید؛ زیرا در آزمون های توانایی، از یادگیرنده می خواهیم که حداکثر توان خود را در پاسخ به سؤال های آزمون نشان دهد؛ اما وسایل سنجش هدف های عاطفی، از آزمون شونده می خواهند، هر آنچه را احساس می کنند، بگویند. مشکل اساسی این وسایل این است که باید موقعیتی را ایجاد کنند تا یادگیرنده رفتار عاطفی مورد علاقه خود را نشان دهند و به او اطمینان دهیم که از نتایج حاصل، به ضرر آن ها استفاده نخواهد شد تا در پاسخ به سؤال ها صادقانه عمل کنند.

روش ها و فنون سنجش هدف های عاطفی

۱- پرسشنامه های بسته پاسخ یا گزینه بایست: که بسیار شبیه به آزمون های عینی در سنجش رفتارهای شناختی اند؛ زیرا در هر دو روش، یادگیرنده ر سنجش رفتارهای باید از بین پاسخ های موجود برای یک سؤال، یک پاسخ را برگزیند.

۲- پرسشنامه های بازپاسخ: که شبیه به آزمون های تشریحی و کوته پاسخ در سنجش هدف های حوزه شناختی اند. در این روش، به آزمون شونده سؤالی را می دهیم و او باید نظر خود را توضیح دهد، معمولا سؤالات را به صورت جمله ناتمام به دانش آموزان می دهیم.

۳ـ روش تفکیک معنایی: (این روش مربوط به چارلز آزگود است). یکی از روش های مهم سنجش نگرش، روش تفکیک معنایی یا افتراق معنایی است. ر این روش، شامل تعدادی صفت دوقطبی است و هر جفت صفت، یک پیوستار هفت بخشی را تشکیل میدهند. در این روش، موضوع نگرش در بالای صفحه قرار می گیرد و از پاسخ دهندگان خواسته می شود، با گذاشتن علامت روی پیوستار، موضوع نگرش را درجه بندی کنند.

مثال:

دانشگاه

خوب 3+ 2+ 1+ 0 1- 2- 3- بد

متنوع 3+ 2+ 1+ 0 1- 2- 3- تکراری

فعال 3+ 2+ 1+ 0 1- 2- 3- منفعل

روش تفکیک معنایی، روشی است برای اندازه گیری معنایی ضمنی مفاهیم. معنای ضمنی مفاهیم، نشان دهنده نگرش فرد است.

۴- روش یادداشت های روزانه: در این روش، دانش آموز احساسات خود را درباره تجارب شخصی و فعالیت های کلاسی به صورت یادداشت های روزانه ابراز می کند.

مثال ۳: ساده ترین سنجش ها در گزینه های زیر کدام است؟

۱) توصیف کمی

۲) طبقه بندی دو جانبه

۳) دسته بندی

۴) درجه بندی کیفی

پاسخ: گزینه «۲» سنجش دو گزینه ای، احتمال انتخاب درست ، است. سنجش، یک اصطلاح کلی است که به گردآوری اطلاعات درست درباره تصمیم گیری اشاره دارد. برای دسته بندی و توصیف کمی و درجه بندی کیفی به تعاریف وسیع تری نیاز داریم که تنها در طبقه بندی دو جانبه این تعریف، محدود و ساده تر می شود.

تست های طبقه بندی شده فصل یازدهم

۱- در کدام یک از روش های ارزشیابی زیر، معلم، ضمن ارزشیابی، آموزش نیز می دهد؟

(سراسری 83)

۱) پورت فولیو

۲) واقعه نگاری

۳) نمونه کار

4) چند گزینه ای

۲- در مصاحبه های گزینشی، کدام یک از عوامل، نتایج را بی اعتبار می سازد؟ .

(سراسری 87)

۱) باهوشی مصاحبه شونده

۲) بی دقتی مصاحبه کننده

۳) خستگی مصاحبه کننده

۴) عوامل هیجانی مصاحبه شونده

۳- در کدام آزمون عملکردی از خودسنجی بیشتر استفاده می شود؟

(سراسری 89)

۱) کارپوشه

۲) نمونه کار

3) آزمون شناسایی

۴) موقعیت های شبیه سازی شده

۴- نوعی از آزمون که در آن آزمودنی باید قطعاتی را روی هم سوار کند، یا طرح معینی را بسازد، چه نام دارد؟

(آزاد 89)

۱) آزمون حرفه ای

۲) آزمون مختلط

۳) آزمون فردی

4) آزمون ابزاری

۵- کدام یکی از ویژگی های آزمون های عملکردی نیست؟

(سراسری ۹۱)

۱) ترغیب تفکر باز

۲) تأکید بر کاربست

۳) استفاده از مسائل واقعی

۴) سنجش غیرمستقیم

6- سنجش عملکردی بیشتر تحت تأثیر کدام مکتب فلسفی قرار دارد؟

(سراسری ۹۳)

۱) ساختن گرایی

۲) رفتارگرایی

۳) ساختارگرایی

۴) مصلحت گرایی

پاسخنامه تست های طبقه بندی شده فصل یازدهم

1- گزینه «1» در روش پورت فولیو (کارپوشه) یادگیرنده برای نشان دادن توانایی ها و پیشرفت خود، نمونه های کار خود را برای ارزشیابی در اختیار معلم سر می گذارد؛ در واقع این نمونه، به «ارزشیابی از خود» یادگیرنده نیز کمک شایانی می کند. استفاده از کارپوشه به عنوان یک روش آموزش و سنجش، نه تنها به پیشرفت تحصیلی دانش آموز و ارزشیابی او از خود کمک می کند، بلکه به گفته «کلنوسکی» مهارت های تفکر، تحلیل انتقادی و راهبردهای فراشناختی نیز گسترش می دهد.

۲- گزینه «۴» مصاحبه، یکی از روش های گردآوری اطلاعات است؛ زیرا عوامل هیجانی مصاحبه شونده باعث می شود که در پاسخ دادن به سؤالات نتواند به طور صحیح عمل کند.

٣- گزینه «1» آزمون های عملکردی با مهارت سر و کار دارند؛ برای مثال، دروس علوم با مهارت های آزمایشگاهی، دروس ریاضی با مهارت های حل مسئله و دروس فارسی و زبان های خارجی با مهارت های برقرار کردن ارتباط سر و کار دارند؛ بنابراین، آزمون های عملکردی به عنوان مکمل آزمون های کتبی، می توانند معلمان را در سنجش بهتر میزان توفیق یادگیرندگان در رسیدن به هدف های متنوع تحصیلی یاری دهند. سالکس (۱۹۹۷) در این باره می گوید:

«آزمون های عملکردی از آزمون شونده می خواهند تا به انجام نوعی رفتار اقدام کند، نه اینکه صرفأ سؤال هایی را جواب دهد.» گزینه ها به انواع مختلف این آزمون اشاره دارد. کارپوشه (portfoio) مجموعه ای از کارهای دانش آموز است که کوشش، پیشرفت و موفقیت تحصیلی او را در یک زمینه خاص نشان میدهد. استفاده از کارپوشه به عنوان یک روش آموزش و سنجش، نه تنها به پیشرفت تحصیلی دانش آموز و ارزشیابی او از خود کمک می کند، بلکه به گفته ی کلنوسکی، مهارت های تفکر، تحلیل انتقادی و راهبردهای فراشناختی را نیز گسترش می دهد. گزینه ۲ (نمونه کار) بنا به شباهت به کارپوشه، می تواند داوطلب را دچار تردید در انتخاب کند.

۴- گزینه «1» هیچ یک از گزینه های سؤال در کتاب های موجود در زمینه سنجش و اندازه گیری، در رابطه با موضوع سؤال نیامده اند؛ اما به طور کلی می توان گفت، آزمون حرفه ای را می توان به عنوان زیر مجموعه ای از آزمون های عملکردی، که با مهارت سر و کار دارند، مهارت در استفاده از فرایندها و شیوه های اجرایی و نیز مهارت در تولید فرایند فرآورده ها به حساب آورد؛ بنابراین گزینه ۱ نزدیک ترین پاسخ به سوال می باشد.

5-گزینه «۴» از ویژگی های آزمون عملکردی، ترغیب تفکر باز، تأکید بر کاربست، استفاده از مسائل واقعی و سنجش مستقیم است.

۶- گزینه «1» نظریه پردازان یادگیری ساختن گرا، درباره اینکه انسان ها چگونه نظام هایی را برای فهم معنی دار جهان و تجربیاتشان طرح ریزی می کنند، به تحقیق و نظریه پردازی پرداخته اند. در مفهوم عام این نظریه، اندیشه های نظریه پردازان و محققان نظام تعلیم و تربیت از جمله دیویی، پیاژه، ویگوتسکی، برونر، کلی و ... بر توسعه انواع متفاوت ساختن گرایی فردی و اجتماعی تأثیر گذاشته است.

نکته: سنجش عملکردی، بیشتر تحت تأثیر مکتب ساختن گرایی قرار دارد.

آزمون فصل یازدهم

۱- آزمون های عملکردی برای سنجش چه چیزی مفیدند؟

1) دانش و معلومات

۲) فرآورده رفتار

۳) فرایند رفتار

۴) فرایند و فرآورده رفتار

۲- آزمون های کتبی برای سنجش کدام یک از حوزه های یادگیری مناسب ترند؟

1) شناختی

۲) عاطفی

۳) روانی - حرکتی

۴) همه موارد

۳- در کدام یک از روش های ارزشیابی، معلم در ضمن ارزشیابی به آموزش هم اقدام می کند؟

1) نمونه کار

۲) واقعه نگاری

۳) آزمون شناسایی

۴) کارپوشه

۴- کدام یک از آزمون های عملکردی به عملکرد واقعی یادگیرنده در محیط های طبیعی نزدیکتر است؟

۱) آزمون شناسایی

۲) نمونه کار

3) انجام عملکرد در موقعیت های شبیه سازی شده

۴) آزمون های کتبی عملکردی

5- خود ارزشیابی، از امتیازات کدام نوع از آزمون های عملکردی است؟

۱) نمونه کار

۲) کارپوشه

۳) آزمون شناسایی

۴) آزمون های کتبی عملکردی

6- اهداف کارپوشه عبارتند از:

1) ارائه بهترین کارها

۲) نشان دهنده رشد تحصیلی یادگیرنده

۳) شناسایی مشکل یادگیرنده

۴) گزینه های ۱ و ۲

۷- روش کار پوشه ابتدا در .................. بکار بسته شد.

۱) تجارت

۲) صنعت

۳) بازرگانی

۴) آموزش و پرورش

8- که این روش، نزدیک ترین سنجش عملکرد واقعی یادگیرنده در محیط طبیعی می باشد؟

1) نمونه کار

۲) عملکرد در موقعیت های شبیه سازی شده

3) آزمون شناسایی

۴) کتبی عملکردی

۹- ساده ترین نوع مقیاس درجه بندی ..................... می باشد؟

۱) مقیاس درجه بندی عددی

۲) مقیاس درجه بندی نگاره ای

3) مقیاس درجه بندی توصیفی

۴) گزینه های ۲ و ۳

۱۰- چارلز آزگود، مبدع کدام یک از روش های سنجش هدف های عاطفی می باشد؟

1) گزینه ی بایست

۲) پرسشنامه باز پاسخ

۳) تفکیک معنایی

۴) یاداشت های روزانه

فصل دوازدهم

« روش های سنجش مستقیم رفتار»

روش های سنجش مستقیم رفتار، روش هایی هستند که در آنها ویژگی های فیزیکی، رفتارهای آشکار یا محصولات رفتاری را می توان به طور مستقیم سنجش کرد. در سنجش مستقیم رفتار، ویژگی ها یا بعدهای رفتار، مشاهده و ثبت می شوند.

ابعاد مختلف رفتار

1- شکل ظاهری رفتار: یعنی رفتار در چه شکلی رخ میدهد؛ مثلا لیلا وقتی عصبانی می شود، ناخنش را می جود.

۲- شدت یا نیروی رفتار: یعنی رفتار با چه شدتی رخ می دهد: زهرا هنگام عصبانیت، طوری جیغ می کشد که صدای او، همکلاسی هایش را ناراحت می کند.

3- نرخ یا فراوانی رفتار: یعنی تعداد دفعات وقوع رفتار. صرفا تعداد دفعات وقوع رفتار را فراوانی می گوییم. تعداد دفعات وقوع رفتار در یک واحد زمانی معین را نرخ رفتار می گویند.

به عنوان نمونه: احمد سه بار از روی درس خواند: فراوانی

احمد در یک جلسه سه تمرین حل کرد: نرخ

۴- طول مدت رفتار: یعنی زمانی که انجام دادن یک کار یا یک پاسخ طول می کشد؛ عصبانیت زهرا ۳۰ دقیقه ادامه داشت.

5- دوره نهفتگی رفتار: مدت زمانی که طول می کشد تا شخص، کار یا رفتاری را که از او خواسته شده است، شروع کند؛ مثال: بعد از این که معلم با زهرا صحبت کرد، ۱۰ دقیقه طول کشید تا او آرام شد.

مکان یا موقعیت انجام رفتار: جایی که رفتار در آن رخ می دهد: زهرا بیشتر در کلاس عربی، عصبانی می شود. انواع روش های سنجش مستقیم رفتار

1- سنجش محصولات ماندنی رفتار یا فرآورده رفتار: برخی از رفتارها از خود، محصولات ماندنی برجای می گذارند که معلم می تواند آن ها را سنجش کند؛ مثلا تصحیح برگه های امتحانی ریاضی دانش آموزان و نمره دادن به آن ها نمونه ای از سنجش محصولات ماندنی رفتار است. این روش، یک روش بعد از واقعه است یعنی اندازه گیری وقتی انجام می شود که رفتار انجام شده و به پایان رسیده باشد. به سنجش محصولات ماندنی رفتار، ثبت بازده هم می گویند. در روش سنجش محصولات ماندنی رفتار، سنجشگر با نتایج ملموس رفتار سر و کار دارد و نیاز ندارد که آن را در حین وقوع، ثبت و اندازه گیری کند؛ در واقع، امتیاز مهم این روش، آن است که به ما امکان می دهد تا بدون حضور فرد، اطلاعاتی درباره او کسب کنیم. روش سنجش محصولات ماندنی رفتار برای تعیین ویژگی ها یا بعدهای رفتاری نیز مفید می باشد.

- سنجش رویدادهای گذرای رفتار با سنجش فرایند رفتار: برخی از رفتارها محصولات ماندنی از خود باقی نمی گذارند؛ برای اندازه گیری این نوع رفتارها لإزم است هنگامی که اتفاق می افتند، به مشاهده و ثبت آن ها بپردازیم، بدین جهت به آن ها روش های ثبت مشاهده ای هم می گویند؛ مثل خارج از نوبت صحبت کردن دانش آموزان یا تنبلی آن ها. در این روش، نمونه هایی از رفتار موردنظر در هنگام وقوع، ثبت و سنجش میشوند. اگر معلم بخواهد تعداد دفعات وقوع رفتار را ثبت کند، از روش ثبت فراوانی بهره می گیرد. روش ثبت فراوانی یا نرخ رفتار، متداول ترین روش ثبت مشاهده ای می باشد. این روش، برای اندازه گیری رفتارهایی که مجزا و مشخص اند، بسیار مفید است؛ یعنی رفتارهایی که آغاز و انجام تعریف شده ای دارند. ساده ترین راه ثبت فراوانی رفتار، استفاده از ماشین حسابگر الکترونیکی است. اگر معلم بخواهد طول مدتی را که صرف رفتار می کند، ثبت کند، از روش طول مدت رفتار استفاده می کند. از رفتارهایی که طول مدت آنها از تعدادشان مهم تر است، می توان زمان توجه کردن کودک، تراشیدن مداد و رؤیای روز را نام برد. اگر او بخواهد طول مدتی را که باید صرف شود تا فرد رفتاری را شروع کند، اندازه بگیرد، از روش ثبت دوره ی نهفتگی استفاده می کند. با این روش، مدت زمان بین لحظه ای که از شخص، انجام رفتاری خواسته شده است و لحظه ای که او عملا آن رفتار را شروع می کند، تعیین می گردد. اگر معلم بخواهد نسبتی از زمان را تعیین کند که در آن رفتار رخ می دهد، از روش ثبت فاصله ای با روش ثبت نمونه گیری زمان استفاده می کند. تفاوت بین روش ثبت فاصلهای رفتار و روش ثبت نمونه گیری زمانی رفتار، این است که در روش ثبت فاصلهای رفتاری در یک زمان معین مثلا 60 دقیقه و بطور پیوسته انجام می گیرد؛ در حالی که در روش ثبت نمونه گیری زمان، رفتار مورد نظر در فاصله های کوتاه مثلا ۶۰ ثانیه مشاهده می شود، اما نه بطور پیوسته بلکه با فاصله های زمانی مثلا ۱۰ ثانیه ای این دو روش برای رفتارهایی که آغاز و انجام قابل تمایز ندارند، مناسب اند.

تست های طبقه بندی شده فصل دوازدهم

۱- برای تشخیص و تصمیم گیری در مورد کم و کیف یادگیری کودکان عقب مانده ذهنی، کاربرد کدام نوع ارزشیابی مناسب تر است؟

(سراسری ۸۱)

1) هنجاری

۲) عملکردی

۳) ملاکی

۴) معیاری

۲- کدام عبارت زیر صحیح تر است؟

(آزاد ۸۴)

1) در امتحان عملی مهارت سنجیده می شود.

۲) در امتحان عملی میزان رغبت ها روشن می گردد

۳) امتحان عملی به صورت فردی اجرا می گردد.

۴) در امتحان عملی معلومات سنجیده میشود.

۳- کدام نوع از آزمون یا وسیله اندازه گیری با سنجش عملکردی همخوانی بیشتری دارد؟

(آزاد ۸۵)

1) مشاهده

۲) پرسشنامه

۳) آزمون عینی

۴) آزمون انشایی

پاسخنامه تست های طبقه بندی شده فصل دوازدهم

1- گزینه «۲» چون کودکان عقب مانده ی ذهنی، خصوصیات منحصر به فرد خود را دارا دارند، نباید با گروه همسالان مقایسه قرار شوند، در واقع، معیار و ملاک ثابتی برای مقایسه آن ها وجود ندارد؛ بنابراین، آن ها باید در مقایسه با نوع عملکرد خود مورد سنجش قرار گیرند. گزینه1 اشاره به نوعی ارزشیابی مطلق دارد که برای کودکان عقب مانده ذهنی کاربردی ندارد.

۲- گزینه «۱» گزینه ۲ اشتباه است؛ چون میزان رغبت ها به وسیله آزمونهایی که مربوط به حوزه عاطفی هستند، سنجیده می شوند. گزینه ۳ اشتباه است؛ چون امتحان عملی را هم به صورت فردی و هم به صورت گروهی می توان اجرا کرد، همچنین گزینه ۴ اشتباه است؛ چون در امتحانات عملی، معلومات فرد سنجیده نمی شود، بلکه توانایی او در انجام درست کار سنجیده میشود.

۳- گزینه «1» بنا به تعریف، آزمون های عملکردی با مهارت سر و کار دارند؛ مهارت در استفاده از فرایندها و شیوه های اجرایی و نیز مهارت در تولید فرآورده ها. از آنجا که مهارت ها قابل مشاهده اند و از طریق آنها می توان فرایندی که فرد طی می کند و نتیجه ای را که به دست می آورد، مشاهده کرد، بهترین روش برای سنجش آن ها، روش مشاهده است.

آزمون فصل دوازدهم

1- نام دیگر سنجش محصولات ماندنی رفتار چیست؟

1) ثبت بازده

۲) ثبت مشاهده ای

۳) ثبت فاصله ای

۴) هیچ کدام

2- اگر معلمی بخواهد برگه های امتحان ریاضی دانش آموزان را تصحیح کند، از کدام روش سنجش مستقیم استفاده خواهد کرد؟

۱) ثبت فراوانی

۲) نمونه گیری زمان

۳) سنجش محصولات ماندنی

۴) ثبت فاصله ای

3- به روش ثبت تمامی موارد یک رفتار در یک زمان مشخص ........... می گوییم.

۱) ثبت فاصله ای

۲) ثبت مداوم

۳) ثبت فراوانی

۴) ثبت فاصله زمان

۴- ثبت فاصله ای برای سنجش کدام ویژگی رفتار مفیدتر است؟

۱) شدت

۲) نرخ

۳) بود یا نبود

۴) کیفیت

5- روش ثبت فاصله ای و نمونه گیری زمان برای چه رفتارهایی مناسب اند؟

1) رفتارهایی که آغاز و انجام قابل تمایزی دارند.

۲) رفتارهایی که آغاز و انجام قابل تمایزی ندارند.

۳) برای رفتارهایی که قابلیت پیش بینی بالایی دارند.

۴) گزینه های 1و۳

6- زمانی که می گوییم «حسن وقتی اضطراب دارد به طور متداول با دستش» روی میز ضربه وارد می کند کدام بعد رفتار را مورد سنجش قرار داده ایم؟

۱) شکل ظاهری رفتار

۲) شدت با نیروی رفتار

3) مکان یا موقعیت رفتار

۴) نرخ یا فراوانی رفتار

۷- نام دیگر روش سنجش رویدادهای گذری رفتار چیست؟

۱) ثبت بازده

۲) ثبت فاصله ای

۳) سنجش فرایند رفتار

۴) ثبت مشاهده ای

8- اگر بخواهیم طول مدتی را که باید صرف شود تا فردی رفتاری را شروع کند، اندازه بگیریم از کدام روش باید استفاده کنیم؟

1) طول مدت رفتار

۲) نرخ یا فراوانی رفتار

3) دوره ی نهفتگی رفتار

۴) هیچ کدام

۹- اگر معلم بخواهد نسبتی از زمان را تعیین کند که در آن رفتار رخ می دهد، از کدام روش باید استفاده کند؟ 1) ثبت فاصله ای یا نمونه گیری زمان

۲) طول دوره ی نهفتگی

۳) طول مدت رفتار

۴) ثبت فراوانی

۱۰- روش های ثبت مشاهده ای به کدام یک از موارد زیر اطلاق می شود؟

1) سنجش محصولات ماندنی رفتار

۲) سنجش رویدادهای گذاری رفتار یک سنجش فرایند رفتار

۲) سنجش مستقیم رفتار

۴) همه موارد

فصل سیزدهم:

« اجرا، نمره گذاری و تحلیل آزمون»

گام نخست در اجرای آزمون های پیشرفت تحصیلی، آماده سازی و فراهم کردن شرایط مناسب برای اجرای آن هاست. تهیه دفترچه ی آزمون و پاسخنامه ترتیب پشت سرهم قرار دادن سؤال های مختلف و در نظر گرفتن زمان و مکان مناسب برای اجرای آزمون، نیاز به دقت فراوان دارند.

ترتیب قرار دادن سؤالات آزمون به دنبال یکدیگر

1- سؤال ها را باید با توجه به نوع آن ها دسته بندی کرد و پشت سرهم قرار داد. بهتر است، سؤال های هرنوع را بر حسب طبقه بندی هدف های آموزشی دسته بندی کرد. برای ردیف کردن سؤالها ترتیب زیر را می توان به کار برد.

الف) صحیح - غلط،

ب) جورکردنی

پ) چند گزینه ای

ت) کوته پاسخ

ث) تشریحی

۲- اگر مرتب کردن سؤال های یک آزمون با توجه به بازده های یادگیری امکان پذیر نشد، سؤال ها را از ساده به دشوار مرتب کنید؛ زیرا دانش آموزان با پاسخ دادن به سر سؤالات ساده اعتماد به نفس پیدا می کنند و کارایی آن ها افزایش می یابد. در این روش، ابتدا سؤال ها بر حسب نوع، دسته بندی می شوند، آن گاه درون هر نوع از

سؤال ها، به ترتیب از ساده به دشوار چیده می شوند.

٣- سؤالات را به ترتیب سازمان اصلی مطالب به دنبال هم مرتب کنید؛ این کار باعث نظم فکری آزمون شونده می شود. در نوشتن دستورالعمل یا راهنمایی آزمون باید دقت کرد تا راهنما مختصر و مفید باشد و به طور آشکار به یادگیرنده بگوید که چه کاری را باید انجام دهد. راهنمای آزمون باید اطلاعات زیر را در اختیار آزمون شونده قرار دهد.

١- زمان لازم برای هر قسمت

2- ارزش هر سؤال: نمرهای که به هر سؤال تعلق می گیرد.

۳- مجاز یا غیرمجاز بودن حدس زدن: یعنی اگر آزمون نمره منفی دارد، باید آزمون شوندگان را مطلع کرد. در اجرای آزمون، باید شرایط را طوری آماده کرد که آزمون شونده از هر لحاظ، آسایش فکری و جسمی داشته باشد؛ همچنین از عوامل تنش زا و اضطراب برانگیز باید خودداری کرد.

مثال1: یکی از علت های مهم اضطراب دانش آموزان از امتحان آن است که:

1) از شکست در امتحان می ترسند.

۲) به اوراق امتحانی، منصفانه نمره داده نمی شود.

۳) به طور مؤثر مطالعه نمی کنند.

۴) رقابت را دوست ندارند.

پاسخ: گزینه «۱» به نظر می رسد این تست برای همه داوطلبان ملموس باشد. بزرگترین علت اضطراب دانش آموزان از امتحان، ترس از شکست در امتحان میباشد. سایر گزینه ها نمی توانند جایگزین انتخاب صحیح شوند؛ هرچند ممکن است که تری از شکست منجر به مطالعه گسسته یا مؤثر نیز شود.

مثال ۲: سؤالات یک آزمون را باید با توجه به .............. آن ها دسته بندی کرد.

۱) نوع

۲) میزان دشواری

۳) طبقه بندی اهداف به کار برده شده

۴) همه موارد

پاسخ: گزینه «4» همه ی موارد صحیح است.

مثال 3: امتیاز مرتب کردن سؤال های یک آزمون براساس سطح دشواری (از ساده به مشکل)، کدام است؟

1) افزایش دقت و پایایی آزمون

2) تشویق فراگیران در پاسخ دهی به سؤال ها

۳) کاهش میزان خطای اندازه گیری

۴) افزایش میزان ضریب تمییز آزمون ها

پاسخ: گزینه «۲» متخصصان (گرانلاند، ولین ۱۹۹۰) در زمینه قرار دادن سؤال های آزمون به دنبال هم، چند پیشنهاد ارائه داده اند:

١- سؤال ها را با توجه به نوع آنها دسته بندی کنید و پشت سر هم قرار دهید.

۲- سؤال ها را از ساده به دشوار مرتبه کنید.

۳- سؤال ها را به ترتیب سازمان اصلی مطالب، به دنبال هم مرتب کنید.

با توجه به متن سؤال که در آیتم شماره ۲ آن را می بینیم، زمانی که سؤال های ساده را در ابتدای آزمون قرار می دهیم، آزمون شوندگان با پاسخ دادن به این سؤال ها، اعتماد به نفس پیدا می کنند و این امر، کارایی آنان را در عملکرد سؤال های پیچیده بعدی نیز افزایش می دهد. در این حالت، حتی ضعیف ترین دانش آموزان نیز خواهند کوشید، که حداکثر توان خود را در پاسخدهی به سؤال ها به کار گیرند. با توجه به این که همهی گزینه های دیگر سؤال، به یک تعریف خاص در سنجش و اندازه گیری اشاره دارد، داوطلبان به راحتی می توانند به جواب درست دست پیدا کنند.

مثال ۴: در نوشتن دستورالعمل یا راهنمای آزمون کدام یک از موارد زیر را باید لحاظ کرد؟

1) زمان لازم برای هر قسمت

۲) ارزش هر سؤال

۳) مجاز یا غیر مجاز بودن حدس

۴) همه موارد

پاسخ: گزینه «۴» در نوشتن دستورالعمل باید این موارد لحاظ شود:

1- زمان لازم برای هر تست

۲- ارزش هر سؤال

۳- مجاز یا غیرمجاز بودن حدس

نمره گذاری نمره گذاری آزمون های تشریحی در فصول قبل توضیح داده شد؛ حال به نمره گذاری آزمون های عینی می پردازیم. در نمره گذاری آزمون های عینی دو روش عمده وجود دارد:

1- محاسبه کلیه ی پاسخ های درست بدون کسر نمره برای حدس زدن (یعنی بدون نمره منفی)

۲- کسر مقداری از نمره برای حدس زدن (یعنی نمره منفی لحاظ شود)

فرمول برای رهایی که نمره منفی لحاط می کنیم:

R : تعداد پاسخ های درست

W: تعداد پاسخ های غلط

N : تعداد گزینه های هر سوال

مثال ۵: اگر در یک آزمون چهار گزینه های ۴۰ سؤال داشته باشیم و دانش آموزی به ۳۴ سؤال پاسخ دهد که ۲۵ پاسخ او صحیح و ۹ پاسخ او غلط می باشد، نمره اصلاح شده این دانش آموز چقدر است؟

N=4

R=25

W=9

1) 20

2) 5/21

3) 23

4) 22

پاسخ: گزینه «4».

مثال ۶: در یک آزمون ۵ گزینه ای که از ۷۰ سؤال تشکیل شده است، دانش آموزی ۶۲ سؤال را صحیح و ۸ سؤال را غلط پاسخ می دهد. نمره ی اصلاح شده او چقدر می باشد؟

1) 58

2) 59

3) 60

4) 61

پاسخ: گزینه «3»

مثال ۷: در یک آزمون چهار گزینه ای که دارای ۱۰۰ سؤال است، دانش آموزی به ۷۸ سؤال پاسخ داده که ۶۰ پاسخ او صحیح و ۱۸ پاسخ او غلط است . نمره اصلاح شده دانش آموز را محاسبه کنید.

N=4

R=60

W=18

1) 52

2) 54

3) 53

4) 51

پاسخ: گزینه «2»

مثال ۸: فردی در یک آزمون ۴۰ سؤالی چهار گزینه ای، به ۲۵ سؤال پاسخ صحیح و به ۱۵ سؤال پاسخ غلط داده است. نمره او پس از اصلاح حدس و گمان، چقدر می شود؟

1) 10

2) 20

3) 21

4) 22

پاسخ: گزینه «2»

مثال ۹: در یک آزمون پنج گزینه ای که دارای هشتاد سؤال است، آزمودنی به ۶۸ سؤال پاسخ صحیح و به ۱۲ سؤال پاسخ غلط داده است. نمره آزمودنی با احتساب نمره منفی چیست؟

1)68

2)67

3)66

4)65

پاسخ: گزینه «۴»

مثال ۱۰: در نمره گذاری آزمون های افشایی، کدام مورد سبب می شود که به پاسخ های طولانی تر، نمره های بیشتری داده شود؟

1) خطای نقل نمره ها

۲) خطای محاسبه

۳) خطای هاله ای

۴) هیچ کدام

پاسخ: گزینه «۴» در نمره گذاری آزمون های انشایی، در زمینه پاسخ طولانی، هیچ گونه خطایی وجود ندارد و سایر گزینه ها به انواع خطاها اشاره دارد که عبارتند از: خطای هاله ای تأثیر سایر خصوصیات آزمون شونده در تصحیح برگهاش، خطای نقل نمره ها تأثیر نمرات سؤال قبل در سؤال جدید، خطای تمایل به مرکز: ارائه نمراتی توسط مصحح، که همه متوسط هستند.

مثال 11: اگر در یک آزمون ۵۰ پرسشی چهار گزینه ای، دانش آموزی به ۴۰ پرسش، درست و به ۶ پرسش، غلط جواب داده باشد، نمره اصلاح شده او بعد از محاسبه خطای حدس زدن چند است؟

1) 30

2) 44

3) 36

4) 38

پاسخ: گزینه «4»

تحلیل سؤال های آزمون

هدف از تحلیل سوال های آزمون، وارسی تک تک سؤال ها و تعیین میزان دقت و نارسایی های آن هاست. تحلیل پاسخ های آزمون شوندگان، اطلاعات تشخیصی لازم را برای بررسی کیفیت یادگیری دانش آموزان و مشکلات آموزش معلمان فراهم می کند. اطلاعات مورد نیاز برای تحلیل سؤالات آزمون، همان پاسخهای آزمون شوندگان هستند؛ بنابراین باید تعیین شود

۱- در هر سؤال، چند نفر گزینه درست را انتخاب کردهاند.

۲- هریک از گزینه های انحرافی چند نفر را به خود جذب کرده است.

3- چند نفر سؤال را بی جواب گذاشته اند.

برای تحلیل سوالات آزمون، ضرورت دارد که گروه قوی (بالا) و گروه ضعیف (پایین) را مشخص کنیم. برای این کار، ابتدا باید برگه های دانش آموزان را

به ترتیب نمره ای که گرفته اند، از بزرگ به کوچک مرتب کنیم.

- اگر تعداد دانش آموزان بین ۲۰ تا ۴۰ نفر است: ۱۰ برگه ی بالا و ۱۰ برگه ی پایین را انتخاب می کنیم.

- اگر تعداد دانش آموزان از ۲۰ نفر کمتر است: کل برگه ها را به دو دسته بالا و پایین تقسیم می کنیم.

- اگر تعداد آزمون شوندگان بیش از ۴۰ نفر باشد، ۲۷ درصد از گروه بالا و ۲۷ درصد از گروه پایین را انتخاب می کنیم؛ البته این برای زمانی است که سے نمرات، دارای توزیع به هنجار هستند و برای آزمون های کلاسی ۲۵ تا ۳۳ درصد مناسب است. بعد از تهیه کردن گروه بالا و پایین باید اطلاعات زیر را برای هر سؤال به دست آورد:

۱) تعداد افراد گروه بالا، که هر یک از گزینه های سؤال را انتخاب کرده یا آن را بی جواب گذاشته اند.

۲) تعداد افراد گروه پایین، که هریک از گزینه های سؤال را انتخاب کرده یا آن را بی جواب گذاشته اند.

مثال ۱۲: هدف از تحلیل سؤال های یک آزمون چیست؟

1) وارسی تک تک آنها و تعیین میزان دقت و نارسایی آنها

۲) مشخص کردن تعداد افرادی که توانسته اند سؤال را صحیح پاسخ دهند.

۳) مشخص کردن تعداد افرادی که نتوانسته اند سؤال را صحیح پاسخ دهند.

۴) گزینه های ۲ و ۳

پاسخ: گزینه «۱» هدف از تحلیل سؤال های یک آزمون، وارسی تک تک آن ها و تعیین میزان دقت و نارسایی آن ها است.

مثال ۱۳: برای تحلیل سوالات یک آزمون به چه اطلاعاتی نیازمندیم؟

۱) تعداد پاسخ های صحیح به هر سؤال

۲) تعداد پاسخ های غلط به هر سؤال سمت

۳) تعداد سؤال های بی جواب

۴) همه موارد

پاسخ: گزینه «۴» برای تحلیل سؤالات یک آزمون به اصطلاحات زیر نیاز داریم.

1- تعداد پاسخ های صحیح هر سؤال

۲- تعداد پاسخ های غلط هر سؤال

۳- تعداد سؤال های بی جواب

مثال ۱۴: اگر تعداد آزمون شوندگان بیش از ۴۰ نفر و نمرات، توزیع بهنجار داشته باشند، گروه قوی و ضعیف را چگونه مشخص می کنیم؟

۱) ۲۵ درصد گروه بالا و ۲۵ درصد گروه پایین را انتخاب می کنیم

۲) ۲۷ درصد گروه بالا و ۲۷ درصد گروه پایین را انتخاب می کنیم.

۳) ۳۰ درصد گروه بالا و ۳۰ درصد گروه پایین را انتخاب می کنیم.

۴) ۳۷ درصد گروه بالا و ۳۷ درصد گروه پایین را انتخاب می کنیم.

پاسخ: گزینه «۲» در این حالت ۲۷ درصد گروه بالا و ۲۷ درصد گروه پایین را انتخاب می کنیم.

مثال ۱۵: اگر تعداد آزمون شوندگان بین ۲۰ تا ۴۰ نفر باشند، گروه قوی و ضعیف را چگونه مشخص می کنیم؟

1) کل برگه ها را به دو دسته پایین و بالا، قوی و ضعیف تقسیم می کنیم.

2) ۴۰ درصد از افراد گروه بام و ۴۰ درصد از افراد گروه پایین را انتخاب می کنیم.

۳) ۱۰ برگه بالا (قوی) و ۱۹ برگه پایین و ضعیف را انتخاب می کنیم.

۴) ۴۵ درصد از افراد گروه بالا و ۴۵ درصد از افراد گروه پایین را انتخاب می کنیم.

پاسخ: گزینه «3» در این حالت ۱۰ برگه ی گروه بالا (قوی) و ۱۰ برگه ی گروه پایین (ضعیف) را انتخاب می کنیم.

با استفاده از اطلاعات بالا می توانیم شاخصهای آماری زیر را محاسبه کنیم.

\_ضریب دشواری سؤال

- ضریب تمییز سؤال

محاسبه ضریب دشواری سؤال

درصد کل آزمون شوندگانی که به یک سؤال پاسخ درست می دهند، ضریب دشواری آن سؤال نام دارد که آن را با حرف P نشان می دهند. اگر در تحلیل یک سؤال، کلیدی برگه های دانش آموزان دخالت داشته باشند، برای محاسبه ضریب دشواری باید تعداد کل افرادی را که به سؤال پاسخ بر درست داده اند، به کل افراد تقسیم و در عدد ۱۰۰ ضرب کرد.

R: تعداد کسانی که پاسخ درست به سوال داده اند.

T: تعداد کل آزمون شوندگان

مثال ۱۶: به نسبت آزمودنی هایی که به سؤال پاسخ درست داده اند، چه گفته می شود؟

۱) انحراف معیار

۲) تمییز

3) آسانی

۴) دشواری

پاسخ: گزینه «۴» همانگونه که در سؤال قبل عنوان شد، ضریب دشواری به درصد کل آزمون شوندگانی که به یک سؤال جواب درست میدهند، گفته میشود؛ بنابراین خاطر نشان میشود که این ضریب با حرف p نشان داده می شود.

گزینه ۳ نام دیگر ضریب دشواری است که در کتب مربوطه به نام ضریب آسانی یا سهولت اشاره شده است؛ اما با توجه به این که ضریب دشواری در تمام کتب آورده شده است، گزینه ۳، گزینه ی انحرافی است.

مثال 17: اگر کل آزمون شوندگان، به یک سؤال پاسخ صحیح بدهند، ضریب دشواری آن ..................... خواهد بود.

1)25/0

2) 5/0

3) 75/0

4) 1

پاسخ: گزینه «۴» اگر کل آزمون شوندگان به یک سؤال پاسخ صحیح بدهند، ضریبا دشواری آن 1خواهد بود.

مثال 18: اگر تعداد پاسخ دهندگان به یک سؤال ۷۰ نفر باشند و از این تعداد ۶۲ نفر به سؤال جواب درست داده باشند، ضریب دشواری این سؤال چند است؟

1) 60

2) 78

3) 89

4) 80

پاسخ: گزینه «3»

اما اگر تعداد آزمون شوندگان زیاد باشد و ما آنها را به گروه بالا و گروه پایین تقسیم کرده باشیم، ضریب دشواری از فرمول زیر به دست می آید:

تعداد افراد گروه بالا + تعداد افراد گروه پایین

مثال ۱۹: در جدول مقابل، اگر گزینه «ب» گزینه صحیح باشد، ضریب دشواری سؤال چقدر است؟

|  |  |  |  |
| --- | --- | --- | --- |
| الف | ب | ج | د |
| 2 | 15 | 0 | 3 |
| 10 | 2 | 1 | 7 |

پاسخ های پیشنهادی

گروه قوی

گروه ضعیف   
1) 42/0

2) 44/0

3) 45

4) 51

پاسخ: گزینه «۱»

نکته ۱: هر چقدر ضریب دشواری یک سال بزرگتر باشد (به 100 نزدیک تر باشد) سؤال، آسان تر است و هر چقدر ضریب دشواری کوچکتر و به صفر نردیک تر می باشد. سؤال. دشوارتر است؛ بنابراین به جای ضریب دشواری می توان از ضریب آسانی با سهولت نام برده اما به طور معمول به آن ضریب دشواری می گویند.

اگر بخواهیم ضریب شد دشواری را به جای این که برحسب یک عدد صحیح دو رقمی نشاں نشیم. به صورت عدد اعشاری نشان دهیم. از فرمول زیر استفاده می کنیم :

نسبت افراد گروه پایین که به سؤال، جواب درست داده اند نسبت افراد گروه بالا که به سؤال، جواب درست داده اند

مثال ۲۰: چنانچه ضریب دشواری یک سؤال ۱ باشد، معنای آن چیست؟

1) عده ی قلیلی از آزمودنی ها به سؤال، پاسخ درست داده اند.

۲) تعداد زیادی از آزمودنی ها به سؤال، پاسخ داده اند.

۳) سؤال خیلی ساده بوده و همه آزمودنی ها به آن پاسخ درست داده اند.

۴) سؤال خیلی دشوار بوده و هیچ کس نتوانسته است به سؤال پاسخ درست دهد.

پاسخ: گزینه «3» ضریب دشواری، بنا به تعریف، درصد کل آزمون شوندگانی را شامل میشود که به یک سؤال، جواب درست می دهند و آن را با حرف p نشان می دهند. هر اندازه که ضریب دشواری یک سؤال، بزرگتر و به ۱۰۰ نزدیکتر باشد، آن سؤال، آسان تر است و هر اندازه که این ضریب کوچکتر باشد (به صفر نزدیکتر باشد)، سؤال، دشوارتر می شود؛ بنابراین به جای ضریب دشواری می توان از ضریب آسانی یا سهولت نیز در این باره نام برد. با توجه به تعریف، گزینه ۲ می تواند برای گزینهی صجیح، رقیب بسیار نزدیکی باشد و در صورتی که داوطلب به متن خواسته شده در سؤال، کم توجه باشد، دست به انتخاب این گزینه می زند.

مثال ۲۱: چنانچه نسبت افراد گروه بالا، که به یک سوال آزمونی پاسخ صحیح داده اند ۵۷/ه و نسبت افراد گروه پایین که به همان سؤال پاسخ اشتباه داده اند، 20/0 باشد، ضریب دشواری سوال چقدر خواهد بود؟

1) 35/0

2) 38/0

3) 40/0

4) 42/0

پاسخ: گزینه «۲»

مثال ۲۲: چنانچه ضریب دشواری یک سؤال برابر با ۸۰/0 باشد ................

۱) این سؤال، یک سؤال دشوار است.

۲) این سؤال، یک سؤال متوسط است.

۳) این سؤال، یک سؤال آسان است.

۴) با این اطلاعات نمی توان نظر داد.

پاسخ: گزینه «3» چون بیشتر دانش آموزان به آن سؤال، جواب صحیح داده اند، پس سؤال آسان بوده است.

مثال ۲۳: در جدول مقابل، اگر «الف» پاسخ صحیح باشد، درجه دشواری سؤال چه قدر است؟

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| پاسخ های پیشنهادی | الف | ب | ج | د |
| گروه قوی | 25 | 0 | 15 | 10 |
| گروه ضعیف | 15 | 20 | 10 | 5 |

1) %۸۰

2) %۴۰

3) %۳۰

4) %۱۰

پاسخ: گزینه «1» با توجه به فرمول محاسبه ضریب دشواری که در سؤال قبل آمده، پاسخ صحیح گزینه ۲ است.

تفسیر ضریب دشواری: در تفسیر ضریب دشواری، اگر با آزمون های وابسته به هنجار سرو کار داشته باشیم، می توانیم این ضریب را از لحاظ آماری بررسی می کنیم. هرچه واریانس نمرات یک آزمون وابسته به هنجار، بزرگتر باشد، آن آزمون، آزمون بهتری است. یکی از راه هایی که از طریر آن می توانیم در مورد مفید بودن سؤال های یک آزمون قضاوت کنیم، این است که ببینیم سؤال، تا چه اندازه به واریانس نمرات آزمون کمک می کند.

واریانس نمرات یک آزمون از دو بخش تشکیل شده است: ۱) واریانس سؤال ها ۲) همبستگی سؤال ها

هر سؤالی که دارای واریانس بزرگتری باشد و با سایر سؤال های آزمون، همبستگی بالاتری داشته باشد به واریانس کل آزمون، کمک بیشتری می کند. ، در مقابل، هر سؤالی که با سایر سؤال ها همبستگی ندارد و دارای واریانس کوچکی است، کمک زیادی به واریانس کل آزمون نمی کند؛ بنابراین:

اگر در تحلیل سؤالات آزمون، سؤالاتی را انتخاب کنیم که همبستگی بالایی با هم دارند و دارای واریانس بزرگتری هم هستند، در مجموع، آزمون بهتری را تدارک دیده ایم. . اگر همه سؤال های یک آزمون با یکدیگر همبستگی کامل داشته باشند و ضریب دشواری همه آن ها 5/0باشد، آن گاه نیمی از آزمون شوندگان در آزمون، نمره صفر و نیم دیگر، نمره کامل خواهند گرفت.

مثال ۲۴: در چه صورتی می توان گفت که نیمی از آزمون شوندگان در یک آزمون، نمره صفر و نیم دیگر نمره ی کامل گرفته اند؟

1) همه ی سؤال ها همبستگی بالایی با هم داشته باشد و ضریب تمییز همه ی آنها ۱ باشد.

۲) همه ی سوال های آزمون، همبستگی کامل با هم داشته باشند و ضریب تمییز همه آنها ۵/0 باشد.

3) سؤال ها همبستگی کمی با هم داشته باشند و ضریب دشواری همه ی آنها ۱ باشد.

۴) همه ی سوال های آزمون، همبستگی کامل با هم داشته و ضریب دشواری همه ی آنها ۵/ه باشد.

پاسخ: گزینه «۴» اگر همه سؤال های آزمون، همبستگی کامل با هم داشته باشند و ضریب دشواری همه آن ها 5/0 باشد، نیمی از آزمون شوندگان در یک آزمون، نمره ی صفر و نیم دیگر نمره کامل گرفته اند.

فرمول برای محاسبه واریانس یک سؤال:

P=(1-P)واریانس سوال

P: نسبت کسانی که به سؤال پاسخ درست داده اند، یعنی همان ضریب دشواری

مثال ۲۵: در صورتی که ضریب دشواری یک سؤال 75/0 باشد، واریانس این سؤال چقدر خواهد بود؟

1) 19/0

2) 22/0

3) 25/0

4) 27/0

پاسخ: گزینه «۱»

واریانس سوال

یک سؤال زمانی دارای حداقل واریانس است که ضریب دشواری آن صفر یا یک باشد؛ یعنی هیچ یک از آزمون شوندگان به سؤال، جواب درست نداده اند یا این که همه ی آزمون شوندگان به آن جواب درست داده باشند.

مثال ۲۶: محققی آزمونی را روی ۵ دانش آموز اجرا کرده است. داده های حاصله برای سوال اول آزمون، به شرح جدول زیر است. واریانس سؤال کدام است؟ (به پاسخ درست، نمره ی ۱ و به پاسخ غلط، نمره ی صفر داده شده است).

پاسخ: گزینه «1» برای پاسخ دادن به این سؤال، ابتدا ضریب دشواری سؤال را به دست می آوریم، سپس آن را در فرمول واریانس سؤال قرار میدهیم تا جواب به دست آید. به ترتیب زیر:

واریانس هر سوال

نکته 2: اگر ضریب دشواری سوال 5/0 P = ،واریانس سوال حداکثر خواهد بود.

= P(1-P) 0/5 (1- 0/5) = 0/25 واریانس سوال

نکته 3: بهتر است سوالاتی را انتخاب کنیم که ضریب دشواری آن ها از کمتر، به 5/0 نزدیک تر و از صفر بیشتر باشد.

البته انتخاب ضریب دشواری مناسب، به نوع سؤال بستگی دارد. در آزمون های صحیح - غلط، سطح دشواری 5/0 P = زمانی به دست می آید که همه در آزمودنی ها با حدس زدن به سؤال، جواب دهید. برای آزمون های چند گزینه ای، سطح بهینه دشواری، اندکی کمتر از فاصله بین یک و سطح موفقیت با حدس زدن است. سطح موفقیت با حدس زدن عبارت است از نسبت پاسخ های درستی که با حدس زدن به دست می آید. برای سؤال های چهار گزینه ای و سطح موفقیت با حدس زدن ۲۵/0 است؛ لذا سطح بهینه دشواری در وسط 25/0 و ۱، یعنی در حدود ۶/0 است.

برای این که دشواری بهینه سؤال را به دقت تعیین کنیم، ابتدا سطح موفقیت ۱۰۰% را از سطح عملکرد با حدس زدن (مثلا سطح عملکرد با حدس زدن و در آزمون چهار گزینه ای: ۲۵%) کم می کنیم و بر ۲ تقسیم می کنیم؛ در نتیجه نقطه ی میانی به دست می آید، سپس نقطه میانی را با سطح عملکرد با حدس زدن، جمع می کنیم تا سطح بهینه دشواری به دست آید:

ضریب دشواری بهینه 625/0 25/0 375/0

یک راه ساده تر هم برای تعیین ضریب دشواری بهینه این است که ۱ را با سطح عملکرد با حدس زدن، جمع کنیم و بعد بر ۲ تقسیم کنیم:

مثال ۲۷: یک سؤال در صورتی بیشترین کمک را به واریانس آزمون می کند که ضریب دشواری آن ................ باشد.

1) صفر

۲) 25/0

۳ 5/0

۴) ۱

پاسخ: گزینه «3» زمانی که ضریب دشواری سؤال 5/0 P= باشد، واریانس سؤال حداکثر خواهد بود.

در نتیجه از لحاظ انتخاب برای گنجاندن در فرم نهایی آزمون، سوال هایی بهتر هستند که ضریب دشواری آنها به ۵/0نزدیک باشند. گزینه شماره 2 به شدت، انتخاب درست را تهدید می کند؛ زیرا داوطلب پس از محاسبه واریانس سؤال، به عدد ۲۵/0 می رسد و به اشتباه دست به انتخاب گزینه ۲ میزند، در حالی که متن سؤال از ضریب دشواری سؤال آزمون پرسیده است.

مثال ۲۸: اگر از ۲۵ نفر گروه قوی، همگی به یک سؤال، پاسخ درست داده باشند و از ۲۵ نفر گروه ضعیف، هیچ کس پاسخ نداده باشد، مقدار انحراف معیار جانسون چه قدر خواهد بود؟

۱) صفر

۲) 5/0

3) 75/0

4) 1

پاسخ: گزینه «۱»

مثال ۲۹- در یک آزمون چند گزینه ای، از ۵۰ دانش آموز گروه قوی، ۴۰ نفر و از ۵۰ دانش آموز گروه ضعیف، ۳۰ نفر به یک سؤال پاسخ صحیح داده اند. سطح دشواری این سؤال چه قدر است؟

1) 70 درصد

2) 40 درصد

3) 35 درصد

4) 30 درصد

پاسخ: گزینه «1»

50 گروه بالا

50 گروه پایین

مفروضات سوال 40 انتخاب درست گروه بالا

30 انتخاب درست گروه پایین

نکته ۴: به طور کلی تمرین های دشواری بین 3/0 تا 7/0 حداکثر تفاوت بین آزمودنی ها را نشان می دهند. با این حال در شرایط ویژه ضریب دشواری، فرق می کند. مثلا آزمونی که به صورت آزمون ورودی فوق لیسان یا دکتری برای انتخاب بهترین داوطلبان به کار می رود، حدود ۱۰% بالای داوطلبان را برمی گزیند که باید شامل سوال های بسیار دشوار باشد، در مقابل: آزمونی که برای انتخاب کودکان برای به دوره آموزش ترمیمی یا اصلاحی مورد استفاده قرار می گیرد، باید شامل سوال های بسیار آسان باشد.

محاسبه ضریب تمییز سؤال

ضریب تمییز که با d نشان داده می شود، قدرت سؤال را در تمییز گذاری بین گروه قوی و ضعیف نشان می دهد؛ یعنی مشخص می کند سؤال تا چه اندازه میتواند گروه ضعیف را از قوی جدا سازد. برای محاسبه ضریب تمییز از فرمول زیر استفاده میشود

انتخاب های درست گروه پایین - انتخابهای درست گروه بالا

d =

تعداد افراد یک گروه (بالا یا پایین)

مثال ۳۰: شاخص دشواری سؤال از لحاظ دامنه ی ارزش بین کدام اعداد می تواند در تغییر باشد؟

1) از صفر تا 1-

2) از صفر تا ۱+

۳) از1+ تا ۱-

۴) از ۳+ تا ۳-

پاسخ: گزینه «۲» ضریب دشواری سؤال، طیفی است که شامل صفر تا ۱+ می باشد. ضریب تمییز نیز از ۱+ تا ۱- را شامل می شود. ضریب دشواری 0 (صفر) زمانی به دست می آید که هیچ یک از آزمون شوندگان نتوانند به سؤال، پاسخ صحیح دهند و ضریب دشواری نیز زمانی به دست می آید که تمام آزمون شوندگان به سؤال مورد نظر پاسخ صحیح داده باشند.

مثال ۳۱: یکی از معایب قدرت تمییز سؤال آن است که:

۱) در محاسبه آن از تمام پاسخها استفاده نمی شود.

۲) تعداد گزینه ها در محاسبه آن تأثیر ندارد.

۳) پایداری لازم در بین سؤال ها را نشان نمی دهد.

۴) اساس ریاضی و آماری ندارد.

پاسخ: گزینه «۱» ضریب تمییز به قدرت سؤال در تمایزگذاری و تشخیص بین گروه قوی و ضعیف آزمون شوندگان اشاره دارد؛ یعنی معلوم می کند که یک سؤال، تا چه اندازه گروه های قوی و ضعیف را از هم جدا می نماید. ضعف ضریب تمییز، در نظر نگرفتن ورقه ها یا نمرات گروه متوسط است که در محاسبات، نادیده گرفته می شود.

مثال ۳۲: در جدول مقابل، اگر گزینه «ج» گزینه صحیح باشد، ضریب تمییز سؤال چقدر است؟

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| پاسخ های پیشنهادی | الف | ب | ج | د |
| گروه قوی | 4 | 2 | 12 | 7 |
| گروه ضعیف | 9 | 6 | 7 | 3 |

1) 35/0

2) 30/0

3) 20/0

4) 45/0

پاسخ: گزینه «3»

مثال ۳۳: جدول مقابل، تحلیل پاسخ های یک سؤال را نشان می دهد. اگر گزینه ی « الف» پاسخ صحیح باشد، «درجه دشواری و قدرت تمییز سؤال»، به ترتیب چقدر است؟

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| پاسخ های پیشنهادی | الف | ب | ج | د |
| 27% اوراق از بالا | 16 | 3 | 1 | 5 |
| 27% اوراق از پایین | 7 | 8 | 1 | 9 |

1) %18 و 46/0

2) %23 و 36/0

3) %46 و 36/0

4) %92 و 18/0

پاسخ: گزینه «3»

نکته: اگر بخواهیم ضریب تمییز را بر اساس نسبت پاسخ های درست و غلط به دست آوریم، از فرمول زیر استفاده می کنیم:

نسبت پاسخ های درست گروه پایین – نسبت پاسخ های درست گروه بالا d=

انتخاب های درست گروه پایین

انتخاب های درست گروه بالا

R= تعداد پاسخ درست

T= کل آزمون شوندگان.

با توجه به متن سؤال، درجه دشواری برابر با 46/0 و قدرت تمییز سؤال برابر با 36/0 می باشد.

مثال ۳۴: در یک آزمون چهار گزینه ای، از ۶۰ نفر گروه قوی، ۴۵ نفر و از ۶۰ نفر گروه ضعیف، ۲۷ نفر به یک سؤال، پاسخ صحیح داده انـد. قـدرت تمییز این سؤال چقدر است؟

1) 15/0

2) 30/0

3) 45/0

4) 60/0

پاسخ: گزینه «2» ابتدا درصد گروه قوی و گروه ضعیف را محاسبه قرار می دهیم:

D ضریب تمییز سوالاتی است که بتواند افراد قوی و ضعیف را از هم جدا کند.

تفسیر ضریب تمییز: هرچه ضریب تمییز بزرگتر باشد، آن سؤال بهتر می تواند آزمون شوندگان قوی و ضعیف را از هم جدا کند؛ اما اگر ضریب تمییز سؤال، پایین باشد، آن سؤال نمی تواند خیلی خوب، آزمون شوندگان قوی و ضعیف را هم جدا کند. اگر ضریب تمییز، صفر باشد، یعنی آن سؤال به هیچ وجه نتوانسته است، گروه قوی و ضعیف را از هم جدا کند. اگر ضریب تمییز، منفی باشنه نشان دهنده ی این است که گروه قوی بدتر از گروه ضعیف عمل کرده اند، که این گونه سؤال ها دارای معایب اساسی اند و باید در آنها تجدیدنظر کرد. اگر سؤال اشکال فنی نداشته باشد، دلیل چنین ضریبی می تواند این باشد که افراد گروه بالا، یا آن سؤال را به طور کامل یاد نگرفته اند یا آن را به غلط آموخته اند؛ بنابراین تمام ضرایب تمییز سؤالات باید مثبت باشند. البته در آزمون های سنجش شخصیت، علاقه و نگرش، چون تمییز دادن بین سنخها یا گروه های مختلف کاری مطلوب است، ضریب تمییز منفی نامناسب نیست.

مثال ۳۵: شاخص تمییز یک سؤال چند گزینه ای برابر با ۱ محاسبه شده است. این بدان معنی است که سؤال:

1) قدرت تشخیص کمی دارد.

۲) دارای بیشینه ی قدرت تشخیص است.

3) قدرت تشخیص بالایی دارد.

۴) فاقد توانایی تمییز دانش آموزان قوی و ضعیف است.

پاسخ: گزینه «۲» طبق تعریف، هرچه قدر ضریب تمییز بزرگ تر باشد، قدرت تشخیص سؤال، بیشتر و هرچقدر این ضریب، کوچک تر باشـد، قـوه تمییز کمتری دارد. ضریب تمییز، قدرت تمایزگذاری بین گروه قوی و گروه ضعیف است که بین عدد ۱+ و ۱- قرار دارد. زمانی که ضریب به دست آمـده عدد ۱ باشد، یعنی سؤال درباره ی تمییز گذاشتن بین گروه قوی و گروه ضعیف از بیشترین قدرت برخوردار است. گزینه ۲ می تواند جایگزین نادرستی برای گزینه ی صحیح باشد.

همبستگی سؤال با کل آزمون

ضریب تمییز سؤال، نوعی ضریب توافق سؤال با کل آزمون است، مثلاً سؤالی را درنظر بگیرید که در آن، افرادی که در کل آزمون، نمره بالایی گرفته اند، آن سؤال هم نمره بالایی گرفته اند یا سؤالی را در نظر بگیرید که درآن، کسانی که در کل آزمون نمره پایینی گرفته اند، در آن سؤال هم نمره پاینی گرفته اند؛ پس این سؤال با کل آزمون همخوانی دارد و دارای ضریب تمییز مثبت و بالاست. در مقابل، سؤالی که در آن افراد توی کی در کل آزمون، نمره خوبی گرفته اند، در آن سؤال نمره ی پایینی گرفته باشند و افرادی که در کل آزمون، نمره پایین گرفته اند، در آن سؤال، نمره بالایی گرفته باشند، آن سؤال دارای ضریب تمییز منفی است.

نکته 5: سؤالات، دارای قوه تمییز بالا، با کل آزمون، همبسته اند، اما سؤال های دارای قوه تمییز صفر یا منفی، با کل آزمون، همبسته نیستند یا همبستگی منفی دارند.

پس یکی دیگر از راه های محاسبه ضریب تمییز سوال های یک آزمون، محاسبه ی ضریب همبستگی بین نمره آزمون شوندگان، در آن سوال ها و نمرات آنان در کل آزمون است که این همبستگی به صورت همبستگی دورشته ای نقطه ای محاسبه می شود، هرچه این همبستگی بالاتر باشد، ضریب تمییز، بالاتر و هرچه کمتر باشد، ضریب تمییز پایین تر است، و یا سایر سؤال های آزمون، همخوانی کمتری دارد اگر، ضریب همبستگی، منفی باشد، یعنی سؤال، با کل آزمون نا همخوان است و چیزی را می سنجد که با آنچه سؤال های دیگر می سنجند، متفاوت است و اگر ضریب همبستگی، صفر باشد، ضریب تمییز هم صفر است. از امتیازهای استفاده از ضریب همبستگی بین سؤال و کل آزمون، به عنوان ضریب تمییز، این است که می توان این ضریب را هم به لحاظ آماری و هم به لحاظ عملی آزمون کرد اگر عود منگی س سؤال و کل آزمون را به توان ۲ برسانیم، می توانیم مشخص کنیم که آن سوال چند درصد از واریانس نمرات را به خود اختصاص داده است. برای مثال اگر همبستگی سؤال با کل آزمون60/0 + باشد. این سؤال ۳۶ درصد واریانس نمرات را به خود اختصاص می دهد، در حالی که ضریب تمییز به دست آمده از فرمول محاسبه ضریب تمییز را نمی توان این گونه تفسیر کرد.

علاوه بر این ها، همبستگی بین سؤال و آزمون مستقیماً به پایایی آزمون مربوط است.

مثال ۳۶: کدام شاخص برای تفاوت گذاشتن بین دارندگان نمرات خوب و بد به کار برد کار برده می شود؟

۱) اعتبار

۲) ضریب پایایی

۳) درجه دشواری

۴) همبستگی دورشته ای نقطه ای

پاسخ: گزینه «۴» همبستگی دو رشته ای نقطه ای به وضوح می تواند بین دو گروه ضعیف و قوی، قدرت تمییز داشته باشد؛ یعنی افراد گروه قوی را از افراد گروه ضعیف جدا کند. گزینه ۳ می تواند داوطلبان را در انتخاب درست دچار تردید نمایـد، امـا بـایـد توجه داشت که ضریب دشواری بـه کـل آزمون شوندگانی اشاره دارد که به یک سؤال، جواب درست می دهند؛ بنابراین با ضریب تمییز کاملاً تفاوت دارد.

مثال ۳۷: در یک تست، همبستگی هر سؤال با بقیه سؤال ها کدام یک از خصوصیات زیر را نشان می دهد؟

۱) درجه دشواری

۲) هماهنگی درونی

۳) قدرت تشخیص

۴) روایی محتوی

پاسخ: گزینه «۳» یکی از راه های محاسبه ضریب تمییز سؤال های یک آزمون، محاسبه ضریب همبستگی بین نمره های آزمـون شـوندگان در سؤال ها و نمرات آنان در کل آزمون است که این ضریب همبستگی، معمولاً به صورت همبستگی دو رشته ای نقطه ای محاسبه می شود؛ در واقع، ضریب توافق سؤال با کل آزمون است که هرچه ضریب همبستگی بیشتر باشد، بدان معنی است که سؤال هم خوانی بیشتری با مجموعه سؤال های دیگر دارد و باعث می شود افراد ضعیف و قوی تمایز بیشتری از هم با جدا شوند.

مثال ۳۸: اگر سؤالی با کل آزمون ۰/۸۰ همبستگی داشته باشد، این سؤال، چند درصد واریانس نمرات را به خود اختصاص می دهد؟

1) 60/0

2) 64/0

3) 70/0

4) 74/0

پاسخ: گزینه «۲»

همبستگی بین سؤال های آزمون

بهتر است در هنگام تحلیل آزمون، همبستگی تک تک سؤال ها بایکدیگر را هم تعیین کنیم، این کار دو فایده دارد:

۱- با در دست داشتن تعداد سؤال های آزمون و میانگین ضرایب همبستگی بین سؤال ها، می توان ضریب پایایی آزمون را محاسبه کرد.

۲ـ می توان از طریق آن به قوه تمییز سؤال دست یافت.

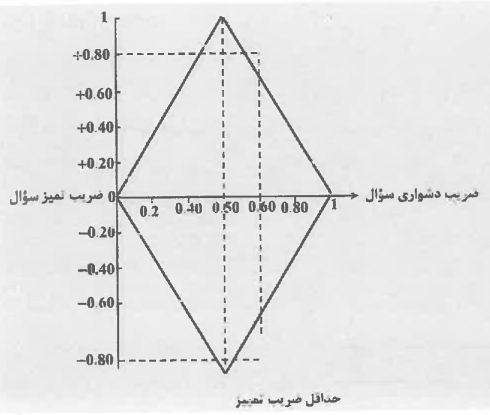
اگر همبستگی بین یک سؤال و کل آزمون کم باشد، دو علت برای آن وجود دارد:

1- ممکن است سؤال مورد نظر با هیچ یک از سؤال های آزمون همبسته نباشد؛ در این حالت یا باید سؤال را بازنویسی کرد یا آن را کاملا کنار گذاشت .

۲- ممکن است این سؤال با بعضی از سؤال های آزمون همبستگی قابل ملاحظه ای داشته باشد، اما با سؤال های دیگر، همبستگی نزدیک به صفر یا حتی منفی داشته باشد. این حالت، زمانی پیش می آید که یک آزمون، دو ویژگی کاملا متمایز را می سنجد؛ مثلا درک کلامی و توانایی ریاضی که طبق نظریه پایایی، این آزمون ها اندازه های با ثباتی را به دست نمی دهند.

رابطه ی بین ضریب دشواری و ضریب تمییز

با استفاد از شکل روبه رو می توان رابطه بین ضریب دشواری و ضریب تمییز را نشان داد حدود چهار ضعلی این شکل، حداکثر و حداقل ضریب دشواری و تمییز را نشان می دهد؛ برای مثال اگر ضریب دشواری برابر 60/0 باشد. ضریب تمیز آن بین 80/0- و 80/0+ خواهد بود. تنها زمانی ضریب تمییز به حداکثر یعنی ۱ می رسد که ضریب دشواری 5/0 باشد. بنابراین در یک آزمون ، سوالات خوب، سوالاتی هستند با ضریب دشواری متوسط و ضریب تمیز بزرگ.



مثال ۳۹: اگر شاخص تمییز (تشخیص) یک سؤال ۱ باشد، شاخص دشواری این سؤال چقدر خواهد بود؟

۱) ۱-

2) 50/0 -

۳) ۵۰/0+

4) 1

پاسخ: گزینه «3» شاخص تمییز سؤال یعنی قدرت سؤال در تمایزگذاری یا تشخیص بین گروه ضعیف و قوی که آن را با حرف a نشان می دهند. شاخص ضریب دشواری که با حرف p نشان داده می شود، میزان آزمون شوندگانی را نشان می دهد که به یک سؤال، جواب درست می دهند. هرچه قوه ی تمییز یک سؤال بالاتر و بیشتر باشد، بیشتر می تواند بین گروه قوی و ضعیف تمایز قایل شود. زمانی ضریب تمییز به حداکثر 1+ می رسد، که ضریب دشواری برابر با ۵/0، یعنی در حد وسط، باشد؛ در نتیجه سؤال های خوب یک آزمون، سؤال هایی هستند که ضریب دشواری متوسط و ضریب تمییز بزرگ دارند. گزینه ۲ به علت مشابهت با گزینه ی ۳ می تواند به پاسخ درست نزدیک باشد؛ اما داوطلب با دقت به علامت منفی توجه می کند و گزینه ی نادرست را انتخاب نمی کند.

مثال ۴۰: در یک آزمون هوش، از دو سؤال که ضریب تشخیص برابر دارند، آن سؤال مناسب تر است که:

1) سطح دشواری آن زیادتر باشد.

۲) سطح دشواری آن کمتر باشد.

۳) دارای سطح دشواری متوسط باشد.

۴) دارای واریانس کوچکتری باشد.

پاسخ: گزینه «3» سؤال خوب در یک آزمون سؤالی است که ضریب تمییز آن بزرگ باشد و بتواند بین گروه قوی و گروه ضعیف تمایز قایل شود، همچنین ضریب دشواری برابر 5/0، یعنی در حد وسط، داشته باشد. تنها در صورتی که ضریب تمییز به حداکثر برسد و ضریب دشواری برابر 5/0 ، یعنی و در حد وسط باشد، سؤال از کیفیت مناسبی برخوردار است. گزینه ۱ می تواند به گزینه ی صحیح، نزدیک قلمداد شود، اما با توجه به این که داوطلب می داند ضریب دشواری، همان ضریب آسانی یا سهولت است، از انتخاب گزینه نادرست خودداری می کند.

مثال ۴۱: در ارزشیابی های هنجاری، سؤالی ایده آل محسوب میشود که مقدار «ضریب دشواری و ضریب تغییز» آن به ترتیب، در چه حدودی باشد؟

1) متوسط - بالا

۲) بالا - متوسط

3) متوسط - متوسط

۴) بالا- بالا

پاسخ: گزینه «۱» سؤال خوب در یک آزمون، سؤالی است که ضریب دشواری متوسط، برابر 5/0P= و ضریب تمییز بزرگ (نزدیک به 1) دارد. ضریب تمییز به قدرت سؤال در تمایزگذاری بین گروه ضعیف و گروه قوی آزمون شوندگان، می گویند.

گزینه ی ۲ به واسطه هم شکلی با گزینه ی صحیح، می تواند انتخاب نادرستی باشد که داوطلبان، با صرف دقت لازم می توانند از این اشتباه جلوگیری کنند.

مثال ۴۲: دانش آموزان یک کلاس، به خوبی به هدف های آموزشی دست یافته اند، در این صورت، انتظار می رود که سؤالات امتحانی دارای ضریب دشواری ........ و ضریب تمییز .............. باشند.

۱) پایین و پایین

۲) بالا و پایین

۳) بالا و بالا

۴) پایین و پایین

پاسخ: گزینه «۲» برای پاسخ گویی به این سؤال، فرض کنید کلاسی داریم که تعداد آنها ۲۰ نفر است و از این ۲۰ نفر، ۱۵ نفر به یک سؤال پاسخ صحیح داده اند. با توجه به فرمول ضریب دشواری ( P)، ضریب دشواری این سؤال می باشد که ضریب دشواری بالایی است و همچنین با توجه به فرمول ضریب تمییز ) اگر از گروه بالا، ۹ نفر به سؤال، جواب صحیح داده باشند و از گروه پایین، ۶ نفر جواب صحیح داده باشند، ضریب تمییز این سؤال

خواهد بود که ضریب تمییز پایین است.

مثال ۴۳: اگر ضریب قوه تمییز (تفکیک) یک سؤال برابر ۱ باشد، ضریب دشواری سؤال چقدر خواهد بود؟

1) 5/0-

۲) صفر

۳) ۵/0

۴) یک

پاسخ: گزینه «3» اگر همه ی سؤال های یک آزمون با یکدیگر همبستگی کامل داشته باشند و ضریب دشواری همه ی آن ها 5/0باشند. آن گاه نیمی از آزمون شوندگان، در آزمون، نمره صفر می گیرند و بقیه، نمره ی کامل را کسب می کنند؛ در نتیجه از لحاظ انتخاب برای گنجانیدن در فرم های آزمون، سؤال های با ضریب دشواری 5/0 از همه بهتر است. در این حالت ضریب تمییز سؤال (یعنی قدرت تمایزگذاری گزوه ضعیف و گروه قوی توسط سؤال)، گزینه ی ۱ است.

تحلیل گزینه های انحرافی

در تحلیل گزینه های انحرافی، باید دقت کرد که هدف از قرار دادن گزینه های انحرافی این است که کسانی که جواب درست را نمی دانند، گمراه شوند. در تحلیل گزینه های انحرافی، قاعده ی کلی این است که هر گزینه انحرافی دست کم باید یک نفر از افراد گروه ضعیف را به خود جلب کند و اگر هر دو گروه ضعیف و قوی را جلب کرد، باید تعداد افراد گروه ضعیف بیشتر از گروه قوی باشد.

تحلیل سؤال های آزمون های تشریحی و آزمون های عملکردی

۱- روش ویتنی و سیبرز: در این روش، ابتدا ۲۵ درصد از گروه بالا و ۲۵ درصد از گروه پایین را انتخاب می کنیم و برای هر سؤال، مجموع نمرات گروه بالا و پایین را حساب می کنیم و در فرمول های زیر قرار می دهیم:

مجموع نمرات برگه های افراد گروه بالا (۲۵ درصد)

مجموع نمرات برگه های افراد گروه پایین (۲۵ درصد)

N تعداد افراد یا برگه های ۲۵ درصد گروه موردنظر

پایین ترین نمره ممکن در سؤال

بالاترین نمره ممکن در سوال

مثال ۴۴: فرض کنید آزمونی تشریحی را برای ۲۰ آزمون شونده اجرا کرده ایم، اگر برای سؤالی و باشد و پایین ترین نمره و بالا ترین نمره ممکن به ترتیب 0 و ۵ باشد، ضریب دشواری این سؤال چند خواهد بود؟

1) 60/0

2) 62/0

3) 65/0

4) 67/0

پاسخ: گزینه «2»

مثال ۴۵: اطلاعات زیر در مورد یک آزمون تشریحی است. ضریب دشواری و تمییز را به دست آورید.

N=30

1) 58/0 – 15/0

2) 67/0 – 13/0

3) 67/0 – 20/0

4) 67/0 – 15/0

پاسخ: گزینه «1»

و

۲- روش نیتکو: در این روش، ضریب دشواری یک سؤال تشریحی یا عملکردی به صورت نمره میانگین گروه آزمون شوندگان برای سؤال، تعریف شده است. برای این که ضریب دشواری را به مقیاس ضریب دشواری سؤال های دارای ارزش صفر و یک، یعنی P، تبدیل کنیم، باید این ضریب را به دامنه ممکن نمرات سؤال تقسیم کنیم:

بنابراین ضریب دشواری سوالی که دامنه نمرات، 1 تا 7 و نمره میانگین، 4/4 است، به صورت زیر محاسبه می شود:

در تفسیر ضریب دشواری 73/0 P = می گوییم به طور متوسط، پاسخ دهندگان به میزان 73%، بالاترین نمره ممکن را در سؤال به دست آورده اند.

برای به دست آوردن ضریب تمییز با روش نیتکو، تفاوت بین میانگین گروه بالا و گروه پایین را برای سؤال محاسبه می کنیم. برای این که این ضریب را به مقیاس ضریب تمییز سؤال های دارای ارزش صفر و یک تبدیل کنیم، آن را به دامنه ممکن نمرات سؤال، تقسیم می کنیم:

مثال ۴۶: اگر نمره میانگین سؤال برای گروه بالا 1/5 و برای گروه پایین 2/3 باشد و دامنه نمرات ۱ تا ۷ باشد، ضریب تمییز را محاسبه کنید.

1) 25/0

2) 32/0

3) 35/0

4) 44/0

پاسخ: گزینه «۲»

نکته 6: از آنجا که ما تفاوت بین نمرات میانگین گروه های بالا و پایین را به دامنه ممکن نمرات سؤال تقدیم می کنیم، می توانیم نتیجه حاصل را این گونه تفسیر کنیم که تفاوت بین نمرات میانگین گروه های بالا و پایین برای این سؤال ۳۲/0 دامنه ممکن سؤال است.

پس از تعیین ضریب تمییز و دشواری، باید سؤال ها را اصلاح کنیم و در مورد سؤالاتی که ضریب دشواری آنها بسیار بالا یا پایین است یا ضریب تمییز آنها خیلی کم است، تجدید نظر کنیم.

تحلیل سؤال های آزمون های وابسته به ملاک

ضرایب دشواری و تمییزی، که تاکنون در مورد آن ها صحبت شد، صرفا برای آزمون های وابسته به هنجار مناسب اند، نه آرمون های وابسته به ملاک. در آزمون های وابسته به ملاک؛ چون نتیجه آرمانی این است که اکثر پاسخ دهندگان، اکثر سؤال ها را درست جواب دهند، ضریب دشواری، بسیار بزرگ و ضریب تمییز، بسیار کوچک می شود. به خاطر همین، محاسبه ضریب دشواری و تمییز خیلی قابل استفاده نیست.

ما در اینجا دو نوع ضریب تمییز را برای آزمون های وابسته به ملاک معرفی می کنیم:

1- تمییز بین پیش و پس از آموزش

آزمون های وابسته به ملاکی که در سنجش یادگیری در حد تسلط (یعنی یادگیری مطالب آموزش داده شده تا یک حد معین) مورد استفاده قرار می گیرند، باید سؤالاتی را شامل شوند که پس از آموزش مطالب درسی، آسان تر از قبل از آموزش آن مطالب جلوه کنند.

فرمول ضریب تمییز پیش و پس از آموزش، یا ضریب حساسیت نسبت به آثار آموزش، که با S یا D نشان داده می شود:

: تعداد کسانی که پس از آموزش به سؤال پاسخ درست داده اند.

: تعداد کسانی که پیش از آموزش به سؤال پاسخ درست داده اند.

T: تعداد کل افراد

چنانچه ضریب محاسبه شده برای یک سؤال ۱+ باشد، آن سؤال، سؤال بسیار خوبی است؛ چون همه دانش آموزان، پیش از آموزش، آن را غلط جواب داده اند ام نداده و پس از آموزش و به آن جواب صحیح داده اند. چنانچه ضریب محاسبه شده صفر (۰) باشد، این سؤال یا سؤال بسیار ساده ای است و همه ی دانش آموزان پیش و پس از آموزش به آن جواب صحیح داده اند، یا سؤال بسیار دشواری است و نتوانسته است آثار آموزش را اندازه گیری کند، با این که آموزش در رابطه با موضوع این سؤال ناکافی بوده است. اگر ضریب محاسبه شده (۱-) باشد، سؤال مورد نظر، سؤال بسیار معیوبی است یا سؤال آسانی است که با آموزش غلط دنبال شده است؛ چون تمام دانش آموزان، پیش از آموزش به آن جواب صحیح داده اند، اما پس از آموزش، آن را اشتباه پاسخ داده اند. در صورتی که ضریب محاسبه شده ۵/0 باشد، سؤال مورد نظر، سؤال کارآمدی است با یک الگوی نسبتا طبیعی؛ زیرا بعضی از دانش آموزان، پیش از آموزش به آن جواب درست داده اند، اما تعداد بیشتری از دانش آموزان، پس از آموزش به آن جواب درست داده اند. شرایط آرمانی برای هر سؤال این است که ضریب تمییز ۱+ با شد، یعنی پیش از آموزش، هیچ یک از دانش آموزان، جواب درست را نمی دانسته اند و در نتیجه ی آموزش، همه دانش آموزان جواب درست را یاد گرفته اند.

مثال ۴۷: چنانچه ضریب تمییز پیش و پس از آموزش یک سؤال ۱ باشد ............

۱) تمام آزمون شوندگان آن را به اشتباه جواب داده اند.

۲) تمام آزمون شوندگان، قبل از آموزش آن را درست ولی بعد از آموزش آن را غلط جواب داده اند.

۳) تمام آزمون شوندگان، هم پیش و هم پس از آموزش، آن را صحیح پاسخ داده اند.

۴) تمام آزمون شوندگان قبل از آموزش آن را غلط ولی بعد از آموزش، آن را درست جواب داده اند.

پاسخ: گزینه «۴» چنانچه ضریب تمییز پیش و پس از آموزش یک سؤال، ۱ باشد، تمام آزمون شوندگان، قبل از آموزش، آن را غلط ولی بعد از آموزش، آن را درست جواب داده اند.

تمییز بین حد تسلط و غیر حد تسلط

چون آزمون های وابسته به ملاک یک نمره مرزی برای حد تسلط دارند، لازم است که معلوم شود آیا سؤال های آزمون در آن نقطه بین آزمون شوندگان قوی و ضعیف، تمییز قائل می شود یا نه. پس ما باید یک ضریب تمییز بین حد تسلط و غیر حد تسلط، محاسبه کنیم که برای محاسبه آن، یک چهارخانه تشکیل می دهیم که در خانه های عمودی آن حد تسلط و غیرتسلط و در خانه های افقی آن، درست و غلط قرار داده شده است.

آزمون

a + c b + d

حد تسلط حد غیر تسلط

|  |  |  |  |
| --- | --- | --- | --- |
| سوال | درست  غلط | b | a |
| d | c |

aتعداد آزمون شوندگانی که سؤال را درست جواب داده اند، اما به حد تسلط نرسیده اند

b: تعداد آزمون شوندگانی که سؤال را درست جواب داده اند، اما به حد تسلط هم رسیده اند.

c: تعداد آزمون شوندگانی که سؤال را غلط جواب داده اند و به حد هم تسلط نرسیده اند.

d: تعداد آزمون شوندگانی که سؤال را غلط جواب داده اند ولی به حد تسلط رسیده اند.

در اینجا با توجه به حد تسلط تعیین شده (مثلا ۹۰ درصد)، ضریب تمییز را برای تک تک سؤال ها محاسبه و در مورد سؤال قضاوت می کنیم.

ضریب تمییز از ۱ تا ۱+ متغیر است. مقدار مثبت D نشان می دهد که سؤال بین کسانی که به حد تسلط رسیده اند و کسانی که به حد تسلط نرسیده اند، تمییز قائل شده است. وقتی که ضریب تمییز سؤالی، منفی باشد، نشان می دهد که سؤال، در جهت خلاف سؤال های دیگر، تمییز قائل می شود.

مثال ۴۸: جدول زیر نشان دهنده پاسخ های درست و نادرست دو گروه از دانش آموزان قوی و ضعیف (که به حد تسلط رسیده یا نرسیده اند) به یک سؤال از آزمون ملاک مرجع است، ضریب تمییز سؤال چقدر است؟

|  |  |  |
| --- | --- | --- |
| تسلط عدم تسلط | | |
| 8 | 3 | درست  نادرست |
| 2 | 7 |

1) 8/0

2) 5/0

3) 9/0

4) 3/0

پاسخ: گزینه «2»

روش هایی که تاکنون برای تحلیل آزمون گفته شد، روش های ساده یا روش های کلاسیک تحلیل آزمون بودند، اما نظریه های جدیدتری هم برای تحلیل آزمون به وجود آمده است؛ از جمله نظریه صفت مکنون با الگوی سؤال - پاسخ (IRT) که در آن ها با استفاده از منحنی ویژگی سؤال، ضرایب دشواری و تمییز سؤال تعیین می شوند.

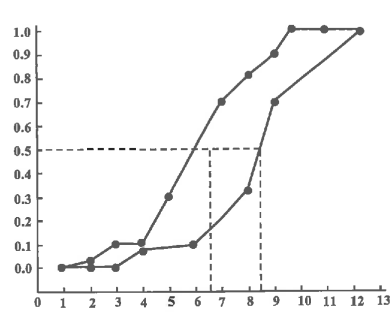
نظریه صفت مکنون بر این فرض استوار است که یک صفت یا ویژگی زیر بنایی وجود دارد که به شخص امکان می دهد تا در یک تکلیف شناختی معین موفق شود؛ از جمله ی این صفات می توان، دانش معنی لغات، استدلال ریاضی یا تجسم فضایی را نام برد. هرچه فرد از آن صفت، بیشتر برخوردار باشد، در آزمون مربوط به آن صفت، عملکرد بهتری خواهد داشت.

در نظریه سؤال - پاسخ به جای تأکید بر نمرات کل آزمون، بر پاسخ های آزمودنی ها به تک تک سؤال های آزمون تهیه می شود. این نظریه، یک معادله ریاضی به دست می دهد که با استفاده از آن می توان احتمال پاسخ درست به یک سؤال آزمون را به عنوان تابعی از توانایی آزمون شونده، و هم چنین برخی ویژگی های سؤال را پیش بینی کرد این ویژگی های سال اطلاعاتی را فراهم می آورند که می توان از آن ها برای قضاوت درباره ی سؤال، طراحی آزمون برای مقاصد ویژه و حل تعدادی از مسائل اندازه گیری آموزشی، که نظریه ی کلاسیک آزمون از عهده ی حل شودن آن ها برمی آیم استفاده کرد. منطق نظریه سؤال - پاسخ منطق ساده ای است. بر طبق این نظریه، احتمال پاسخ درست دادن به یک سوال (مثل ریاضی) با افرایش دانش ریاضی آزمون شونده افزایش می یابد و این، بالقوه یک بیان منطقی است.

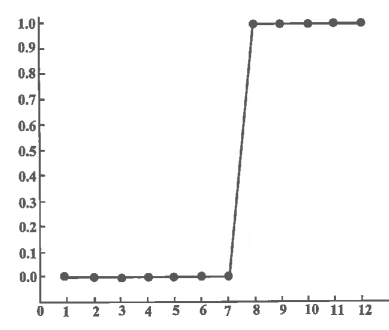
استفاده از منحنی ویژگی سؤال برای تحلیل سؤال آزمون

منحنی ویژگی سؤال، یک بازنمایی نموداری از رابطه ی بین احتمال پاسخ درست دادن به یک سؤال و موقعیت آزمون شونده در صفت مورد اندازه گیری توسط آزمون است. برای تهیه ی منحنی ویژگی سؤال، نسبت با درصد کسانی که آن سؤال را درست جواب داده اند، مشخص می شود. با استفاده از منحنی ویژگی سؤال، می توان ضریب های دشواری و تمییز سؤال را تعیین کرد. ضریب دشواری، عبارت است از نمره ی معیاری که در آن ۵۰ درصد آزمون شوندگان، سؤال را درست پاسخ داده اند. ضریب تمییز، برابر است با شیب منحنی ویژگی سؤال. اگر همه ی آزمون شوندگان به یک سؤال جواب اور صحیح بدهند، منحنی ویژگی آن سؤال، یک خط افقی واقع در بالای محور عمودی خواهد بود که از نسبت ۱ می گذرد؛ در این صورت، ضریب دشواری سؤال ۱P= است. برعکس، اگر همه ی آزمون شوندگان به یک سؤال جواب غلط بدهند، منحنی ویژگی آن سؤال نیز یک خط افقی خواهد بود که از پایین محور عمودی و از نسبت صفر می گذرد. ضریب دشواری این سؤال 0 = P است. با توجه به این توضیحات، هرچه سطح زیر منحنی و ویژگی سؤال بیشتر باشد، ضریب دشواری سؤال بزرگتر است.

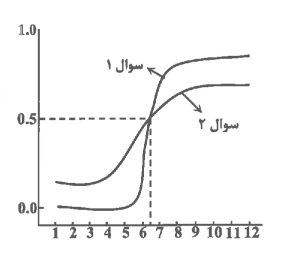
شکل مقابل، منحنی های ویژگی ۲ سؤال و ضریب دشواری آن ها را نشان می دهد. مهم ضریب دشواری سؤال 1، برابر است با 5/6 و ضریب دشواری سؤال ۲، برابر است با 5/8 به نقطه ی تقاطع خط افقی برآورد شده از نسبت ۵/0 منحنی ویژگی سؤال، هم که نقطه ی عطف منحنی است، آستانه می گویند؛ زیرا در این نقطه، جهت منحنی و تغییر پیدا می کند. افرادی که نمره کل آزمون آن ها پایین این نقطه است، احتمال جواب درست دادن به سؤال برای آن ها در حد تصادف است، اما کسانی که نمره ی کل آزمون آن ها بالاتر از این نقطه است، به احتمال زیاد به سؤال جواب درست خواهند داد؛ به این سبب گفته میشود که عملکرد آزمون شونده در پاسخ به سؤال، به توانایی او وابسته است.



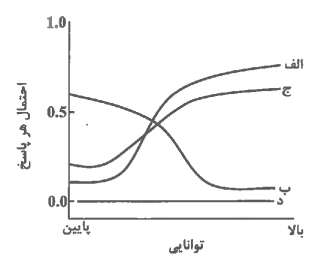
هرچه منحنی ویژگی سؤال، حالت یکسانی داشته باشد، همبستگی بین آن سؤال و کل آزمون بیشتر است و ضریب تمییز آن بهتر می تواند بین افراد گروه قوی و ضعیف تمایز قایل شود. برای توضیح بیشتر، به شکل مقابل است کنید. این منحنی، دارای شیب تند است. سؤالی که این منحنی، معرف آن است آزمون شوندگان دارای نمرات بالای ۸ و پایین ۸ را به خوبی از یکدیگر متمایز می کند.



ویژگی های دشواری و تمییز سؤال های یک آزمون، گرچه به هم وابسته اند، اما تاحدودی از یکدیگر مستقل اند. برای روشن شدن مطلب، به منحنی های ویژگی های ۲ سؤال، که در شکل زیر آمده است دقت کنید. در این شکل، نشان داده شده است که سؤال ۱ دارای قدرت نمییز بیشتری نسبت به سؤال ۲ است، زیرا شیب آن تندتر و به حالت یکسانی نزدیک تر است. این در حالی است که آستانه یا ضریب دشواری هر دو سؤال، برابر است.



از منحنی ویژگی سؤال می توان برای تحلیل گزینه های انحرافی نیز استفاده کرد؛ برای توضیح این مطلب به شکل زیر دقت کنید. گزینه ی درست یا گزینه ی الف دارای قدرت تمییز مثبت و بالاست. گزینه ی ج نیز دارای قدرت تمییز مثبت و نسبتا بالاست؛ اما گزینه ی را افراد کمتری انتخاب کرده اند و قدرت تمییز ندارد. گزینه ی دارای قدرت تمییز منفی است؛ یعنی پاسخ دهندگان ضعیف بیشتر از پاسخ دهندگان قوی، آن را انتخاب کرده اند.



مثال ۴۹: هرچه منحنی ویژگی سؤال بیشتر به صورت یکسانی باشد، این سؤال ......................

۱) ضریب دشواری بالایی دارد.

۲) ضریب تمییز بالایی دارد.

۳) ضریب دشواری پایینی دارد.

۴) ضریب تمییز پایینی دارد.

پاسخ: گزینه «۲» هرچه منحنی ویژگی سؤال، بیشتر به صورت یکسانی باشد، این سؤال ضریب تمییز بالایی دارد.

تست های طبقه بندی شده فصل سیزدهم

1- در یک آزمون هنجار - مرجع، سؤال هایی خوب تلقی می شوند که دارای ضریب دشواری ................... و ضریب تشخیص ..................... باشند.

(سراسری ۸۱)

۱) متوسط - پایین

۲) پایین به بالا

۳) بالا - متوسط

۴) متوسط - بالا

۲- از بین ۱۰ دانش آموز، که در آزمونی به حد تسلط رسیده اند، ۸ نفر به یکی از سؤال ها پاسخ صحیح داده اند و از میان ۵ دانش آموز که به حد تسلط نرسیده اند، فقط ۲ نفر به سؤال فوق پاسخ صحیح داده اند. شاخص تمییز این سؤال چند است؟

(سراسری ۸۱)

1) 20/0

2) 30/0

3) 40/0

۴) 50/0

3- مقادیر ضریب دشواری سؤال در کدام یک از حالات ممکن، بهترین شرایط را برای قدرت تشخیص سؤال آزمون های هنجار - مرجع ایجاد می کند؟

(سراسری ۸۱)

1) 40/0 > P

2) 60/0 < P

3) 80/0 ≥ P ≥ 20/0

4) 60/0 ≥ P ≥ 40/0

۴- در یک آزمون هوش از دو سؤال که ضریب تشخیص برابر دارند، آن سؤال مناسب تر است که .............. باشد.

(سراسری ۸۱)

1) سطح دشواری آن زیادتر

۲) سطح دشواری آن کمتر

3) دارای واریانس کوچکتر

۴) دارای سطح دشواری متوسط

5- کدام یک از سیستم های نمره گذاری، هم آسان تر انجام می گیرد و هم اطلاعات دقیق تر در اختیار می گذارد؟

(سراسری ۸۲)

۱) صفر تا ۲۰

۲) صفر تا ۱۰۰

3) خوب، متوسط، ضعیف

4) الف، ب، ج، د، ه

6- سطح دشواری 5/0برای کدام نوع از سؤال ها مناسب تر است؟

(سراسری ۸۲)

۱) تشریحی

۲) صحیح - غلط

3) چهار گزینه ای

۴) کوتاه پاسخ

۷- در سطح دشواری ملاک - مرجع، کارکرد یادگیرنده با ........................ سنجیده می شود.

(سراسری ۸۳)

۱) جدول نرم

۲) یک معیار

۳) با گروه خودش

۴) مقیاس نمره گذاری

۸- در یک آزمون ۴۰ سؤالی پنج گزینه ای که به پاسخ های غلط نمره ی منفی تعلق می گیرد، دانش آموزی به ۳۲ سؤال پاسخ صحیح، و به ۱۲ سؤال، هم پاسخ غلط داده است. اگر هر سؤال یک نمره داشته باشد، نمره ی اصلاح شده ی وی کدام است؟

(سراسری ۸۲)

1) 20

2) 27

3) 28

4) 29

9- در جدول مقابل، اگر گزینه ی «الف» صحیح باشد، درجه ی دشواری و قدرت تمییز سؤال به ترتیب (از راست به چپ) چقدر خواهد شد؟ (سراسری ۸۲)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| جدول تحلیل سوال | | | | |
| پاسخ های پیشنهادی | الف | ب | ج | د |
| گروه قوی | 15 | 2 | 0 | 3 |
| گروه ضعیف | 2 | 10 | 0 | 8 |

1) 5/42 % و 65/0

2) 65 /0 و 5/42 %

3) 85 % و 85/0

4) 85/0 و 85 %

۱۰- اگر معلمی ۶۰ ورقه ی انشایی را پشت سرهم تصحیح کند، سپس میانگین ۴ گروه ورقه ی ۱۵ تایی را جداگانه به دست آوردم میانگین کدام گروه ۱۵ تایی بیشتر از بقیه خواهد شد؟

(سراسری ۸۲)

۱) اول

2) دوم

3) چهارم

۴) چهارم

۱۱- سؤال خوب، سؤالی است که منحنی ویژگی آن در نقطه برش ................ شیب دارد.

(سراسری ۸۲)

۱) حداقل

۲) حداکثر

3) 50%

4) 68%

۱۲- همبستگی هر سؤال با سایر سؤال ها در یک آزمون، بیانگر چیست؟

(سراسری ۸۲)

1) همگنی درونی

۲) روایی محتوا

۳) قدرت تشخیص

۴) درجه ی دشواری

۱۳- دانش آموزی در یک آزمون پنج گزینه ای به ۶۰ سؤال، صحیح و ۳۰ سؤال، غلط پاسخ داده است. نمره او با احتساب نمره منفی کدام است؟

(آزاد ۸۲)

1) 54

2) 5/52

3) 50

۴) هیچ کدام

۱۴- در جدول مقابل، اگر پاسخ صحیح گزینه «ج» باشد، قدرت تمییز سؤال کدام است؟

(آزاد ۸۳)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| پاسخ های پیشنهادی | الف | ب | ج | د |
| گروه بالا | 0 | 10 | 25 | 15 |
| گروه پایین | 20 | 5 | 15 | 10 |

1) 10/0

2) 20/0

3) 25/0

4) 40/0

۱۵- در نظریه کلاسیک آزمون وقتی تعداد مشاهده ها افزایش یابد، میانگین نمره های مشاهده شده، برآورد ................... میانگین نمره های واقعی می باشد.

(آزاد ۸۲)

۱) وزن های رگرسیون

۲) دقت اندازه گیری

۳) خطای معیار

۴) بدون سوگیری

۱۶- در آزمونی که ۶۰ نفر در آن شرکت کرده اند، ۱۵ نفر به سؤالی پاسخ غلط می دهند. ضریب دشواری این سؤال کدام است؟

(سراسری ۸۳)

1) 15/0

2) 25/0

3) 45/0

4) 75/0

۱۷- اگر ضریب تشخیص (قوه تمییز) یک سؤال ۱۰۰ درصد باشد، ضریب دشواری سؤال چنددرصد خواهد بود؟

(سراسری ۸۳)

1) صفر

2) 25

3) 50

4) 100

که ۱۸- افزودن تعدادی سؤال با ضرایب دشواری ..... و ضریب، تمییز .........، پایایی سؤالات آزمون را افزایش می دهد.

(سراسری ۸۳)

۱) زیاد - متوسط

۲) زیاد - بالا

۳) متوسط - متوسط

۴) متوسط - بالا

۱۹- برای تعیین قدرت تمییز یا تشخیص سؤال، کدام یک از ضرایب زیر مناسب تر است؟

(آزاد ۸۳)

۱) پیرسون

۲) دو رشته ای

۳) دو رشته ای نقطه ای

۴) اسپیرمن براون

۲۰- سؤالی که هیچ کدام از پاسخ دهندگان، به آن پاسخ درسته نداده اند، دارای ضریب دشواری ................. است.

(آزاد ۸۳)

۱) صفر

۲) یک

۳) بین صفر و یک

۴) بیشتر از صفر

۲۱- مناسب ترین ضریب دشواری برای سوال های صحیح و غلط، مساوی کدام یک از مقادیر زیر است؟

1) 68/0 – 62/0

۲) بالاتر از ۷۵/0 و کمتر از 90/0

3) 75/0

4) بالاتر از 90/0

۲۲- کدام درجه دشواری سؤال، اطلاعات افتراقی بیشتری را به دست می دهد؟

(سراسری ۸۴)

1) 25/0 P =

2) 50/0 P =

3) 75/0 P =

4) 00/1 P =

۲۳- نسبت پاسخ درست آزمودنی های گروه ملاک بالا (قوی) به یک سؤال آزمون (ضعیف)2۷/0 است، 3۷/0 و نسبت پاسخ درست آزمودنی های گروه ملاک پایین ضریب تمییز این سؤال کدام است؟

(سراسری ۸۴)

۱) 05/0 -

2) 10/0 -

3) 32/0 +

4) 64/0 +

۲۴- درصورتی که آزمودنی های یک آزمون، از جهت خصیصه مورد اندازه گیری متجانس باشند، انتظار می رود تغییر درجه دشواری سؤال ها چگونه باشد؟

۱) محدود

۲) گسترده

۳) در حد متوسط

۴) در حد بالا

۲۵- جدول زیر نشان دهنده پاسخ های درست و نادرست دو گروه از دانش آموزان قوی و ضعیف (دانش آموزانی که به حد تسلط رسیده اند و دانش آموزانی که به حد تسلط نرسیده اند) به یک سؤال از آزمون ملاک مرجع است. قدرت تمییز این سؤال چقدر است؟

(سراسری ۸۴)

|  |  |  |
| --- | --- | --- |
| تسلط | عدم تسلط |  |
| 9 | 2 | پاسخ درست |
| 2 | 5 | پاسخ نادرست |

1) 43/0

2) 53/0

3) 63/0

4) 73/0

۲۶- ارزش تمییز قبل و بعد از آموزش سؤالی که تمام دانش آموزان در پیش آزمون و در پس آزمون، صحیح پاسخ داده باشند، کدام است؟

(آزاد ۸۴)

۱) صفر

2) 5/0

3) 1

۴) هیچ کدام

۲۷- دانش آموزی در یک آزمون ۱۰۰ ماده ای از نوع پنج گزینه ای به ۷۴ سؤال، صحیح و ۱۶ سؤال، غیرصحیح پاسخ داده است. نمره او با استفاده از فرمول جبران حدس و در مبنای ۲۰ کدام است؟

(آزاد ۸۴)

1) 5/13

2) 5/12

3) 13

4) 14

۲۸- در جدول داده شده، اگر «ب» پاسخ صحیح باشد، قدرت تمییز سؤال کدام است؟

(آزاد 84)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| پاسخ های پیشنهادی | الف | ⁎ب | ج | د |
| گروه قوی | 0 | 50 | 30 | 20 |
| گروه ضعیف | 40 | 30 | 20 | 10 |

1) 10/0

2) 20/0

3) 25

4) 40/0

۲۹- بین سطح دشواری و واریانس یک سؤال رابطه وجود دارد. براین اساس می توان گفت:

(سراسری ۸۵)

1) سؤالات دشوار و سؤالات آسان، واریانس بزرگتری دارند.

۲) هراندازه که دشواری سؤال متوسط باشد، واریانس آن بیشتر است.

۳) با افزایش سطح دشواری سؤال، واریانس آن همواره افزایش می یابد.

۴) با افزایش سطح دشواری سؤال، واریانس آن همواره کاهش می یابد.

۳۰- در تصحیح اوراق دانش آموزان، «اثر پیگمالیون» عبارت است از:

(سراسری ۸۵)

۱) اثر انتظارات معلم بر موفقیت دانش آموزان خود

۲) کمتر از حد نمره دادن به ورقه دانش آموز تنبل

۳) بیش از حد نمره دادن به ورقه ی دانش آموز ساعی

۴) تأثیر منشأ قومی دانش آموزان بر نمرات اوراق امتحانی

۳۱- در جدول زیر، وضعیت پاسخگویی گروه های ضعیف و قوی کلاس به یک آزمون نشان داده شده است. پاسخ صحیح با علامت ⃰ مشخص شده است. ضریب دشواری سؤال چند درصد است؟

(آزاد ۸۵)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | الف | ب | ج | د |
| گروه قوی | 2 | 10 | 4 | 0 |
| گروه ضعیف | 5 | 6 | 2 | 3 |

1) 5/12 %

2) 50 %

3) 25 %

4) 5/62 %

۳۲- در جدول مندرج در سؤال ۱۱۳ (قبلی) ضریب تمییز سؤال چقدر است؟

(آزاد ۸۵)

1) 25/0

2) 125/0

3) 50/0

4) 625/0

۳۳- نظریه صفت مکنون در کدام نظریه اندازه گیری، مطرح شده است؟

(آزاد ۸۵)

1) کلاسیک

۲) سنجش کیفی

۳) سؤال - پاسخ

۴) فرایند مدار

۳۴- کدام یک از گزینه های زیر، در مورد رابطه ی شاخص های دشواری و تمییز، درست نیست؟

(سراسری ۸۶)

١) اگر ضریب دشواری سؤالی ۵/0 باشد، ضریب تمییز آن ۱ خواهد بود.

۲) اگر ضریب تمییز سؤالی 1 باشد، ضریب دشواری آن ۵/0 خواهد بود.

۳) اگر ضریب دشواری سؤالی ۱ باشد، ضریب تمییز آن صفر خواهد بود.

۴) اگر ضریب دشواری سؤالی صفر باشد، ضریب تمییز آن صفر خواهد بود.

۳۵- در یک آزمون ۵۵ سؤالی پنج گزینه ای، دانش آموزی به ۳۵ سؤال پاسخ درست و به ۱۸ سؤال، پاسخ غلط داده است و بقیه سؤالات را بدون پاسخ گذاشته است. نمره اصلاح شده او با احتساب نمره منفی چقدر است؟

(سراسری ۸۶)

1) 17

2) 29

3) 30

4) 5/30

۳۶- فردی در یک آزمون ۶۰ سؤالی که روی ۵۰ نفر اجرا شده، نمره ۴۵ را به دست آورده است. درصد عملکرد او در کل آزمون چقدر است؟

(سراسری ۸۶)

1) 67 %

2) 75 %

3) 83 %

4) 90 %

۳۷ـ در بحث مربوط به تفسیر و نمره گذاری آزمون های فرافکن، منظور از «خطا یا اثر هاله ای» چیست؟

(سراسری 86)

۱) ویژگی های مثبت آزمون شونده بزرگنمایی میشوند.

۲) ویژگی های منفی تحت تاثیر ویژگی های مثبت نادیده گرفته میشوند.

۳) ویژگی های مثبت، تحت تاثیر ویژگی های منفی نادیده گرفته میشوند.

۴) یک ویژگی برجسته، سایر خصوصیات فرد آزمون شونده را تحت الشعاع قرار می دهد.

۳۸- یک سؤال با درجه دشواری ۸/0 خیلی ........... از سؤالی است که درجه دشواری آن 3/0 است.

(سراسری 86)

۱) قویتر

۲) ضعیف تر

۳) دشوارتر

۴) آسان تر

۳۹ـ در یک آزمون چهارگزینه ای از ۵0 نفر شرکت کننده، ۳۰ نفر به سؤال شماره ۱۰ پاسخ صحیح و بقیه پاسخ غلط داده انـد. ضریب دشواری سؤال یاد شده چه مقدار است؟

(آزاد ۸۶)

1) 60/0

2) 75/0

3) 17/0

4) 30/0

۴۰ـ فرض کنید سؤالی دارای ضریب تمییز منفی باشد، براساس این ضریب تمییز می توان پیش بینی کرد، دانش آموزی که به این سؤال جواب درست تمییر منفی داده است در آزمون نمره:

(آزاد ۸۶)

۱) بالایی گرفته است.

2) پایینی گرفته است.

3) متوسطی گرفته است.

4) صفر گرفته است.

۴۱- کدام یک از ضرایب دشواری زیر معرف یک سؤال دشوار است؟

(آزاد 86)

1) 90/0

2) 35/0

3) 11/0

4) 50/0

۴۲- برای تعیین سطح دشواری لازم سؤالات تست، بهتر است چه ترتیبی اتخاذ شود؟

(سراسری 87)

۱) سؤالات، اغلب با سطح دشواری بالا مدنظر قرار گیرد و تعدادی سؤال متوسط هم داده شود.

۲) سؤالات اغلب با سطح دشواری متوسط مدنظر قرار گیرد و تعدادی سؤال با سطوح دیگر داده شود.

۳) سؤالات همگی با سطح دشواری کم مدنظر قرار گیرد به جز موارد استثنایی.

۴) نوع تصمیمی که توسط نتیجه ی آن لازم است گرفته شود، مورد توجه قرار گیرد.

۴۳ـ نسبت افرادی که به سؤال پاسخ صحیح می دهند ........... ضریب نام دارد.

(سراسری 87)

۱) پیوستگی

۲) تشخیص

۳) دشواری

۳) همبستگی

۴۴ـ در جدول زیر، پاسخ های داده شده توسط گروه های قوی و ضعیف یک کلاس به یک سؤال چهارگزینه ای ارائه شده است. در این آزمون گزینه ج گزینه درست بوده است. ضریب دشواری سؤال چقدر است؟

(سراسری ۸۷)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| گزینه ها | الف | ب | ج⁎ | د |
| 27% بالا | 4 | 4 | 9 | 3 |
| 27% پایین | 6 | 4 | 3 | 7 |

1) 15/0

2) 30/0

3) 45/0

4) 60/0

۴۵- اگر ۳۰ درصد از افراد گروه قوی و ۱۵ درصد از افراد گروه ضعیف به سؤال چهار گزینه ای، پاسخ صحیح بدهند، ضریب تشخیص سؤال چقدر است؟

(سراسری ۸۷)

1) 15/0

2) 30/0

3) 45/0

4) 075/0

۴۶- اگر درجه دشواری سؤالی برابر با 5/0 باشد، واریانس سؤال برابر کدام مورد است؟

(سراسری 87)

۱) یک

۲) صفر

۳) حداکثر

4) حداقل

۴۷ـ بالا بودن « ضریب دشواری»، نشان دهنده .......... سؤال است.

(سراسری 88)

۱) نامناسب بودن

۲) مناسب بودن

3) مشکل بودن

4) آسان بودن

۴۸- در تهیه فرم نهایی یک آزمون، معمولاً سؤالاتی را نگه می دارند که شاخص های دشواری آن ها ....... است.

(سراسری 88)

1) بین 3/0 تا 7/0

2) بین 20/0 تا 80/0

3) بین 40/0 تا 60/0

4) کمتر از 20/0 و بیشتر از 75/0

۴۹- شاخص دشواری هورست» در پی:

(آزاد ۸۸)

1) تعیین تعداد فراوان ترین جواب غلط هر سؤال است.

۲) تعیین مجموع سؤال های صحیح و غلط است.

۳) تعیین تعداد سؤال های صحیح است.

۴) تعیین تعداد سؤال های بی جواب است.

۵۰- اگر دانش آموزی در یک امتحان ۶۰ سؤالی از نوع چهار گزینه ای، ۴۵ سؤال را صحیح و ۱۵ سؤال را غلط جواب داده باشد، نمره خام او در این امتحان عبارتست از:

(آزاد ۸۸)

1) 40

2) 45

3) 35

4) 30

51- آزمونی دشوار و دارای کجی به سمت چپ، دارای کدام ویژگی است؟

(سراسری 89)

۱) اثر سقف

۲) تأثیر کف

۳) ضریب تشخیص بالا

4) انحراف استاندارد بالا

۵۲- کدام توزیع از آزمونی که سوال های آن دارای دشواری متوسط ولی قدرت تمییز بالاست، تشکیل می شود؟

(سراسری 89)

1) طبیعی

2) مسطح

۳) نامتقارن

4) u - شکل

53- شیب زیاد خم ویژه سؤال، حاکی از کدام مورد است؟

(سراسری 89)

1) دشواری بالا

۲) قدرت تمییز پایین

۳) دشواری پایین

4) قدرت تمییز بالا

54- اضافه کردن کدام سؤالات، توزیع دارای کجی منفی را به نرمال تبدیل می کند؟

(سراسری 89)

1) سؤالات دشوار

۲) سؤالات با دشواری متوسط

۳) سؤالات آسان

۴) نوع سؤال ارتباطی با توزیع نمرات آزمون ندارد.

55- نمودار روبه رو خم ویژه یک سؤال است، با توجه به آن می توان گفت: این سؤال ...............

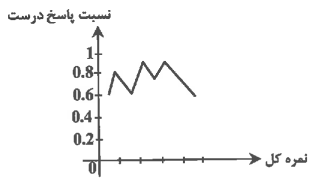
(سراسری ۹۰)

۱) قدرت تمییز ندارد.

۲) وضعیت نسبی و نمره کلی را نشان می دهد.

۳) تفاوت های بین نمرات دانش آموزان را به خوبی نشان میدهد.

۴) قدرت تمییز خوبی دارد ولی دشواری آن نوسان دارد.



56- مطابق نظریه کلاسیک آزمون، هنگامی که شاخص تمییز یک سؤال چهارگزینه ای برابر یک باشد، شاخص دشواری آن سؤال کدام است؟

(سراسری ۹۰)

۱) صفر

2) 25/0

3) 1

4)50/0

۵۷- در کدام یک از شرایط زیر، نتایج تحلیل های کلاسیک سؤال، شباهت بیشتری به نتایج تحلیل های IRT دارد؟

(سراسری 90)

1) سؤالات، شبیه به هم باشند.

۲) توربع درجات دشواری نرمال باشند

۳) سؤالات، تا حد ممکن از هم متفاوت بوده و برای توانایی های مختلف مناسب باشند.

۴) سؤالات از شیب مناسب و متفاوتی برخوردار باشند.

۵۸- اگر مقدار D برابر با 4/0 وPu برابر با ۶/0 باشد، درجه دشواری سؤال چقدر است؟

(سراسری ۹۰)

1) 4/0

2) 5/0

3) 1

4) 6/0

۵۹- شیب خم ویژه سؤال (ICC) بیانگر کدام ویژگی سؤال است؟

(سراسری 90)

۱) حس پذیری سؤال

2) قدرت تشخیص سؤال

3) همبستگی سوال با سوال دیگر

4) میزان دشواری سوال

۶۰- همبستگی بین نمره یک سؤال و نمره کل آزمون چه نام دارد؟

(سراسری 91)

۱) ثبات درونی

۲) قدرت آزمون

۳) قدرت تمییز

۴) درجه دشواری

61- ضریب تمییز منفی سؤال، بیانگر کدام یک از حالات است؟

(سراسری ۹۱)

۱) هر دو گروه به یک اندازه در دادن پاسخ صحیح به سؤال موفق بوده اند.

۲) گروه ضعیف در یافتن پاسخ صحیح عاجز مانده است،

3) گروه قوی از گروه ضعیف، بدتر عمل کرده است.

۴) گروه قوی بهتر از گروه ضعیف به سؤال پاسخ داده است.

۶۲- توزیع پاسخ های دو گروه قوی و ضعیف در جدول زیر آمده است. اگر گزینه «ب» پاسخ درست سؤال باشد، ضریب دشواری سؤال چقدر است؟

(سراسری ۹۱)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | الف | ب | ج | د |
| گروه قوی | 6 | 12 | 6 | 6 |
| گروه ضعیف | 9 | 8 | 9 | 4 |

1) 2/0

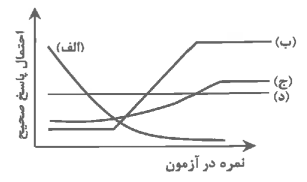
2) 33/0

3) 8/0

4) 66/0

۶۳- منحنی پاسخ صحیح چهار سؤال در شکل زیر آمده است. کدام سؤال از همه مناسب تر است؟

(سراسری ۹۱)



1) (د)

۲) (ج)

۳) (الف)

۴) (ب)

۶۴- وجود کدام ویژگی در آزمون از اهمیت بیشتر برخوردار است؟

(سراسری ۹۲)

1) تعداد مناسب گزینه ها

۲) کوتاه بودن طول آزمون

۳) روشن و واضح بودن سؤالات

۴) سهولت اجرا و نمره گذاری

65- در یک امتحان چهار گزینه ای که از ۵۰ سؤال تشکیل شده است، دانش آموزی به ۴۶ سؤال پاسخ گفته است. در صورتی که جواب شش سؤال، صحیح نباشد، نمره اصلاح شده وی با احتساب نمره منفی برای پاسخ های غلط چقدر است؟

(سراسری ۹۲)

1) 38

2) 40

3) 42

4) 44

۶۶- توزیع پاسخ های دو گروه قوی و ضعیف در جدول زیر آمده است. به شرط درست بودن گزینه «ج»، دشواری سؤال چقدر است؟

(سراسری ۹۲)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| گزینه  گروه | الف | ب | ⁎ج | د |
| قوی | 4 | 5 | 10 | 1 |
| ضعیف | 6 | 6 | 7 | 1 |

1) 075/0

2) 425/0

3) 85/0

4) 5/0

67- سؤالی با کدام ویژگی را برای گنجانیدن در آزمون انتخاب می کنید؟

(سراسری ۹۲)

۱) واریانس آن زیاد باشد.

۲) با سایر سؤالات آزمون، همبستگی بالایی داشته باشد. ا

۳) دشواری آن زیاد ولی واریانس آن کم باشد.

۴) سؤال آسانی باشد ولی واریانس بالایی داشته باشد.

۶۸- مهم ترین مرحله ی تهیه آزمون های روان شناختی کدام است؟

(سراسری ۹۳)

۱) نوشتن سؤالات

۲) تعیین روایی و اعتبار

۳) تعریف صفت مورد اندازه گیری

۴) تهیه نرم و ملاک قضاوت در مورد عملکرد آزمودنی ها

۶۹- اگر هدف از اجرای آزمون، گزینش درصد کمی از داوطلبان باشد، استفاده از کدام سؤالات مناسب تر است؟

(سراسری ۹۳)

۱) دشوار

۲) آسان

۳) نه آسان و نه دشوار

۴) نرخ گزینش، ارتباطی با درجه ی دشواری سؤال ندارد.

70- چنانچه فراوانی پاسخگویی به یک سؤال در گروه بالا نسبت به گزینه های مختلف مساوی باشد، احتمالا کدام حالت اتفاق افتاده است؟

(سراسری ۹۳)

1) جذابیت نسبتا بالای گزینه ها

۲) قدرت تمییز بالای سؤال

۳) ارائه ی پاسخ از روی حدس و گمان

۷۱- چنانچه آزمون دارای اثر کف باشد، منحنی توزیع نمرات دارای چه شکلی است؟

(سراسری ۹۳)

1) مسطح

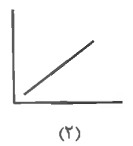
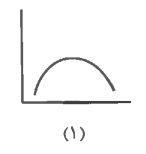
2) متقارن

۳) چولگی مثبت

۴) چولگی منفی

۷۲- کدام شکل، رابطه ی میان دشواری و واریانس سؤالات چند گزینه ای را بهتر نشان می دهد؟

(سراسری ۹۳)

۷۳- اگر قدرت تمییز سؤالی برابر با یک باشد، نسبت پاسخ صحیح به سؤال در گروه ضعیف (پایین) چقدر است؟

(سراسری 93)

1) 5/0 -

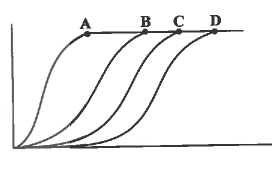
2) 0

3) 5/0

4) 1

۷۴- نمودار روبه رو، منحنی ویژه (ICC) چهار سؤال را نشان میدهد. کدام سؤال از دشواری بیشتری برخوردار است؟

(سراسری ۹۳)



1) A

2) B

3)C

4)D

۷۵- نسبت پاسخگویی به گزینه های یک سؤال چهارگزینه ای مطلوب، به تفکیک گروه قوی و ضعیف در جدول زیر ارائه شده است. کدام مورد درست است؟

(سراسری ۹۴)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| گزینه ها  گروه | الف | ب | ج | د |
| قوی | 15/0 | 60/0 | 10/0 | 15/0 |
| ضعیف | 25/0 | 30/0 | 18/0 | 27/0 |

1) الف

2) ب

3) ج

4) د

۷۶- اگر میانگین افرادی باشد که به یک سؤال پاسخ درست، میانگین افرادی باشد که به سؤال پاسخ نادرست و میانگین کل افراد باشد، از تساوی چه نتیجه ای به دست می آید؟

(سراسری ۹۴)

1) روایی سؤال بالا است.

۲) سؤال به نفع هیچ گروهی سوگیری ندارد.

۳) سؤال با نمره کل همبستگی ندارد.

۴) بین سؤال و نمره کل رابطه مثبت وجود دارد.

۷۷ - اساس نظریه کلاسیک در اندازه گیری صفات روانی بر کدام مورد استوار نیست؟

(سراسری 94)

1) خطای درون فردی

۲) تفاوت های بین فردی

۳) خطای تصادفی

۴) کنترل آزمایشی

۷۸- کدام مورد، از ویژگی های اصلی برای سؤال مناسب در آزمون های روان شناختی استاندارد شده است؟

(سراسری ۹۴)

1) برخورداری از دشواری پایین

۲) برخورداری از ضریب تمییز متوسط

۳) همبستگی متوسط سؤال با نمره آزمون

۴) امکان افزایش واریانس آزمون

پاسخنامه تست های طبقه بندی شده فصل سیزدهم

1- گزینه «۴» سؤال های خوب در یک آزمون، سؤال هایی هستند که با ضریب دشواری متوسط و ضریب تمییز بزرگ (بالا) طرح شوند. متخصصان سنجش و اندازه گیری، معتقدند که یک سؤال خوب، ضریب تمییز حداکثر ۱+ و ضریب دشواری ۵/0 در حد وسط دارد. ضریب تمییز به قدرت سوال در تمایزگذاری یا تشخیص بین گروه قوی و گروه ضعیف اشاره دارد. ضریب دشواری به درصد کل آزمون شوندگانی اشاره دارد که توانسته اند به یک سؤال، جواب درست دهند. گزینه ی ۳، بسیار نزدیک به گزینه ی صحیح می باشد که می تواند داوطلب را به اشتباه بیندازد.

۲- گزینه «٣» برای پاسخ گویی به این سؤال ابتدا جدول چهار خانه ای مربوط به تمییز بین حد تسلط و غیر حد تسلط را ترسیم می کنیم، سپس بر اساس فرمول محاسبه تمییز بین حد تسلط و غیر حد تسلط، پاسخ را به دست آوریم.

|  |  |  |  |
| --- | --- | --- | --- |
| سوال | درست  غلط | حد تسلط حد غیر تسلط | |
| b | a |
| d | c |
| سوال | درست  غلط | حد تسلط حد غیر تسلط | |
| 8 | 2 |
| 2 | 3 |

۳- گزینه «۴» علامت P به ضریب دشواری اشاره دارد. در سؤال های چهارگزینه ای، سطح موفقیت با حدس زدن ۲۵/0 است و سطح بهینه دشواری در وسط ۲۵/0 و یعنی در حدود 6/0 است. با توجه به گزینه های موجود، می توان گفت گزینه ی ۲، نزدیکترین گزینه به پاسخ درست می باشد که میتواند داوطلب را دچار اشتباه کند؛ زیرا همان طور که در گزینه 1 می بینیم علامت ≥ باعث شده است، بین گزینه ۲ و گزینه ۴ تمایز به وجود آید.

۴- گزینه «۴» ضریب دشواری، درصد کل افرادی را که به یک سؤال، جواب درست می دهند، نشان می دهد؛ بنابراین در فرم نهایی آزمون، سؤال هایی بهتر هستند که ضریب دشواری آن ها کمتر از 1 و بیشتر از صفر و به 5/0 نزدیک باشد، پس برای انتخاب سؤال مناسب تر، سؤالی را که سطح دشواری متوسط دارد، انتخاب می نماییم. گزینه ۱، با توجه به نوع سؤال می تواند به جای گزینه ی درست انتخاب شود، که البته انتخاب درستی نمی باشد.

۵- گزینه «1» به دلین آن که سیستم نمره گذاری صفر تا ۲۰ یکی از رایج ترین، آسان ترین و در عین حال دقیق ترین سیستم نمره گذاری می باشد، که طیف نمره خاصی را در برمی گیرد؛ ضمن آن که صفر آن، نشان دهنده صفر مطلق نمی باشد و اگر دانش آموزی در یک درس، این نمره را کسب کند، نشانگر بی اطلاعی کامل او در آن درس نیست. نمره صفر تا ۱۰۰ در گزینه ۲ به دلیل آن که شامل طیف بسیار گسترده ای می باشد و روش نمره گذاری را سخت می کند، نمی تواند سیستم جایگزینی برای صفر تا ۲۰ محسوب شود.

6- گزینه «۲» از لحاظ انتخاب برای استفاده از سؤال در فرم نهایی آزمون، سؤال هایی مناسب تر هستنم که ضریب دشواری آن ها از ۱ کمتر و از صفر، بیشتر و به ۵/0 نزدیک باشد. البته ذکر این نکته لازم است که انتخاب سؤال هایی با ضرایب دشواری مناسب به نوع سؤال های آزمون بستگی دارد. در آزمون های صحیح - غلط، ضریب دشواری 5/0 P = زمانی به دست می آید که همه آزمودنی ها با حدس زدن به سؤال ها جواب دهند. گزینه 3 می تواند داوطلب را دچار اس خطا کند، در سؤال های چهار گزینه ای، سطح موفقیت با حدس زدن، حدود ۲۵/0 است و سطح بهینه دشواری در حد وسط بین 1 و ۲۵/0 قرار دارد، که در حدود 6/0 خواهد بود. ضریب دشواری اشاره به درصد کسانی دارد که به سؤال، جواب درست داده اند.

۷- گزینه «۲» ارزشیابی وابسته به ملاک یا ارزشیابی ملاکی (ملاک - مرجع)، معیاریا ملاکی جهت ارزشیابی است که از قبل تعیین می شود و عملکرد یادگیرنده با توجه به آن ملاک، مورد قضاوت قرار می گیرد. در این نوع ارزشیابی، که باملاک مطلق از قبل تعیین شده صورت می گیرد، از آزمون شوندگان انتظار میرود که به هدف های آموزشی معینی دست یابند یا به سطح به خصوصی در عملکرد برسند. گزینه ی ۳ (با گروه خودش) به ارزشیابی هنجاری اشاره دارد، که ملاک نسبی مورد نظر بوده که عملکرد دانش آموزان را با یکدیگر مورد مقایسه قرار می دهد، اما نه با یک ملاک که از پیش تعیین شده است.

۸- گزینه «۴» مفروضات سؤال:

تعداد سؤالات آزمون = سؤال ۴۰ و ۵ = N نمره اصلاح شده

32 =R= پاسخ صحیح

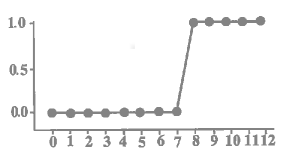
12 =W = پاسخ اشتباه

9- گزینه «۱»

باید به این نکته، به دقت توجه کرد که در سؤال، ابتدا به درجه دشواری، سپس به قدرت تمییز اشاره شده است؛ بنابراین داوطلب پس از انجام جای گذاری اعداد در فرمول و کسب نتیجه، باید در زمینه انتخاب گزینه صحیح دقت کافی داشته باشد.

۱۰- گزینه «۱» بنابر نظر متخصصان در کتاب های سنجش و اندازه گیری، زمانی که مصحح، پشت سرهم تعداد زیادی برگه را تصحیح کند، دقت و حوصله ی او در هنگام تصحیح، بر نمره ی آزمون شونده بسیار تأثیرگذار است. در آزمون تشریحی، نمره گذاری ها یکنواخت نمی باشد و بخش اعظم اندازه گیری توانایی نوشتاری آزمون شونده به شرایط مصحح، در حین تصحیح برگه ها، بستگی دارد. برای کم شدن این دخالت ذهنی، به قواعد نه گانه ای اشاره شده است. با توجه به آنچه بیان شد زمانی که مصحح، تصحیح را شروع می کند، میانگین ۱۵ برگه اول نسبت به سایر برگه ها بالاتر می باشد که این موضوع، به حوصله و دقت نظر مصحح در حین تصحیح اوراق اشاره دارد.

۱۱- گزینه «۲» زمانی که سؤال، در نقطه برش دارای حداکثر شیب باشد، منحنی ویژگی آن به صورت یکسانی است و بهتر می تواند بین گروه قوی و ضعیف تمایز قائل شود. برای نزدیکی بیشتر ذهن به پاسخ سوال به شکل روبه رو دقت کنید:



۱۲- گزینه «1» همبستگی هر سؤال با سؤال دیگر با ضریب فی به دست می آید که نشان دهنده همسانی درونی آزمون است. اگر همبستگی هر سؤال با کل آزمون مدنظر باشد، ضریب دو رشته ای نقطه ای به دست می آید که می تواند نوعی ضریب تشخیص باشد.

۱۳- گزینه «۲» باتوجه به فرمول زیر، جواب صحیح را می توان به راحتی به دست آورد.

تعداد گزینه ها N

تعداد پاسخ های غلط W

تعداد پاسخ های درست R

۱۴- گزینه «۲» بر اساس فرمول زیر می توان پاسخ صحیح را به راحتی به دست آورد:

15- گزینه ۴» هرچه تعداد مشاهدات، بیشتر باشد، نمرات بزرگ و کوچک، همدیگر را خنثی می کنند و میانگین نمرات مشاهده شده، برآورد بدون اریبی بدون سوگیری از میانگین نمره های واقعی می باشد.

۱۶- گزینه «۴» ضریب دشواری به درصد کل آزمون شوندگانی اطلاق می شود که توانسته اند به یک سؤال، جواب درست دهند. فرمول به شرح زیر می باشد: مفروضات سؤال

۴۵ = ۶۰ - ۱۵ = تعداد پاسخ درست افراد = R

تعداد کل آزمون شوندگان =T

۱۷- گزینه ۳» ضریب دشواری، عبارت است از درصد کل آزمون شوندگانی که به یک سؤال، جواب درست میدهند. ضریب تمییز، قدرت تمایزگذاری یا تشخیص بین گروه قوی و گروه ضعیف آزمون شوندگان را مشخص می کند، یعنی معلوم می نماید یک سؤال، تا چه اندازه می تواند گروه قوی را از گروه ضعیف جدا می کند. ضریب تمییز ۱0۰، یعنی حداکثر، که در این صورت، آزمون به طور کامل، گروه قوی و ضعیف را از هم جدا ساخته است و این بدان معناست که کل گروه قوی به سؤال، پاسخ درست داده اند و کل گروه ضعیف، به سؤال، پاسخ نادرست داده اند. با این توصیف، تنها نیمی از آزمون شوندگان، پاسخ سؤال را درست جواب داده اند؛ بنابراین، زمانی که نیمی از آزمون شوندگان، به یک سؤال، پاسخ درست می دهند، ضریب دشواری آن برابر با ۵۰ درصد خواهد بود.

۱۸- گزینه «۴» عوامل مؤثر بر افزایش پایایی آزمون (وابسته به هنجار) شامل ۶ مؤلفه می باشند که برام پاسخ به این سؤال به ۲ مولفه می توان اشاره کرد:

با افزودن سؤالاتی با ضریب تمییز زیاد، پایایی آزمون افزایش می یابد.

با افزودن تعدادی سؤال با ضریب دشواری متوسط به یک آزمون، پایایی آن افزایش می یابد.

۱۹- گزینه «۲» همبستگی دو رشته ای نقطه ای، نشان دهنده توافق سؤال بال آزمون است؛ هرچه ضریب همبستگی بیشتر باشد، بدان معنی است که سؤال با مجموعه سؤال های دیگر همخوانی بیشتری دارد؛ لذا بین افراد قوی و ضعیف تمایز بیشتری قائل می شود. هر چه این ضرایب کمتر باشند، مقدار این همخوانی کمتر است و لذا آن سؤال، بین افراد قوی و ضعیف تمایزی کمتری قایل می شود. برای نمونه سؤالی را در نظر بگیرید که افرادی که در آن نمره های در بالایی گرفته اند و کسانی که در کل آزمون نمره هایو پاپینی گرفته اند، در آن سؤال، نمره پائینی گرفته اند. این سؤال، با کل آزمون همخوانی دارد و ضریب

تمییز آن مثبت و بالاست. ضریب همبستگی پیرسون، روشی است که برای تعیین همبستگی بین دو متغیر پیوسته به کار برده می شود. ضریب همبستگی حسر اسپیرمن نیز روشی است که برای تعیین همبستگی بین دو متغیر رتبه ای به کار برده می شود. ضریب همبستگی دو رشته ای، ضریبی است که برای تعیین همه ضریب همبستگی بین دو متغیر پیوسته (متغیر اول) و دو ارزشی ساختگی (متغیر دوم) به کار برده می شود، اما ضریب همبستگی دو رشته ای نقطه ای، اس روشی است که برای تعیین همبستگی بین دو متغیر پیوسته (مانند نمره ای که یک فرد در کل آزمون میگیرد) و دو ارزشی واقعی (مانند نمره ای که یک فرد در یک سؤال می گیرد) به کار می رود.

۲۰- گزینه «۱» برای روشن شدن این سؤال فرض کنید، تعداد گروه بالا در یک آزمون، ۱۰ نفر و تعداد افراد گروه پایین، نیز ۱۰ باشد و هیچ کدام از این افراد، پاسخ صحیح به سؤال مورد نظر نداده باشند. طبق فرمول محاسبه ضریب دشواری می توان نتیجه گرفت که ضریب دشواری سؤال 0 میباشد.

۲۱- گزینه «۲» با توجه به فرمول زیر می توان عدد 75/0را، که همان ضریب دشواری بهینه برای سوال های صحیح - غلط است، به دست آورد.

سطح عملکرد مبتنی بر حدس زدن در آزمون های مختلف، متفاوت است. در آزمون های چهارگزینه ای مساوی با ۲۵/0، در آزمون های صحیح - غلط مساوی با 50/0 و در سؤال های ۵ گزینه ای مساوی با ۲۰/0 می باشد. برای به دست آوردن سطح عملکرد مبتنی بر حدس زدن، ۱۰۰% را بر تعداد گزینه های سؤال مورد نظر تقسیم می کنیم.

۲۲- گزینه «۲» طبق تعریف، برای اینکه آزمون از سؤال های بهتری تشکیل شود، باید ضریب دشواری آن از ۱ کمتر و از صفر بیشتر و به 5//0 نزدیک باشد؛ البته این مورد برای سوال های چهار گزینه ای، ۶/0 است، یعنی حدفاصل بین 25/0 تا ۱، همچنین می توان گفت زمانی که ضریب دشواری 50/0 باشد، ضریب تمییز برابر با او کامل است.

۲۳- گزینه «۲»

۰/۱۰ – 37/0 – 27/0 = نسبت گروه ضعیف - نسبت گروه قوی

ضریب تمییز سؤال = ۱۰/۰ - =d

۲۴- گزینه «۱» در صورتی که آزمودنی های یک آزمون از نظر سن یا درجه تحصیلی یا جنس، با یکدیگر متجانس باشند، دامنه تغییر درجه دشواری سؤال ها به صورت محدود خواهد بود؛ یعنی طیف تغییر سؤالات، از ساده به مشکل یا برعکس، یکنواخت است. بنابر تعریف، درصد کل آزمون شوندگانی که به یک سؤال، جواب درست می دهند، ضریب دشواری آن سؤال نام دارد. در این حالت، افراد باتوجه به تجانسی که با یکدیگر دارند، در پاسخ به سؤالات نزدیک به یکدیگر عمل می کنند.

۲۵- گزینه «۲» ضریب تمییز، قدرت سؤال را در تشخیص دو گروه قوی و گروه ضعیف آزمون شوندگان مشخص می کند و یا این فرمول محاسبه می شود:

|  |  |
| --- | --- |
| 9  b | 2  a |
| d  2 | c  5 |

۲۶- گزینه «1» برای پاسخ به این سؤال فرض کنید که آزمونی را روی ۵ نفر اجرا کرده ایم، که هم در پیش از آموزش توانسته اند به آن پاسخ صحیح بدهند و هم پس ای آموزش با دراختیار داشتن فرمول محاسبهی ضریب تمییز پیش و پس از آموزش و جایگزین کردن اعداد، جواب صحیح، که همان صفر است، به دست می آید.

تعداد کسانی که پیش از آموزش به سؤال، جواب صحیح داده اند

تعداد کسانی که پس از آموزش، به سؤال جواب صحیح داده اند

تعداد کسانی که هم پیش و هم پس از آموزش، به سؤال جواب داده اند یعنی تعداد کل افراد = T

۲۷- گزینه «۴» برای پاسخ به این سؤال، ابتدا نمرهی اصلاح شدهی فرد را طبق فرمول زیر به دست می آوریم و سپس آن را با استفاده از یک طرفین - وسطین ساده، بر مبنای ۲۰ محاسبه می کنیم.

تعداد پاسخ غلط

تعداد گزینه ها

۲۸- گزینه «۲» با قراردادن اعداد داده شده در فرمول ضریب تمییز، 20/0، که همان جواب صحیح سؤال می باشد، به دست می آید.

۲۹- گزینه «۲» ضریب دشواری، میزان آزمون شوندگانی را شامل می شود که به یک سؤال، جواب درست می دهند؛ بنابراین ضریب دشواری متوسط، یعنی سؤالی که نه آسان است و نه مشکل ( 5/0 P =) برای محاسبه واریانس یک سؤال از این فرمول استفاده می کنیم:

(p – 1) p واریانس سؤال

P ضریب کسانی که به سؤال، جواب درست داده اند ضریب دشواری

زمانی که ضریب دشواری سؤال را متوسط 5/0 p = درنظر بگیریم، واریانس سؤال حداکثر خواهد بود. ضریب دشواری سؤال، بین عدد ۱ و عدد قرار دارد.

۳۰- گزینه «۱» زمانی که در آزمون انشایی - تشریحی، مصحح اوراق را تصحیح می کند، دچار (bias) سوگیری نیز می شود؛ چنان که از دانش آموز قوی و . بااستعداد خود انتظار دارد که مطابق امتحانات گذشته، نمره ی خوبی به دست آورد و از دانش آموز ضعیف کلاس خود انتظار چشمگیری ندارد. این مورد را اثر پیگمالیون نامیده اند.

۳۱- گزینه «۲» با توجه به فرمول ضریب تمییز، پاسخ صحیح، که همان 25/0 است، به دست می آید:

۳۲- گزینه «۱» با توجه به فرمول ضریب تمییز نیز پاسخ صحیح گم همان 25/0 است به دست می آید:

۳۳- گزینه «٣» نظریه سؤال - پاسخ بر نظریه یا الگوی صفت مکنون استوار است و نظریه صفت مکنون بر این فرض استوار است که یک صفت یا ویژگی زیر بنایی وجود دارد که به شخص امکان می دهد تا در یک تکلیف شناختی معین، موفق شود؛ از جمله ی این صفات می توان دانش معنی لغات، استدلال ریاضی یا تجسم فضایی را نام برد. افزون بر این، چنین فرض می شود که هر چه شخص بیشتر از این صفت برخوردار باشد، در آزمون مربوط به آن صفت عملکرد بهتری خواهد داشت.

۳۴- گزینه «1» ضریب دشواری بندر صد کل آزمون شوندگانی اشاره دارد که به یک سؤال، جواب درست می دهند. ضریب تمییز به قدرت تشخیص بین گروه قوی و گروه ضعیف آزمون شوندگان اشاره دارد. با توجه به توضیحات بالا، در صورتی ضریب تمییز به حل اکثر یا ۱+ می رسد، که ضریب دشواری در حد وسط یا برابر 5/0 باشد، اما اگر ضریب دشواری 5/0 باشد، ضریب تمییز حتمأ یک نیست.

۳۵- گزینه «۴» فرمول نمره اصلاح شده به این شرح می باشد:

=Rتعداد پاسخ های درست

=Wتعداد گزینه های هر سؤال

=Nتعداد پاسخ های غلط

۳۶- گزینه «۲» نمره یادگیرنده را بر تعداد سؤالات تقسیم می کنیم:

حاصل این تقسیم، درصد عملکرد او را در آزمون نشان می دهد.

۳۷- گزینه «۴» خطای هاله ای یا اثر هاله ای:

خطایی است که پژوهشگر را به اشتباه می اندازد، او را به قضاوت نادرست هدایت می کند و به یک تعمیم نادرست به نمونه یا جامعهی آماری منجر می شود. بنابراین تعریف، گزینه ی شماره ۴ صحیح است که یک ویژگی برجسته در آن، سایر خصوصیات فرد آزمون شونده را تحت الشعاع قرار می دهد.

۳۸- گزینه «۴» ضریب دشواری به درصد کل آزمون شوندگانی که به یک سؤال جواب درست می دهند، گفته می شود؛ بنابراین، هرچه ضریب دشواری یک سؤال بالاتر باشد، آن سؤال آسان تر و هرچه ضریب دشواری سؤال پایین تر باشد، آن سؤال سخت تر می باشد. یک سؤال با درجه دشواری ۸/0 خیلی آسان تر از سؤالی است که درجه دشواری آن ۳/0 است.

۳۹- گزینه «۱» باتوجه به فرمول، ضریب دشواری ۶۰/0، که همان گزینه ۱ می باشد، به دست می آید.

۴۰- گزینه «۲» زمانی که سؤال ضریب تمییز منفی دارد، به صورت منفی بین افراد قوی و ضعیف تمایز قائل می شود؛ به صورتی که گروه قوی که در کل آزمون عملکرد بهتری دارند، در این سؤال، عملکرد ضعیفی دارند و افراد ضعیف، که در کل آزمون، نمرهی پائینی گرفته اند، در این سؤال، عملکرد بهتری دارند.

۴۱- گزینه «٣» ضرایب دشواری 90/0 و 50/0 و 35/0 و 11/0 به ترتیب نشان می دهند که ۹۰ درصد، ۵۰ درصد، ۳۵درصد و ۱۱ درصد آزمون شوندگان توانسته اند به سؤال موردنظر، پاسخ صحیح بدهند؛ بنابراین ضریب دشواری 11/0، که تنها ۱۱ آزمون شونده توانسته اند به آن پاسخ درست دهند، از بقیه گزینه ها ضریب دشواری بیشتری دارد. |

۴۲- گزینه «۴» یکی از راه های قضاوت درباره مفید بودن سؤال های آزمون این است که ببینیم سؤال، تاچه اندازه به واریانس نمرات کمک می کند. ضریب دشواری، برای تحلیل یک سؤال کلیه ی افراد (کلیه برگه های امتحانی) کاربرد دارد. این ضریب به تصمیم گیری درباره نحوه آزمون کمک می کند؛ بنابراین به نتیجه آزمون بستگی دارد.

۴۳- گزینه «۳» وقتی در یک آزمون، درصد خاصی از آزمون شوندگان به یک سؤال، پاسخ درست دهند، ضریب دشواری آن سؤال مشخص می شود. بهتـر است که به جای ضریب دشواری، ضریب سهولت یا آسانی به کاربرد، اما اصطلاح معمول، همان ضریب دشواری است.

۴۴- گزینه «۲»

۴۵- گزینه «۱»

۴۶- گزینه «۳» برای محاسبه واریانس یک سؤال از این فرمول استفاده می شود:

= P(1-P) واریانس سوال

Pدرصد کسانی است که جواب درست به سؤال داده اند، یعنی همان ضریب دشواری. یک سؤال، زمانی دارای حداقل واریانس است که ضریب دشواری آن صفر یا ۱ است. وقتی دشواری سؤال 5/0 = P باشد، واریانس حداکثر خواهد بود.

۴۷- گزینه «۴» ضریب دشواری، درصد کل آزمون شنوندگانی را نشان می دهد که به یک سؤال، جواب درست می دهند؛ بنابراین هرچه ضریب دشواری یک سؤال، بالا باشد نشان می دهد که آن سؤال، آسان تر است؛ پس میتوان ضریب آسانی یا سهولت را به جای ضریب دشواری به کار برد، اما معمولاً ضـریب دشواری رواج بیشتری دارد.

۴۸- گزینه «۱» به طور کلی، وقتی ضریب دشواری بین 3/0 تا ۷/0 باشد، حداکثر اطلاع در مورد تفاوت بین آزمودنی هـا بـه دست می آید؛ بنابراین، این سؤال مشخص کننده تعداد افرادی است که اطلاعات کامل تر و بیشتری نسبت به کل آزمون شوندگان دارند و می توانند در آزمون نهایی لحاظ شوند.

۴۹- گزینه «۳» باتوجه به این که در فرمول ضریب دشواری، انتخاب های درست گروه بالا و پایین، مبنای تعیین ضریب دشواری قرار می گیرد، می توان گفت که گزینه ۳ صحیح می باشد.

۵۰- گزینه «۱» با توجه به فرمول زیر جواب به دست می آید:

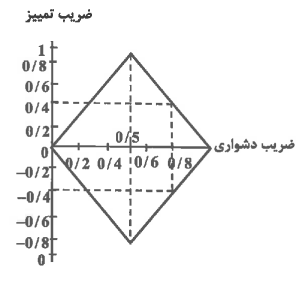
تعداد پاسخ های غلط تعداد پاسخ های درست

تعداد گزینه ها

۵۱- گزینه «۱» آزمون دشوار، یعنی تعداد اندکی از افراد توانسته اند پاسخهای درست ارائه کنند. توزیع در کجی مثبت، نامتقارن متمایل می شود.

فراوانی های بزرگتر در انتهای پایین تر و فراوانی های کوچک تر در انتهای بالای توزیع قرار می گیرند؛ در این صورت، می توان گفت که سطح آزمـون، دشـوار است. اثر سقف، به نمرات بالا اشاره می کند و تأثیر کف، به نمرات پایین اشاره دارد. با توجه به آنچه بیان شد، در آزمون دشوار، ویژگی اثر سقف بارز تر است. ضریب تشخیص بالا (گزینه ۳) به آزمون های آسان اشاره دارد که همه افراد می توانند به آن پاسخ درست بدهند.

۵۲ـ گزینه «۴» در شکل روبه رو رابطه بین ضریب دشواری و ضریب تمییز سؤال را مشاهده کنیم؛ برای مثال اگر ضریب دشواری 8/0 باشـد، ضریب تمییز بین 4/0 + و 4/0 - خواهد بود. ضریب تمییز درصورتی به حداکثر (یعنی ۱+) می رسد که ضریب دشواری برابر 5/0 باشد، یعنی درحد وسط؛ در نتیجه، سؤال های خوب یک آزمـون، سؤال هـایی بـا ضریب دشواری متوسط و ضریب تمییز بزرگ هستند که درنتیجه به نرمال شدن توزیع جواب آزمون شوندگان منجر میشوند (منحة (منحنی u شکل).



۵۳- گزینه «۴» زمانی که در یک مجموعه از سؤالات، با شیب زیاد خم ویژه سؤال روبه رو می شویم، نتیجه می گیریم که سؤال، از قدرت تمییز بالا توسط آزمون شونده برخوردار است.

هرقدر ضریب تمییز، بزرگتر باشد، قوه تمییز آن سؤال بیشتر می شود. ضریب تمییز، تفاوت بین گروه قوی و گروه ضعیف را مشخص می کند و قدرت سؤال را در تمایزگذاری یا تشخیص بین گروه قوی و گروه ضعیف آزمون شوندگان مشخص می کند؛ به عبارت دیگر، مشخص می کند که سؤال، تا چه اندازه می تواند گروه قوی را از گروه ضعیف جدا کند.

54- گزینه «1» کجی منفی در صورتی رخ می دهد که میزان سؤالات آسان در آزمون زیاد باشد و اکثر افراد بتوانند به میزان بیشتری از سؤالات پاسخ دهند؛ در این صورت بار وارد کردن سؤالات دشوار، میزان حدس و گمان یا پاسخگویی به سؤالات توسط افراد کاهش می یابد و توزیع به حالت نرمال برمی گردد؛ زیرا اگر آزمونی دشواری نداشته باشد، یعنی ضریب دشوارتر آزمون برابر با یک باشد، آزمون به هدف مورد نظر نمی رسد. با این تعریف، گزینه ۲ هم می تواند برای داوطلبان تردید ایجاد کند و آن ها را برای انتخاب دچار مشکل کند، اما چون در متن سؤال، صراحتا از کجی منفی صحبت شده است، که آزمون آسان را شامل می شود برای بالابردن اطمینان از صحت آزمون و تبدیل کردن نمودار کجی منفی به سمت نرمال، فقط باید از سؤالات دشوار استفاده کرد.

۵۵- گزینه «۱» از آنجا که این سؤال هیچ گونه رابطه ی منظمی بین نسبت پاسخ های درست و نمرات کل آزمون نشان نمی دهد و نوسان دارد، به خوبی عمل نمی کند و به هیچ وجه نمی تواند بین آزمون شوندگان مختلف تمییز قائل شود. به طور کلی هرچه منحنی ویژگی سؤال، حالت پلکانی بیشتری داشته باشد، همبستگی بیشتر می تواند بین افراد گروه قوی و گروه ضعیف تمایز قائل شود.

56-گزینه «۴» هنگامی که ضریب تمییز یک سؤال چهار گزینه ای ۱ باشد، بدین معنی است که این سؤال، به طور کامل بین افراد گروه قوی و گروه ضعیف تمایز قائل شده است و کل افراد گروه قوی، آن سؤال را صحیح و کل افراد گروه ضعیف، آن را اشتباه پاسخ داده اند. حال فرض کنید که تعداد افراد گروه قوی ۱۵ و تعداد افراد گروہ ضعیف نیز ۱۵ باشد و کل افراد گروه قوی به سؤال موردنظر پاسخ صحیح داده باشند، با جایگزین کردن این اعداد در فرمول ضریب دشواری، عدد ۵۰/0 به دست می آید.

۵۷- گزینه «3» نظریه صفت مکنون یا Item - responsetheory) IRT) بر این فرض استوار است که یک صفت یا ویژگی زیربنایی وجود دارد که به شخص امکان می دهد تا در یک تکلیف شناختی معین، موفق شود. از جمله این صفات می توان به «دانش معنی لغات» یا «استدلال ریاضی» اشاره کرد؛ به این صورت که فرض می شود، هرچه شخص از این صفت، بیشتر برخوردار باشد، در آزمون مربوط به آن صفت، عملکرد بهتری خواهد داشت.

کاربرد این نظریه به این مساله اشاره می کند که به جای تأکید بر نمرات کل آزمون، بر پاسخ های آزمودنی ها به تک تک سؤال های آزمون تکیه میشود. بر طبق این نظریه می توان به قضاوت در مورد آزمون شوندگان پرداخت.

۵۸- گزینه «۱» برای به دست آوردن ضریب دشواری این سؤال، به ۲ نسبت نیاز داریم: نسبت افرادی از گروه بالا که به سؤال، پاسخ صحیح داده اند و نسبت در افرادی از گروه پایین، که به سؤال پاسخ صحیح داده اند: اولی را با و دومی را با نشان می دهند. با قراردادن مقدار (60/0) و مقدار ضریب تمییز (40/0) در فرمول ضریب تمییز می توانیم مقدار را نیز به دست آوریم.

اگر بپذیریم که ضریب تمییز برابر است با و سپس مقادیر D و را به جای آن جایگزین کنیم، مقدار برابر با ۲۰% خواهد بود. با جایگزین کردن آن ها در فرمول ضریب دشواری، مقدار آن به صورت روبه رو به دست می آید:

۵۹- گزینه «۲» از روی منحنی ویژه سؤال، می توان ضریب دشواری و تمییز سؤال را تعیین کرد. ضریب دشواری عبارت است از: نمره ی معیاری که در آن ۵۰ درصد آزمون شوندگان، پاسخ درست به سوال داده اند. ضریب تمییز نیز برابر است با شیب منحنی ویژه ی سؤال.

60- گزینه «٣» قدرت تمییز سؤال، نوعی ضریب توافق سؤال با کل آزمون است.

61-گزینه «3» ضریب تمییز منفی نشان می دهد که در آن سؤال، گروه قوی بدتر از گروه ضعیف عمل کرده است.

۶2- گزینه «۲» فرمول ضریب دشواری:

63- گزینه «۴» سؤال «ب» نسبت به سؤال های دیگر ضریب تمییز بهتری دارد؛ یعنی بهتر می تواند نمره کل آزمون شونده را پیش بینی کند.

۶۴- گزینه «3» آزمون، وسیله ای است عینی و استاندارد شده که برای اندازه گیری نمونه ای از رفتاری خصایص آدمی به کار می رود. مراد از عینی،آن است که روش اجرا، نمره دادن و تعبیر و تفسیر نتایج آزمون براسنایی قواعدی معین و مشخص صورت می گیرد و قضاوت و نظر شخصی در آن ها بی تأثیر است. برای هر آزمون دستورالعمل خاصی در مورد نحوه ی اجرا، طول مدت آزمایشی دستورهای شفاهی، نحوه ی ارائه مثال ها و روش تفسیر نمره ها وجود دارد. اصطلاح استانداردشده بدین معناست که آزمون قبلا در مورد گروه نمونه ای از افراد مورد نظر در بوته ی آزمایش گذاشته شده، نتایج پژوهش های مربوط به آن، از راه روش های آماری، مورد تجزیه و تحلیل قرار گرفته و روایی و اعتبار آن تعییتر شده است؛ به علاوه، جداول ترم یا هنجارهایی دارد که نمره های خام آزمون بر مبنای آن تعبیر و تفسیر می شوند.

تعداد مناسب گزینه های روشن و واضح بودن سؤالات و سهولت اجرا و نمره گذاری، از ویژگی های یک آزمون مناسب است؛ در این بین، روشن و واضح بودن . سؤالات اهمیت بیشتری دارد؛ زیرا این مبحث به مفهوم روایی (validity) مربوط است. روایی به هدفی اشاره می کند که آزمون برای تحقق بخشیدن به آن درست شده است. برای اینکه آزمونی روا باشد، باید سؤالات مرتبط به آن آزمون، به صورت روشن و واضح نوشته شده باشند تا آزمودنی ها دقیقا متوجه سؤال بشوند و به دلیل عدم درکشان از سؤال، نمره ای را از دست ندهند؛ بنابراین، کیفیت سؤال های آزمون، موجب افزایش روایی می شود (روایی، یکی از سه ویژگی های مهم آزمون است). کوتاه بودن طول آزمون با میزان پایایی آزمون، رابطه معکوس دارد. با افزودن تعداد سؤالات یک آزمون، پایایی آن نیز افزایش در می یابد؛ از طرف دیگر روایتی از پایایی تأثیر می پذیرد، پس هر چه تعداد سؤال های آزمون بیشتر باشد، نمونه ی بهتری از محتوای مورد اندازه گیری را شامل خواهد بود و به همان نسبت به روایی آزمون افزوده میشود.

65- گزینه «۱» در نمره گذاری آزمون های عینی، ۲ روش عمده وجود دارد:

۱) محاسبه ی کلیدی پاسخ های درست بدون کسر نمره برای حدس زدن.

۲) کسر مقداری از نمره برای جبران حدس زدن. روش اول، روشی متداول تر و ساده تر است، اما اگر معلم بخواهد از روش دوم استفاده کند باید هنگام اجرای آزمون آزمون شوندگان را آگاه کند که به ازای پاسخ های غلطی که می دهند، مقداری از نمره ی آنان کم خواهم شد.

R = تعداد پاسخ های درست

W = تعداد پاسخ های غلط

N = تعداد گزینه های هر سوال

۶۶- گزینه «۲»

هر اندازه ضریب دشواری یک سؤال، بزرگتر باشد، آن سؤال، آسان تر است.

۶۷- گزینه «۲» یکی از انتظارات ما از آزمون های وابسته به هنجار این است که برای آزمون شوندگان مختلف، نمراتی به دست دهد که در طول یک پیوست و پراکنده باشند، هر چه این پراکندگی بیشتر باشد، بهتر است؛ به سخن دیگر، هرچه واریانس نمرات حاصل از یک آزمون وابسته به هنجار بزرگتر باشد، آن آزمون، آزمون بهتری است؛ بنابراین یکی از راه های قضاوت درباره ی مفید بودن سؤال های یک آزمون، آن است که ببینیم سؤال، چه اندازه به پراکندگی یا واریانس نمرات کمک می کند.

واریانس نمرات گروهی پاسخ دهنده به یک آزمون، از ۲ قسمت تشکیل می یابد. ۱) واریانس سؤال ها ۲) همبستگی بین سؤال ها؛ بنابراین هر سؤالی که دارای واریانس بزرگتری است و یا سایر سؤال های آزمون همبستگی بیشتری دارد، به واریانس کل آزمون، کمک بیشتری می کند. در مقابل، سؤالی که با سایر سؤال ها همبستگی ندارد و واریانس کوچکی دارد، کمک زیادی به واریانس کل آزمون نمی کند؛ بنابراین، اگر در تحلیل آزمون، سؤال هایی را انتخاب کنیم به که همبستگی زیادی با یکدیگر دارند و در ضمن، دارای واریانس های بزرگ و قوی هستند، در مجموع، آزمون بهتری را تدارک دیده ایم. اگر همه سؤال های یک آزمون با یکدیگر همبستگی کامل داشته باشند و ضریب دشواری آن ها ۵/0 باشد، آن گاه نیمی از آزمون شوندگان در آزمون، نمره صفر و نیمی دیگر، نمره ی کامل خواهند گرفت.

px(1-p)= واریانس سؤال

68- گزینه «3» مهم ترین مرحله تهیه آزمون های روان شناختی، مرحله تعریف صفت مورد اندازه گیری است؛ زیرا تا تعریف ها، مشخصات و ویژگی های صفت مورد نظر بررسی نشود، نمی توان آزمونی با درجه روایی و پایایی بالا تهیه کرد.

۶۹- گزینه «1» اگر هدف از اجرای آزمون، گزینش تعداد کمی از داوطلبان باشد، استفاده از سؤالات دشوار، مناسب تر است.

۷۰- گزینه «٣» چنانچه فراوانی پاسخگویی به یک سؤال در گروه بالا نسبت به گزینه های مختلف، مساوی باشد، احتمالا آزمودنی ها، پاسخ هایی از روی حدس و گمان ارائه کرده اند و اگر سؤالات در گروه بالا نسبت به گزینه های مختلف، متفاوت باشد، نشان دهنده قدرت تمایز بالای سؤال است.

۷۱- گزینه «3» چنانچه آزمون، اثر کف داشته باشد، منحنی توزیع نمرات، دارای شکل چولگی مثبت است.

۷۲- گزینه «۱» شکل مقابل، رابطه میان دشواری سؤال و واریانس سؤالات چندگزینه ای را نشان می دهد.

۷۳- گزینه «۲» ضریب تشخیص یا قوه تمییز هر سؤال، تفاوت سطح دشواری آن را در افراد گروه قوی و ضعیف نشان می دهد؛ بنابراین با فرمول قابل محاسبه است. در این سؤال، اگر قدرت تمییز سؤالی برابر با ۱ باشد، نسبت پاسخ صحیح به سؤال در گروه پایین، صفر است. به این ترتیب:

= D ضریب تشخیص

= گروه بالا

= گروه پایین

۷۴- گزینه «۴» محاسبه سطح دشواری هر سؤال در گروه قوی ()، یعنی نسبت افرادی که از گروه قوی به آن سؤال، پاسخ داده اند . محاسبه سطح دشواری هر سؤال در گروه ضعیف ()، یعنی نسبت افرادی که از گروه ضعیف به آن پاسخ درست داده اند .

در شکل ارائه شده، سؤال D دارای بیشترین دشواری است.

۷۵-گزینه «۲» طبق جدول ارائه شده و بحث تمییز گروه ضعیف و قوی، هرچقدر تفاوت آشکارتری در نمرات بین گروه قوی و ضعیف باشد، می توان آن آزمون را مناسب تلقی کرد. ضریب تمییز، تفکیک مناسبی از گروه های ضعیف، متوسط و قوی آزمودنی ها به دست می دهد و هر چقدر این تفکیک، متمایزتر باشد، آزمون، استاندارد تر است.

۷۶- گزینه «3» با توجه به اینکه میانگین افرادی که پاسخ صحیح داده اند با میانگین افرادی که پاسخ صحیح نداده اند، برابر است، نمی توان بین افراد قوی و ضعیف تمایز قائل شد و این، بدان معناست که همبستگی سوال با مجموعه سوالات، مفهومی ندارد؛ هرچقدر سؤالی با سؤالات دیگر همخوانی داشته باشد، ضریب همبستگی آن بیشتر است.

۷۷- گزینه «1» اساس نظریه کلاسیک در اندازه گیری صفات روانی بر پایه ی ۱- تفاوت های بین فردی ۲- خطای تصادفی ۳- کنترل آزمایشی استوار است.

۷۸- گزینه «۴» یکی از ویژگی های اصلی سؤالات مناسب در آزمون های روان شناختی استاندارد امکان افزایش واریانس آزمون است.

آزمون فصل سیزدهم

1- فردی در یک آزمون چهارگزینه ای که ۴۰ سؤال دارد، به ۳۰ سؤال، پاسخ صحیح و به ۹ سؤال، پاسخ غلط داده است. نمره او پس از اصلاح حدس زدن چقدر است؟

1) 21

2) 30

3) 31

4) 27

۲- دشواری سؤال از لحاظ دامنه بین کدام ۲ عدد قرار دارد؟

1) صفر تا ۱-

۲) صفر تا 1+

3) 1- تا 1+

4) 2- تا 2+

3- ضریب دشواری یعنی:

1) درصد کل آزمون شوندگانی که به سؤال، پاسخ غلط داده اند.

۲) میزانی که گزینه های انحرافی یک سؤال، نتوانسته دانش آموزان ضعیف را گمراه کند.

۳) درصد کل آزمون شوندگانی که به سؤال، پاسخ درست داده اند.

۴) درصد کل آزمون شوندگان قوی که به سؤال، پاسخ درست داده اند.

۴- در جدول مقابل اگر گزینه الف، گزینه صحیح باشد، ضریب دشواری و تمییز سؤال به ترتیب چقدر است؟

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| پاسخ های پیشنهادی | الف | ب | ج | د |
| گروه قوی | 11 | 0 | 3 | 1 |
| گروه ضعیف | 4 | 2 | 6 | 3 |

۱) 5/0 و 32/0

2) 41/0 و 47/0

3) 33/0 و 56/0

4) 5/0 و 47/0

5- در تحلیل سؤالات آزمون باید سؤالاتی را انتخاب کرد که:

1) با سایر سؤال های آزمون، همبستگی پایینی دارند و دارای واریانس کوچکتری هستند.

۲) با سایر سؤال های آزمون، همبستگی بالایی دارند و دارای واریانس بزرگتری هستند.

۳) با سایر سؤال های آزمون همبستگی بالایی دارند و دارای واریانس کوچکتری هستند.

۴) با سایر سؤال های آزمون، همبستگی پایینی دارند اما دارای واریانس بزرگتری هستند.

۶- هر چقدر ضریب دشواری سؤال کوچکتر و به صفر نزدیکتر باشد، آن سؤال:

1) دشوارتر است.

۲) آسان تر است.

۳) بیشتر دانش آموزان به آن جواب داده اند.

۴) هیچ کدام

۷- یک سؤال وقتی دارای حداقل واریانس است که ضریب دشواری آن ......... باشد.

1) 5/0

2) 1

3) صفر یا 1

4) صفر

8- وقتی واریانس سؤالی حداکثر خواهد بود که ضریب دشواری آن ......... باشد.

1) 5/0

2) صفر

3) 1

4) صفر یا 1

9- در سؤال های چهارگزینه ای، سطح موفقیت با حدس زدن چقدر است؟

1) 45/0

2) 25/0

3) 10/0

4) 5/0

۱۰- سؤالی که درجه دشواری آن 6/0 است از سؤالی که درجه دشواری ۹/0 است، .......

۱) آسان تر است.

۲) تفاوتی ندارد.

۳) دشوارتر است.

۴) هیچ کدام

۱۱- سطح دشواری مطلوب برای یک سؤال سه گزینه ای چقدر است؟

1) 70%

2) 30%

3) 50%

4) 66/0

۱۲- اگر همبستگی سؤال با کل آزمون ۵/0 باشد، این سؤال چند درصد از واریانس نمرات را به خود اختصاص می دهد؟

۱) ۲۵ درصد

۲) ۴۰ درصد

3) ۵۰ درصد

۴) ۳۵ درصد

۱۳- در یک آزمون، سؤالات خوب، سؤالاتی هستند:

۱) با ضریب دشواری بالا و ضریب تمییز بزرگ

۲) با ضریب دشواری متوسط و ضریب تمییز بزرگ

۳) با ضریب دشواری پایین و ضریب تمییز بزرگ

۴) با ضریب دشواری بالا و ضریب تمییز کوچک

۱۴- از بین ۱۵ دانش آموزی که به حد تسلط رسیده اند، ۱۰۰ نفر به یکی از سؤال ها پاسخ درست داده اند و از میان ۶ دانش آموزی که به حد تسلط نرسیده اند، ۳ نفر به سؤال، پاسخ درست داده اند. شاخص تمییز این سؤال چقدر است؟

1) 4/0

2) 47/0

3) 30/0

4) 025/0

۱۵- جدول زیر نشان دهنده پاسخ های درست و نادرست دو گروه از دانش آموزان قوی و ضعیف (که به حد تسلط رسیده اند یا نرسیده اند) به یک سؤال از آزمون های ملاک مرجع است. ضریب تمییز سؤال چقدر است؟

|  |  |  |
| --- | --- | --- |
| تسلط عدم تسلط | | |
| درست  نادرست | 9 | 4 |
| 1 | 6 |

1) 30/0

2) 20/0

3) 50/0

4) 45/0

۱۶- اگر علی دریک آزمون پنج گزینه ای که ۴۵ سؤال دارد به ۳۳ سؤال، پاسخ صحیح و به ۷ سؤال، پاسخ غلط داده باشد، نمره ی اصلاح شده او ................ می باشد.

1) 31

2) 25/31

3) 32

4) 5/32

۱۷- ضریب دشواری بهینه یک آزمون چهار گزینه ای چند می باشد؟

1) 60/0

2) 61/0

3) 62/0

4) 63/0

۱۸- ضریب دشواری بین ...................... تا .................. حداکثر تفاوت بین آزمودنی ها را به دست می دهند.

1) 25/0 تا 60/0

2) 30/0 تا 60/0

3) 25/0 تا 70//0

4) 30/0 تا 70/0

۱۹- اگر سوالی دارای ضریب دشواری ۱ باشد، به این معنی است که ....................

۱) تمام آزمون شوندگان به آن سؤال، پاسخ صحیح داده اند.

۲) نیمی از آزمون شوندگان به سؤال، پاسخ صحیح داده اند.

۳) هیچکدام از آزمون شوندگان نتوانسته اند به این سؤال، پاسخ صحیح بدهند.

۴) با این اطلاعات نمی توان نظر داد.

۲۰- در چه صورتی ضریب تمییز یک سؤال 5 خواهد بود؟

۱) در صورتی که ضریب دشواری ۲۵/0 باشد.

۲) درصورتی که ضریب دشوار ۵/0 باشد.

۳) در صورتی که ضریب دشواری ۷۵/0 باشد.

۴) در صورتی که ضریب دشواری 1 باشد.

۲۱- در جدول زیر، اگر گزینه ی ج گزینه صحیح باشد، ضریب دشواری و تمییز سؤال به ترتیب ........خواهد بود.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| پاسخ های پیشنهادی | الف | ب | ج | د |
| گروه قوی | 0 | 3 | 1 | 16 |
| گروه ضعیف | 3 | 5 | 2 | 10 |

1) 55/0 – 40/0

2) 60/0 – 35/0

3) 075/0 – 05/0

4) 75/0 – 4/0

۲۲- اگر ضریب دشواری سؤالی ۵۸/0باشد، واریانس آن سؤال چقدر خواهد بود؟

1) 20/0

2) 24/0

3) 30/0

4) 34/0

۲۳- کدام یک از موارد زیر نشان دهنده ی ضریب تمییز بالا برای یک سؤال می باشد؟

۱) همبستگی پایینی بین سؤالات یک آزمون

۲) همبستگی بالا بین سؤالات یک آزمون

۳) همبستگی پایین بین سؤال با کل آزمون

۴) همبستگی بالا بین سؤال و کل آزمون

فصل سیزدهم: اجرا، نمره گذاری و تحلیل آزمون

۲۴- کدام یک از موارد زیر نشان دهنده ی سؤال دشوار تری است؟

1) 25/0 p =

2) 40/0 p =

3) 67/0 p =

4) 92/0 p =

25- یک آزمون تشریحی را برای ۳۵ نفر اجرا کرده ایم، چنانچه یک سؤال از این آزمون و نمره ممکن در این آزمون به ترتیب 0 و ۴ باشد، ضریب تمییز این سؤال چقدر خواهد بود؟

1) 07/0

2) 70/0

3) 08/0

4) 80/0

۲۶- اگر سؤالی دارای دامنه ۱ تا ۵ باشد و نمره میانگین آن نیز 7/3 باشد، ضریب دشواری آن را محاسبه کنید.

۱) ۸۸/0

2) 90/0

3) 92/0

4) 98/0

2۷- اگر همه ی آزمون شوندگان به یک سؤال، جواب صحیح بدهند، منحنی ویژگی آن سؤال ..................

1) یک خط افقی در پایین محور عمودی خواهم بود.

۲) به صورت یکسانی خواهد بود.

3) از نسبت صفر (0) می گذرد.

۴) یک خط افقی در بالای محور عمودی خواهد بود.

۲۸- نقطه تقاطع خط افقی برآمده از نسبت ................... را که نقطه ی عطف منحنی است، آستانه می گویند.

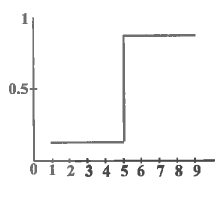
1) 25/0

2) 5/0

3) 75/0

4) 1

۲۹- اگر منحنی ویژگی یک سؤال به صورت زیر باشد، کدام گزینه زیر صحیح می باشد؟



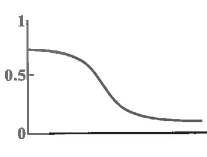
1) سؤال دارای ضریب دشواری پایینی می باشد.

۲) سؤال دارای ضریب تمییز متوسطی می باشد

۳) سؤال دارای ضریب دشواری بالایی می باشد.

۴) سؤال دارای ضریب تمییز بالایی می باشد.

۳۰- اگر منحنی ویژگی سؤال برای یک گزینه انحرافی به صورت زیر باشد بدین معنی است که ...................



1) سؤال به طور کامل بین آزمون شوندگان قوی و ضعیف تمایز قائل می شود.

۲) سؤال به صورت مثبت بین آزمون شوندگان قوی و ضعیف تمایز قائل می شود.

۳) سؤال نمی تواند بین افراد قوی و ضعیف تمایز قائل شود.

۴) سؤال به صورت منحنی بین گروه قوی و ضعیف تمایز قائل می شود.

فصل چهاردهم

«استفاده از روش های آماری برای تفسیر نمره های آزمون»

برای تفسیر و معنی دادن به نتایج آزمون های پیشرفت تحصیلی، لازم است با پاره ای از مفاهیم و روش های آماری آشنا شویم. در این فصل روش های ساده آماری را، که برای تفسیر نمرات امتحانی لازم هستند، معرفی می کنیم.

تعاریف

متغیر: به هرگونه ویژگی یا صفتی گفته می شود که می تواند مقادیر مختلفی را به خود بگیرد.

فراوانی: دفعات وقوع یک متغیر را فراوانی آن متغیر می گویند.

توزیع فراوانی: به تعدادی نمره گفته می شود که به ترتیب بزرگی مرتب شده باشند و فراوانی آن ها نیز مشخص شده باشد. بنا به تعریف، توزیع فراوانی، جدولی است که نشان می دهد هریک از نمرات در گروه نمرات چندبار تکرار شده است.

دو نوع توزیع فراوانی وجود دارد:

1- توزیع فراوانی طبقه بندی نشده: که در آن نمره ها به ترتیب کوچکی با بزرگی پشت سرهم قرار گرفته اند و فراوانی هر نمره در مقابل آن قرار دارد.

|  |  |
| --- | --- |
| جدول 1:توزیع فراوانی نمرات طبقه بندی نشده | |
| فراوانی | نمره |
| 1  5  6  3  2  3  20 N = | 20  19  18  17  16  15 |

در نظام نمره گذاری آموزشگاه های ایران، که نمرات کلأ۲۰ طبقه از صفر تا ۲۰ را تشکیل می دهند، طبقه بندی نمرات فایده ای ندارد و بهتر است با نمرات طبقه بندی نشده جدول توزیع فراوانی را تهیه کنیم.

نکته 1: مجموع فراوانی ها با تعدات کلی برابر است.

۲- توزیع فراوانی طبقه بندی شده: وقتی تعداد نمرات، زیاد باشد، آن ها را طبقه بندی می کنیم، برای این کار باید فاصله طبقاتی مناسب را تعیین کنیم. برای به دست آوردن فاصله طبقاتی مناسب، بهتر است که بزرگترین داده را از کوچکترین داده کم کنیم، عدد حاصل را بر ۱۵ تقسیم کنیم و خارج قسمت این تقسیم را فاصله ی طبقاتی در نظر بگیریم؛ چنانچه این خارج قسمت، عددی اعشاری باشد، نزدیکترین عدد به آن را به عنوان فاصله طبقه انتخاب می کنیم.

جدول ۲ توزیع فراوانی، نمرات خام ۳۰ دانش آموز را در درس ریاضی به شرح زیر نشان می دهد.

12,14,13,18,20,19,25,26,25,38,24,28,33,14,19,23,37,34,16,18,21,31,25,36,15,18,33,22,30,24

۲۶ = ۱۲- ۳۸ = فاصله طبقاتی

2 73/1 15 ÷ 26

چون ۲، عدد خیلی کوچکی است و فاصله طبقات را ۳ در نظر می گیریم.

فراوانی تراکمی: یعنی فراوانی هر طبقه به علاوه فراوانی طبقه ما قبل.

|  |  |  |
| --- | --- | --- |
| جدول 2- توزیع فراوانی طبقه بندی شده | | |
| فراوانی تراکمی | فراوانی | طبقات |
| 30  27  23  21  20  15  12  6  4 | 3  4  2  1  5  3  6  2  4 | 38-36  35-33  32-30  29-27  26-24  23-21  20-18  17-15  14-12 |
| 30 N = | | |

نمودار

نمودار، به رسم یا شکلی گفته میشود که برای تبدیل داده های آماری استفاده می شود.

اطلاعات مندرج در توزیع فراوانی را می توان به صورت نمودار نشان داد. نمودار، کار تحلیل و تفسیر نمرات را آسان می کند. نمودارها انواع مختلفی دارند؛ بے مثل نمودار هیستوگرام، نمودار چندضلعی، نمودار تراکمی، نمودار ستونی، نمودار دایره ای و نمودار تصویری.

1- نمودار هیستوگرام: در همه نمودارها از جمله هیستوگرام، توزیع فراوانی نمرات با استفاده از محورهای مختصات نشان داده می شود. برای ترسیم سر نمودار هیستوگرام، فاصله طبقات را روی محور افقی و فراوانی ها را روی محور عمودی مشخص می کنیم، سپس روی هر طبقه، ستون یا مربعی می سازیم؟ به گونه ای که عرض آن برابر با فاصله طبقه و ارتفاعش اندازهی فراوانی همان طبقه باشد.

نکته ۲: برای ترسیم هیستوگرام نمرات طبقه بندی نشده. فاصله طبقه ۱ فرض می شود که از 5/0 نمره قبل از هر نمره تا 5/0 که نمره بعد از آن ادامه می یابد. مثلا فاصله طبقه ۱۰ از 5/9 تا 5/10 است.

۲- نمودار چندضلعی: اگر وسط قاعده ی فوقانی هر یک از مربع – مستطیل های نمودار هیستوگرام را به هم وصل کنیم، شکل حاصل به صورت یک چندضلعی است که نمودار چندضلعی نام دارد. نمودار هیستوگرام و نمودار چندضلعی فرق عمده ای با یکدیگر ندارند و انتخاب یکی از آن ها به سلیقه ی شخصی معلم بستگی دارد.

نکته ۳: نمودارهای هیستوگرام و چندضلعی، مخصوص دانه های پیوسته حاصل از مقیاس های فاصله ای و نسبی هستند.

3- نمودار تراکمی: هر نقطه، معرف هر طبقه در نمودار تراکمی است که پس از اضافه کردن فراوانی طبقه مورد نظر به طبقات قبل از آن تعیین می شود؛ بنابراین، نمودار تراکمی هرگز به پایین نمیرسد: یا بالا می رود یا محور افقی موازی باقی می ماند.

۴- نمودار ستونی: زمانی که با داده های مربوط به مقیاس اسمی یا ترتیبی سروکار داریم، می توانیم از نمودار ستونی یا میله ای استفاده کنیم. رسم این نمودار شبیه نمودار هیستوگرام است، اما با دو تفاوت: ۱- در رسم این نمودار هیچ ترتیبی در نظر گرفته نمی شود. ۲- در این نمودار، ستون با میله ها را باید جدا از یکدیگر رسم کرد. تفاوت نمودار داده های مقیاس ترتیبی با نمودار ستونی مقیاس اسمی این است که در نمودار ستونی مربوط به مقیاس اسمی، هیچ گونه ترتیبی درنظر گرفته نمی شود، اما نمودار ستونی مربوط به مقیاس ترتیبی، نوعی ترتیب دارد که هنگام رسم باید در نظر گرفته شود.

5- نمودار دایره ای: منظور، یک منحنی دایره شکل است که نسبت های مختلف یک کلی را نشان میزدهد.

۶- نمودار تصویری: در نمودار تصویری از تعدادی تصویر یا شکل استفاده قرار می شود؛ به گونه ای که هر کدام معرف فراوانی با درصد وقوع یک رویداد است. نمادهایی که در نمودار تصویری به کار می روند، باید ساده اما معنی دار باشند؛ در ضمن، در هر ردیف بیشتر از حدود ۱۵ تصویر نباید باشد؛ زیرا موجب سردرگمی خواهد شد.

اندازه های گرایش مرکزی

زمانی که بخواهیم با استناد به یک نمره ی واحد، چگونگی تمرکز نمرات را مشخص کنیم، از اندازه های گرایش مرکزی استفاده می کنیم. این اندازه ها عبارتند از: نما، میانه، میانگین. هرچقدر از سمت نما به سمت میانگین حرکت می کنیم، ثبات اندازه گیری بیشتر می شود.

نما

نما نمره ای است که بیشترین فراوانی را دارد. در جدول ۱، عدد ۱۸ نما است؛ چون فراوانی آن از بقیه ی نمرات بیشتر است.

در توزیع های طبقه بندی شده، نما نقطه ی وسط طبقه ای است که بزرگترین فراوانی را دارد؛ مثلاً در جدول ۲، طبقه ۲۰ - ۱۸ بزرگ ترین فراوانی را دارد، که نقطه ی وسط این طبقه یعنی عدد ۱۹، نما است. اگر بزرگترین فراوانی، فقط یک بار در توزیع تکرار شده باشد، آن توزیع، توزیع تک نمایی است؛ اگر بزرگ ترین فراوانی ۲ بار تکرار شده باشد، توزیع انمـایی است و اگر بزرگ ترین فراوانی چندبار تکرار شده باشد، توزیع چندنمایی است.

میانه

میانه نقطه ای روی توزیع نمرات است که توزیع را به ۲ قسمت مساوی تقسیم می کند. به میانه نقطه ۵۰ درصدی می گویند؛ زیرا ۵۰ درصد نمره ها در بالای آن و ۵۰ درصد نمره ها در پایین آن قرار دارند.

محاسبه میانه برای توزیع های طبقه بندی نشده:

زمانی که توزیع، طبقه بندی نشده است و فراوانی هر نمره هم ۱ است، دو حالت داریم:

۱ـ وقتی تعداد نمرات یک توزیع عدد فرد باشد، در اینجا عدد وسطی به عنوان میانه انتخاب می شود.

بهتر است اعداد را یکی از اول و یکی از آخر خط بزنیم، عدد وسطی میانه است.

میانه

۲- وقتی تعداد نمرات یک توزیع، عدد زوج است، در اینجا دو نمره میانی را با هم جمع و بر ۲ تقسیم می کنیم. عدد حاصل، میانه است.

زمانی که توزیع، طبقه بندی نشده است و ما عدد تکراری داریم روش محاسبه ی میانه به شکل زیر است:

۱ـ ابتدا اعداد را به ترتیب از بزرگ به کوچک مرتب می کنیم؛ مثلاً اعداد

13,15,12,11,10,12,12,13,12,17,18,20

20 18 17 15 13 12 12 12 12 12 11 10

۲- تعداد کل اعداد توزیع را بر ۲ تقسیم می کنیم؛ مثلاً توزیع بالا ۱۲ عدد دارد:

۳- حد پایین و حد بالای عدد تکراری را مشخص می کنیم؛ مثلاً اینجا: 5/11 و 5/12 است. (منظور، آن عدد تکراری است که توزیع را به ۲ قسمت مساوی تقسیم می کند.)

۴- در مرحله بعد: حد پایین را با کسری جمع می کنیم که مخرج آن، تعداد اعداد تکراری است و صورت آن نصف تعداد کل اعداد، منهای تعداد اعدادی است که از عدد تکراری کوچک ترند و قبل از آن قرار گرفته اند: مثلاً در توزیع بالا: 4 = ۲ - ۶، عدد حاصل از این جمع، میانه است.

مثال 1: میانه در توزیع روبه رو کدام است؟

10,11,12,13,13,13,13,14,15,16,17,20

1) 5/12

2) 7/13

3) 25/12

4) 25/13

پاسخ: گزینه «4»

حد پایین → 5/12

حد بالا → 5/13

مثال 2: میانه در توزیع مقابل چقدر است؟

5 6 6 7 8 8 8 8 11 13 15

1) 25/7

2) 8

3) 78/7

4) 5/7

پاسخ: گزینه «۳» در این مثال، هم عدد ۸ تکراری است و هم ۶؛ اما چون میانه، نقطه ای است که توزیع را به ۲ قسمت مساوی تقسیم می کند، میانه بین ۸ اول و دوم قرار دارد و به ۶ کاری نداریم.

15 13 11 8 8 8 7 6 6 5

محاسبه میانه برای توزیع های طبقه بندی شده:

برای محاسبه میانه در توزیع های طبقه بندی شده از فرمول زیر استفاده می کنیم.

فاصله طبقات حد پایین طبقه ای که میانه در آن قرار دارد M =

یا به شکل خلاصه تر

برای این که بفهمیم میانه در کدام طبقه قرار گرفته است، باید ابتدا را حساب کنیم، سپس در ستون فراوانی تراکمی از پایین به سمت بالا حرکت می کنیم، اولین طبقه ای که فراوانی تراکمی آن را پوشش داده طبقه ای است که میانه در آن قرار دارد.

به عنوان مثال: میانه جدول ۲ به قرار مقابل است:

میانه در طبقه ۲۳ -۲۱ قرار دارد که حد واقعی این طبقه 5/23 و 5/20 حد پایین و 5/23 حد بالا نام دارد.

میانگین

میانگین تعدادی نمره عبارت است از مجموع نمرات تقسیم بر تعداد آنها:

مثال 3: میانگین نمرات مقابل چند است؟

15 13 12 9 8 7 6

1) 9

2) 10

3) 12

4) 11

پاسخ: گزینه «2»

اگر اعداد تکراری داشته باشیم، هر عدد را در فراوانی اش ضرب می کنیم، سپس بر تعداد تقسیم می کنیم.

زمانی که می خواهیم میانگین یک توزیع طبقه بندی شده را محاسبه کنیم، حد وسط هر طبقه را محاسبه و سپس در فراوانی ضرب می کنیم پس از آن، مجموع را بر تعداد کل تقسیم می کنیم.

مثال ۴: برای جدول ۲ میانگین را محاسبه می کنیم.

|  |  |  |  |
| --- | --- | --- | --- |
| فراوانی حد میانی | فراوانی | حد میانی | طبقات |
| 111= 3×37  136= 4×34  62= 2× 31  28= 1×28  125= 5×25  66= 3×22  114= 6×19  32= 2×16  52 =4×13 | 3  4  2  1  5  3  6  2  4 | 37  34  31  28  25  22  19  16  13 | 38-36  35-33  32-30  29-27  26-24  23-21  20-18  17-15  14-12 |
| 726ƩFX= | | | |

از طریق روش فرضی هم می توانیم میانگین را به همان دقت روش اصلی به دست آوریم؛ بدون این که به محاسبه اعداد بزرگ نیاز داشته باشیم.

فرمول محاسبه میانگین به روش فرضی:

x: حد میانه طبقه موردنظر، که به آن ارزش صفر داده ایم. اگر از طبقات نزدیک به وسط توزیع فراوانی انتخاب کنیم، کار محاسبه میانگین کمتر خواهد بود.

d: x تبدیل یافته N = تعداد کل i: فاصله طبقاتی

برای محاسبه میانگین از طریق روش فرضی، به انتخاب خودمان به یکی از طبقات، ارزش صفر می دهیم، از طبقات پایین صفر، یک واحد، کسر می کنیم و به طبقات بالای صفر ۱ واحد اضافه می کنیم؛ یعنی x را به d تبدیل می کنیم.

به عنوان نمونه: میانگین توزیع مثال قبل را به روش فرضی به دست می آوریم.

|  |  |  |  |
| --- | --- | --- | --- |
| fd | d | f | طبقات |
| 12  12  4  1  0  3-  12-  6-  16- | 4  3  2  1  0  1-  2-  3-  4- | 3  4  2  1  5  3  6  2  4 | 38-36  35-33  32-30  29-27  26-24  23-21  20-18  17-15  12-14 |
| 8- Ʃfd = | | | |

ویژگی های میانگین

1- میانگین، بهترین اندازه ی گرایش مرکزی است، اما مشکل عمده ی آن، این است که نسبت به نمرات خیلی بزرگ یا کوچک حساس است؛ یعنی نمرات خیلی زیاد یا خیلی کم می توانند به شکل چشم گیری، میانگین را بالا یا پایین ببرند. در چنین حالتی به جای میانگین، معمولا از میانه استفاده می شود.

۲- اگر نمره ثابتی را با تک تک، نمرات توزیع جمع کنیم یا از آن کم کنیم، همان عدد به میانگین هم اضافه یا کم می شود؛ مثلا اگر به همه نمرات توزیع، ۲نمره اضافه شود، با میانگین هم ۲ نمره اضافه می شود.

مثال ۵: بهترین اندازه ی گرایش مرکزی کدام است؟

۱) مد

۲) میانه

۳) میانگین

۴) انحراف معیار

پاسخ: گزینه «3» میانگین بهترین اندازه گرایش مرکزی است.

مثال ۶: مشکل عمده ی میانگین چیست؟

١) اگر توزیع نمرات یک توزیع زوج باشد، میانگین، عدد بالایی به دست می آید.

۲) نسبت به نمرات خیلی بزرگ یا کوچک حساس است.

۳) نمی تواند به تنهایی معرف گروه نمرات باشد.

۴) اگر عدد ثابتی بر آن ضرب یا تقسیم شود، تغییر محسوس در آن رخ نمی دهد.

پاسخ: گزینه «۲»

اندازه های تغییرپذیری

برای معرفی یک توزیع، هم به یکی از اندازه های گرایش مرکزی نیاز داریم و هم به یکی از اندازه های تغییرپذیری.

اندازه های تغییر پذیری عبارت اند از: دامنه تغییر، واریانس و انحراف معیار.

دامنه تغییر

دامنه تغییر ساده ترین اندازه تغییر پذیری است و برابر است با فاصله بین بزرگ ترین و کوچک ترین نمره ی گروه؛ مثلا دامنه ی تغییر در مثال ۲ برابر است با:

26= 12-38

برخی از کتاب های آماری، دامنه ی تغییرات با استفاده از فرمول زیر نوشته شده است:

کوچکترین عدد بزرگترین عدد

دامنه تغییرات، بی ثبات ترین شاخص تغییرپذیری است؛ زیرا در تعیین آن، تنها ۲ نمره دخالت دارند.

واریانس و انحراف معیار

وایانس و انحراف معیار، بهترین اندازه های تغییر پذیری اند.

فرمول محاسبه واریانس، ارتباط تنگاتنگی با میانگین دارد:

انحراف معیار، جذر واریانس است:

برای محاسبه انحراف معیار تمرات طبقه بندی شده، می توان از فرمول زیر استفاده کرد:

= SD انحراف معیار ,

= CI فاصله طبقاتی ,

= Ʃfd حاصل جمع نمرات انحراف از میانگین ,

= حاصل جمع مجذورات نمرات انحراف

که روش محاسبه آن شبیه به روش میانگین قرضی است، فقط ستون، را هم اضافه می کنیم و در فرمول قرار می دهیم.

هرچه واریانس و انحراف معیار بزرگتر باشند، یعنی پراکندگی نمرات، زیادتر است و بالعکس. به عبارت دیگر، ما از بزرگی و کوچکی مقدار واریانس و انحراف معیار به زیادی و کمی پراکندگی نمرات پی می بریم.

مثال ۷: بی ثبات ترین شاخص تغییرپذیری کدام است؟

1) واریانس

۲) انحراف معیار

۳) میانگین

۴) دامنه تغییر

پاسخ: گزینه «۴» بی ثبات ترین شاخص تغییر پذیری، دامنه تغییر می باشد.

همبستگی بین نمرات

در تعبیر و تفسیر نمرات، گاهی لازم است رابطه بین دو یا چند دسته نمره را تعیین کنیم. میزان رابطه بین ۲ دسته نمره را، که متعلق به یک گروه واحد هستند، از طریق محاسبه ضریب همبستگی مشخص می کنند. متداول ترین ضریب همبستگی، ضریب همبستگی پیرسون است.

ضریب همبستگی مقادیر از ۱+ تا ۱- را شامل می شود.

وقتی ضریب همبستگی 1+ است همبستگی مستقیم و کامل است؛ یعنی با افزایش یک متغیر، متغیر دیگر هم افزایش می یابد یا با کاهش یک متغیر، متغیر دیگر هم کاهش می یابد. وقتی ضریب همبستگی1- است همبستگی، معکوس و کامل است؛ یعنی با افزایش یک متغیر، متغیر دیگر کاهش می یابد یا با کاهش یکی، دیگری افزایش می یابد.

وقتی ضریب همبستگی صفر است یعنی بین ۲ متغیر، همبستگی وجود ندارد. وقتی ضریب همبستگی بین ۱- و0 قرار دارد یعنی همبستگی معکوس و غیر کامل است. وقتی ضریب همبستگی بین ۱+ و 0 قرار دارد یعنی همبستگی مستقیم و غیر کامل است.

انواع ضریب همبستگی

1- ضریب همبستگی پیرسون: ضریب همبستگی پیرسون، متداول ترین روش همبستگی است و زمانی استفاده می شود که هر دو متغیر از نوع فاصله ای یا نسبی باشند، مثلا وقتی می خواهیم رابطه ی بین، نمرات هوش و پیشرفت تحصیلی را تعیین کنیم.

فرمول محاسبه ضریب همبستگی پیرسون از راه انحراف از میانگین:

فرمول محاسبه ضریب همبستگی پیرسون با استفاده از اعداد خام:

تعداد X ها و Y ها N =

مجموع نمرات خام گروه ƩX=X

مجموع نمرات خام گروه ƩY=Y

مجموع حاصل ضرب های X و Y ƩXY=

مجموع مجذورات X ها

مثال ۸: برای محاسبه رابطه ی بین هوش کلامی و هوش غیرکلامی، استفاده از کدام یک از ضرایب همبستگی ارجح تر است؟

۱) فی

۲) دورشته ای

۳) گشتاوری پیرسون

۴) دورشته ای نقطه ای

پاسخ: گزینه «3» متداول ترین روش همبستگی، همبستگی گشتاوری پیرسون است. این ضریب، شاخص معتبری برای تعیین رابطه بین متغیرهاست و مفروضاتی به این شرح دارد:

1- در رابطه بین دو متغیر، خطی باشد. ۲- توزیع ها دارای شکل مشابه باشند. ۳- نمودار پراکندگی یکسان باشد. با توجه به جمله ی « محاسبه رابطه بین هوش کلامی و غیرکلامی»، اولویت انتخاب با همبستگی گشتاوری پیرسون می باشد.

۲- ضریب همبستگی اسپیرمن: ضریب همبستگی اسپیرمن زمانی مورد استفاده قرار می گیرد که هر دو متغیر از نوع رتبه ای باشند؛ مثلا وقتی افراد را از نظر طبقه اجتماعی به فقیر، متوسط و غنی تقسیم می کنیم و افراد را از نظر سطح درآمد به بالا، متوسط و پایین تقسیم می کنیم و می خواهیم رابطه بین ۲ متغیر طبقه اجتماعی و سطح درآمد را تعیین کنیم، از ضریب همبستگی اسپیرمن استفاده می کنیم.

تفاوت بین رتبه های هر جفت از نمرات d =

مثال۹: اگر مقدار برای نمرات ۴ نفر، برابر ۶ باشد، مقدار ضریب Prho چه قدر خواهد بود؟

1) 4/0

2) 5/0

3) 6/0

4) 7/0

پاسخ: گزینه «۱» در جایی که تعداد نمرات کم است و نمرات به صورت رتبه ای هستند، می توان از ضریب همبستگی رتبه ای اسپیرمن استفاده کرد. فرمول محاسبه ی آن به این ترتیب است:

مفروضات سؤال: ۴ = N

همبستگی دو رشته ای نقطه ای: زمانی استفاده می شود که یک متغیر پیوسته و متغییر دوم، دو مقوله ای یا دو ارزشی باشند

دو ارزشی یعنی فقط یکی از دو مقدار صفر یا یک را شامل شود؛ مثل مرد/ زن مقدار این ضریب همبستگی غالبا از ۱ کمتر است.

میانگین نمرات تعمیر پیوسته برای آزمون شوندگانی که در آزمون دو ارزشی، نمره ۱ گرفته اند.

انحراف معیار نمرات در متغیر پیوسته

میانگین سوالات آزمون شوندگان در متغییر پیوسته

نسبت کسانی که در متغییر دو ارزشی 0 گرفته اند.

نسبت کسانی که در متغییر دو ارزشی 1 گرفته اند. p =

مثال ۱۰: همبستگی دو رشته ای نقطه ای به کدام منظور به کار برده می شود؟

1) درجه دشواری

۲) تعیین همبستگی بین سؤال ها

۳) قدرت تمیز

۴) محاسبه ضریب همبستگی بین نمرات دانش آموزان

پاسخ: گزینه «3» ضریب تمیز به ضریب توافق سؤال با کل آزمون اشاره دارد. این ضریب می تواند افراد قوی و افراد ضعیف را که یک آزمون از هم تفکیک کند. گزینه ی ۲ می تواند ذهن داوطلب را دچار تردید کند، اما لازم به یادآوری است که برای تعیین ضریب همبستگی بین سؤال ها از ضریب دیگری به نام ضریب فی استفاده می شود.

مثال ۱۱: نوع ضریب پیوستگی بین نمره سؤال و نمره کل آزمون، کدام است؟

۱) فی

۲) پیرسون

۳) چهارخانه ای یا تتراکوریک

۴) دورشته ای نقطه ای

پاسخ: گزینه «۴» هنگام تحلیل آزمون، علاوه بر محاسبه ضریب همبستگی بین هر سؤال و کل آزمون، بهتر است ضریب همبستگی تک تک سؤال ها نیز محاسبه شود. در تعریف همبستگی دو رشته ای نقطه ای، توافق سؤال با کل آزمون است و هر چه ضریب همبستگی بیشتر باشد، یعنی سؤال با مجموعه سؤال ها تطابق کامل داشته؛ همخوان است و می تواند بین افراد قوی و ضعیف تمیز قایل شود. هر چه ضریب همبستگی کمتر باشد، مقدار همخوانی سؤال با مجموعه سؤال ها کمتر است. سؤالاتی که در این روش استفاده می شود، تنها یکی از دو مقدار یک یا صفر را شامل می شود؛ مانند بله/ نه، زن / مرد، قبول / غیرقبول. برای محاسبه نیز به یکی نمره ی صفر (۰) و به دیگری نمره ی (یک) داده می شود.

۴- همبستگی دورشته ای: زمانی استفاده می شود که یک متغیر، پیوسته و متغیر دیگر، دو ارزشی ساختگی باشد. دو ارزشی ساختگی یعنی متغیر در اصل پیوسته بوده است اما ما آن را به دو ارزشی تبدیل کرده ایم؛ مثلا نمرات 0 تا 20، که در مقیاس فاصله ای یا پیوسته قرار دارد، ما ۱۰ تا ۲۰ را قبول و از ۱۰ به پایین را مردود اعلام کنیم. این جا مقیاس فاصله ای را به دو ارزشی تبدیل کرده ایم.

آن مقدار از منحنی بهنجار دارای مساحت ۱ که با p مطابق است Z =

مقدار این همبستگی گاهی از ۱+ بیشتر و گاهی از ۱- کمتر می شود.

مثال ۱۲: مطالعه همبستگی بین جنسیت و نمرات پیشرفت تحصیلی ریاضی مستلزم کدام نوع همبستگی است؟ ۱) تتراکوریک

۲) دورشته ای

3) دورشته ای نقطه ای

۴) فی

پاسخ: گزینه «3» جنسیت اعم از زن و مرد، متغیر و ارزشی واقعی است. نمره پیشرفت تحصیلی، متغیر پیوسته است. بنابر تعریف متخصصان، زمانی که و یکی از دو متغیر مورد نظر، پیوسته است و متغیر دوم نیز دو ارزشی است، از روش همبستگی دورشته ای نقطه ای استفاده می کنیم. منظور از متغیر دو ارزشی در این همبستگی متغیری است که تنها یکی از دو مقدار را شامل می شود؛ مانند صفر و یک، بله و خیر، زن و مرد، مورد سؤال نیز به جنسیت زن و مرد اشاره دارد.

5- همبستگی فی ϕ : زمانی استفاده می شود که هر دو متغیر، دو ارزشی واقعی باشند، مثلا همبستگی بین جنسیت و قبول یا رد شدن در یک آزمون. بهترین مورد استفاده از این نوع ضریب همبستگی محاسبه همبستگی بین سوالی های یک آزمون است.

نسبت کسانی که در متغیر y نمره ۱ گرفته اند

نسبت کسانی که در متغیر x نمره 1 گرفته اند

نسبت کسانی که در ۲ متغیر نمره ۱ گرفته اند

مثال ۱۳: برای محاسبه همبستگی بین دو سؤال صحیح - غلط، کدام ضریب همبستگی مطلوب تر است؟

1) دو رشته ای

۲) پیرسون

۳) فی (فای)

۴) دو رشته ای نقطه ای

پاسخ: گزینه «3» زمانی که پرسش ها به صورت صحیح - غلط دسته بندی شده اند، از ضریب همبستگی فی (فای) برای تعیین همبستگی بین پرسش های آزمون استفاده میشود. چنانچه داوطلبان به مبحث همبستگی مسلط باشند، می دانند که از ضریب همبستگی در رشته ای نقطه ای برای سؤالات چهار گزینه ای استفاده می شود؛ همچنین همبستگی دورشته ای، برای موارد تشریحی کاربرد دارد، که نمره ی از صفر تا بیست را شامل می شود.

همبستگی چهارخانه ای با تتراکوریک: زمانی استفاده قرار می شود که هر ۲ متغیر دو ارزشی ساختگی باشند و زمانی به کار می رود که بتوان تو متغیر را در یک نقطه ی معین نصف کرد، که بهتر است ۲ متغیر را در حوالی میانه به ۲ بخش تقسیم کنیم.

با استفاده از روش های همبستگی دو رشته ای و دو رشته ای نقطه ای می توانیم برآورد کنیم که یک سؤال تا چه اندازه همان متغیری را می سنجد که سؤال های دیگر می سنجند، یعنی هم خوانی یک سؤال را با سؤال های دیگر آزمون مشخص می کنیم. اگر بخواهیم تعیین کنیم که سوال های مختلف، چه اندازه متغیر واحدی را نشان می دهند، باید جفت جفت روابط بین سوال های یک آزمون را محاسبه کنیم.در اینجا از همبستگی تتراکوریک استفاده می کنیم.

نکته ۴: ضرایب همیشگی پیرسون، اسپیرمن. دو رشته ای نقطه ای و فی از یک خانواده اند و تفاوت آن ها فقط در مقیاس اندازه گیری است، اما ضریب همبستگی دو رشته ای و تتراکوریک ازخانواده ضریب همبستگی پیرسون نیستند.

مثال ۱۴: کدام ضریب همبستگی نشان می دهد که نمره های کوچکتر از میانگین در یک متغیر با نمره های بزرگتر از میانگین در متغیرهای دیگر، همبسته شده اند؟

1) 73/0 -

2) 54/0

3) 64/0

4) 89/0

پاسخ: گزینه «۱» گزینه ی متن سؤال اشاره به همبستگی منفی دارد؛ یعنی همبستگی نمرات کوچکتر از میانگین در یک متغیر با نمره های بزرگ تر از میانگین در متغیری دیگر همبستگی منفی را نشان می دهد. یعنی افزایش یک نمره در یک متغیر به کاهش یک نمره در متغیر دیگر منجر می شود.

تفسیر ضریب همبستگی

1- رابطه بین ۲ متغیر ممکن است از جمعیتی به جمعیت دیگر فرق کند.

۲- هر چه افراد در متغیرها نامتجانس تر باشند، ضریب همبستگی، بزرگتر است.

٣- وجود همبستگی بین ۲ متغیر، ممکن است به خاطر یک متغیر سوم باشد.

۴- ضریب همبستگی بین ۲ متغیر را نمی توان به صورت درصد تفسیر کرد، اما اگر ضریب همبستگی را به توان ۲ برسانیم، یعنی مجذور ضریب همبستگی نشان دهنده ی درصد مشترک بین ۲ متغیر است؛ مثلا اگر بین ۲ آزمون پیشرفت تحصیلی، ضریب همبستگی ۸۰/0 وجود داشته باشد، می توان گفت ۶۴ درصد مطالب بین این ۲ آزمون، مشترک است.

۵- علامت مثبت و منفی در ضریب همبستگی جهت رابطه را نشان می دهد، به شدت رابطه؛ مثلا همبستگی 60/0 - از ضریب همبستگی 50/0+ بزرگتر است، فقط جهت آن معکوس است.

مثال ۱۵: متداول ترین روش ضریب همبستگی، روش همبستگی .................. می باشد و زمانی به کار برده می شود که هر دو متغیر از نوع فاصله ای یا نسبی باشند.

1) پیرسون

2) اسپیرمن

۳) دورشته ای

۴) دورشته ای نقطه ای

پاسخ: گزینه «۱» متداول ترین روش ضریب همبستگی، روش همبستگی پیرسون است و زمانی به کار برده می شود که هر دو متغیر از نوع فاصله ای یا نسبی باشند.

مثال ۱۶: بهترین ضریب مورد استفاده برای همبستگی بین سؤال های یک آزمون، کدام مورد می باشد؟

1) چهارخانه ای یا تتراکودیک

۲) فای

۳) دورشته ای

۴) دورشته ای نقطه ای

پاسخ: گزینه «۲» زمانی که پرسش ها به صورت صحیح - غلط دسته بندی شده اند، از ضریب همبستگی فی (فای) برای تعیین همبستگی بین پرسش های آزمون استفاده می شود.

تست های طبقه بندی شده فصل چهاردهم

۱- در گزارش نتایج یک آزمون، مقدار انحراف استاندارد « 5/1-» ذکر شده است. بنابراین می توان گفت:

(سراسری 82)

١) دامنه ی نمرات، خیلی زیاد بوده است.

۲) واریانس نمرات منفی بوده است. است.

۳) محاسبات انجام شده، غلط بوده است.

۴) جمع نمرات مثبت و منفی، صفر نبوده است.

۲- چنانچه همبستگی یک آزمون پیشرفت تحصیلی و یک آزمون هوشی مساوی 60/0 باشد، آزمون پیشرفت تحصیلی چه مقدار از تغییرات هوش را تبیین می کند؟

(آزاد ۸۳)

۱) ۶۰ درصد

2) 6/3 درصد

3) 36 درصد

4) 64 درصد

۳- برای محاسبه رابطه بین نمره های درس دستور زبان و درک زبان انگلیسی، که هر یکه توزیعی دو ارزشی دارند، کدام یک از ضرایب همبستگی ارجح است؟

(سراسری ۸۴)

1) ضریب دو رشته ای

۲) ضریب فی

۳) ضریب تتراکوریک

۴) ضریب دو رشته ای نقطه ای

۴- هنگامی که یافته های آماری، مبتنی بر نمونه کوچکی باشد:

(آزاد ۸۴)

1) معمولا از دقت بالایی برخوردار است.

۲) پیش بینی کننده مفیدی است

۳) سطح عملکرد را دقیق تر نشان میدهد.

۴) معمولا ناپایدار خواهد بود.

۵- به کدام دلیل، گاهی ضریب همبستگی کمتر از اندازه واقعی برآورد می شود؟

(سراسری 85)

۱) حجم جامعه

۲) توزیع نمونه گیری

۳) مقیاس اندازه گیری

۴) محدودیت در دامنه

6- به منظور ارزشیابی میزان همبستگی یک سوال آزمون با کل آزمون، کدام یک از روش های ضریب همبستگی مناسب تر است؟

(سراسری ۸۵)

1) فی

2) تتراکوریک

۳) دو رشته ای

۴) دو رشته ای نقطه ای

۷- در ارتباط با محاسبه همبستگی بین سؤالات یک آزمون، کدام یک از ضرایب همبستگی، مبنای نظری کاملا متفاوتی با ضرایب همبستگی گشتاوری پیرسون دارد؟

(سراسری ۸۵)

1) فی

۲) کندال

۳) دو رشته ای

۴) دو رشته ای نقطه ای

8- در صورتی که بخواهیم ارتباط بین یک سؤال را با نمره کل بسنجیم و توزیع متوسط و توزیع نرمال باشد، کدام روش همبستگی مناسب است؟

(آزاد ۸۵)

1) بای سریال

۲) بای سریال نقطه ای

۳) پیرسون

۴) اسپیرمن

۹- برای مطالعه همبستگی میان دو سؤال از یک تست روانشناختی، که ارائه پاسخ به آن ها به صورت (بله / خیر)است، بهترین روش مطالعه:

(سراسری ۸۶)

1) تتراکوریک

۲) دورشته ای

۳) دورشته ای نقطه ای

۴) فای دورشته ای

۱۰- هنگامی که تعداد داده ها کم باشد و بر حسب مقیاس رتبه ای، اندازه گیری شده باشند، کدام روش محاسبه ضریب همبستگی را توصیه می کنید؟

(آزاد ۸۶)

1) پیرسون

۲) اسپیرمن

3) فای

۴) دو رشته ای

۱۱- اگر همه دانش آموزان در آزمونی نمره یکسانی به دست آورند، واریانس نمره آزمون چقدر است؟

(سراسری 88)

۱) 0

2) 1

3) n

4) 1- n

۱۲- در کدام یک از روش های محاسبه ضریب همبستگی، هدف محقق محاسبه ضریب همبستگی بین دو متغیر است که اندازه یا مقیاس این دو متغیر از نوع مقیاس های ارزشی و کیفی است؟

(آزاد ۸۸)

۱) ضریب همبستگی اسپیرمن

۲) ضریب همبستگی بای سریال

۳) ضریب همبستگی پیرسون

۴) ضریب همبستگی فای

۱۳- برای بدست آوردن ضریب همبستگی اسپیرمن، اندازه متغیرها باید از چه نوع مقیاسی باشد؟

(آزاد 88)

1) نسبتی

2) فاصله ای

۳) رتبه ای

۴) ارزشی - اسمی

۱۴- دامنه تغییرات نمرات در یک آزمون عبارت است از ۹۰-۷۰ و در آزمون دیگری عبارتست از ۱۰۰- 0. درحالی که این دو آزمون بر یک گروه اجرا شده باشند:

(آزاد ۸۹)

۱) اصولاً نمی توان ضریب همبستگی بین دو آزمون را محاسبه کرد.

۲) این دو آزمون به صورت خطی با هم ارتباط ندارد.

۳) ممکن است بتوان ضریب همبستگی بین دو آزمون را محاسبه کرد.

۴) میانگین هر دو آزمون یکی نیست و احتمالاً رابطه ندارند.

۱۵- روش مناسب برای بررسی همبستگی میان دو سؤال زیر کدام است؟

الف) من رشته ام را دوست دارم. بلی □ خیر □

ب) دوست دارم تغییر رشته بدهم. بلی □ خیر □

(سراسری 90)

۱) دو رشته ای

2) تتراکوریک

۳) دو رشته ای نقطه ای

۴) فای دو رشته ای

۱۶- اگر نمره گذاری یک سؤال چهار گزینه ای به صورت (۳ ۰۱۲) باشد، میانگین سؤالات به طور متوسط چه مقدار باید باشد تا توزیع نمرات کل تست، نرمال گردد؟

(سراسری ۹۰)

1) 1

2) 5/1

3) 5/2

4) 2

۱۷- روش همبستگی مناسب برای بررسی رابطه میان سؤال «من به مطالعه علاقه دارم: بلی / خیر» و نمره کل مقیاس چیست؟

(سراسری ۹۱)

۱) دو رشته ای

۲) فای

۳) تتراکوریک

4) پیرسون

18ـ در یک توزیع از نمرات، کدام شاخص می تواند به عنوان واحد اندازه گیری مورد استفاده قرار گیرد؟

(سراسری 94)

1) (میانگین)

2) Md (میانه)

3) (واریانس)

4) S (انحراف استاندارد)

پاسخنامه تست های طبقه بندی شده فصل چهاردهم

1- گزینه «3» در تعاریف واریانس و انحراف داریم: برای از بین بردن علامت های منفی نمرات انحراف، آنها را مجذور می کنیم؛ یعنی مقیاس نمرات را به هم میزنیم، بنابراین برای برگرداندن عدد حاصل به مقیاس قبلی، یعنی همان مقیاس اعداد اصلی، لازم است که از واریانس، جذر بگیریم. جذر واریانس، انحراف معیار نام دارد. با توجه به این که می خواهیم علامت منفی را با مجذور از بین ببریم، گزینه ۳ صحیح می باشد. گزینه ۲ به این موضوع اشاره دارد که واریانس نمرات منفی بوده است.

۲- گزینه «3» ضریب همبستگی را نباید به صورت درصد تعبیر کرد، اما مجذور همبستگی (همبستگی به توان2) نشان دهنده ی درصد وجه مشترک بین دو متغیر است؛ بنابراین در این سؤال، که ضریب همبستگی بین دو آزمون ۶۰/0 است، می توان گفت که ۳۶ درصد ( 36/0 = 60/0 × 60/0 ) آزمون پیشرفت تحصیلی تغییرات هوش را تبیین می کند.

۳- گزینه «۲» زمانی که هر دو متغیر موردنظر، متغیرهای دوارزشی واقعی هستند، از ضریب همبستگی فی (فای) استفاده می شود؛ به طور مثال از کاربردهای همبستگی فی می توان تعیین همبستگی بین پرسش های یک آزمون را نام برد که به صورت صحیح و غلط دسته بندی شده اند.

۴- گزینه «۴» نتیجه ای که از یک نمونه ی کوچک به دست می آید می تواند با افزودن نمراتی که از پراکندگی بالایی برخوردارند، تغییر کند؛ بنابراین، پایداری زیادی ندارد، درصورتی که اگر نمونه بزرگتر باشد، پراکندگی آن ها کمتر است و حساسیت کمتری نسبت به نمره های کوچک و بزرگ جدید دارفور و ناپایداری کمتری دارد. هرچه حجم نمونه کمتر باشد، آزمون موردنظر نمی تواند پارامتر جامعه را برآورد کند؛ چون نمونه، معرف جامعه نیست.

5- گزینه «۴» مهم ترین عاملی که ضریب همبستگی را به شدت کاهش می دهد، محدودیت در دامنه یک متغیر است.

6-گزینه «۴» به منظور تعیین میزان همبستگی یک سؤال آزمون با کل آزمون، از همبستگی دو رشته ای نقطه ای استفاده میشود.

7- گزینه «۲» هنگامی که با بیش از دو مجموعه رتبه، روبه رو می شویم و مایلیم بدانیم بین رتبه هایی که توسط m داور به n فرد یا شیء داده شده است، تا چه حد توافق وجو دارد، می توان از ضریب توافق کندال استفاده کرد. مبنای نظری سایر گزینه ها براساس مقیاس فاصله ای است ولی کندال براساس مقیاس رتبه ای است.

8- گزینه «۲» همبستگی بای سریال نقطه ای (دو رشته ای نقطه ای) را زمانی به کار می بریم که متغیر اول، پیوسته و متغیر دوم، دو ارزشی واقعی باشد؛ از آنجا که نمره ی کل آزمون، یک متغیر پیوسته و نمره ی یک ماده از یک آزمون، یک متغیر دو ارزشی است، می توان از این ضریب همبستگی استفاده کرد.

9- گزینه «1» وقتی دو متغیر، دو ارزشی واقعی هستند، از ضریب همبستگی فای استفاده می کنیم. سؤالات بلی - خیر هم دو ارزشی واقعی هستند.

۱۰- گزینه «۲» ضریب همبستگی پیرسون را زمانی به کار می بریم که متغیرهای هر دو، پیوسته باشند. ضریب همبستگی فای را نیز زمانی به کار می بریم که هر دو متغیر، دو ارزشی واقعی باشند که بهترین مورد استفاده از آن، محاسبه همبستگی بین سؤال های یک آزمون می باشد. همبستگی دو رشته ای نقطه ای را زمانی به کار می بریم که با یک متغیر پیوسته و یک متغیر دو ارزشی واقعی سروکار داشته باشیم؛ مانند همبستگی بین یک سؤال با کل آزمون. اما ضریب همبستگی اسپیرمن را زمانی به کار می بریم که تعداد نمرات، کم و بر اساس مقیاس رتبه ای اندازه گیری شده باشند.

۱۱- گزینه «1» واریانس، یک شاخص پراکندگی است که از طریق محاسبه ی انحراف نمره ها از میانگین، محاسبه می شود؛ اگر همه دانش آموزان در آزمون نمره یکسانی کسب نمایند. میانگین نیز، همان نمره میشود؛ بنابراین، انحرافی از میانگین وجود ندارد، پس واریانس نمرات صفر می شود.

۱۲- گزینه «۴» چون متغیرها در ضریب همبستگی اسپیرمن، رتبه ای هستند نه کیفی (اسمی). در ضریب همبستگی دورشته ای (بای سریال) نیز یکی از متغیرها اسمی و دیگری پیوسته است. در ضریب پیرسون، هر دو متغیر، پیوسته هستند؛ فقط در ضریب همبستگی فای است که هر دو متغیر، کیفی (اسمی) هستند.

۱۳- گزینه «3» زمانی که تعداد نمرات، کم است و نمرات به صورت رتبه ای هستند، یا می توان آن ها را به صورت رتبه ای تنظیم کرده می توان ضریب همبستگی رتبه ای اسپیرمن را برای آن ها به کار برد.

۱۴- گزینه «۴» چون نمرات آزمون اول، نمرات متمرکزی هستند و نمرات خیلی با هم تفاوت ندارند، اما نمرات آزمون دوم، با یکدیگر تفاوت دارند و دامنه تغییراتشان بالاست؛ پس میانگین در آزمون، یکی نیست و احتمالا رابطه ای با هم ندارند.

۱۵- گزینه «۴» چون هر دو سؤال الف وب، دو ارزشی واقعی اند. بهترین مورد استفاده از ضریب همبستگی فی (فای) نیز زمانی است که هر دو متغیرها دو ارزشی واقعی باشند؛ بنابراین، مناسب ترین روش برای بررسی همبستگی میان این دو سؤال، ضریب همبستگی فای است.

۱۶- گزینه «۲» میانگین سؤالات به طور متوسط باید باشد تا توزیع نمرات کل تست، نرمال شود.

۱۷- گزینه «1» زمانی که یکی از دو متغیر موردنظر، پیوسته است، اما متغیر دوم، یک متغیر دو ارزشی غیر حقیقی یا ساختگی است، یعنی نمرات به دو مقوله تقسیم شده اند، از روش همبستگی دو رشته ای استفاده می شود.

۱۸- گزینه «۴» به طور معمول، انحراف استاندارد (S) می تواند واحد اندازه گیری مناسبی برای بررسی توزیع نمرات باشد. انحراف استاندارد است که نشان می دهد به طور میانگین، داده ها چه مقدار از مقدار متوسط فاصله دارند. برای محاسبه انحراف معیار، کافی است ریشه دوم (جذر) واریانس را حساب کنیم.

آزمون فصل چهاردهم

1- کدام یک از روش های همبستگی زیر جزء خانواده روش همبستگی پیرسون نیست؟

1) اسپیرمن

۲) دو رشته ای نقطه ای

۳) دو رشته ای

۴) فی

۲- برای محاسبه ضریب همبستگی بین نمرات یک سؤال صحیح - غلط و نمرات کل آزمون، کدام روش تعیین همبستگی بهترین است؟

1) رتبه ای اسپیرمن

۲) پیرسون

۳) دورشته ای

۴) دورشته ای نقطه ای

3- اگر هر ۲ متغیر از نوع دو ارزشی واقعی باشند، کدام روش همبستگی مناسب تر است؟

1) تتراکوریک

۲) فی

۳) دورشته ای

۴) اسپیرمن

۴- برای محاسبه همبستگی بین سؤالات یک آزمون، کدام روش همبستگی بهتر است؟

1) دورشته ای

۲) پیرسون

۳) فی

۴) دورشته ای نقطه ای

5- کدام یک از ارقام زیر معرف ضریب همبستگی بزرگتری است؟

1) 70/0+

2) 55/0+

3) 75/0-

4) 80/0-

6- اگر بخواهیم رابطه بین جنسیت و انتخاب رشته علوم تجربی را تعیین کنیم، کدام ضریب همبستگی مناسب تر است؟

1) فی

۲) تتراکوریک

۳) دورشته ای

۴) پیرسون

۷- هرگاه ضریب همبستگی پیرسون بین صفر و ۱- باشد، همبستگی چگونه است؟

1) مستقیم کامل

۲) معکوس کامل

۳) مستقیم ناقص

۴) معکوس ناقص

۸- اگر هر ۲ متغیر از نوع دو ارزشی ساختگی باشند، از کدام ضریب همبستگی استفاده می کنیم؟

1) دو رشته ای نقطه ای

۲) فی

۳) تتراکوریک

۴) اسپیرمن

۹- وقتی یک متغیر، پیوسته و متغیر دیگر دو ارزشی واقعی باشد، از کدام ضریب همبستگی استفاده می کنیم؟ ١) دو رشته ای نقطه ای

۲) پیرسون

3) فی

4) تتراکوریک

۱۰- کدام یک از موارد زیر جزء خواص میانگین است؟

1) میانگین، تحت تأثیر اعداد خیلی بزرگ و خیلی کوچک قرار ندارد.

۲) اگر عددی بر همه اعداد بیافزایم، همین عدد بر میانگین آن ها نیز افزوده می شود.

۳) میانگین، نقطه ۲۵ درصدی است.

۴) گزینه های او۲

فصل پانزدهم «هنجارها و نیمرخ ها»

هنجارها و نیم رخ ها روش هایی هستند که به ما کمک می کنند سطح موفقیت دانش آموزان را در درس های مختلف با یکدیگر مقایسه کنیم و میزان توفیق آنان را نسبت به یکدیگر بسنجیم.

نکته 1: مقایسه نمرات دانش آموران با یکدیگر مخصوص آزمون های وابسته به هنجار است، با آزمون های وابسته به ملاک، عملکرد دانش آموران با یک ملاک از پیش تعیین شده مقایسه می شود.

مثال ۱: هنگامی که ما یک آزمون را برگزار می کنیم و هدفمان این است که آزمودنی ها را با هم مقایسه کنیم، استفاده از کدام نوع آزمون بهتر است؟

1) وابسته به هنجار

۲) وابسته به ملاک

۳) گزینه های ۱ و ۲

۴) هیچ کدام

پاسخ: گزینه «۱» از آزمون های وابسته به هنجار برای مقایسه آزمودنی ها استفاده میشود.

تعریف هنجار

نرم یا هنجار عبارت است از متوسط عملکرد گروه نمونه ای از آزمودنی ها، که به روش تصادفی از یک جامعه ی تعریف شده، انتخاب می شوند. اگر برای تفسیر نمره یک دانش آموز، نمره او را با میانه یا میانگین نمره های یک گروه مرجع مقایسه کنیم، به آن گروه مرجع، گروه هنجار می گویند.

گروه مرجع شبیه به دانش آموزان موردنظر هستند. فایده هنجار این است که برای قضاوت درباره سطح نسبی عملکرد یک فرد به عنوان معیار به کار می رود. هنجاریابی یکی از مراحل اصلی در استاندارد کردن آزمون های وابسته به هنجار است.

مثال ۲: یکی از مراحل اصلی استاندارد کردن آزمون های وابسته به هنجار ................... می باشد.

۱) طرح سؤالات

۲) تدوین فرضیه

۳) گزینش آزمودنی

۴) هنجاریابی

پاسخ: گزینه «۴» یکی از مراحل اصلی استاندارد کردن آزمون های وابسته به هنجار، هنجاریابی است.

مثال 3: کدام جمله، هنجار (نرم) را بهتر توصیف می کند؟

۱) حد متوسط نمره افراد یک گروه

۲) هدف آموزشی برای دانش آموزان متوسط

3) شاخصی که برای وصول به آن، همه افراد باید کوشش کنند.

۴) شاخصی برای رد یا قبول کردن دانش آموزان

پاسخ: گزینه «۱» اگر برای تفسیر نمره ی یک دانش آموز، آن نمره را با میانگین نمره ی یک گروه مرجع مقایسه کنیم، به آن گروه مرجع، گروه هنجار می گوییم. گروه مرجع از کسانی تشکیل شده است است که به گونه ای، شبیه به دانش آموز موردنظر عمل می کنند.

مثال ۴: هدف تبدیل نمرات خام آزمون ها به شاخص های نسبی چیست؟

۱) جدا سازی گروه ضعیف

۲) ارزیابی دقیق تر نمره آزمون

۳) ارزش دادن به نمره خام آزمون

۴) تعیین جایگاه فرد در مقایسه با نمونه هنجاریابی

پاسخ: گزینه «۴» به طور کلی، آزمون با هدف مقایسه یک فرد با فرد دیگر یا گروه ها گروه دیگر انجام می شود و برای تشخیص میزان ارتباط، نمرات خام به نمرات استاندارد تبدیل می شوند. با توجه به گزینه ۲، در هنگام انتخاب گزینه های صحیح باید دقت زیادی شود؛ زیرا تعیین جایگاه فرد در مقایسه با نمونه با توجه به ارزیابی دقیق نمره آزمون انجام می شود. در واقع ابزار اجرای کار، گزینه ۲ است ولی هدف طرح سؤال گزینه ۴ است.

پس از انتخاب گروه هنجار، نمرات آن گروه را در آزمون موردنظر تعیین و یکی از شاخص های گرایش مرکزی را برای آن ها حساب می کنیم. این شاخص، که معمولا میانه نمرات است، هنجار موردنظر است.

نمرات هنجار، که به آن ها نمرات مشتق هم می گویند، عملکرد بهنجار یا نوعی را گزارش می کنند، نه عملکرد بسیار خوب و مطلوب را. منظور این است که نمراتی که به دست آمده است، حاصل اجرای آزمون روی تعداد خاصی از افرادی است که آزمون برای آن ها میزان شده است؛ بنابراین، قرار نیست که تمام افراد دیگر هم این نتایج و نمرات را کسب کنند.

هنجارها در پاسخ دادن به سؤال های زیر به ما کمک می کنند:

1- عملکرد یک دانش آموز در یک آزمون در مقایسه با عملکرد یک دانش آموز دیگر چگونه است؟

۲- عملکرد یک دانش آموز در یک آزمون در مقایسه با عملکرد او در یک آزمون دیگر چگونه است؟

۳- عملکرد یک دانش آموز در یک فرم آزمون در مقایسه با عملکرد او در فرم دیگر آزمون، که قبلا اجرا شده است، چگونه است؟

فایده نمرات هنجار یا مشتق از نظر گرانلاند و لین:

١- پیش بینی موفقیت احتمالی دانش آموز در زمینه های مختلف

٢- تشخیص نقاط ضعف و قوت عملکرد دانش آموز

۳- اندازه گیری رشد آموزشی

انواع هنجارها

نمرات هنجاره انواع مختلفی دارند که عبارتند از: هنجارهای سنی یا معادل های سنی، هنجارهای کلاسی یا معادل های کلاسی، هنجار های درصدی و نمرات معیار. در هنجار سنی و هنجارهای کلاسی، عملکرد یک دانش آموز با عملکرد یک گروه مرجع مقایسه می شود، اما در هنجارهای درصدی و نمرات معیاری موقعیت نسبی یک دانش آموز هم گروهی که خود عضوی از آن است، نشان داده می شود.

مثال ۵: در کدام یک از انواع هنجارها، موقعیت نسبی یک دانش آموز در گروهی که خود، عضو آن است، نشان داده می شود؟

1) هنجار سنی

۲) هنجار کلاس

۳) معادل های کلاسی

۴) هنجار نمرات معیار

پاسخ: گزینه «۴» در هنجار نمرات معیار، موقعیت نسبی یک دانش آموز در گروهی که خود، عضو آن است، نشان داده می شود.

هنجار سنی

هنجار سنتی یا معادل سنی، نمره میانی یک آزمون است که افراد گروه سنی معینی، نمره ای برابر آن گرفته اند؛ مثلا اگر از دانش آموزانی که سن آن ها ۹سال و ۴ ماه است، یک آزمون گرفته شود و میانه یا میانگین نمرات آن ها مثلا عدد ۱۷ بشود، به عدد ۱۷، معادل سنی ۴-۹ تعلق می گیرد. البته برای تهیه اسم یک هنجار سنی برای یک ویژگی خاص، باید نمونه کاملی از افراد آن سن را بررسی کنیم. بسیاری از آزمون های میزان شده ی پیشرفت تحصیلی دارای هنجارهای سنی هستند که در این آزمون ها، معادله ای سنی با عنوان موضوع های درسی نامگذاری می شوند؛ مثلا سن خواندن، سن حساب و ..... اگر دانش آموزی سن خواندن ۶-۱۲ و سن حساب ۰-۱۲ داشته باشد، نمره او در آزمون خواندن برابر است با نمره متوسط کودک ۱۲ سال و ۶ ماه و نمره او

در آزمون حساب، معادل نمره کودکی است که در آغاز ۱۲ سالگی است.

در یکی از محدودیت های هنجار سنی این است که عملکرد آزمون در آن در قالب واحدهای نامساوی معرفی می شود؛ چون تغییرات الگوهای رشد از یک اور توانایی به توانایی دیگر، سبب می شوند که واحدهای سنی، معنی یکسانی نداشته باشند. ممکن است یک توانایی خاص در یک سال رشد، پیشرفت بسیار چشمگیرتر یا ناچیزتری نسبت به سال دیگر داشته باشد.

بهترین مرحله کاربرد هنجار سنی در سطح دبستان است؛ زیرا در این دوران، رشد ذهنی و تحصیلی، بیش از سال های دیگر یکنواخت و منظم است.

مثال ۶: بهترین سطح کاربرد هنجاری سنی کدام سطح می باشد؟

۱) دبستان

۲) راهنمایی

۳) دبیرستان

۴) دانشگاه

پاسخ: گزینه «۱» بهترین کاربرد هنجار سنی، سطح «دبستان» می باشد.

هنجار کلاسی

نرم کلاسی، نمره میانی یک آزمون است که دانش آموزان پایه تحصیلی معین، نمره ای برابر آن گرفته اند. هنجارهای کلاسی به هنجارهای سنی شبیه هستند؛ با این تفاوت که در هنجارهای سنی، گروه های سنی، گروه های مرجع اند، اما در هنجارهای کلاسی، گروه های کلاسی، گروه مرجع اند؛ برای مثال اگر از دانش آموزان پایه پنجم آزمونی بگیریم و بعد میانه این آزمون را حساب کنیم و حاصل آن عدد ۱۹ شود، بـه ایـن نـمـره خـام، معـادل کلاسی 0/5 داده می شود. معادل های کلاسی را به صورت دورقمی نشان می دهند: رقم اول نشان دهنده سال و رقم دوم نشان دهنده ماه است. رقـم دوم معادل های کلاسی، فقط تا ۹ را شامل می شود؛ زیرا سال تحصیلی ۹ ماه است؛ مثلاً معادل کلاسی 4/7 نشان دهنده پیشرفت در سطح ماه چهارم از کلاس هفتم است و معادل ای است. 0/4 نشان دهنده پیشرفت در آغاز کلاس چهارم است. هنجارهای کلاسی هم، ، مانند هنجارهای سنی، عملکرد متوسط دانش آموزان را منعکس کنند؛ یعنی ۵۰ درصد دانش آموزان گروه مرجع در بالای معادل کلاسی و ۵۰ درصد دیگر در پایین آن قرار دارند. معـادل کلاسی را نبایـد نـمـره ای دانست که همه دانش آموزان یک کلاس خاص باید آن را بگیرند.

به دلیل این که پیشرفت در موضوعات درسی آموزشگاه، به طور نامساوی صورت می گیرد، هنجارهای کلاسی هم، واحدهای نامساوی به دست می دهند. هنجارهای کلاسی هم برای سطح دبستان و هم درس هایی که تأکید آموزشی یکسانی بر آن ها می شود، مثل حساب و خواندن، بیشتر استفاده قرار می شود.

مثال ۷: کاربرد هنجار کلاسی در چه مقطعی بهترین نتیجه را به دست می دهد؟

۱) راهنمایی

۲) دانشگاه

۳) دبستان

۴) دبیرستان

پاسخ: گزینه «۳» هنجار کلاسی مربوط به سطوح اولیـه سـال هـای تحصیلی است که می تواند در میان گروه سنی پایین رقابت ایجـاد کـنـد. بنابراین نمی تواند برای سطوح بالاتر تحصیلی مناسب باشد.

استفاده از هنجار سنی در سطح دبستان، بهترین مرحله ی کاربرد این نوع هنجار است؛ زیرا در این سال ها، رشد ذهنی و تحصیلی، بیش از سال های دیگر یکنواخت و منظم است.

هنجار سنی نمایی

هنجار سنی نمایی، همان هنجار کلاسی است؛ با این تفاوت که چون دانش آموزان یک دامنه معین، سنی بزرگی را تشکیل می دهنـد، نمـره دانش آموزانی را که از لحاظ سنی، دانش آموزان خیلی بزرگ یا خیلی کوچک محسوب هستند، کنار می گذاریم و نمره هنجار را فقط برای دانش آموزان ســن موردنظر برای آن کلاس محاسبه می کنیم. این هنجار سنی محدودشده را، هنجار سنی نمایی می گویند.

هنجار سن ذهنی

هنجار سن ذهنی یا سن عقلی، که آلفرد بینه آن را اولین بار معرفی کرد، نوعی هنجار سنی است که در آزمون های هوش به کار می رود، روش محاسبه آن هم شبیه به محاسبه معادل های سنی است. اگر نمره خام یک دانش آموز در یک آزمون هوش، برابر با میانه نمرات خام کودکان نه ساله باشد، سن ذهنی او ۹ خواهد بود.

مثال ۸: ابداع کننده هنجار سن ذهنی، کدام گزینه زیر می باشد؟

۱) گوتگ

۲) آلفرد بینه

3) آزوبل

4) جان اسمیت

پاسخ: گزینه «۲» آلفرد بینه ابداع کننده هنجار سن ذهنی است.

هوشبهر

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

به عنوان مثال: اگر سن عقلی دانش آموزی ۹ و سن تقویمی او ۱۱ باشد، هوشبهر او ۸۱ است.

مثال 9: سن عقلی دانش آموزی ۱۲ و سن تقویمی او ۱۰ می باشد. هوشبهر او چند می باشد؟

1) 110

2) 90

3) 120

4)140

پاسخ: گزینه «۳»

مثال ۱۰: چنانچه سن عقلی کودک ۸ ساله ای به اندازه ی کودک ۱۰ ساله باشد، بهره هوشی او کدام یک از اعداد زیر خواهد بود؟

1) 130

2) 125

3) 120

4) 80

پاسخ: گزینه «۲» باتوجه به مفروضات سؤال، اعداد را در فرمول هوشبهر قرار می دهیم:

10 = سن عقلی

8 = سن تقویمی

هنجار درصدی

در هنجار کلاسی یا سنی، فرد با گروه سنی با کلاسی خودش مقایسه می شود، اما در هنجارهای درصدی از رتبه های درصدی یا صدکها استفاده می شود. رتبه درصدی، وضعیت نسبی فرد در گروه را برحسب کسانی که نمره پایین تر از او گرفته اند، مشخص می کند. هر نمره خام دارای یک رتبه درصدی است؛ مثلا نمره خام ۱۷ دارای رتبه درصدی ۸۵ است؛ به شرط این که نمره خام ۸۵ درصد دانش آموزان گروه، کمتر از ۱۷ باشد، یعنی نمره این دانش آموز از ۸۵ درصد گروه هنجار، بیشتر است.

رتبه درصدی و صدک، دو اصطلاح مکمل یکدیگرند، مثلا اگر نمره خام دانش آموزی ۱۷ باشد و رتبه درصدی او ۸۵ باشد، نمره او در صدک هشتاد و پنجم قراردارد؛ در واقع، رتبه درصدی، عملکرد نسبی آزمون شوندگان را نشان می دهند. صدکها به نقاطی روی توزیع نمرات اشاره می کنند. صدک برخلاف رتبه بندی است و شمارش آن از پایین به بالا است؛ یعنی در رتبه بندی، به بهترین فرد، رتبه ۱ داده می شود، اما در صدک برعکس است.

مثال ۱۱: نمره خام دانش آموزی ۱۹ می باشد و رتبه درصدی او ۶۵ می باشد. نمره او در کدام صدک قرار دارد؟

1) 75

2) 90

3) 65

4) 95

پاسخ: گزینه «۳» رتبه درصدی ۶۵ در صدک ۶۵ قرار دارد.

صدک یا نقطه درصدی را با علامت p نشان می دهند. صدک ۵۰ یا 50P همان میانه است که عملکرد متوسط را نشان می دهد. صدکهای کوچکتر از۵۰ نشان دهنده عملکرد پایین تر اگر متوسط و صدکهای بزرگتر از ۵۰ معرف عملکرد بالاتر از متوسط هستند. صدک ۲۵، نقطه چارکی اول و صدک ۷۵، نقطه چارکی سوم هستند.

مثال ۱۲: در کدام یک از انواع نرم ها یا هنجارهای زیر، فرد با گروهی که منطقة عضو آن استه، مقایسه می شود؟

1) سنی

2) درصدی

۳) کلاسی

۴) استاندارد

پاسخ: گزینه «۲» در هنجارهای سنی و کلاسی، گزینه های 1و3، از راه تعیین یک گروه سنی یا کلاسی، که نمره فرد معادل نمرات متوسط آن است به نمره فرد معنی داده می شود؛ اما در هنجار درصدی، فرد با گروه سنی یا کلاسی خود، یعنی گروهی که منطقأ می توان او را عضوی از آن دانست، مقایسه می شود. رتبه یا نمره درصدی، وضعیت نسبی فرد در گروه را بر حسب کسانی که نمره پایین تر از او گرفته اند، مشخص می کند. هر نمره خام دارای یک رتبه درصدی است.

مثال ۱۳: نمره خام دانش آموزی ۱۷ و رتبه درصدی اش ۷۵ می باشد. یعنی اینکه:

۱) نمره او از نیمی از نمره ها بالاتر است.

۲) نمره درصدی اش بالای ۹۵ میباشد.

۳) نمره او از ۷۵ درصد دانش آموزان بالاتر است.

۴) نمره او معادل ۷۵ درصد دانش آموزان می باشد.

پاسخ: گزینه «3» از ۷۵ درصد دانش آموزان بالاتر است و از ۲۵ درصد دانش آموزان پایین تر است بر روش محاسبه صدک ها شبیه به روش محاسبه میانه است و از طریق فرمول زیر محاسبه می شود:

صدک موردنظر یا نقطه درصدی = P

فراوانی طبقه ای که PN در آن قرار دارد. fw =

حد پایین طبقه ای که P×N در آن قرار دارد. L =

فراوانی تراکمی طبقه ماقبل که PN در آن قرار دارد. cf =

فاصله طبقاتی i =

مثال ۱۴: مفهوم عبارت «دانش آموز در آزمون پیشرفت تحصیلی، رتبه ۶۸ درصدی را کسب کرده است.» چیست؟

۱) از ۶۸ درصد افراد شرکت کننده بیشتر گرفته است.

۲) به ۶۸ درصد سؤال های آزمون پاسخ صحیح داده است.

۳) نمره فرد، دو انحراف معیار از میانگین بالاتر است.

۴) نمره فرد، یک انحراف معیار از میانگین بالاتر است.

پاسخ: گزینه «۱» رتبه ی درصدی یا نمره ی درصدی، وضعیت نسبی فرد در گروه را بر حسب کسانی که نمره ی پایین تر از او گرفته اند، مشخص می کند. با توجه به متن سؤال، گزینه ی اول با این تعریف هماهنگی دارد.

مثال ۱۵: اگر گفته شود رتبه درصد فراگیری ۱۷% است، چنین استنباط می شود که این فراگیر:

1) از ۸۳ نفر بالاتر است.

۲) به ۱۷% سؤالات، پاسخ صحیح داده است.

۳) در بین همکلاسان خود نفر ۱۷ است.

4) از ۱۷% افراد گروه خود بالاتر است.

پاسخ: گزینه «۴» رتبه درصدی یا نمره ی درصدی، وضعیت نسبی فرد را در گروه، برحسب کسانی که نمره ی پایین تری از او گرفته اند، مشخص می کند. هر نمره ی خام، یک رتبه ی درصدی است. درباره ی این سؤال، رتبه درصدی ۱۷ نشان دهنده ی این است که او از ۱۷ درصد افراد گروه بهتر و از ۸۳ درصد افراد گروه خود پایین تر و بدتر عمل کرده است.

به عنوان مثال، نقطه ۷۵ درصدی را برای جدول ۲-۱۴ محاسبه می کنیم.

یعنی نمره ی خام 75/31 معادل نقطه ۷۵ درصدی است:

رتبه درصدی را با علامت PR نشان می دهیم. در اینجا نمره خام به ما می دهند و ما رتبه درصدی متناسب با آن را پیدا می کنیم که از فرمول زیر به دست می آید:

فراوانی طبقه ای که نمره خام به آن تعلق دارد. = fw

فراوانی تراکمی طبقه ماقبلی که نمره خام به آن تعلق دارد. = Cf

تعداد = N

اگر در مثال قبل بخواهیم مشخص کنیم، کسی که نمره خام ۲۱ دارد، رتبه درصدی اش چند است، به شکل مقابل، عمل می کنیم:

بنابراین، کسی که نمره خام او ۲۱ است، رتبه درصدی اش ۴۵ است؛ یعنی نمره او از ۴۵ درصد دانش آموزان بالاتر است.

یکی از معایب هنجارهای درصدی، مانند هنجارهای سنی و کلاسی این است که واحدهای مربوط به صدک ها، در طول مقیاس مساوی نیستند.

مثال ۱۶: یکی از خصوصیات هنجار(نرم) درصدی این است که:

۱) در تمام طول مقیاس، ثابت است.

۲) با یکدیگر قابل مقایسه است.

۳) استاندارد است.

۴) در وسط، فشرده و در طرفین، گسترده است.

پاسخ: گزینه «۴» از خصوصیات هنجار درصدی این است که در وسط، فشرده و در طرفین، گسترده است.

مثال ۱۷: یکی از محدودیت های مهم « نرم درصدی»، کدام است؟

1) دشواری تفسیر آن

۲) دشواری محاسبه و فهم آن

3) عدم انطباق آن با منحنی نرمال

۴) نابرابری واحدها در طول مقیاس

پاسخ گزینه «۴» در هنجارهای سنتی و کلاسی، از راه تعیین یک گروه سنی یا کلاسی، که نمره ی فرد، معادل نمرات متوسط آن است، به نمره او سر معنی داده می شود. در هنجارهای درصدی، فرد با گروه سنی یا کلاس خودش، یعنی گروهی که منطقأ می توان او را عضوی از آن دانسته مقایسه می شود. در یکی از معایب هنجار درصدی، این است که واحدهای مربوط به صدک ها، در طول مقیاس، مساوی نیستند. تفاوت صدک های اواسط مقیاس (مثلا صدک های ۴۵ تا ۵۵) خیلی کمتر از تفاوت صدک های دو انتهای مقیاس (مثلا ۸۵ تا ۹۵) است. علت آن مشخص است، نمراتی که تعداد زیادی از افرام گروه کسب اس می کنند، در حدود وسط مقیاس قرار دارند؛ در حالی که نمره های تعداد نسبتا کمی از افراد خیلی عالی یا خیلی ضعیف است؛ زمانی که یک افزایش کلی بر تمام نمرات لحاظ می شود، کسی که نمره وسط یا کمتر از میانگین دارد، به صورت محسوسی در منحنی جابه جا می شود، اما کسی که نمره به فرض 20 گرفته است، اصلا از این تغییر سودی نمی کند.

مثال ۱۸: نمرات یک آزمون، پس از استخراج نمودار آن، کشیدگی منفی پیدا کرده است. کدام یک از شاخص های هنجاری زیر می تواند تا حدودی شکل نمودار فوق را شبیه به توزیع طبیعی کند؟

1) نمرات تی

2) نمرات زد

3) نمرات صدکی

۴) هوشبهر انحرافی

پاسخ: گزینه «3» زمانی که طبق متن سؤال، نمودار دارای کشیدگی منفی می شود، فقط نمرات صدکی (درصدی) می توانند، شکل نمودار را تا حدودی به شکل طبیعی بازگردانند. لازم به ذکر است که سایر گزینه ها به نمره های T, Z و IQ اشاره می کنند که تنها در زمان نرمال بودن منحنی کاربرد دارند و درصورت کجی منفی، غیرنرمال است.

هنجارهای مربوط به نمرات معیار

گفتیم که واحدهای نمره های درصدی یا رتبه های درصدی، در طول مقیاس، نامساوی است و این عدم تساوی، مقایسه عملکرد افراد با یکدیگر را مشکل می کند؛ برای رفع این مشکل از نمرات معیار استفاده می شود. نمرات معیار موقعیت نسبی یک دانش آموز در گروه را با مشخص کردن فاصله نمره او از میانگین نشان می دهند. معروف ترین نمرات معیار عبارتند از:

۱- نمره Z: نمرات Z به آن دسته از نمرات معیار گفته میشود که دارای توزیعی با میانگین صفر و انحراف استاندارد یک هستند.

برای تبدیل نمرات خام به نمره Z از فرمول مقابل استفاده :

نمره خامX =

میانگین نمرات

انحراف معیار نمراتSD =

به عنوان مثال: اگر نمره خام دانش آموزی در یک آزمون، که دارای میانگین ۴۳ و انحراف معیار ۳ است، ۴۰ باشد، نمره Z او برابر است با:

نمره Z به ما نشان می دهد که یک نمره خام، چند واحد انحراف معیار از میانگین بالاتر یا پایین تر است؛ مثلاً دانش آموزی که نمره خام او 1 انحراف معیار از میانگین بالاتر باشد، نمره Z او برابر با ۱+ است و دانش آموزی که نمره خام او ۱ انحراف معیار از میانگین پایین تر باشد، نمره Z او برابر با ۱- خواهد بود و به همین ترتیب اگر نمره خام دانش آموزی ۲ انحراف معیار از میانگین بالاتر باشد، نمره Z او ۲+ خواهد بود و الی آخر.

مثال ۱۹: اگر نمره ی خام دانش آموزی ۳۵ و میانگین نمرات آن ها ۴۵ و انحراف معیار نمرات 5 باشید، نمره Z این دانش آموز چند است؟

1) 3

2) 2-

3) 1

4) 3-

پاسخ: گزینه «2»

مثال۲۰: فردی در سه آزمون با این مشخصات شرکت کرده است: آزمون اول با میانگین ۸۵ و انحراف استاندارد ۱۶، آزمـون دوم با میانگین ۹۰ و انحراف استاندارد ۱۵، آزمون سوم با میانگین ۱۰۰ و انحراف استاندارد ۵. اگر نمرات مشاهده شده در سه آزمون، یکسان و برابر ۱۰۸ باشد، در کدام آزمون وضع بهتری داشته است؟

۱) آزمون سوم

2) آزمون اول

۳) در آزمون های ۱ و ۲ نمره یکسانی داشته است.

4) آزمون دوم

پاسخ: گزینه «۱» مفروضان سؤال:

بنابراین محاسبات، مشاهده می شود که وضعیت آزمون سوم (گزینه ی اول) فرد نسبت به بقیه حالت بهتری دارد.

= 108 نمره مشاهده شده فرد در سه آزمون

مثال ۲۱: در کدام یک از انواع هنجارها یا مقیاس های سنجش، واحد نمره ها در طول مقیاس نابرابرند؟

۱) رتبه های درصدی

۲) نمره های T

۳) نمره های CEEB

۴) هوشبهر انحرافی

پاسخ: گزینه «۱» رتبه درصدی یا نمره درصدی، وضعیت نسبی به گروه نشان می دهد و گروهی را که نمره پایین تر از او گرفتـه انـد، مـعـین کند؛ در واقع رتبه های درصدی، عملکرد نسبی آزمون شوندگان را نشان می دهد در نرم درصدی نسبی و کلاسی، واحد نمـره هـا در طول مقیاس نابرابر است، ولی در نمرات استاندارد به خاطر صفر قراردادی و قرار گرفتن نمره ها در مقیاس فاصله ای، واحد نمره ها در طول مقیاس برابر می باشد؛ بنابراین گزینـی صحیح، رتبه های درصدی می باشد و سایر گزینه ها شامل نمرات استاندارد است که پاسخ سؤال نیست.

مثال ۲۲: اگر نمره ی خام دانش آموزی ۲ انحراف معیار بالاتر از میانگین باشد، به این معنی است که نمره Z او ..........

1) نمره ی ۴ است

۲) ۲- می باشد

۳) ۲+ می باشد

4) 4-

پاسخ: گزینه «3»

مثال ۲۳: چنانچه هوشبهر (IQ) یک کودک در آزمون وکسلر برابر با 5/137 باشد، نمره معیار وی (Z) چقدر است؟

1) 5/2-

2) 5/1-

3) 5/1+

4) 5/2+

پاسخ: گزینه «۴» با توجه به فرمول استاندارد نمره ی Z داریم:

5/137 x =

15S =

100

۲- نمره T :

از آنجا که نمرات Z گاهی به صورت منفی و گاهی به صورت اعشاری نشان داده می شوند نمرات T برای حل مشکلات آن ها ابداع شده اند. نمرات T با ضرب کردن نمرات Z در ۱۰ و جمع کردن حاصل با ۵۰ به دست می آیند؛ بنابراین، آن دسته از نمرات معیار، که دارای توزیعی یا میانگین ۵۰ و انحراف معیار ۱۰ هستند، نمرات T نام دارند. فرمول محاسب نمرات T به شرح مقابل است:

مثال ۲۴: با توجه به این که بعضی اوقات نمره Z اعشاری و گاهی منفی است، بهتر است از ................. استفاده شود.

۱) نمرات ۹ بخشی

۲) نمرات SAT

۳) نمره SAS

۴) نمره T

پاسخ: گزینه «۴»

مثال ۲۵: اگر دانش آموزی در مقیاسZ ، نمره ۵/0 بگیرد، نمره او در مقیاس T چند است؟

1) 55

2) 50

3) 60

4) 45

پاسخ: گزینه «۱»

مثال ۲۶: اگر دانش آموزی در مقیاس Z، نمره ۲- بگیرد، نمره او در مقیاس T چند است؟

1) 25

2) 30

3) 40

4) 60

پاسخ: گزینه «۲»

مثال ۲۷: اگر نمره دانش آموزی در مقیاس Z 1+ باشد، نمره T او چند می باشد؟

1) 52

2) 50

3) 55

4) 48

پاسخ: گزینه «۱»

مثال ۲۸: اگر نمره دانش آموزی در مقیاس Z 1+ باشد، نمره T او چند است؟

۱) ۴۰ وابسته به هنجار

2) 60

3) 50

۴) 51

پاسخ: گزینه «۲»

چون نمرات T دارای میانگین ۵۰ و انحراف معیار ۱۰ هستند، کسی که نمره T او برابر ۵۵ است، نمره او به اندازه نصف انحراف معیار، بالاتر از میانگین است و کسی که نمره T او ۳۰ است، به اندازه ۲ انحراف معیار، نمره او پایین تر از میانگین است و به همین ترتیب الی آخر.

حال اگر میانگین ۵۰۰ و انحراف معیار ۱۰۰ باشد، نمره ۵۵۰ ، نیم انحراف معیار بالاتر از میانگین است؛ بنابراین نمرات معیار، نمراتی با میانگین و انحراف معیار دلخواه هستند.

نمرات SAT

کلمه SAT از آزمون استعداد تحصیلی کشور آمریکا گرفته شده است. نمرات SAT دارای توزیعی بهنجار با میانگین ۵۰۰ و انحراف معیار ۱۰۰ هستند.

مثال ۲۹: اگر نمره Z دانش آموزی ۲+ باشد، نمره SAT او چند است؟

1) 300

2) 400

3) 700

4) 650

پاسخ: گزینه «3»

مثال ۳۰: چنانچه دانش آموزی در یکی از درس هایی که دارای میانگین کلاس ۱۲ و انحراف استاندارد ۲ است نمره ۱۸ بگیرد، نمره استاندارد وی چیست؟

1) 3-

2) 5/1-

3) 5/1+

4) 3+

پاسخ: گزینه «۴» برای نمره استاندارد از فرمول z استفاده می کنیم:

18 x = نمره

12 میانگین

2S = انحراف معیار

نمرات 9 بخشی

نمرات 9 بخشی، به صورت نمرات معیار یک رقمی نشان داده می شود. به این مقیاس، ۹ بخشی گفته می شود؛ چون توزیع این نمرات به ۹ بخش تقسیم می شود. این مقیاس دارای میانگین ۵ و انحراف معیار ۲ است.

نیتکو، نمرات ۹ بخشی را برای نمره گذاری آزمون های تشریحی، که ارزشیابی آن ها جنبه کیفی دارد، به ویژه آزمون های خواندن و فهمیدن، به شرح زیر پیشنهاد می کند:

9- بسیار عالی ( ۴% بالا)

8- عالی (۷% بعدی)

۷- خیلی خوب ( ۱۲/ه بعدی)

۶- خوب (۱۷% بعدی)

۵- متوسط (۲۰% وسط)

۴- نسبتأ ضعیف (۱۷% بعدی)

۳- ضعیف (۱۲% بعدی)

۲- خیلی ضعیف (۷% بعدی)

1- بیش از حد ضعیف ( ۴% آخر)

مثال ۳۱: نمره ی استاندارد «9 بخشی» (Stanine score) دانش آموزی با هوشبهر ۱۰۰ کدام است؟

1) 4

2) 5

3) 6

4) 7

پاسخ: گزینه «۲» از انواع نمرات هنجار، می توان به نمرات و بخشی اشاره کرد. در این نمرات، میانگین ۵ و انحراف معیار ۲ می باشد. در این نمرات، که برای نمره گذاری آزمون های تشریحی استفاده می شود، بیشتر به جنبه کیفی توجه می شود؛ در واقع چون میانگین هوشبهر ۱۰۰ است و میانگین نمره و بخشی ۵ است، پس این دو برابر هستند.

هوشبهر انحرافی (DIQ)

هوشبهر انحرافی، یک هنجار مربوط به نمرات معیار است که دارای میانگین ۱۰۰ و انحراف معیار ۱۵ یا (۱۶) است و فرمول آن به قرار زیر است:

100 Z 15 = هوشبهر انحرافی

بنابراین، اگر هوشبهر انحرافی دانش آموزی ۷۰ باشد، دانش آموز ۲ انحراف معیار پایین تر از میانگین است یا اگر هوشبهر انحرافی دانش آموزی ۱۱۵ باشد، دانش آموز، ۱ انحراف معیار بالاتر از میانگین می باشد.

مثال ۳۲: چنانچه نمره 7 استاندارد یک آزمودنی برابر با ۲- یاشد، نمره خام ۷۰، متناظر با کدام یک از «نمرات استاندارد برای این آزمودنی است؟

۱) رتبه درصدی

۲) نمره ی معیار T

3) هوشبهر انحرافی

۴) استاناین یا نمره ۹ بخشی

پاسخ: گزینه «3» با توجه به مفروضات سؤال، نمره استاندارد، هوشبهر انحرافی می باشد.

هوشبهر انحرافی، یک نمره معیار است و به این اعتقاد قدیمی، که سن ذهنی و سن تقویمی کودکان باید برابر باشد، مبتنی نمی باشد؛ برای مثال، براساس مفهوم قدیمی هوشبهر، کودکی که سن تقویمی و سن ذهنی او ۵ سال کامل باشد، هوشبهر معادل ۱۰۰ دارد، اما هوشبهر واقعی این کودک در آزمون استنفورد بینه، جز این است.

مثال ۳۳: با توجه به توزیع بهنجار، در یک آزمون هوشی، که دارای انحراف معیار ۱۵ است، هوشبهر ۱۴۵ برابر چند در مقیاس T است؟

1) 50

2) 60

3) 70

4) 80

X = ۱۴۵, x=۱۰۰ , S=۱۵

پاسخ: گزینه «۴» از فرمول تعیین نمره استاندارد Z استفاده می نماییم. مفروضات سؤال:

15s =

145x =

سپس از فرمول نمره T استفاده و نمره 2 را جایگزین می کنیم:

مثال ۳۴: میانگین و انحراف معیار نمرات خرده آزمون های مقیاس هوش و کسلر، به صورت استاندارد چقدر است؟

1) 10و3

2) 50 و 10

3) 100 و 15

4) 100 و 16

پاسخ: گزینه «۱» برای آنکه بتوان آزمون را تفسیر کرد، باید نمرات خام حاصل از آزمون را به نمرات استاندارد تبدیل نمود که این نمره دارای میانگین ۱۰ و انحراف استاندارد ۳ می باشد. گزینه های دیگر برای داوطلبانی تهیه شده است که با حدس و گمان به این سؤال جواب میدهند. در میانگین و انحراف معیار آزمون وکسلر ۱۰۰ و ۱۵ و میانگین و انحراف استاندارد خرده آزمون های وکسلر ۱۰ و ۳ میباشد.

مثال ۳۵: مقیاس اندازه گیری «هوشبهر انحرافی» کدام است؟

۱) اسمی

2) رتبه ای

3) فاصله ای

۴) نسبتی

پاسخ: گزینه «3» برای پاسخ دادن به این سوال باید بسیار هوشمندانه رفتار کنیم. چون گزینه های ۲ و ۳ در این سؤال بسیار به هم نزدیک می باشند. ابزارهای اندازه گیری نظیر پرسشنامه های نگرش سنج، دارای سؤالاتی هستند که برای پاسخ به آن ها باید از یک طیف کاملا موافق تا کاملا مخالف را در نظر گرفت؛ در واقع هیچ سؤالی دارای جواب غلط نیست و مقیاس مورد استفاده ی آن ها، مقیاس فاصله ای می باشد. برای جمع بندی این پرسشنامه ها میانگین نمرات پاسخ دهندگان به یک سؤال محاسبه قرار می شود. یکی از مثال های کتاب آقای دکتر علی اکبر سیف درباره ی مقیاس فاصله ای، به این مسئله مستقیما اشاره دارد: هوشبهر (IQ) قهرمان سریال موردعلاقه خود را حدس بزنید (مورد سؤال: هوشبهر).

نمره سنی معیار SAS

نمره سنی معیار یا SAS، نمراتی هستند با انحراف معیار ۱۶ و میانگین ۱۰۰. بنابراین اگر نمره SAS فردی ۱۳۲ باشد، نمره او ۲ انحراف استاندارد از میانگین بالاتر است و اگر نمره فردی ۷۶ باشد، نمره او انحراف استاندارد از میانگین پایین تر است.

مثال ۳۶: نمره Z دانش آموزی ۳ می باشد. نمره سنی معیار SAS او چند می باشد؟

۱) ۱۳۸

۲) ۱۲۸

۳) ۱۴۸

۴) ۱۳۲

پاسخ: گزینه «3»

مثال ۳۷: نمره Z فردی ۲- می باشد. نمره سنی معیار SSA او چند می باشد؟

1) 58

2) 132

3) 116

4) 68

پاسخ: گزینه «۴»

مثال ۲۸: اگر توزیع ها را نرمال فرض کنیم، کدام یک از مقادیر زیر، نشان دهنده نمره بالاتر در آزمون خواهد بود؟

1) 2z =

۲) 7=9 بخشی

3) 64T =

4) 120IQ =

پاسخ: گزینه «۱» در این سؤال باید نمرات مانند هم باشند تا بتوان مقایسه ی درستی انجام داد؛ بنابراین نمرات را به نمره ی Z تبدیل می کنیم و از همه گزینه ها برای به دست آوردن پاسخ استفاده می کنیم:

|  |  |  |
| --- | --- | --- |
| انحراف استاندارد | میانگین M | نمره استاندارد |
| 1  2  10  15 | 0  5  50  100 | Z  9 بخشی  T  IQ |

در مقایسه بین ۲ = Z و Z در سایر نمره ها که محاسبه شده اند، نتیجه می گیریم که گزینه اول از وضعیت استانداردتری برخوردار است.

نمره ی معادل منحنی بهنجار: یا NCE

نمره NCE که در آزمون های میزان شده پیشرفت تحصیلی به کار می رود؛ به همان منظوری استفاده می شود که هنجارهای کلاسی به کار می روند، اما معایب آن ها را ندارند. نمرات NCE دارای میانگین ۵۰ و انحراف معیار۰۶/21 هستند.

50 + 06/21 NCE =

مثال ۳۹: کدام یک از نمرات تراز شده (استاندارد) زیر نشان دهنده وضعیت بهتر آزمودنی در مقایسه با سایر افراد گروه می باشد؟ (با فرض نرمال بودن توزیع نمرات).

1) 5 Stanine =

2) 1+ Z =

3) 55T =

4) 100IQ =

پاسخ گزینه «۲» در اینجا باید همه نمرات استاندارد را به 2 تبدیل کنیم. هر کدام از نمرات که 2 بزرگتری داشته باشد، وضعیت بهتری را نشان می دهد؛ یعنی:

نکته ۲: نظام های گوناگون نمره گذاری، راه های مختلف بیان مطلبی واحد هستند، یعنی ما می توانیم به سادگی از هر مقیاس به مقیاس دیگر برویم، به شرطی که توزیع نمرات ما، یک توزیع بهنجار باشد.

ویژگی های منحنی بهنجار

1- در منحنی بهنجار، نما، میانه و میانگین روی هم قرار دارند.

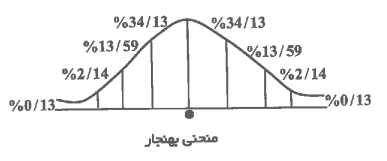
٢- دو قسمت منحنی، نسبت به خط وسط آن قرینه است.

۳- روی این منحنی، ۲ نقطه وجود دارد که در آن ها جهت منحنی تغییر می کند. این دو نقطه در فواصل 1± انحراف معیار از میانگین قرار دارند.

۴- اگر تمام سطح منحنی بهنجار را ۱ یک واجد فرض کنیم، آنگاه:

26/68% از کل مساحت منحنی بین ۱+ و ۱- انحراف معیار قرار دارد، 44/95% از کل مساحت زیر منحنی بین ۲+ و ۲- انحراف معیار قرار دارد و 74/99% از مساحت زیر منحنی بین ۳ و ۳- انحراف معیار قرار دارد.

۵- در منحنی بهنجار، اکثر نمره ها در وسط توزیع انباشته شده اند.



نیمرخ

از جمله فواید نمرات هنجاری یا مشتق این است که با استفاده از آن ها می توان عملکرد دانش آموزان را در آزمون های مختلف مستقیما با هم مقایسه کرد که این کار از طریق نیمرخ آزمون امکان پذیر است. به کمک نیم رخ، که یک بازنمایی نگاره ای از نمرات آزمون است، می توان نقاط قوت و ضعف دانش آموزان را مشخص کرد؛ همچنین با یک نظر اجمالی می توانم برآوردی کلی از وضعیت پیشرفت دانش آموزان به دست آورد. هرگاه تعدادی نمره ی حاصل از آزمون های مختلف با آزمون های فرعی یک آزمون کلی را، برای یک دانش آموز به یکی از هنجارها تبدیل کنیم و این نمرات را در یک نمودار قرار دهیم، نیمرخ عملکرد آن دانش آموز را درست کرده ایم.

تست های طبقه بندی شده فصل پانزدهم

1- هرگاه در یک آزمون 40 سؤالی، ۴۸ درصد افراد به بیش از ۱۵ سؤال پاسخ درست دهند، در این صورت، نمره خام ۱۵ معادل صدک چند خواهد بود؟

(سراسری ۸۱)

1) 15

2) 25

3) 48

4) 52

۲- نمره ۶۵ در مقیاس T، معادل کدام نمره در مقیاس Z است؟

(سراسری۸1 )

1) 1

2) 5/1

3) 2

4) 5/2

۳- نمره فراگیری در مقایس 7 برابر ۴/0 می باشد. نمره همین فرد در مقیاس دیگری با میانگین ۵۰ و انحراف معیار ۱۰ چقدر است؟

(سراسری 81)

1) 20

2) 50

3) 54

4) 90

۴- کدام گروه از نمره های استاندارد زیر، معادل یکدیگر هستند؟

(سراسری ۸۱)

1)

2)

3)

4)

5- کدام هنجارها، از تبدیل های غیرخطی به دست می آیند؟

(سراسری ۸۱)

۱) سن ذهنی و صدک ها

۲) سن ذهنی و نمرات Z

3) نمرات معیار و هوشبهر انحرافی

۴) نمرات معیار و سن ذهنی

۶- در یک توزیع بهنجار، با میانگین ۳۴ و انحراف استاندارد ۴، نمره ی خام ۴۴ معادل کدام نمره در مقیاسTاست؟

(سراسری ۸۲)

1) 65

2) 70

3) 75

4) 80

۷- منظور از استاندارد بودن آزمون در روان سنجی، عبارت است از یکسان کردن:

(سراسری 82)

1) سؤالات و آزمودنی ها

۲) سؤالات آزمون بر اساس محتوا

3) نمرات بر اساس نمره ی Z

4) مواد آزمون، اجراء و نمره گذاری

۸- رتبه ی درصدی دانش آموزی که نمره ی استاندارد او در توزیع برابر با ۱+ محاسبه شده است، چقدر است؟

(سراسری ۸۲)

1) 16%

2) 34%

3) 50%

4) 84%

۹- در آزمونی ۲۰۰ دانش آموز شرکت داشته اند. میانگین این آزمون 5/14و انحراف استاندارد آن 5/2 است. اگر نمره احمد ۱۷ باشد، نمره چند نفر از او بیشتر است؟

(سراسری ۸۲)

1) 32

2) 74

3) 122

4) 168

۱۰- کدام یک از مقادیر زیر نشان دهنده ضعیف ترین عملکرد در یک آزمون است؟

(آزاد ۸۲)

1) رتبه درصدی

۲) نمره استاندارد 5/1-

۳) نمره استاندارد شده 30

4) نمره نه گانه ۲

۱۱- میانگین نمرات جغرافیای یک کلاس 5/14 و انحراف معیار آن 25/1است. اگر نمره معیار دانش آموزی در این درس (۲-) باشد، نمره او در امتحان چقدر بوده است؟

(آزاد ۸۳)

1) 17

2) 75/15

3) 25/13

4) 12

۱۲- در صورتی که گفته شود علی به ۴ سؤال یک آزمون، پاسخ درست داده است و بهره هوشی او برابر ۴ سال است، از کدام یک از نرم های زیر استفاده شده است؟

(آزاد ۸۳)

۱) جنس

۲) سنی

۳) درصدی

4) عددی

۱۳- برای مقایسه نمرات یک دانش آموز در دروس مختلف، نمرات باید به کدام یک از صورت های زیر باشد؟

(آزاد ۸۳)

1) خام

۲) محاسبه شده

۳) استاندارد

۴) درصد

۱۴- هنجار عبارت است از: عملکرد ...............

(سراسری ۸۴)

1) متوسط

۲) بالاتر از میانگین

۳) قابل قبول

۴) قابل انتظار

۱۵- کدام نمره، یک نمره تراز شده واقعی نیست؟

(سراسری ۸۴)

1) نمره Z

۲) نمره T

۳) نمره Stanine

۴) نمره CEEB

۱۶- نمره استاندارد (Z) دانش آموزی ۵/0 است. اگر نمره خام این دانش آموز ۱۵ و میانگین کلاس ۱۴ باشد، انحراف استاندارد نمرات کلاس چند خواهد بود؟

(سراسری ۸۴)

۱) 1

2) 5/1

3) 2

4) 5/2

۱۷- کدام هنجارها از راه تبدیل خطی به دست می آید؟

(سراسری ۸۴)

1) هوشبهر انحرافی و نمره Z

۲) نمرات معیار و صدک ها

۳) هوشبهر انحرافی و هوشبهر نسبی

۴) معادل های کلاسی و هوشبهر نسبی

۱۸- در منحنی Z صدک ۳۵ روی کدام انحراف معیار قرار دارد؟

(سراسری ۸۴)

۱) بین ۰ و ۱-

۲) بین 0 و 1+

3) بین ۱- و ۲-

۴) بین ۲- و 3-

۱۹- در یک آزمون ۱۰۰ سوالی با توزیع نرمال و انحراف معیار ۱۵ و میانگین ۵۰، تقریبا چند درصد افراد، نمره ای بین ۵ و ۲۰ دارند؟

(سراسری ۸۴)

۱) ۲ درصد

۲) ۵ درصد

3) ۸ درصد

۴) ۱۶ درصد

۲۰- نمره 5/1- Z=، کدام نمره CEEB است؟

(آزاد ۸۴)

۱) ۵۰

2) 150

3) 350

4) 450

۲1- حسن در آزمون پیشرفت ۱۰۰ ماده ای، نمره خام ۸۴ کسب کرده است. کدام تفسیر زیر بیشتر توجیه پذیر است؟

(آزاد ۸۴)

1) نمره ۹ بخشی او ۷ است.

۲) رتبه صدکی او در آزمون ۸۴ است.

۳) او در نیمه ی بالای گروه قرار دارد.

۴) او نمره صحیح ۸۴ درصدی دارد.

۲۲- تفسیر ملاکی، (ملاک - مرجع)، متمرکز است بر:

(آزاد ۸۴)

1) هم بستگی ها با ملاک های مرتبط

۲) استفاده از رتبه های صدکی

۳) توصیف های عملکرد

۴) جایگاه نسبی در گروه

۲۳- یکی از محاسن نمره های T در مقایسه با نمره های Z آن است که:

(آزاد ۸۴)

1) آن را می توان جمع و تفریق کرد.

۲) خیلی آسان به نمره صدکی تبدیل می شود.

۳) همیشه انحراف معیار یکسان دارد.

۴) فقط نمره های مثبت را شامل می باشد.

۲۴- «معرف بودن» یک هنجار به .................. مربوط می شود.

(سراسری ۸۵)

1) حجم و روش نمونه گیری

۲) تعریف دقیق و عینی هنجارها

۳) تازگی و جدید بودن هنجاریابی

۴) مربوط و مناسب بودن گروه هنجاری

۲۵- در نمره استاندارد و بخشی (Stanine)، از کدام نوع مقیاس اندازه گیری استفاده می شود؟

(سراسری ۸۵)

1) اسمی

۲) رتبه ای

۳) فاصله ای

4) نسبی

۲۶- وقتی نمره فردی را در یک آزمون استاندارد شده ی پیشرفت تحصیلی به صورت رتبه ی درصدی ارائه می نماییم، معنی آن این است که:

(سراسری ۸۵)

۱) درصد پاسخ های صحیح فرد را تعیین می کنیم.

۲) فر را با افراد گروه هنجار به صورت درصدی مقایسه می کنیم.

۳) درصد افرادی را که از او نمره بیشتر کسب کرده اند، معین می کنیم.

4) درصد پرسش های درست او را نسبت به پرسش های نادرست تعیین می کنیم.

۲۷- کدام دسته از سؤالات زیر را می توان دارای مقیاس فاصله ای دانست؟

(آزاد ۸۵)

۱) هنجار کلاسی

۲) نمرات تراز شده

۳) هنجار درصدی

۴) هنجار سنی

۲۸- نمره فردی در مقیاس T با میانگین ۵۰ و انحراف استاندارد ۱۰ معادل ۵۷ است، نمرات معادل این فرد در مقیاس دیگری با میانگین ۲۰۰ و انحراف استاندارد ۲۰ چقدر خواهد بود؟ |

(آزاد ۸۵)

1) 210

2) 207

3) 214

4) 186

۲۹- نیمرخ زیر نشان دهنده وضعیت نمره های چهار درس است. آیا با حدود 96/0 اطمینان اختلاف کدام دروس معنی دار نیست، حال آن که با حدود اطمینان معنی دار است. هر سمت هاشور زده یک خطای معیار اندازه گیری است.

(آزاد ۸۵)

1) فارسی و دیکته

۲) دیکته و علوم

۳) فارسی و علوم

۴) فارسی و ریاضیات



۳۰- توزیع نمرات هنجار کلاسی یک گروه دانش آموز به شرح زیر است. مجموع این گروه در چه وضعیتی است؟ در صورتی که باید پایه چهارم و انتهای ماه چهارم سال تحصیلی، تحصیل کند.

(آزاد ۸۵)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 2/4  2 | 3/4  2 | 4/4  4 | 6/4  3 | نمره ی هنجاری کلاس  فراوانی |

1) قوی

۲) ضعیف

۳) با این اطلاعات، قابل تعیین نیست.

۴) متوسط

۳۱- در تبدیل نمرات آزمون های شخصیت، کدام یک از مقیاس های ذیل بیشتر مورد استفاده قرار می گیرد؟

(سراسری 86)

1) میانگین ۵ انحراف معیار تقریبا ۲

۲) میانگین ۵۰۰ انحراف معیار ۱۰۰

۳) میانگین ۱۰۰ انحراف معیار ۱۵

۴) میانگین ۵۰ انحراف معیار ۱۰

۳۲- کدام عبارت، توصیف بهتری از هنجار به دست می دهد؟

۱) عملکرد متوسط

۲) عملکرد ایده آل

۳) عملکرد قابل قبول

۴) عملکرد قابل انتظار

۳۳- برای تعیین موقعیت یک دانش آموز در کلاسی که خودش جزء آن می باشد، مناسب ترین روش کدام است؟)

۱) ارزشیابی

۲) اجرای آزمون استاندارد

3) محاسبه نمره خام دانش آموز

۴) تعیین نرم

۳۴- نمره ۶۵ در مقیاس T معادل چه نمره ای در مقیاس نمرات هوش با (میانگین ۱۰۰ و انحراف معیار برابر با 15) می باشد؟

(سراسری 87)

1) 5/92

2) 100

3) 5/107

4) 5/122

۳۵- نمره فردی در مقیاس t برابر ۶۵ است. نمره این فرد در مقیاسی با میانگین ۱۰ و انحراف استاندارد ۴ کدام است؟

(سراسری ۸۸)

1) 16

2) 12

3) 8

4) 4

۳۶- نمرات کدام هنجار می تواند در مقیاس فاصله ای تلقی شود؟

(سراسری ۸۸)

۱) هنجار سنی

۲) هنجار درصدی

3) هنجار کلاسی

۴) هنجار نمرات معیار

37- هرگاه، نمونه ای که معرف دختران ۸ ساله باشد فراهم سازیم و اندازه قد هر یک از آنان را معین کنیم و آن گاه میانگین این اندازه ها را محاسبه نماییم، مشخص کنید این ملاک و معیار به کدام مورد اشاره دارد؟

(آزاد ۸۸)

۱) نرم سنی

۲) نرم کلاسی

۳) نمره استاندارد

۴) شاخص درصدی

۳۸- کدام مورد، جایگاه نسبی فرد در گروه را نمایش می دهد؟

(سراسری ۸۹)

1) میانگین

۲) دهک ها

۳) رتبه درصدی

۴) نقطه درصدی

۳۹- مهم ترین دلیل روی آوردن به هوشبهر انحرافی کدام است؟

(سراسری ۸۹)

۱) نمود بیشتر تفاوت عملکرد ذهنی در سنین بالا

۲) ضعف مبنای نظری هوش به طور کلی

۳) سرعت متفاوت رشد سن عقلی در سنین مختلف

۴) رابطه ضعیف میان کشش های ذهنی و سن عقلی در سنین پایین

۴۰- نمره ۲ در توزیع نمرات معیار (Z) معادل چه نمره ای در توزیع (T) می باشد؟

(سراسری۸۹ )

1) 70

2) 50

3) 115

4) 130

۴۱- نمره احمد در گروهی که جمعا ۶0 نفر هستند، از امتحانی شامل ۱۰۰ سؤال در نقطه ۸۰ درصدی قرار گرفته است. این بدان معنی است که او ...............

(آزاد ۸۹)

۱) از ۸۰ نفر گروه در این امتحان بهتر است.

۲) از ۴۸ نفر گروه در این امتحان بهتر است.

۳) به ۸۰ سؤال پاسخ صحیح داده است.

۴) بهتر از آن چه انتظار داشته نمره آورده است.

۴۲- در صورتی که منحنی توزیع داده ها دارای انحنای منفی است، میانگین داده ها در چه نقطه درصدی قرار می گیرد؟

(آزاد ۸۹)

۱) کمتر از ۵۰ درصدی

2) 50 درصدی

۳) بیشتر از ۵۰ درصدی

۴) اطلاعات بیشتری مورد نیاز است.

۴۳- در مورد ملاک نسبی می توان گفت که:

(آزاد ۸۹)

۱) موقعیت هر فرد در مقایسه با عملکرد سایر همکلاسی ها مشخص می شود.

۲) در آن هیچ دانش آموزی با دانش آموزان دیگر مقایسه نمی شود.

3) تعیین کردن سطحی است که هر یک از دانش آموزان به آن رسیده باشد، قبول محسوب می شود.

۴) امتحان های پایان هر کلاس در دوره های مختلف تحصیلی از این نوع است.

۴۴- کدام هنجار، در مقیاس فاصله ای قرار دارد؟

(سراسری ۹۰)

1) سنی

۲) کلاسی

۳) نمرات معیار

4) درصدی

۴۵- محدودیت نرم درصدی کدام است؟

(سراسری ۹۰)

۱) دشواری تفسیر آن

۲) دشواری محاسبه و فهم آن

3) نابرابری واحدها در طول مقیاس

۴) عدم انطباق با منحنی نرمال

۴۶- اگر اندازه انحراف استاندارد توزیع گروهی از نمرات کاهش یابد، به شرط یکسانی سایر شرایط، چه تغییری در Z مربوط به هر نمره ایجاد می شود؟

(سراسری ۹۰)

1) منفی می شود.

2) مثبت می شود.

۳) کاهش می یابد.

۴) افزایش می یابد.

۴۷- نرم - مرجع و یا ملاک - مرجع بودن آزمون به کدام عامل وابسته است؟

(سراسری ۹۰)

۱) چگونگی اجرا

۲) نحوه نمره گذاری

۳) نحوه ی قضاوت در مورد فرد

۴) نوع سؤالات

۴۸- یکی از اشکالات رتبه درصدی، آن است که تغییر ناچیز در ناحیه میانی توزیع نمرات ..................

(سراسری ۹۱)

1) تغییری به وجود نمی آورد.

۲) موجب تغییر چشمگیر در آن می شود.

۳) تغییر اندکی در رتبه درصدی ایجاد می کند.

۴) موجب گمراهی در اندازه گیری میشود.

۴۹- در کدام یک از انواع نرم ها، فاصله نمره از میانگین بر حسب انحراف معیار بیان می شود؟

(سراسری ۹۱)

1) نرم های سنی

۲) نرم های درصدی

۳) نرم های کلاسی

۴) نرم های استاندارد

50- اگر رتبه درصدی نمره یک فرد در آزمونی که در آن ۸۰ نفر شرکت کرده اند، برابر با ۵۰ باشد، کدام نتیجه گیری منطقی تر است؟

(سراسری ۹۲)

۱) پنجاه نفر از شرکت کنندگان از این فرد نمره بیشتری کسب کرده اند.

۲) نمره این فرد از نمره پنجاه نفر از شرکت کنندگان بیشتر بوده است.

3) این فرد از پنجاه درصد افراد نمره بیشتری کسب کرده است.

۴) این فرد به پنجاه درصد سؤالات درست پاسخ داده است.

۵۱- در تبدیل نمرات خام آزمون به میانگین برابر ۵۰ و انحراف معیار برابر 10، کدام حالت رخ می دهد؟

(سراسری ۹۲)

1) شکل توزیع نمرات تغییر نمی کند.

2) به میانگین قبلی ۵۰ نمره اضافه می شود.

۳) نمرات شرکت کنندگان همگن تر می شود.

۴) افرادی که نمره پایینی در این آزمون داشتند از این تغییر، سود بیشتری می برند.

۵۲- در توزیع نرمال، چند درصد افراد، بالاتر از میانگین به اضافه یک انحراف استاندارد قرار می گیرند؟

(سراسری ۹۲)

1) 16

2) 40

3) 880

4) 84

۵۳- در توزیع T (50 و 10SD =) نمره فردی که ۹۵% افراد از او نمره بیشتری گرفته اند، چقدر است؟

(سراسری ۹۳)

1) 4/30

2) 66/33

3) 4/66

4) 6/69

۵۴- انگیزه تحصیلی در یک کلاس، دارای توزیع نرمال با میانگین ۱۷۸ و واریانس ۱۰۰ است. چنانچه انگیزه تحصیلی یک دانشجو برابر با ۱۸۸ باشد، رتبه درصدی او حدودأ چند است؟ |

(سراسری ۹۴)

1) 34

2) 44

3) 84

4) 94

۵۵- نمرات استاندارد یک آزمودنی در چهار آزمون مختلف به صورت زیر است:

2z = و 4=9 بخشی و 115IQ = و 480SAT =

(سراسری ۹۴)

کدام نمره بیانگر عملکرد بهتر آزمودنی است؟

1) 2z =

2) 4=9 بخشی

3) 115IQ =

4) 480SAT =

پاسخنامه تست های طبقه بندی شده فصل پانزدهم

۱- گزینه «۴» باتوجه به این که ۴۸ درصد افراد به بیش از ۱۵ سؤال پاسخ داده اند، پس ۵۲ درصد افراد به ۱۵ یا کمتر از ۱۵ سؤال، پاسخ درست داده اند.

۲- گزینه «۲» با توجه به این که میانگین نمره T برابر با ۵۰ و انحراف استاندارد آن ۱۰ است، خواهیم داشت:

3-گزینه «۳» برای تبدیل نمره Z به نمرات استاندارد دیگر از فرمول مقابل استفاده می کنیم.

۴- گزینه «۴» در ابتدا کلیه ی نمرات استاندارد را به Z تبدیل می کنیم و مشاهده می کنیم که گزینه آخر صحیح است.

|  |  |  |
| --- | --- | --- |
| نمرات استاندارد | میانگین | انحراف معیار |
| Z  t  IQ | 0  50  100 | 1  10  15 |

1Z =

5-گزینه «۱» در تبدیل نمره خام به نمرات معیار، شکل توزیع تغییر نمی کند؛ زیرا برآورد، خطی است و در تبدیل نمرات خام به رتبه درصدی، فرم سنی و کلاسی، برآورد خطی وجود ندارد؛ زیرا واحد ها در طول مقیاس، برابر نیستند.

۶- گزینه «۳» ابتدا نمره ی Z را به دست می آوریم، سپس با توجه به مفروضات سؤال، این اعداد را در فرمول قرار می دهیم. پس از آن، این عدد حاصل را در فرمول T جایگزین می کنیم.

مفروضات سؤال: ۴۴ x = ، ۳۴ و ۴ = SD

۷- گزینه «۴» استاندارد بودن آزمون یعنی رعایت شرایط خاص قبل، حین و بعد از آزمون، طبق قواعد و مقررات. این شرایط استاندارد مؤلفه های معینی دارند که عبارتند از: ۱ـ دستورالعمل اجرای استاندارد ۲- مدت زمان تعیین شده برای انجام آزمون یا سؤالات آزمون ۳- نمره گذاری استاندارد آزمون.

۸- گزینه «۴» با توجه به منحنی زیر و مقایسه درصد ها و نمرات معیار در توزیع بهنجار، گزینه (۴) صحیح است.

رتبه درصدی = 84% ، هوشبهر = 115، نمرات Z = 1 ، نمرات T= 60

9- گزینه «۱» باتوجه به این که Z این نمره برابر ۱ است و از نمره ۱ = Z حدود ۱۶ درصد افراد بدتر عمل کرده اند، برای تعیین این که در ۲۰۰ نفر، چنـد نفر از او بدتر عمل کرده اند، عدد ۱۶ را در ۲ ضرب می کنیم که ۳۲ به دست می آید.

۱۰- گزینه «۱»

۱۱ـ گزینه «۴» هر دو طرف معادله را در 25/1 ضرب می کنیم.

۱۲- گزینه «۲» هنجار سنی یا معادل سنی بر میانگین یا میانه نمراتی مبتنی است که دانش آموزان در سنین مختلف کسب می کنند و به صورت معادل های سنی نشان داده می شود. در متن این سؤال گفته شد که علی به ۴ سؤال یک آزمون، که یک نرم برای کودکان ۴ ساله است، پاسخ داده است؛ بنابراین بهره هوشی او با هنجار سنی که فرد در آن قرار دارد، نشان داده شده است. از دیگر هنجارهایی که از آن ها می توان برای نشان دادن نمره یک فرد استفاده کرد، می توان به هنجار کلاسی (از گروه های کلاسی به عنوان گروه های مرجع استفاده می شود) و هنجار درصدی (که در آن فرد با گروه سنی یا گروه کلاسی خود، یعنی گروهی که منطقا می توان او را عضوی از آن دانست، مقایسه می شود) اشاره کرد.

۱۳- گزینه «٣» نمرات استاندارد، نمرات خام را قابل مقایسه می کنند.

۱۴- گزینه «۱» در تعریف داریم: اگر برای تفسیر نمره یک دانش آموز، نمره او را با میانه یا میانگین نمره یک گروه مرجع مقایسه کنیم به آن گروه مرجع، گروه هنجار می گویند. گروه مرجع از کسانی تشکیل می شود که به گونه ای شبیه به دانش آموز مورد نظر هستند.

۱۵- گزینه «3» نمره نه گانه (stanine) فقط ۲ انحراف معیار بالا و پایین، قدرت تغییر پذیری دارد و برای مقیاس رتبه ای کاربرد دارد؛ پس واقعی نیست.

۱۶- گزینه «3»

۱۵= نمره خام = x

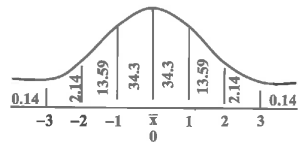
۱۴ = نمره میانگین =

x= انحراف معیار

5/0 = z

۱۷- گزینه «۱» در این سؤال فقط هوشبهر انحرافی و نمره Z جزء گروه هنجارهای مربوط به نمرات معیار هستند. نمرات معیار، موقعیت نسبی یک دانش آموز در یک گروه را نشان می دهد که فاصله نمره او را با نمره میانگین سایر نمرات معلوم می کند.

۱۸- گزینه «۱» صدک (رتبه درصدی) با سطح زیر منحنی برای داده های تا ۱- انحراف استاندارد 34/13یا تقریبأ ۳۵ است.



۱۹- گزینه «۱» ابتدا نمرات هر کدام را حساب می کنیم :

مساحت زیر منحنی بین ۲- تا ۳- برابر است با 14/2 درصد

۲۰- گزینه «3» نمره های CEEB که معمولا در آزمون های ورودی دانشگاه ها به کار می روند، نمره هایی هستند که میانگین آن ها ۵0۰ و انحراف معیار آن ها ۱۰۰ می باشد؛ بنابراین برای تبدیل نمره های Z به نمره های CEEB کافی است اندازه های را در ۱۰۰ ضرب کرده، سپس حاصلضرب را با ۵۰۰ جمع کنیم. این رابطه را می توان به صورت زیر تعریف کرد:

برای این سؤال خواهیم داشت:

۲۱- گزینه «۱» گزینه ۲ اشتباه است؛ چون ما به شرطی می توانیم بگوییم رتبه صدکی او ۸۴ است که بدانیم، نمره خام ۸۰ درصد دانش آموزان از او پایین تر است. گزینه ی ۳ نیز اشتباه می باشد؛ چون ما باز هم از نمره ی سایر افراد خبر نداریم و ممکن است سایر افراد، همگی نمره ی خام بالای ۸۴ گرفته باشند، بنابراین او در نیمه ی پایین گروه قرار داشته باشد. همچنین گزینه ۴ اشتباه است، نمره ی صحیح فرد، معادل همان نمره ی خام او نیست و باید با توجه به فرمول تصحیح، نمره صحیح او را ابتدا به دست آورد و سپس در مورد آن نظر داد.

۲۲- گزینه «۱» در ارزشیابی ملاکی یا ارزشیابی وابسته به ملاک، معیار یا ملاک ارزشیابی از پیش تعیین می شود و عملکرد یادگیرنده با توجه به آن ملاک به هنجار می باشند. گزینه ۳ نیز که تنها به توصیف می پردازد و

قضاوت میشود.گزینه های ۲ و ۴ اشتباه هستند؛ چون مربوط به ارزشیابی های هیچ گونه تفسیر یا ارزشیابی از عملکرد فرد به عمل نمی آورد، بنابراین این گزینه نیز اشتباه می باشد.

۲۳- گزینه «۴» چون نمرات 2 گاه به صورت منفی و گاه به صورت اعشار به دست می آید، برای رفع این مشکل از نمرات T استفاده می کنیم که همواره نمره های مثبت را شامل می شود.

۲۴- گزینه «۱» اگر برای تفسیر نمره یک دانش آموز، نمره او را با میانه یا میانگین یک گروه مرجع مقایسه کنیم، به آن گروه مرجع، گروه هنجار می گویند. در تعیین هنجار برای یک آزمون، به این سؤال می پردازیم که آزمون در مورد چه گروهی به کار خواهد رفت؟ پـس نـوع گروه و حجم آن موردنظر است. بنابراین معرف بودن یک هنجار به حجم و روش نمونه گیری مربوط می شود.

۲۵- گزینه «۲» نمرات ۹ بخشی، نوعی از گروه نمرات هنجار است و علت نامگذاری آن به ۹ بخشی این است که توزیع ایـن نـمـرات بـه ۹ بخـش مـحـدود میشود. این مقیاس دارای میانگین ۵ و انحراف معیار ۲ است. این نمره گذاری برای آزمون های تشریحی کاربرد دارد که ارزشیابی آن ها جنبه کیفی دارد. مقیاس رتبه ای، افراد یا اشیا را با توجه به صفت مرتب میک کند. نمر رتبه بندی مشخص خاص خود را از طیف بسیار عالی تا بیش از حد ضعیف دارد.

۲۶- گزینه «۲» رتبه درصدی یا نمره درصدی، وضعیت نسبی فرد را در گروه برحسب کسانی که نمره پایین تر از آن فرد گرفته اند، مشخص می کند. در هنجـار درصدی، فرد با گروه سنی یا کلاس خودش یعنی گروهی که منطقاً می توان او را عضوی از آن دانست، مقایسه می شود.

۲۷ـ گزینه «۳» در مقیاس فاصله ای، فاصله های بین واحدها معلوم و برابر است و هیچ نوع اطلاعی از مقدار مطلق صفت مورد نظر برای هیچ یک از اشیا در دست نیست. هنجار درصدی نیز وضعیت نسبی فرد در گروه را بر حسب کسانی که نمره ی پایین تر از او گرفته اند، مشخص کند.

۲۸- گزینه «۳» برای پاسخ به این سؤال ابتدا نمره Z را به دست می آوریم.

۲۹ـ گزینه «۳» چون هرچه قسمت های هاشورزده با هم همپوشی داشته باشند یا به هم نزدیک باشند، تفاوت کمتری میان آن دو آزمـون اسـت؛ یعنی تفاوتشان معنادار نیست.

۳۰ـ گزینه «۱» باتوجه به این که نرم این گروه ۴/۴ است و فراوان گروه نیز ۴ نفر است، هم چنین ۳ نفر از ایـن گـروه، نمـره هـای بـالاتر از هنجار (6/4) گرفته اند و فقط چهار نفر نمره های پایین تر از هنجاری آن قرار دارند گرفته اند، می توان گفت گروه وضعیتی و قوی دارد.

۳۱ـ گزینه «۴» نمره معیار T-scores دارای میانگین ۵۰ و انحراف معیار ۱۰ میباشد و هر نمره T به تنهایی و به طور مستقیم قابل تفسیر و استفاده است. در بعضی سازمان ها و مؤسسات از نمرات معیار دارای میانگین ۵۰۰ و انحراف ١٥٥ استفاده می کنند. در واقع، نمرات معیار با انحراف معیار و میانگین دلبخواهی می باشند، اما نمره T رایج تر است. باید به این نکته توجه کرد که نمره میانگین ۵۰۰ و انحراف معیار ۱۰۰، نمره SAT می باشند.

۳۲- گزینه «۱» اگر برای تفسیر نمره یک دانش آموز آن نمره را با میانه یا میانگین نمره گروه مرجع مقایسه کنیم، به آن گروه مرجع گروه هنجار می گویند. گروه مرجع از کسانی تشکیل می شود که به گونه ای شبیه به دانش آموز مورد نظر است و فایده هنجار این است که برای قضاوت درباره سطح نسبی عملکرد یک فرد، به عنوان معیار به کار می رود.

33- گزینه «4» در هنجارهای درصدی و نمرات معیار، موقعیت نسبی یک دانش آموز در گروهی که خود عضوی از آن است، نشان داده میشود. بنابراین در این سؤال تعیین نرم (گزینه ۴) مناسب ترین روش می باشد.

۳۴- گزینه «۴»

۳۵- گزینه «۱»

برای بررسی میزان انحراف استاندارد از شیوه ی مقابل بهره می گیریم:

۳۶- گزینه «۴» تعریف هنجار: اگر برای تفسیر یک نمره، آن را با میانه یا میانگین نمره یک گروه مرجع مقایسه کنیم به آن گروه مرجع، گروه هنجار می گویند. مقیاس فاصله ای یعنی اینکه افراد (اشیاء) از لحاظ صفت معین چقدر با یکدیگر فاصله دارند. معروف ترین نمرات هنجار عبارت اند از هنجارهای سنی، هنجارهای کلاسی، هنجارهای درصدی و نمرات معیار (استاندارد). واحدها در طول مقیاس فقط در نمرات معیار برابرند و مقیاس فاصله ای است.

۳۷- گزینه «1» هنجار (نرم) سنی یا معادل سینی بر میانگین یا میانه نمراتی مبتنی است که دانش آموزان در سنین مختلف کسب می کنند و به صورت معادل های سنی نشان داده می شود.

۳۸- گزینه «3» رتبه درصدی، نمره های خام را به نمره هایی تبدیل می کند که به آسانی تفسیر می شوند. رتبه درصدی به سرعت وضعیت فرد را فقط درون یک جامعه معین، مشخص می کند و وقتی دارای معنی است که ماهیت مقایسه گروهی معلوم باشد؛ اما نقاط درصدی، نقاطی در نمره های خام هستند سر که نشان می دهند چند درصد نمره ها پایین تر از نمره مورد نظر قرار می گیرد. میانگین، می تواند داوطل را دچار تردید کند؛ اما با توجه به متن سؤال، که به جایگاه نسبی فرد اشاره دارد، تنها گزینه سوم صحیح است. گزینه ی میانگین در صورتی صحیح بود که به گروه و جایگاه نسبی گروه اشاره می کرد؛ چون در میانگین به نمرات یک گروه توجه شده است و این نمرات مورد بررسی قرار می گیرد.

39-گزینه «3» هوشبهر انحرافی یک هنجار مربوط به نمرات معیار است که دارای میانگین ۱۰۰ و انحراف معیار ۱۵ (یا ۱۶) است. تعریف هوشبهر با بهره هوشی (IQ) یک خارج قسمت است که از فرمول مقابل حساب می شود:

هوشبهر انحرافی از تبدیل نمرات خام به نمرات تراز شده به دست می آید. ابتدا نمرات خام به نمرات استاندارد Z تبدیل می شود، سپس در عدد ثابت ۱۵ (انحراف استاندارد مقیاس وکسلر) یا نمره ۱۶ (انحراف استاندار مقیاس بینه) ضرب و با عدد ۱۰۰ (میانگین وکسلر و بینه) جمع می شود. هوشبهر انحرافی، یک نمره معیار است و بر این عقیده قدیمی که سن ذهنی و سن تقویمی کودکان باید برابر باشند، مبتنی نیست.

براساس مفهوم قدیمی هوشبهر، کودکی که سن تقویمی و سن ذهنی او ارسال کامل است، هوشبهری معادل ۱۰۰ خواهد داشت، اما هوشبهر واقعی این کودک در آزمون استنفورد بینه جز این است.

۴۰-گزینه «1» نمره استاندارد T با میانگین ۵۰ و انحراف استاندارد ۱۰، مقیاسی مثبت و بدون اعشار است که با این فرمول محاسبه می شود:

با توجه به نمره ۲ در توزیع نمرات معیار (Z) جایگزین می کنیم

۴۱- گزینه «۲» رتبه درصدی یا نمره درصدی، وضعیت نسبی فرد را در گروه بر حسب کسانی که نمره پایین تر از او گرفته اند، مشخص می کند. گزینه ۱ به اشتباه است، چون تعداد گروه ۶۰ نفر می باشد. گزینه ۳ نیز اشتباه است، چون وضعیت احمد را با توجه به تعداد سؤالی که پاسخ داده است بیان کرده، در صورتی که گفتیم رتبه درصدی وضعیت فرد را با توجه به گروهی که در آن قرار دارد بیان می کند. به همین دلیل گزینه ۴ نیز اشتباه می باشد. بنابراین تنها گزینه صحیح گزینه ۴ می باشد. این که نمره احمد از ۴۸ نفر گروه بهتر است را نیز با یک طرفین وسطین ساده می توان به دست آورد.

۴۲- گزینه «3» منحنی های دارای کجی (انحنا) را بر اساس دنباله (کشیدگی) آن ها به دو دسته تقسیم می کنند. در صورتی که دنباله راست منحنی طولانی تر از دنباله چپ آن باشد، منحنی دارای کجی مثبت یا کجی به سمت راست می باشد. زمانی که منحنی دارای چنین کجی باشد، فقط تعداد کمی از به افراد گروه دارای نمره بالا هستند و اکثر گروه نمره پایینی دارند. توزیع زمانی دارای کجی منفی است که دنباله چپ آن طولانی تر (کشیده تر) از دنباله راست آن باشد. زمانی که منحنی دارای چنین کجی باشد فقط تعداد کمی از افراد نمره پایینی گرفته اند و نمره اکثر افراد بالاست. در این سؤال نیز چون منحنی دارای کجی منفی است، بنابراین نمره اکثر افراد بالاست و میانگین داده ها در نقطه درصدی بیشتر از ۵۰ درصدی قرار دارد.

۴۳- گزینه «۱» در ارزشیابی هنجاری و آزمون های مبتنی بر ملاک نسبی که در سنجش هوش و استعداد و نیز شخصیت و سایر جنبه های عاطفی به کار می روند، عملکرد فرد با توجه به عملکرد سایر افراد و رتبه ای که در مقایسه با سایر افراد کسب می کند، مورد قضاوت قرار می گیرد. در آزمون های ورودی دانشگاه ها که در آن نمره هر فرد در مقایسه با سایر افراد سنجیده می شود، از این ملاک (نسبی) استفاده می شود.

۴۴- گزینه «3» در هنجارهای سنی و کلاسی و درصدی، واحدها در طول مقیاس برابر نیستند، اما در نمرات معیار واحدها در طول مقیاس برابرند و مقیاس فاصله ای است.

۴۵- گزینه «3» یکی از معایب هنجارهای درصدی مانند هنجارهای کلاس و سنی این است که واحدهای مربوط به صدک ها در طول مقیاس مساوی نیستند. تفاوت صدک های وسط مقیاس (مثلا تفاوت صدک های ۴۵ تا ۵۵) خیلی کمتر از تفاوت صدک های دو انتهای مقیاس مثلا (۸۵ تا ۹۰) است. علت آن است که مقدار زیادی از افراد هر گروه نمراتی بدست می آورند که حوالی وسط مقیاس قرار می گیرند. در حالی که تعداد نسبتا کمی از افراد نمره هایشان خیلی زیاد یا خیلی کم است. در نتیجه با افزایش اندکی در نمره خام دانش آموزی که نمره اش نزدیک به میانگین است، ممکن است مقدار زیادی بر رتبه اش افزوده شود، اما دانش آموزی است که نمره کمتری دارد برای اینکه همان مقدار به نمره اش افزوده شود، مقدار زیادتری به نمرهی خام او اضافه می گردد. برای رفع این مشکل متخصصان روان سنجی از نمرات معیار که موقعیت نسبی یک دانش آموز در یک گروه را با نشان دادن فاصله ی نمره او از میانگین نمرات معلوم می کند، استفاده می نمایند.

۴۶- گزینه «۴» زمانی که نمره های خام را از میانگین کم می کنیم و تقسیم بر انحراف معیار می کنیم، نمره ی Z به دست می آید حال فرض کنید نمره ی خام دانش آموزی در یک آزمون ۵۸، میانگین آن ۵۶ و انحراف معیار ۴ باشد، نمره ی Z برابر ۵/0 می شود (5/0 = = Z). اکنون اگر انحراف معیار را کاهش دهیم (مثلا به ۲ برسانیم)، نمره Z به 1 افزایش می یابد (1 = = Z : نمره Z با مقدار انحراف معیار، رابطه عکس دارد.

۴۷- گزینه «3» ارزشیابی های آموزشی با توجه به ملاک مورد استفاده در آن ها به دو دسته ی هنجار (گروه) مرجع محور و ملاک محور تقسیم می شود که هر سر دوی اینها عملکرد یادگیرنده را براساس یک ملاک (هنجار یا ملاک مطلق) قضاوت می کنند. در ارزشیابی ملاک یا معیار مرجع، نوعی ملاک مطلق (اهداف از قبل تعیین شده) مورد استفاده می شود و از آزمون شوندگان انتظار می رود که به هدف های معینی دست یابند. در ارزشیابی هنجاری یا ترم مرجع، به جای یک ملاک مطلق، ارزشیابی به نوعی ملاک نسبی وابسته است. اگر بخواهیم پیشرفت گروهی دانش آموزان را ارزشیابی کنیم، عملکرد آنها را با یکدیگر مقایسه می کنیم، نه با یک ملاک معین از قبل تعیین شده.

۴۸- گزینه «۲» یکی از معایب هنجارهای درصدی، مانند هنجارهای کلاسی و سنی، این است که واحدهای مربوط به صدک ها در مقیاس مساوی نیستند. تفاوت صدک های اواسط مقیاس، خیلی کمتر از تفاوت صدک های انتهای مقیاس است. علت آن است که تعداد زیادی از افراد هر گروه نمراتی به دست می آورند که در حوالی وسط مقیاس قرار دارند، در حالی که نمره های تعداد نسبتا کمی از افراد، خیلی زیاد یا خیلی کم است. در نتیجه با افزایش اندک نمره خام دانش آموزی که نمره اش نزدیک میانگین است، ممکن است مقدار زیادی بر رتبه اش اضافه شود؛ اما دانش آموزی که نمره بالایی دارد، برای این که همان مقدار به رتبه اش اضافه شود، باید بر نمره خام او مقدار زیادی افزوده شود.

۴۹- گزینه «۴» نرم های استاندارد، موقعیت نسبی یک دانش آموز را در یک گروه با نشان دادن فاصله نمره او از میانگین نمرات، معلوم می کنند. نرم های استاندارد، عملکرد هر دانش آموز را در یک آزمون بنابر اختلاف عملکرد او از میانگین گروه بر حسب واحد انحراف معیار نشان می دهند.

۵۰- گزینه «۲» رتبه ی درصدی یا نمره درصدی، وضعیت نسبی فرد در گروه را بر حسب کسانی که نمره ی پایین تر از او گرفته اند، مشخص می کند. هر نمره ی خام یک رتبه ی درصدی دارد؛ مثلا نمره ی خام ۱۵ یک دانش آموز دارای رتبه ی درصدی ۶۵ خواهد بود، به شرط اینکه ۶۵ درصد دانش آموزان سرگروه، تمره خام کمتر از ۱۵ گرفته باشند. به بیان دیگر، نمره ی این دانش آموز (۱۵)، از نمره ی ۶۵ درصد گروه هنجار بیشتر است. در رتبه درصدی و ملاک، دو اصطلاح مکمل یکدیگرند. رتبه های درصدی، عملکرد نسبی آزمون شوندگان را نشان می دهند، در حالی که ملاکها به سر نقاطی روی توزیع نمرات اشاره می کنند. ملاک را می توان نوعی رتبه در یک گروه ۱۰۰ نفری تلقی کرد که شمارش آن از پایین شروع می شود. و هرچه ملاک کوچک تر باشد، به همان نسبت جایگاه فردی که ملاک معرف نمره خام اوست، در گروه پایین تر است. در اینجا چون رتبه درصدی فرد ۵۰ است، یعنی این فرد از ۵۰ درصد افراد نمره بیشتری کسب کرده است.

51- گزینه «۱» تبدیل نمره های خام به دو صورت اساسی انجام می شود: خطی و غیرخطی. تبدیل خطی از طریق معادله خطی از نوعb y = ax تعریف می شود. در اینجا a و b ثابت هستند و x وy به نمره تبدیل شده است. در تبدیل خطی شکل توزیع نمره های تبدیل شده همانند شکل نمره های خام است. تبدیل خطی، اندازه همبستگی را نیز تغییر نمی دهد. ساختن نمره معیار و نمره های معیار شده به تبدیل خطی نمره های خام نیاز دارد.

تبدیل غیرخطی، به صورت معادله خطی بیان نمی شود. برای مثال یک تبدیل غیرخطی از X است. تبدیل غیرخطی، شکل توزیع نمره ها و همبستگی ها را نیز تعبیر می دهد. درصدها، نمرات سنی و کلاسی، جدول انتظار، نمره های بهنجار شده و مقیاس های فاصله ای مساوی، جزء تبدیل خطی هستند.

نمره ی T یکی از نمرات معیار است؛ با میانگین ۵۰ و انحراف معیار ۱۰.

گزینه ۳ یک گزینه کاملا انحرافی است. گرچه قرار دادن نمرات در یک توزیع طبیعی آن ها را قابل مقایسه می کند، هیچ گاه نمرات را همگن (متجانس) نمی سازد. در واقع تبدیل نمرات به یک نمره طبیعی همچون Z تغییری در واریانس ایجاد نمی کند.

۵۲- گزینه «۱» برای ۱ = Z، سطح بین میانگین و ۱= Z، 3413/0 است. نسبت سطح پایین میانگین، ۵/0 است. نسبت کل سطح پایین 1=Z ، 3413/0 5/0است. نسبت بام های این نقطه نیز برابر است با

1587/0 = 8413/0- 1

۵۳- گزینه «۲» در توزیع T (میانگین: ۵۰ و انحراف معیار: ۱۰) نمره فردی که ۹۵% افراد از او نمره بیشتری گرفته اند، 66/33 است.

۵۴- گزینه «۳» رتبه درصدی چیست؟ گاهی لازم است موقعیت فرد را در ویژگی مورد نظر نسبت به مجموع افراد بدانیم، به همین دلیل رتبه درصدی ایجاد شد. طبق معادله رتبه درصدی، نمره حدودی به دست آمده ۸۴ است.

در منحنی نرمال استاندارد، ۸۴ درصد نمرات زیر1 Z= قرار دارد.

۵۵- گزینه «1» طبق اعداد پیشنهادی سؤال، نمره Z در این آزمون عملکرد بهتر آزمودنی را نشان می دهد.

آزمون فصل پانزدهم

۱- نمره ۳۰ در مقیاس T معادل چه نمره ای در مقیاس Z است؟

1) 5/0

2) 1-

3) 5/1+

4) 2-

۲- در یک آزمون هوشی که دارای انحراف معیار ۱۵ است، هوشبهر ۸۵ برابر با چه نمره ای در مقیاس T است؟

1) 50

2) 40

3) 70

4) 60

۳- اگر دانش آموزی در آزمون پیشرفت تحصیلی، رتبه ۸۸ درصدی را کسب کرده باشد یعنی:

1) دانش آموز به ۸۸ درصد سؤال های آزمون پاسخ درست داده است.

2) نمره فرد و انحراف معیار از میانگین بالاتر است.

3) یعنی از ۸۸ درصد افراد شرکت کننده نمره بیشتری گرفته است.

۴) گزینه های ۲ و ۳

۴- اگر سن عقلی کودکی ۸ ساله به اندازه کودک ۱۲ ساله باشد، بهره هوشی او کدام است ؟

1) 80

2) 75

3) 120

4) 150

5- در کدام یک از هنجارهای زیر واحد نمره ها در طول مقیاس نابرابرند ؟

۱) نمره های T

2) نمره SAT

۳) رتبه های درصدی

۴) هوشبهر انحرافی

6- مفهوم عبارت «علی در آزمون پیشرفت تحصیلی رتبه ۵۸ را کسب کرده است» چیست؟

1) على از ۵۸ درصد افراد شرکت کننده نمره بیشتری گرفته است.

۲) به ۶۸ درصد سؤال های آزمون پاسخ صحیح داده است

۳) نمره على 3 انحراف استاندارد از میانگین بالاتر است.

۴) نمره علی ۲ انحراف معیار بالاتر است.

۷- اگر مریم دارای رتبه درصدی ۲۰% باشد، چنین برداشت می شود که ...........

۱) مریم به ۲۰% سؤالات پاسخ صحیح داده است.

۲) مریم از ۸۰ نفر از همکلاسی هایش بالاتر است.

۳) در بین همکلاسی های خود، مریم نفر ۲۰ می باشد.

۴) از ۲۰% افراد گروه خود بالاتر است.

۸- اگر فردی در آزمونی با میانگین ۱۵ و انحراف استاندارد ۵ نمره ۲۰ بگیرد، نمره استاندارد وی چیست؟

1) 2

2) 1-

3) 1

4) 2-

۹- اگر فردی دارای سن عقلی ۱۰ و سن تقویمی ۲۰ باشد، هوشبهر او چند است؟

1) 100

2) 80

3) 85

4) 50

۱۰- قوی ترین عملکرد در این آزمون مربوط به کدام گزینه می باشد؟

1) 5/1z = +

2) 3- z =

3) 1- z =

4) 2z =

۱۱- نمره ۷۰ در مقیاس T معادل کدام نمره در مقیاس Z می باشد؟

1) 1

2) 2-

3) 2

4) 1-

۱۲- در یک توزیع بهنجار با میانگین ۲۵ و انحراف استاندارد ۵ نمره ۵۰ معادل کدام نمره در مقیاس «T» می باشد؟

1) 65

2) 70

3) 45

4) 80

۱۳- نمره استاندارد ۹ بخشی و هوشبهر انحرافی به ترتیب جزء کدام مقیاس می باشند؟

1) رتبه ای - فاصله ای

۲) رتبه ای – رتبه ای

۳) فاصله ای - رتبهای

۴) فاصله ای - فاصله ای

۱۴- نمره فردی در آزمون T 40 می باشد. نمره این فرد در مقیاس با میانگین ۳۰ و انحراف استاندارد ۵ چند می باشد؟

1) 23

2) 25

3) 4

4) 50

۱۵- نمرات کدام مورد در مقیاس فاصله ای قرار می گیرد؟

1) هوشبهر

۲) ۹ بخشی

۳) هنجار کلاسی

۴) هنجار سنی

۱۶- اگر سن عقلی کودکی ۱۰ ساله به اندازه کودکی ۵ ساله باشد، بهره هوشی او چیست؟

1) 70

2) 50

3) 120

4) 200

۱۷- کدام یک از روش های نمره گذاری، جنبه کیفی دارد؟

1) نمره هوشبهر

۲) نمره NEC

۳) نمرات ۹ بخشی

۴) نمره SAS

۱۸- به کمک کدام معیار می توان به نقاط قوت و ضعف دانش آموزان پی برد؟

۱) نمره SAS

۲) نمره NCE

۳) نیمرخ

۴) IQ

۱۹- ویژگی های یک منحنی بهنجار عبارتند از این که ....................

۱) نما، میانه، میانگین روی هم قرار می گیرند.

2) دو قسمت منحنی به خط وسط آن قرینه است.

۳) 26/68% از کل مساحت منحنی بین ۱+ و ۱- قرار دارد.

۴) هر سه گزینه

۲۰- نمره Z فردی برابر ۲۵ می باشد، نمره SAT این فرد چند است؟

1) 700

2) 650

3) 750

4) 800

فصل شانزدهم

«روایی آزمون»

هر آزمون با هر قصد و منظوری که تهیه شود، باید دارای ویژگی هایی مثل عینیت، سهولت اجرا و نمره گذاری، عملی بودن و روایی و پایایی باشد که روایی و پایایی از تمام این ویژگی ها مهم ترند.

تعریف روایی و انواع آن

اصطلاح روایی به هدفی که آزمون برای تحقق بخشیدن به آن درست شده است، اشاره دارد. آزمونی دارای روایی است که آنچه را در نظر دارد، اندازه بگیرد؛ برای مثال آزمونی که برای انتخاب داوطلبان یک شغل به کار می رود، در صورتی دارای روایی است که نمره آن با عملکرد آتی آزمودنی ها در آن شغل رابطه قوی داشته باشد، یا مثلا یک آزمون هوش روا، فقط میزان بهره هوشی را اندازه می گیرد. یک آزمون ممکن است از جهتی روا باشد اما از جهات دیگر روا نباشد، پس روایی امری نسبی است؛ نه یک ویژگی که هست یا نیست.

مثال 1: اگر یک ابزار اندازه گیری نتواند خصیصه مورد نظر را اندازه گیری کند، فاقد کدام ویژگی است؟

۱) روایی

2) پایایی

۳) قدرت تشخیص

۴) نرم

پاسخ: گزینه «1» آزمونی دارای روایی است که آنچه را در نظر دارد اندازه بگیرد.

به طور کلی روایی به سه دسته: محتوایی، ملاکی و سازه تقسیم می شود که هر کدام با یکی از سوالات زیر مطابقت دارد:

1- آیا سؤالات آزمون از محتوای درس نمونه کاملی ارائه می دهند: روایی محتوایی

۲- آیا نمرات آزمون عملکرد فعلی یا آتی دانش آموز را پیش بینی می کنند: روایی ملاکی

3- آیا نمرات آزمون به مفاهیم نظری یا سازه هایی که آزمون برای سنجش آنها درست شده است، مربوط اند: روایی سازه روش های تعیین روایی محتوایی که روایی صوری هم مشتقی از آن است، به روش های قضاوتی معروف اند؛ زیرا در آن ها خود ابزار اندازه گیری قضاوت می شود تا تعیین شود ابزار اندازه گیری تا چه حد می تواند ویژگی های موری نظر را بسنجد.

مثال ۲: در کدام یک از انواع روایی، خود ابزار اندازه گیری مورد قضاوت قرار می گیرد و به روش قضاوتی معروف می باشد؟

1) ملاکی

۲) محتوایی

۳) سازه س

۴) گزینه های1و ۲

پاسخ: گزینه «۲» در روایی محتوایی، خود ابزار اندازه گیری قضاوت می شود و به روش قضاوتی نیز معروف است. روش های تعیین روایی ملاکی که خود به دو دسته روایی پیش بین و روایی همزمان تقسیم می شود و روایی سازه، روش های تجربی نام دارند؛ زیرا در آن ها ماهیت پاسخ هایی که به سؤالات آزمون داده می شود، مورد بررسی قرار می گیرد تا تعیین شود که این پاسخ ها تا چه حد می توانند شواهد لازم را در باره این که آیا ابزار موردنظر واقعا ویژگی موردنظر را می سنجد، فراهم می کنند.

مثال 3: اگر یک ابزار اندازه گیری (مانند آزمون) نتواند خصیصه ی مورد نظر را اندازه گیری کند، فاقد کدام ویژگی است؟

۱) روایی (Validity)

۲) پایایی (Reliability)

۳) همسانی درونی (Internal Consistency)

۴) قابلیت ثبات (Stability)

پاسخ: گزینه «۱» اصطلاح روایی به هدفی اشاره دارد که آزمون برای تحقق بخشیدن به آن ترتیب داده شده است. آزمونی دارای روایی است که برای اندازه گیری آنچه مورد نظر است کافی و مناسب باشد. آزمون روا، محتوا و اهداف مورد نظر را به درستی اندازه گیری می کند. تفاوت روایی و پایایی در تعاریف آن ها می باشد. پایایی به میزان دقت و تکرارپذیری آزمون اشاره دارد؛ بنابراین با روایی ارتباط دارد. یک آزمون باید ابتدا دقیق و قابل اعتماد باشد (پایایی) و سپس اهداف و محتوا را به درستی اندازه بگیرد (روایی).

روایی محتوایی

روایی محتوایی یعنی نمونه سوال های مورد استفاده در یک آزمون تا چه حد معرف کل جامعه سؤال های ممکنی است که می توان از محتوا یا موضوع مورد نظر تهیه کرد. هر چه آزمون از این لحاظ بهتر باشد، دارای روانی بیشتری است. یعنی آیا سؤال های یک آزمون، نمونه کاملی از هدف ها و محتوای درس هستند که استفاده از جدول مشخصات بهترین راه برای انجام این کار است از روایی محتوایی بیشتر در آزمون های پیشرفت تحصیلی استفاده می شود.

روش تعیین روایی محتوایی: قضاوت متخصصان است. در تعیین این روایی از محاسباات آماری استفاده نمی شود. با این حال قضاوت های مربوط به روایی محتوایی نه قطعی هستند و نه آزمون ها از سطوح مختلف روایی برخوردارند و متخصصان در قضاوت های خود همیشه با هم توافق نظر ندارند.

مثال ۴: در آزمون های پیشرفت تحصیلی، کدام روش برای تعیین روایی (Validity) به کار می رود؟

1) محتوایی

۲) سازه

۳) همزمان

۴) پیش بین

پاسخ: گزینه «۱» روایی اصطلاحی است که به هدفی که آزمون برای تحقق بخشیدن به آن درست شده است اشاره می کند. اگر یک آزمون محتوا و اهداف مطلب مورد نظر را به خوبی اندازه گیری کند، به آن آزمون روا می گویند. انواع روایی عبارت است از: روایی محتوایی - روایی ملاکی و روایی سازه. برای تعیین روایی محتوایی که به روش های قضاوتی معروف اند، خود ابزار اندازه گیری قضاوت می شود. روایی محتوایی به این موضوع اشاره می کند که نمونه سؤال های مورد استفاده در یک آزمون تا چه اندازه معرف کل جامعه سؤال های ممکن است که میتوان از موضوع یا محتوای مورد نظر تهیه کرد. آزمون های پیشرفت تحصیلی در یک زمان محدود به کار می روند و از آنجا که هیچ آزمونی نمی تواند تمام محتوا و کلیه هدف های یک درس را شامل شود پس بهترین روش، روش روایی محتوایی می باشد.

مثال ۵: روابی محتوایی درباره کدام یک از آزمون های زیر بیشتر معنا پیدا می کند؟

1) شخصیت

۲) رغبت

3) پیشرفت تحصیلی

۴) استعداد

پاسخ: گزینه «3» از آنجا که همه آزمون های پیشرفت تحصیلی در یک زمان محدود به کار می روند و از آنجا که هیچ آزمونی نمی تواند تمام محتوا و کلیه هدف های یک درس را شامل می شود، روایی محتوایی به این مطلب اشاره می کند که نمونه سؤال های مورد استفاده در یک آزمون تا چه حد معرف کل جامعه ی سوال های ممکن است که از این محتوا با موضوع می توان تهیه کرد. مثلا اگر یک معلم فیزیک بخواهد از میزان یادگیری شاگردان خود اطلاع پیدا کند، می تواند یک آزمون پیشرفت تحصیلی روا بسازد.

مثال ۶: کدام روش بررسی روایی، بیشتر در آزمون های فراگیری به کار گرفته می شود؟ .

۱) سازه

۲) صوری

3) محتوایی

4) ملاکی

پاسخ: گزینه «3» روایی کاربردی در آزمون های فراگیری، روایی محتوایی می باشد برای تعیین روایی محتوایی یک آزمون از قضاوت متخصصان درباره ی این موضوع که سؤال های آزمون تا چه حد معرف محتوا و اهداف برنامه ها هستند، استفاده می شود. در روایی محتوایی به این مطلب اشاره می شود که نمونه سؤال های مورد استفاده در یک آزمون تا چه حد معرف کل جامعه ی سؤال های ممکن است؛ بنابراین در آزمون های پیشرفت تحصیلی از روایی محتوایی استفاده می شود.

روایی صوری (نوعی از روایی محتوایی است)

روایی صوری: یعنی سوال های آزمون تا به حد در ظاهر شبیه به موضوعی هستند که برای اندازه گیری آن تهیه شده اند. روایی صوری به درجه ای از منطقی بودن یک آزمون، از نظر آزمون شوندگان گفته می شود. مثلا اگر آزمونی در مورد مباحث، حسابداری است. اگر سؤال ها درباره ترازنامه و دفتر حساب باشد، آزمون شوندگان آن آزمون را منطقی تلقی می کنند و این آزمون دارای روایی صوری نمی تواند، نوعی روایی باشد. بلکه تنها یک ویژگی آزمون است، که در پاره ای، مواقع وجود آن مفید به نظر می رسد. روایی صوری روی انگیزش آزمون شوندگان تأثیر می گذارد و گاهی وقت ها هم مثل تشخیص بیماری های روانی، استفاده از روایی صوری جایز نیست.

روش تعیین روایی صوری: چون روایی صوری هم نوعی روایی محتوایی است، پس روش تعیین روایی صوری هم قضاوت متخصصان است.

مثال 7: این نوع روایی به درجه ای از منطقی بودن یک آزمون از نظر آزمون شوندگان گفته می شود.

۱) ملاکی

۲) سازه

۳) محتوایی

4) صوری

پاسخ: گزینه «۴» روایی صوری به درجه ای از منطقی بودن یک آزمون از نظر آزمون شوندگان گفته می شود.

مثال 8: در کدام یک از روش های رواسازی (validation) آزمون از محاسبه ضریب همبستگی استفاده نمی شود؟

1) پیش بین

۲) سازه

3) ملاکی

4) محتوا

پاسخ: گزینه «۴» ابتدا باید به تأکید سؤال درباره استفاده نشدن توجه کرد. بنابراین موضوع سؤال استفاده نکردن از ضریب همبستگی در رواسازی است. با این مقدمه، تنها در روایی محتوایی و روایی صوری، عملیات آماری کاربردی ندارد. شایان ذکر است در روایی محتوایی از نظر متخصصان استفاده می شود و در روایی صوری نظر آزمودنی ها مهم می باشد. ضمن آن که روایی صوری از روایی محتوایی مشتق است؛ یعنی در مجموع از نظر متخصصان در تعیین روایی استفاده می شود.

مثال9: پیشرفت تحصیلی بیشتر به کدام نوع روایی ارتباط پیدا می کند؟

۱) صوری

۲) پیش بینی

3) محتوایی

4) همزمان

پاسخ: گزینه «۳» روایی محتوایی

روایی ملاکی

منظور از روایی ملاکی، میزان ارتباطی است که بین نمرات حاصل از یک آزمون با نمرات حاصل از یک آزمون یک وسیله از اندازه گیری دیگر وجود دارد. هرگاه از نمرات یک آزمون برای پیش بینی عملکرد افراد در آزمونی دیگر استفاده کنیم، با روایی ملاکی سروکار داریم. آزمون دوم که عملکرد فرد در آن پیش بینی می شود، ملاک نام دارد. هر قدر رابطه بین نمرات آزمونی که میخواهیم روایی ملاکی آن را بسنجیم با نمرات آزمون ملاک بیشتر باشد، روایی ملاک هم به همان نسبت بیشتر می شود.

مثال ۱۰: در کدام یک از انواع روایی (اعتبار)، رابطه عملکرد در یک آزمون با یک ضابطه مورد بررسی قرار می گیرد؟

۱) صوری

۲) سازه ای

۳) محتوا

۴) ملاکی

پاسخ: گزینه «۴» روایی ملاکی به میزان ارتباط بین نمرات حاصل از یک آزمون بـا نـمـرات حاصل از یک آزمـون یـا وسیله ی اندازه گیری دیگر اشاره دارد؛ یعنی هر زمان از نمره آزمون فیزیک برای پیش بینی عملکرد آزمون شیمی فردی استفاده شود، با روایی ملاکی سرو کار داریم. روایی صوری به نوع خاصی از روایی محتوایی اشاره دارد و به ارتباط شکل سؤالات آزمون و موضوع مورد آزمون اشاره دارد.

روایی محتوایی نیز به این مطلب اشاره دارد که: نمونه سؤال های مورد استفاده در یک آزمون تا چه حد معرف کل جامعه سؤال های ممکن است که می توان از محتوا یا موضوع مورد نظر تهیه کرد.

روایی ملاکی به ۲ نوع روایی پیش بینی و روایی همزمان تقسیم می شود:

روایی پیش بینی:

هر زمان از نمرات یک آزمون برای پیش بینی عملکرد افراد در آزمون دیگر استفاده کنیم و این پیش بینی برای آینده باشد، یعنی نمرات آزمـون مـلاک را بعد از گذشت یک فاصله زمانی از آزمون اول گردآوری کنیم، به آن روایی پیش بینی می گوییم. روایی پیش بینی، ویژه آزمون هایی است که آن آزمون ها برای پیش بینی موفقیت افراد در امور تحصیلی یا شغلی به کار می روند. مثلا اگر از نمرات دوره دبیرستان فردی برای بیش بینی موفقیت او در دانشگاه استفاده کنیم.

روش تعیین روایی پیش بینی: برای تعیین روایی پیش بینی ۲ آزمون ضروری هستند که یکی از آزمون ها، یعنی آزمـون مـلاک در آینـده اجـرا می شود. ضریب همبستگی بین نمرات حاصل از آزمون اول و آزمون ملاک، شاخص روایی پیش بینی آزمون موردنظر است و هر چقدر این ضریب همبستگی بزرگ تر باشد، روایی بیشتر خواهد بود و بالعکس.

در تعیین روایی پیش بینی هم می توان از چندین ملاک استفاده کرد تا روایی را افزایش داد، مثلاً برای تعیین موفقیت یک معلم می توان از چند ملاک مثـل نظر مدیر، همکاران، دانش آموزان و میزان پیشرفت تحصیلی دانش آموزان استفاده کرد، هم می توان از چندین آزمون پیش بینی استفاده کرد تا روایی افزایش یابد، مثلاً در پیش بینی موفقیت افراد در دانشگاه علاوه بر نمرات دوره دبیرستان آن ها از نمرات آن ها در آزمون های هوش هم می توان استفاده کرد.

مثال۱۱: برای تعیین «روایی پیش بینی» عملکرد یادگیرندگان در یک هدف آموزشی بر اساس هدف آموزشی دیگر، از کدام روش آماری استفاده می شود؟

۱) آزمون t

۲) تحلیل واریانس

۳) برآورد میانه ها

۴) ضریب همبستگی

پاسخ: گزینه «۴» روایی پیش بینی ویژه آزمون هایی است که برای موفقیت افراد در امور تحصیلی یا شغلی کاربرد دارند. آزمـون هـایی کـه بـرای ایـن منظور به کار می روند، مختلف هستند. به دلیل آنکه در روایی پیش بینی همیشه ملاکی مورد نظر است که نمرات آزمون پیش بینی کننده را می توان با آن مقایسه کرد، به آن روایی ملاکی نیز می گویند. روایی ملاکی به فن تجربی مطالعه ی رابطه بین نمرات آزمون و نوعی اندازه بیرونی مستقل (ملاک) مربوط می شود. در نتیجه برای تعیین روایی پیش بینی یک آزمون، دو آزمون ضروری هستند. ضریب همبستگی بین نمرات حاصل از این دو آزمون، شاخص روایی پیش بینی آزمون مورد نظر است. ممکن است تعیین ضریب همبستگی در بعضی آزمون ها کار ساده ای و در بعضی آزمون ها کار دشواری باشد.

مثال ۱۲: روایی پیش بین را از چه طریقی می توان سنجید؟

1) همبستگی نتایج آزمون با نتایج آینده

2) قضاوت داوران متخصص

3) روش دو دو نیمه کردن

4) آزمون های معادل

پاسخ: گزینه «۱» روایی پیش بین ویژه آزمون های موفقیت افراد در امور تحصیل و یا شغلی به کار می روند. این آزمون ها برای مقاصد پیش بینی به کار می روند. مثلا می توان از آزمون های ورودی دانشگاه ها به عنوان آزمون پیش بینی نام برد که می تواند نمرات فرد در آزمون کنکور را ملاکی برای پیش بینی نمرات او در دانشگاه و امتحانات پایان ترم قرار دهد.

مثال ۱۳: ضریب همبستگی بین نمره های امتحانات ورودی دانشگاه و نمره های دروس دانشگاهی پذیرفته شدگان شاخصی است از:

۱) پایایی (reliablity) امتحانات دروس دانشگاهی

۲) روایی همزمان امتحانات ورودی

۳) روایی (Validity) پیش بینی امتحانات ورودی

۴) روایی پیش بینی امتحانات دروس

پاسخ: گزینه «۳» زمانی که بین نمره ورودی دانشگاه و نمره های کسب شده در امتحانات دانشگاه ضریب همبستگی برقرار باشد از روایی پیش بین استفاده می شود. روایی پیشبین به این نکته توجه دارد که نمرات یک آزمون تا چه اندازه نمرات آزمـون دیگر (آزمـون مـلاک) را پیش بینی می کنند. روایی پیش بینی ویژه آزمون هایی است که برای مقاصد پیش بینی کاربرد دارند و از موفقیت های فعلی یا قبلی برای پیش بینی پیشرفت تحصیلی استفاده می کنند که متن سؤال به یکی از این موارد اشاره دارد.

مثال ۱۴: تست های استعداد بیشتر به کدام نوع روایی ارتباط پیدا می کنند ؟

1) صوری

۲) محتوایی

3) همزمان

۴) پیش بینی

پاسخ: گزینه «۴» روایی پیش بینی.

روایی همزمان:

در روایی همزمان به جای تعیین رابطه ۲ آزمون پس از گذشت یک فاصله زمانی، رابطه بین ۲ آزمون به طور همزمان تعیین می شود. اگر به جای مقایسـه نـمـرات ر یک آزمون با نمرات آزمون دیگری که در آینده اجرا خواهد شد، نمرات آزمون اول را با نمرات آزمونی که همزمان یا با فاصله زمانی کوتاهی اجرا شده است ، مقایسه کنیم، به آن روایی همزمان می گویند. هدف از تعیین روایی همزمان این است که تعیین کنیم آیا می توان یک آزمون را به جای آزمون دیگر استفاد کرد.

روش تعیین روایی همزمان: محاسبه ضریب همبستگی بین نمرات آزمون اول و نمرات آزمون ملاک که به شکل همزمان با آزمون اول به دست آمـده اسـت. رابطه حاصل بین این دو آزمون که همزمان اجرا شده، نشان دهندهی روایی همزمان است.

مثال ۱۵: ضریب همبستگی بین نمره های کنکور دانشجویان رشته روانشناسی و نمره های دانشگاهی آنان در دروس روان سنجی، بیانگر کدام نـوع روایی (Validity) است؟

۱) همزمان آزمون کنکور

۲) پیش بین آزمون کنکور

3) همزمان آزمون درس روان سنجی

4) پیش بین آزمون درس روان سنجی

پاسخ: گزینه «۴» به طور کلی روایی پیش بینی ویژه آزمون هایی است که برای پیش بینی موفقیت افراد در امور شغلی یا تحصیلی کاربرد دارد. آزمون هایی که به این منظور به کار می روند، مختلف هستند. ممکن است آزمون پیشرفت تحصیلی فعلی و یا قبلی برای پیش بینی پیشرفت تحصیلی آینـده به کار گرفته شوند. مثلا ممکن است زمانی از نمرات ورودی دانشگاه (آزمون کنکور) برای پیش بینی پیشرفت تحصیلی او در جریان دوران تحصیل در دانشگاه استفاده میشود یا از نمرات کسب شده فرد در دوره راهنمایی برای پیش بینی موفقیت هـای بعدی او در دبیرستان استفاده شـود. بنابراین ایـن آزمون ها باید از خاصیت روایی پیش بینی برخوردار باشند. گزینه ی ۲ میتواند باعث انتخاب اشتباه شود؛ چون به پیش بین اشاره دارد ولی چون به مورد آزمون کنکور اشاره دارد، نمی تواند گزینه ی درست باشد. بنابراین نتیجه می گیریم اگر نمره های درس روان سنجی فرد در دانشگاه مانند نمره کسب شده در کنکور بالا باشد، بین نمرات روان سنجی و آزمون کنکور همبستگی وجود دارد.

مقاسیه روایی پیش بینی با روایی همزمان:

در روایی پیش بینی، داده های مربوط به ملاک در آینده به دست می آیند، به این خاطر نمرات آزمون واقعا برای پیش بینی به کار می روند، اما در روایی همزمان چون داده های پیش بینی کننده و پیش بینی شونده به طور همزمان جمع آوری می شوند، هدف آزمون پیش بین، پیش بینی ملاک نیست، بلکه هدف جانشین کردن آزمون پیش بین با آزمون ملاک است و این جانشین کردن زمانی مفید است که بین نمرات آزمون و نمرات آزمون ملاک همبستگی بالایی وجود داشته باشد و این جانشینی از لحاظ اقتصادی و زمانی به صرفه باشد. رابطه حاصل بین این دو آزمون اجرا شده نشان دهنده ی روایی همزمان است.

مثال ۱۶: اگر هدف این باشد که بتوانیم یک آزمون را به جای آزمون دیگری استفاده کنیم از چه نوع روایی استفاده خواهیم کرد؟

۱) پیش بینی

۲) سازه

۳) محتوا

۴) همزمان

پاسخ: گزینه «۴» انواع روایی ها به سه دسته ملاکی، محتوایی و سازه تقسیم بندی می شوند که روایی ملاکی شامل روایی همزمان و پیش بینی می باشد . تعریف روایی همزمان این است که آیا یک آزمون یا وسیله ی اندازه گیری، جانشین مناسبی برای آزمون یا وسیله ی اندازه گیری دیگری هست یا خیر. هدف روایی همزمان این است که آیا می توان در بین دو آزمون، یکی از دو آزمون را به جای آزمون دیگری مورد استفاده قرار دارد یا خیر؟

با توجه به سؤال گزینه ۴ صحیح است. با توجه به اینکه روایی ملاکی شامل روایی همزمان و پیش بین است.

مثال 1۷: اگر هدف یک روان سنج این باشد که یک آزمون را جایگزین آزمون دیگر کند، از کدام نوع روایی بهتر است، استفاده کند؟

1) محتوا

2) سازه

۳) همزمان

۴) پیش بین

پاسخ: گزینه «3» متخصصان روان سنجی، دو نوع روایی را با اصطلاح روایی ملاکی معرفی می کنند: روایی پیش بینی و روایی همزمان این دو با یکدیگر شباهت زیادی دارند اما به نوعی می توان بین آن ها تمایز قائل شد. روایی پیش پین به موقعیتی اشاره می کند که در آن داده های مربوط به ملاک در آینده به دست می آیند. بنابراین، نمرات آزمون واقعا برای انجام نوعی پیش بینی به کار می رونکرد، اما در روایی همزمان، داده ها در یک زمان جمع آوری تر می شوند و هدف آنها پیش بینی نمی باشد. برای مثال هدف از تعیین روایی همزمان این است که مثلا پرسشنامه تشخیص بیماری روانی یک فرد که توسط یک دانشجوی جوان پر می شود، می تواند جای معاینات روانپزشکی معروف را برای تشخیص نوع بیماری بگیرد یا خیر؟ این مثال بسیار واضح، می تواند در ذهن داوطلب باقی بماند تا زمان روبه رو شدن با این تست، گزینه ی ۴، پیش بین را انتخاب نکند و به سوی انتخاب گزینه ی صحیح هدایت شود.

ویژگی های مطلوب ملاک به ترتیب اهمیت

١- ربط داشتن: یک آزمون ملاکی، با توجه بر قابلیت انطباق آن با موقعیت واقعی شخص قضاوت می شود؛ برای این کار، هیچ گونه شواهد تجربی وجود ندارد و باید به قضاوت تخصصی افراد اطمینان کنیم و ممکن است میان قضاوت های افراد هم اختلاف نظر وجود داشته باشد.

٢- بی طرفی: یعنی همه افراد در کسب نمره خوب، فرصت مساوی داشته باشند. از مثال های مربوط به عدم بی طرفی می توان به تفاوت بین تجهیزات و شرایط کار برای کارگران یک کارخانه، یا تفاوت بین کیفیت آموزش داده شده به دانش آموزان تا به آنجا که در ارزشیابی از کار با کیفیت آموزش، نمره حاصل به عوامل مربوط به شرایط کار مربوط می شود تا شایستگی کارگر، یا به کیفیت آموزش مربوط می شود تا به شایستگی دانش آموز. همچنین بی طرفی زمانی نقض می شود که نمره ملاک تحت تأثیر دانش مربوط به نمره ی پیش بینی قرار می گیرد.

۳- پایایی: یک آزمون یا وسیله اندازه گیری باید باثبات و قابل تکرار باشد.

۴- در دسترس بودن: دستیابی به اندازه ملاک عملی و میسر باشد. درباره این ویژگی با توجه به زمان، هزینه و امکانات موردنیاز قضاوت می شود؛ بنابراین، ملاکی که سهولت بیشتر در چهار چوب امکانات سهل تر قابل دسترسی باشند، ملاک بهتری است.

روایی سازه

یک آزمون وقتی دارای روایی سازه است که نمرات حاصل از اجرای آن به مفاهیم یا سازه های نظریه موردنظر مربوط باشد. برای مثال، یک آزمون با پرسشنامه اضطراب در صورتی دارای روایی سازه است که نمرات حاصل از آن به سازه هایی که در اضطراب آمده اند، ارتباط داشته باشد. روایی سازه بیشتر از روایی محتوایی و روایی پیش بینی، جنبه نظری دارد. سازه یک مفهوم است، اما در این مبحث سازه یک مفهوم اضافی بر مفهوم دارد و آن این است که سازه برای مقاصد ویژه علمی و به طور عمد و از روی آگاهی ابداع می شود.

روش تعیین روایی سازه:

برای تعیین روایی سازه، نخست باید سازه یا متغیر موردنظر را تعریف کرد. پس از تعریف کردن سازه برای اندازه گیری آن آزمونی تهیه می شـود. سپس با توجه به پیش بینی های نظریه، نمرات حاصل از این آزمون باید با متغیرهای دیگر موردبحث در نظریه، روابط پیش بینی شده ای را نشان دهند؛ اگر وجود ایـن روابط نشان داده شد، آزمون دارای روایی سازه است. داده شد، آزمون راه های مختلف برای تعیین روایی سازه است.

راه های مختلف برای تعیین روایی سازه

۱- تعیین همبستگی آزمون با سایر آزمون های روا: یعنی همبستگی بین نمرات آزمون موردنظر را با آزمونی که برای همان سازه ساخته شده و دارای روایی هم هست، تعیین کنیم؛ اگر این همبستگی بالا بود فرض می شود این 2 آزمون، سازه واحدی را اندازه می گیرند و آزمون جدید هـم، بـرای سـازه مـورد نـظـر رواست که به این روایی، روایی همگرا می گوییم. از سوی دیگر آزمون موردنظر باید با آزمون های دیگری که سازه های متفاوتی را اندازه گیـری مـی کننـد، همبستگی نداشته باشد. این روایی را، روایی واگرا یا تمییزی می گوییم.

۲- تمایز سنی: مفهوم تمایز سنی این است که اگر یکی از ویژگی های انسان با بالارفتن سن او پیشرفت نشان می دهد، نمرات آزمونی که برای اندازه گیری این ویژگی درست شده است، این مطلب را باید تأیید کنند؛ مثلا با بالا رفتن سن، کودکان می توانند به سؤالات بیشتری در موش پاسخ دهند و آزمون موردنظر هم باید این ویژگی را نشان دهد. البته اگر سن یک عامل مهم ویژگی یا رفتار مورد اندازه گیری نباشد، در آن صورت در تعیین روایی سازه و نباید تمایزی در نظر گرفت.

۳- تحلیل عاملی: تحلیل عاملی یک روش آماری است که به وسیله آن تعداد و ماهیت متغیرهایی را که یک آزمـون انـدازه می گیرد، مشخص میک این کار از طریق ادغام تعداد زیادی از ماده ها یا متغیرها با همدیگر و ایجاد تعداد معدودی متغیر به نام عامل انجام می شود. این روش، روابط درونـی مـیـان داده های موردتحلیل را تعیین می کند و برای ایجاد سهولت، متغیرها را به خوشه ها یا عوامل کاهش می دهد، سپس این خوشه ها یا عوامل را با توجه به آنچه به وسیله سؤال های آزمون اندازه گیری می شوند، نام گذاری می کنند.

۴- همسانی درونی: در این روش، ساخت درونی آزمون بررسی می شود و ملاک مورداستفاده برای بررسی، همسانی درونی، نمره کل آزمون است. برای ایـن کار باید همبستگی بین نمرات آزمون های فرعی یا خرده آزمون ها با نمره کل آزمون محاسبه شود و هر یک از خرده آزمون ها که با نمره کل آزمون همبستگی اندکی را نشان داد، حذف شود. نمرات خرده آزمون ها باید با نمره کل آزمون همبستگی زیادی داشته باشند؛ زیرا فرض می شود که همه آن ها روی هم حـوزه رفتاری واحدی را اندازه می گیرند. خرده آزمون ها نباید با هم همبستگی زیادی داشته باشند؛ زیرا در این صورت فرض می شود که همه آن ها یک چیز را اندازه می گیرند، لذا تکراری هستند. در این روش به جای خرده آزمون ها از ماده ها یا سوالات آزمون هم می توان استفاده کرد که در این صورت همبستگی بین ماده های آزمون با کل آزمون باید محاسبه شود.

مثال ۱۸: فرض کنید برای تعیین میزان روایی (Validity) یک آزمون استعداد ادبی، همبستگی این آزمون را با یک آزمـون اسـتعداد حسابداری محاسبه کرده ایم، نام روایی به دست آمده کدام است؟

۱) افتراقی

۲) همگرا

3) همزمان

4) ملاکی

پاسخ: گزینه «۱» یکی از راه های تعیین روایی سازه این است که همبستگی آرمون موردنظر را با یک آزمون موجود و دارای روایی کـه بـرای هـمـان سازه ساخته شده است، تعیین کنیم. اگر آزمون جدید با آزمون موجود همبستگی زیادی داشته باشد، فرض می شود که این دو آزمون، سازه واحدی را اندازه می گیرند. در نتیجه آزمون جدید نیز برای سازه مورد نظر یک آزمون رواست که به این نوع روایی، روایی همگرا می گویند. از سوی دیگر آزمـون مـورد نـظـر باید با آزمون های دیگری که سازه های متفاوتی را اندازه گیری می کنند، همبستگی نداشته باشد. این گونه روایی را روایی واگرا یا روایی تمییزی (افتراقـی) مینامند. با توجه به این که آزمون مورد نظر در این سؤال آزمون استعداد ادبی است و آزمون مربوط به استعداد حسابداری، سازه دیگری را اندازه گیری می کند، می توان گفت که همبستگی بین این دو آزمون از نوع روایی افتراقی است.

مثال ۱۹: مهم ترین و دقیق ترین نوع روایی کدام است؟

1) پیش بینی

2) سازه

3) محتوایی

4) همزمان

پاسخ: گزینه «۲» مهم ترین و دقیق ترین نوع روایی، روایی سازه می باشد.

مثال۲۰: کدام یک از انواع روایی بیشتر جنبه نظری دارد؟

1) سازه

2) ملاکی

3) محتوایی

4) ملاکی

پاسخ: گزینه «۱» روایی سازه بیشتر جنبه نظری دارد.

مثال ۲۱: اگر محتوای دو آزمون هوش متفاوت باشند، احتمالاً می توان گفت که:

۱) یکی از این دو آزمون فاقد اعتبار (reliablity) است.

۲) یکی از آن ها فاقد روایی (Validity) برای سنجش هوش است.

۳) سازندگان برای هنجاریابی هوش از گروه های متفاوت استفاده کرده اند.

۴) تعریف سازندگان آن ها از هوش متفاوت بوده است.

پاسخ: گزینه «۴» هوش یک سازه است که براساس مفهوم ساخته انسان تعریف می شود؛ بنابراین در آزمون های هوش اگر محتوای دو آزمـون تـفـاوت داشته باشد، می توان گفت که تعاریف سازندگان این آزمون ها از مولفه های هوش می تواند متفاوت باشد. مفهوم هوش می تواند توسط افراد و متخصصان گوناگون با ارائه ی تعاریف مختلفی بیان شود.

مثال ۲۲: افزایش میانگین نمره های یک آزمون هوش با افزایش سن آزمودنی ها، یکی از شواهد کدام نوع از روایی آزمون است.

۱) محتوایی

۲) هم زمان

۳) پیش بینی

۴) سازه

پاسخ: گزینه «۴» هوش یک مفهوم انتزاعی وکاملا نظری است و عینیت بیرونی و خارجی ندارد؛ بنابراین برای سنجش هوش افراد و تعیین روایی آن از روایی سازه بهره می گیریم. هوش یک مفهوم غیرقابل مشاهده است که نظریه های روانشناختی درباره آن بیان شده است و آزمون های آن با روایـی سـازه سنجیده می شود. سازه مفهومی است که برای مقاصد ویژه علمی به طور عمد و از روی آگاهی ایجاد می شود.

مثال ۲۳: اگر یک آزمون روانی، خصیصه ای را اندازه گیری کند که برای آن ملاک خارجی وجود نداشته باشد، با کدام روش می توان روایی آن را برآورد کرد؟

۱) همزمان

2) محتوا

3) سازه

۴) عاملی

پاسخ: گزینه «۳» روایی سازه جنبه نظری دارد و یک مفهوم می باشد. سازه برای مقاصد ویژه ی علمی به طور عمـد و از روی آگاهی ابـداع مـی شـود. مثال هایی که می توان در روانشناسی آورد عبارتند از: هوش، انگیزه، اضطراب و ... .

منظور از سازه، استنباط از چیزی است که غیرقابل مشاهده است؛ برای مثال سازه ی هوش اگر چه یک مفهوم است که قابل مشاهده نیست امـا نظریـه هـای روان شناختی زیادی درباره آن بیان شده اند. برای تعیین روایی سازه می توان از تحلیل عاملی، همسانی درونی و... استفاده کرد.

روایی آزمون های وابسته به ملاک

هر سه نوع روایی معرفی شده علاوه بر آزمون های وابسته به هنجار، در مورد آزمون های وابسته به ملاک هم می توانند به کار روند.

روایی محتوایی آزمون های وابسته به ملاک

در روایی محتوایی آزمون های وابسته به ملاک، وابستگی سؤال های آزمون به محتوا و هدف های مشخص یک موضوع، بیشتر از روایی محتوایی آزمـون هـای وابسته به هنجار است؛ زیرا در آزمون های وابسته به هنجـار، هـدف هـای دقیق آموزشی مدنظر نیست، در حالی که در آزمـون هـای وابسـته بـه مـلاک، کاربرد هدف های دقیق آموزشی ضرورت دارد. این آزمون ها درصورتی روا هستند که هدف های آموزشی درس یا موضوع مدنظر را به دقت بسنجند. بـرای داشتن روایی محتوایی، هم در آزمون های وابسته هم در آزمون های وابسته به ملاک، کاربرد جدول مشخصات ضروری است. در جدول مشخصات آزمون های وابسته به هنجار، فقط سطوح طبقات یادگیری (دانش، فهمیدن، کاربرد.....) منظور می گردد، در جدول مشخصات آزمون های وابسته به ملاک، هم سطوح یادگیری و هم هدف های دقیق آموزشی باید در نظر گرفته شود.

روش تعیین روایی محتوایی آزمون های وابسته به ملاک

برای تعیین روایی محتوایی آزمون های وابسته به ملاک، باید از تعدادی از صاحب نظران خواسته شود تا سؤال های آزمون را بخوانند و آن ها را با توجه به میزان دقتی که هدف های آموزشی و محتوای موردنظر را می سنجند، درجه بندی کنند. هرچه ویژگی های موضوع مورد آزمون دقیق تر باشند، بر سر این که آیا یک سؤال هـدف معینی را به خوبی می سنجد یا نه، میان داوران توافق بیشتری وجود خواهد داشت، اگر توافق وجود نداشته باشد، لازم است در سؤال های آزمون تجدیدنظر شود.

روش لاشه برای تعیین روایی محتوایی آزمون های وابسته به ملاک

یکی از روش هایی که قضاوت کیفی داوران درباره روایی محتوایی را به کمیت تبدیل می کند، روش لاشه است که به آن ضریب روایی محتوایی CVR گویند. روش محاسبه این ضریب به این شکل است که از تعدادی متخصص خواسته می شود تا هریک از سؤال های آزمون را بررسی کنند و درباره آن هـا نظر دهند و هر سؤال را به سه شکل ضروری، مفید یا غیرضروری قضاوت کنند.

تعداد کل داوران = N، تعداد داورانی که سؤال را ضروری تشخیص داده اند = ne

مقدار این ضریب بین ۱+ و ۱- قرار دارد و هر قدر مقدار آن بیشتر باشد، روایی هم بیشتر است.

به عنوان مثال، در یک آزمون ۵۰ سؤالی از ۱۶ متخصص خواسته شده است تا سؤالات را بخوانند و نظر خود را اعلام کنند. یکی از سؤالات را ۱۲ داور ضروری، 3 داور مفید و یک داور غیر ضروری تشخیص داده اند. ضریب روایی آن سؤال عبارت است از:

روایی آموزشی

یکی از مفاهیم جدید که به روایی محتوایی وابسته است، مفهوم روایی آموزشی است که به آن، روایی برنامه درسی می گویند.

روایی محتوایی با این پرسش در ارتباط است که آیا محتوا موضوع درسی مورنظر را به خوبی پوشش می دهد و در مقابل، روایی آموزشی یا روایی برنامه درسی در ارتباط با این پرسش است که آیا محتوای آزمون یا وسیله سنجش، عملا آموزش داده شده است؟

روایی پیش بینی آزمون های وابسته به ملاک

روایی پیش بینی در آزمون های وابسته به هنجار به این منظور به کار می رود که نشان دهد عملکرد آزمون شونده در یک آزمون، تا چه حد عملکرد او را در آزمون دیگر (ملاک) پیش بینی می کند. در آزمون های وابسته به ملاک روایی پیش بینی به دو منظور به کار می رود:

۱- پیش بینی عملکرد یادگیرنده در یک هدف آموزشی براساس هدف آموزشی دیگر. یعنی عملکرد آزمون شونده در یک هدف، تا چه اندازه عملکرد او را در هدف دیگری که وابسته به هدف اول است، پیش بینی می کند. ممکن است هدف اول برای هدف دوم پیش نیاز باشد. مثلا برای یادگیری ضرب، یادگیری جمع پیش نیاز است؛ یعنی آیا ما میتوانیم از روی عملکرد فرد در جمع، عملکرد او را در ضرب پیش بینی کنیم؟

برای تعیین روایی پیش بینی یادگیرنده در یک هدف آموزشی بر اساس هدف آموزشی دیگر، باید ضریب همبستگی عملکرد دانش آموزان را در سؤال های اور مربوط به هدفهای موردنظر تعیین کنیم. یعنی بین نمرات سؤالاتی که مربوط به هدف شماره ۱ هستند و نمرات سؤالاتی که مربوط به هدف شماره ۲هستند، ضریب همبستگی پیرسون را محاسبه کنیم و اگر مقدار این همبستگی زیاد باشد، یعنی از نمرات افراد در هدف شماره ۱ به عنوان عامل پیش بینی توفیق دانش آموزان در یادگیری هدف شماره ۲ می توان استفاده کرد و می توان نتیجه گرفت هدف شماره ۱ پیش نیاز یادگیری هدف شماره ۲ است.

مجموع نمرات خام گروه ƩY=Y

مجموع نمرات خام گروه ƩX=X

مجوع مجذورات X ها

مجموع حاصل ضرب های X در Y ƩXY

مجموع حاصل جمع X ها

مجموع مجذورات Y ها

تعداد X ها و Y ها N =

مجوع حاصل جمع Y ها

۲- پیش بینی یادگیری در حد تسلط: که برای تعیین این روایی، بهتر است درصد دانش آموزانی که عملکرد حد تسلط و غیر حد تسلط آن ها در هدف های او ۲ یکسان است، تعیین شود.

روایی سازه آزمون های وابسته به ملاک

سازه به صفت با ویژگی اطلاق می شود که به طور مستقیم قابل مشاهده نیست، ولی از نظریه های روان شناختی قابل استنباط است. مثلا انگیزش، اضطراب، هوش. ما سازه ها را نمی توانیم به طور مستقیم سنجش کنیم، بلکه عملکرد ها با رفتارهایی را که منعکس کننده سازه هستند می سنجیم، ، مثلا نمرات حاصل از یک پرسشنامه احساس خودکار آمدی باید با نمرات به دست آمده از یک آزمون مفهوم خود همبسته باشد و میانگین نمرات مفهوم خود پسران و دختران کلاس پنجم هم تقریبا برابر باشند، در این صورت ما می توانیم ادعا کنیم آزمون مفهوم خود دارای روایی است؛ یعنی همان چیزی را میسنجد که ما قصد اندازه گیری ن را داریم. تعیین روایی سازه آزمون های وابسته به ملاک مستلزم دو نوع تحلیل منطقی و آماری است.

تحلیل منطقی: این تحلیل بسیار شبیه به تحلیل روایی محتوایی است؛ به علت این که روایی سازه با دیگر روایی ها بویژه روایی محتوایی همپوشی دارد. تحلیل آماری- تفاوت های گروهی: ما می توانیم آزمون خود را با گروه های مختلف افراد انجام دهیم و نتایج را محاسبه کنیم و ببینیم که پیش بینی های ما درست در می آیند یا نه. اگر معلوم گشت که نتایج حاصل، پیش بینی های نظری ما را تأیید می کنند، آنگاه اطلاعات بیشتری درباره روایی آزمون در دست خواهیم داشت، یعنی که آزمون مه آنچه را که به قصد اندازه گیری اش درست شده است، می سنجد.

مثال ۲۴: روشی که قضاوت کیفی داوران را در مورد روایی محتوایی آن به کمیت تبدیل می کند ...............نامیده می شود.

1) سازه

۲) لاشه

۳) ملاکی

۴) هنجاری

پاسخ: گزینه «۲» «لاشه» روشی است که قضاوت کیفی داوران در مورد روایی محتوایی را به کمیت تبدیل می کند.

عوامل مؤثر بر روایی آزمون

عوامل درونی مؤثر بر روایی

1- دستورالعمل یا راهنمای آزمون: اگر به شکل نامناسب تهیه شده باشد، روایی را کاهش می دهد.

۲- کیفیت سؤال های آزمون: اگر سؤالات آزمون طبق قواعد صحیح سؤال نویسی تهیه نشده باشند، روایی آزمون را کاهش می دهند.

٣- ترتیب قرار گرفتن سؤال های آزمون به دنبال یکدیگر: اگر سؤالات دشوار در آغاز آزمون قرار گیرند یا اگر سؤالات از نظر درجه دشواری متنوع نباشند یا اگر سؤالات آزمون خیلی دشوار یا خیلی ساده باشند، روایی آزمون کاهش می یابد.

۴- طول یک آزمون با تعداد سؤالات یک آزمون: افزایش سؤالات یک آزمون باعث می شود که نمونه بهتری را از محتوای مورد اندازه گیری انتخاب کنیم و روایی افزایش می یابد.

عوامل بیرونی مؤثر بر روایی

1- اجرای نامناسب آزمون مثل کمبود زمان، فضای نامناسب، مکان نامناسب و غیره که بر عملکرد آزمون شوندگان و در نتیجه در این آزمون تأثیر نامطلوبی می گذارد.

۲- اشکالات موجود در تصحیح پاسخ های آزمون شوندگان و نمره گذاری آن ها به ویژه در آزمون های تشریحی.

3- ترکیب آزمون شوندگان از لحاظ توانایی مورد اندازه گیری؛ هر چه آزمون شوندگان از لحاظ توانایی مورد اندازه گیری نامتجانس تر باشند، روایی آزمون بیشتر خواهد شد، اما تجانس بیشتر باعث می شود روایی کاهش پیدا کند.

۴- ویژگی های روانی آزمون شوندگان مثل انگیزش آن ها با علاقه مندی و بی علاقه بودن آنها می تواند بر روایی تأثیر داشته باشد. علاقه مندی و کم بودن اضطراب باعث می شود روایی بالا رود اما بی علاقگی و اضطراب زیاد باعث می شود روایی کاهش یابد.

تست های طبقه بندی شده فصل شانزدهم

۱- به منظور ارزیابی و تعیین میزان روایی یک آزمون، آزمون ساز تصمیم می گیرد که ضریب همبستگی این آزمون را با یک آزمون مشابه محاسبه کند. نام روایی حاصله چیست؟

(سراسری ۸۱)

۱) سازه

۲) افتراقی

۳) همگرا

۴) همزمان

۲- اولین قدم در تعیین روایی (Validity) سازه چیست؟

(سراسری 82)

1) تعریف تازه مورد نظر

2) مشخص کردن نوع آزمون

3) تعیین اندازه پذیر بودن سازه

۴) تعیین اندازه گیری سازه

۳- در تعیین کدام یک از انواع روایی (Validity) از محاسبات آماری استفاده نمی شود؟

(سراسری ۸۲)

1) همزمان

۲) پیش بین

۳) محتوا

۴) سازه

4- تفاوت اساسی روایی همزمان و روایی پیش بین در چیست؟

(سراسری 82)

1) نوع آزمون

۲) هدف های اندازه گیری

3) نمره گذاری آزمون

۴) زمان اجرای آزمون

۵- بین خطای پیش بینی و همبستگی و بین آزمون پیش بینی کننده و پیش بینی شونده چه رابطه ای وجود دارد؟

(آزاد ۸۲)

۱) آشکار

2) مکنون

۳) مستقیم

۴) معکوس

6- اعتبار وابسته به ملاک به کدام یک از صورت های زیر طبقه بندی می شود؟

(آزاد ۸۲)

1) پیش آزمون و پس آزمون

۲) پیش بین

3) سازه

4) عاملی

۷- آزمون های پیشرفت تحصیلی بیشتر با کدام یک از انواع روایی (Validity) ارتباط دارد؟

(سراسری 83)

1) سازه

2) محتوا

۳) همزمان

۴) پیش بین

۸- مطابق برخی نظریه های یادگیری، درماندگی آموخته شده بر اثر شکست های پیایی حاصل می شود. اگر بخواهیم برای درماندگی آموخته شده یک آزمون بسازیم، این آزمون باید از کدام نوع روایی برخوردار باشد؟

(سراسری ۸۳)

1) سازه

۲) محتوایی

3) صوری

۴) پیش بین

9- کدام یک از روش های تعیین اعتبار، متغیرهای ملاک و پیش بینی کننده با هم اندازه گیری می شوند؟

(آزاد ۸۳)

۱) پیش بین

۲) وابسته به ملاک

۳) همزمان

۴) عاملی

۱۰- کدام یک از انواع روش های تعیین اعتبار، آزمون رفتار را پیش بینی می کند؟

(آزاد ۸۳)

۱) محتوا

2) منطقی

۳) سازه

۴) وابسته به ملاک

۱۱- آزمونی که انگیزه درونی یادگیری را می سنجد، در موقعیت حذف پاداش های بیرونی، دانش آموزان با سطح بالای یادگیری را از دانش آموزان با سطح پایین تفکیک می کند. لذا می توان گفت این آزمون دارای روایی (Validity) ................... است.

(سراسری ۸۴)

1) صوری

۲) محتوی

3) ملاکی

۴) سازه

۱۲- کدام عبارت حداکثر مقدار روایی یک آزمون را نشان می دهد؟

(آزاد ۸۴)

1)

2)

3)

4)

۱۳- کدام یک از انواع روایی (Validity) شکل پیچیده تری از روایی صوری است؟

(سراسری (۸)

۱) منطقی

۲) ملاکی

۳) پیش بین

4) سازه

۱۴- در تحقیقی برای انطباق آزمون هوشی وکسلر برای استفاده در ایران، رابطه ی نمرات حاصل از اجرای این آزمون را با نمرات حاصل از آزمون استنفورد بینه که قبلا برای استفاده در ایران مورد انطباق قرار گرفته، به دست آورده ایم. این کار کوششی در جهت احراز کدام گونه از روایی بوده است؟

(سراسری ۸۵)

1) صوری ((Face Validity

۲) ملاکی (Criterion related validity)

۳) محتوا (Content validity)

۴) سازهای (Construct validity)

۱۵- تعیین روایی پیش بین در کدام یک از انواع آزمون ها مناسب تر است؟

(سراسری 85)

1) شخصیت

۲) مهارت

۳) پیشرفت

۴) استعداد

۱۶- کدام روش های تعیین روانی (validity) با یکدیگر شباهت بیشتری دارند؟

(آزاد ۸۵)

1) محتوی و همزمان

2) پیش بین و سازه

3) همزمان و پیش بین

۴) محتوا و سازه

۱۷- روایی یک مجموعه سؤال مربوط به یک خصوصیت چند عاملی در صورتی برقرار است که سؤالات...

(آزاد ۸۵)

۱) عوامل مختلف همبستگی بالایی داشته باشند.

۲) هر عامل، همبستگی کم ولی سؤالات عوامل مختلف همبستگی بالا داشته باشند.

۳) هر عامل دارای همبستگی بالا و سؤالات هر دو عامل همبستگی کمی داشته باشند.

۴) عوامل مختلف، همبستگی کمی داشته باشند.

۱۸- جدول مشخصات برای تأمین کدام نوع روایی مورد استفاده قرار می گیرد؟

(آزاد ۸۵)

1) محتوی

۲) سازه

۳) همزمان

۴) پیش بین

۱۹- مهم ترین ویژگی متغیری که به عنوان ملاک در تعیین روایی یک آزمون مورد استفاده قرار می گیرد ................. می باشد.

(سراسری 86)

1) در دسترس بودن

۲) پایایی داشتن

۳) بیطرفی یا غیر سودار بودن

۴) مربوط بودن

۲۰- اگر به طور نظری بپذیریم که استعدام ریاضی از خلاقیت متفاوت است، مشاهده همبستگی بالا بین نمرات یک آزمون ریاضی و آزمون خلاقیت نشان دهندهی چیست؟

(سراسری ۸۶)

1) پایایی پایین

۲) پایایی بالا

۳) روایی پایین

۴) روایی بالا

۲۱-اگر آزمونی بتواند رفتار و شایستگی مورد ادعای خود را به گونه ای مناسب شناسایی و اندازه گیری کند، این آزمون دارای چه نوع روایی است؟

(آزاد ۸۶)

1) صوری

2) پیش بین

۳) سازه

۴) محتوا

۲۲- مهم ترین ویژگی یک سوال آزمون آن است که:

(سراسری ۸۷)

۱) افراد بی اطلاع را از افراد با اطلاع جدا کند.

۲) دانش آموز را به فکر کردن وادار کند.

3) رفتار خاصی را اندازه گیری کند.

۴) هدف های مورد نظر را بسنجد.

۲۳- در آزمون های تشریحی مشکل معرف نبودن سؤالات را به کدام صورت می توان حل کرد؟

(سراسری ۸۷)

1) اجرای آزمون های متعدد

۲) افزایش تعداد سؤال

۳) افزودن سؤالات با دشواری کم

۴) دادن وقت بیشتر

۲۴- کدام ویژگی، تعیین می کند که سؤال آنچه در نظر دارد را می سنجد؟

(سراسری ۸۷)

1) نرم آزمون

۲) شرایط اجرای آزمون

۳) روایی آزمون (Validity)

۴) اعتبار آزمون (Reliability)

۲۵- نقض کدام یک از انواع روایی (Validity) تفسیرپذیری نمره ها را بیشتر تحت تأثیر قرار می دهد؟

(سراسری ۸۷)

1) پیش بین

۲) سازه

۳) محتوا

۴) همزمان

۲۶- هر چه نمرات آزمودنی ها در متغیر ملام به هم شبیه تر باشد ...................

(سراسری ۸۷)

۱) اعتبار آزمون بیشتر می شود.

۲) اعتبار آزمون کمتر می شود.

۳) روایی آزمون بیشتر میشود.

۴) روایی آزمون کمتر میشود.

۲۷- برای تعیین کدام یک از انواع روایی می توان از تفاوت های گروهی استفاده کرد؟

(سراسری ۸۷)

۱) پیش بین

۲) سازه

۳) محتوا

۴) همزمان

۲۸- روش تحلیل عوامل به منظور برآورد کدام یک از انواع روایی به کار می رود؟

(سراسری ۸۸)

1) محتوا

2) ملاکی

۳) سازه

4) ثبات درونی

۲۹- کدام دو نوع روایی (Validity) با هم شباهت بیشتری دارند؟

(سراسری ۸۸)

1) محتوا و همزمان

۲) محتوا و پیش بین

۳) پیش بین و همزمان

۴) پیش بین و سازه

30- به کدام نوع روایی صرفا به ویژگی های آنی و ظاهری امتحان نظر دارد؟

(آزاد 88)

1) روایی همزمان

۲) روایی صوری

۳) روایی محتوا

۴) روایی سازه یا ساختار

۳۱- روایی آزمون بیشتر به کدام ویژگی نمرات ارتباط دارد؟

(سراسری ۸۹)

1) دقت

۲) معنی

۳) حساسیت

۴) تکرارپذیری

۳۲- سازه (Construct) در واقع مفهومی......... درباره جنبه ای از رفتار است که مستقیما" ......... آن را اندازه گیری یا مشاهده می کند.

(سراسری ۸۹)

۱) عینی - می توان

۲) عینی - نمی توان

۳) فرضی - نمی توان

۴) فرضی - میتوان

۳۳- تشخیص وجود کدام روایی ها در یک آزمون یا امتحان از طریق قضاوت داوران متخصص مقدور است؟

(آزاد ۸۹)

1) روایی محتوا و روایی همزمان

2) روایی صوری و روایی محتوا

3) روایی صدری و روایی همزمانی

۴) روایی ساختار و روایی همزمان

۳۴- کدام مورد نشان دهنده میزان روابی داخلی بیشتر سؤال مورد تحلیل است؟

(آزاد ۸۹)

1) هر چه مقدار D به عدد منفی یک نزدیک تر باشد.

۲) هر چه مقدار Di به صفر نزدیکتر باشد.

۳) هر چه مقدار D به عدد یکه نزدیکتر باشد.

۴) هر چه مقدار D از عدد یک بزرگتر باشد.

۳۵- در مطالعه روایی آزمون، مهمترین شرط متغیر ملاک کدام است؟

(سراسری ۹۰)

۱) حساس بودن

2) دقیق بودن

۳) نرمال بودن

۴) مربوط بودن

۳۶- بیشینه ضریب روایی (validity) برابر کدام کمیت زیر است؟

(سراسری ۹۰)

1)

2)

3)

4)

۳۷- کدام موزد از عوامل درونی مؤثر بر روایی آزمون قلمداد نمی شود؟

(سراسری ۹۱)

۱) کمبود زمان برای پاسخ دهی به سؤالات

۲) ترتیب قرار گرفتن سؤالات آزمون

۳) کاربرد واژگان و جمله بندی پیچیده در سؤال

۴) طول آزمون و تعداد سؤالات

۳۸- در تعیین روایی سازه آزمون ها کدام تکنیک کاربرد بیشتری دارد؟

(سراسری ۹۱)

۱) تحلیل واریانس

۲) مقایسه میانگین ها

۳) تعیین خطای معیار اندازه گیری

۴) بررسی همبستگی ها

۳۹- مهم ترین ویژگی متغیر ملاک در مطالعه روایی آزمون کدام است؟

(سراسری ۹۱)

۱) اعتبار (Reliability)

۲) بی طرفی

3) یکنواخت بودن

۴) ربط داشتن

۴۰- برای تعیین کدام شواهد مربوط به روایی، تدوین فرضیه ضروری است؟

(سراسری ۹۲)

1) پیش بین

۲) سازه

۳) همزمان

4) منطقی

۴۱- کدام شاخص به کاربرد «داشتن ابزار در تصمیم گیری در مورد موضوع مورد نظر» اشاره دارد؟

(سراسری ۹۲)

۱) اعتبار (Reliability)

۲) همسانی درونی

3) تکرارپذیری

۴۲- در تعیین روایی، وجود کدام ویژگی در متغیر ملاک از اهمیت بیشتری برخوردار است؟

(سراسری ۹۳)

1) مربوط بودن

۲) دقیق بودن

۳) معروف بودن

۴) سهل الوصول بودن

۴۳- برای تحلیل روایی تشخیصی سؤالات یک آزمون، به محاسبه ی کدام همبستگی نیاز است؟

(سراسری ۹۳)

۱) هر سؤال با عضویت گروهی

۲) نمرات کل آزمون با عضویت گروهی

۳) نمرات هر سؤال با نمرات کل آزمون

۴) نمرات کل آزمون با متغیر ملاک دیگری

۴۴- تفسیرپذیری نتایج آزمون مربوط به کدام یک از ویژگی های روان سنجی می باشد؟

(سراسری ۹۳)

۱) نرم آزمون

2) اعتبار

۳) روایی

۴) دقت

۴۵- کدام روایی به روایی سازه مربوط می باشد؟

(سراسری ۹۴)

۱) همزمان

2) همگرا

۳) ملاکی

۴) محتوایی

۴۶- ۱۶ درصد از واریانس عملکرد شغلی کارکنان یک سازمان، از طریق نمرات آزمونی که به منظور استخدام آن ها اجرا شده بود، قابل تبیین است. روایی ملاکی این آزمون استخدامی چقدر است؟

(سراسری ۹۴)

1) 16/0

2) 40/0

3) 60/0

4) 84/0

پاسخنامه تست های طبقه بندی شده فصل شانزدهم

1- گزینه «3» این سؤال در مفهوم به روایی سازه اشاره دارد. اگر آزمونی بسازیم و آن آزمون را با آزمونی که برای همان سازه ساخته شده است، از لحاظ همبستگی بسنجیم، در صورتی که در آزمون با همدیگر همبستگی زیادی داشته باشند و نتیجه بگیریم که آزمون جدیگر نیز برای سازه و مفهوم مورد نظر یک آزمون روا است، به این نوع روایی، روایی همگرا می گویند.

۲- گزینه «۱» روایی سازه دارای جنبه ی نظری بالایی نسبت به روایی محتوایی و روایی پیش بینی است. روایی سازه در صورتی در یک آزمون وجود دارد که نمرات حاصل از اجرای آن آزمون به مفاهیم مورد نظر نزدیک باشد. در نتیجه تعریف این مفهوم بسیار مهم و اساسی است گزینه ی ۳ با توجه به متن سؤال می تواند ذهن داوطلب را مشغول کند. این گزینه به تحلیل عاملی که یک روش آماری است، اشاره دارد که از طریق آن تعداد و ماهیت متغیرهایی را که یک آزمون اندازه می گیرد، مشخص می کند، اما با محتوای سؤال تطبیق ندارد.

۳- گزینه «3» انواع روایی عبارتند از:

۱- روایی محتوایی

۲- روایی ملاک شامل: روایی پیش بینی، زوایی همزمان

3- روایی سازه

روایی ملاکی شامل گزینه های ۱ و ۲ می باشد و میزان ارتباط بین نمرات حاصل از یک آزمون با آزمون دیگر را نشان می دهد؛ پس نتیجه می گیریم که چون در این دو روایی (پیش بین و همزمان) از محاسبات آماری استفاده می شود، این دو گزینه کاملا انحرافی می باشند. روایی محتوایی به روش های قضاوتی سر معروف است و در آن خود ابزار اندازه گیری مورد قضاوت و داوری قرار می گیرد و هیچ گونه روش آماری برای تعیین ضریب روایی محتوایی وجود ندارد. در در تعیین روایی یک آزمون از قضاوت متخصصان حوزهی مورد آزمون استفاده می شود و آن ها نظر می دهند که سؤال های آزمون تا چه میزانی معرف محتوا و اهداف برنامه هستند. با این حال این قضاوت ها قطعی و نهایی نمی باشند و متخصصان در یک زمینه، با هم توافق کامل ندارند.

۴- گزینه «۴» روایی پیش بینی به موقعیتی اشاره می کند که در آن داده های مربوط به ملاک در آینده به دست می آیند. بنابراین، نمرات آزمون واقعأ برای انجام نوعی پیش بینی به کار می روند؛ اما در روایی همزمان، داده های پیش بینی کننده (پیش بینی شونده) همزمان جمع آوری می شوند. هدف آزمون پیش بین، پیش بینی ملاک نیست؛ زیرا داده های ملاک در زمان اجرای آزمون پیش بین وجود دارند و هدف آن جانشین ساختن آزمون پیشبین با آزمون ملاک است. بنابراین تفاوت این دو نوع روایی به زمان اجرای آزمون برمی گردد.

5- گزینه «۴» برای تعیین روایی پیش بینی یک آزمون، دو آزمون ضروری است که ضریب همبستگی بین نمرات حاصل از این دو آزمون شاخص روایی پیش بینی آزمون مورد نظر است. هر چه این ضریب بزرگتر باشد، آزمون دارای روایی بیشتری خواهد بود، در نتیجه آزمون توانایی پیش بینی بالایی دارد و ضریب خطای آن پایین است؛ بنابراین بین آنها رابطه معکوس وجود دارد.

۶- گزینه «۲» از بین گزینه های موجود نزدیک ترین پاسخ به سؤالی گزینه ۲ می باشد، البته این گزینه کامل نیست و باید دارای پیش بینی همزمان باشد که در این گزینه ها تنها به روایی پیش بینی اشاره شده است؛ اما از آنجا که گزینه ۲ از بین بقیه گزینه ها به پاسخ صحیح نزدیک تر است، گزینه صحیح میباشد. گزینه ۱ نیز از آنجایی که به پیش آزمون اشاره کرده است و در روایی ملاکی نیز ۲ آزمون اجرا می شود که یکی آزمون ملاک و دیگری آزمون پیش بینی نام دارد می تواند با گزینه صحیح ۲ اشتباه گرفته شود.

۷- گزینه «۲» آزمون های پیشرفت تحصیلی در یک زمان محدود به کار می روند. این آزمون از مجموعه ای از سؤال ها به عنوان نمونه ای از تمام محتوا و هدف های آموزشی درس منتخب تشکیل می شود. روایی محتوا به این مطلب اشاره می کند که نمونه سوال های مورد استفاده در یک آزمون تا چه حد معرف کل جامعه سؤال های ممکن است که می توان از محتوا با موضوع مورد نظر تهیه کرد. برای تهیه آزمون های پیشرفت تحصیلی روا باید تلاش شود تا سؤال های آزمون نمونه کاملی از اهداف و محتوای درس ها باشد.

۸- گزینه «۱» احساس عجز اکتسابی یا درماندگی آموخته شده به دلیل رویدادهای خارجی که برای فرد رخ می دهد، کسب می شود. این احساس، جنبه نظری دارد و یک مفهوم است. برای اندازه گیری این حالت باید متخصص سنجش، ابتدا یک تعریف از درماندگی آموخته شده ارائه دهد، سپس از آن یک آزمون بسازد و رابطه بین این آزمون را فرضا با پیشرفت تحصیلی و سرعت یادگیری و ... تعیین کند.

۹- گزینه «3» یکی از روش های تعیین روایی (اعتبار) روایی ملاکی است که منظور از آن میزان ارتباط بین نمرات حاصل از یک آزمون با نمرات حاصل از که یک وسیله یا آزمون اندازه گیری دیگر است هنگامی که تمزات حاصل از متغیر ملاک بس از گذشت یک فاصله زمانی گردآوری شود با روایی پیش بینی و هنگامی که نمرات حاصل از متغیر ملاک بدون فاصله و همزمان با متغیر پیش بینی کننده به کار برده شود با روایی همزمان سر و کار داریم.

۱۰- گزینه «۴» منظور از روایی ملاکی میزان ارتباط بین نمرات حاصل از یک آزمون با نمرات حاصل از یک آزمون یا وسیله اندازه گیری دیگر است. بنابراین هر زمان که از نمرات یک آزمون برای پیش بینی عملکرد افراد در آزمون دیگر استفاده کنیم، با روایی ملاکی سروکار داریم.

۱۱- گزینه «۴» روایی سازه بیشتر از روایی های دیگر جنبه نظری دارد. معنی سازه به «مفهوم» نزدیک است. یک آزمون در صورتی دارای روایی سازه است که نمرات حاصل از اجرای آن به مفاهیم یا سازه های نظریه مورد نظر مربوط باشد. بنابراین پس از حذف پاداش بیرونی نیز در سطح یادگیری را در دانش آموزان از هم تفکیک می کند.

۱۲- گزینه «۴» حداکثر مقدار ضریب پایایی برای یک آزمون برابر است با توان ۲ پایایی آن.

۱۳- گزینه «۴» روایی سازه، کامل ترین نوع روایی است که سایر انواع روایی را تحت پوشش خود قرار می دهد، به طوری که سایر روایی ها نوعی از همان روایی سازه هستند.

۱۴- گزینه «۴» یکی از روش های تعیین روایی سازه روایی همگرا است که در این روش، نمرات آزمون جدید باید با نمرات آزمون قبلی که برای سازه ساخته شده است و همان صفت را می سنجد، همبستگی بالایی داشته باشد.

۱۵- گزینه «۴» روایی پیش بینی، ویژه آزمون هایی است که برای پیش بینی موفقیت افراد در امور تحصیلی و شغلی به کار می روند. گاهی از آزمون های پیشرفت تحصیلی فعلی یا قبلی برای پیش بینی پیشرفت تحصیلی آینده استفاده می شود. از این موارد می توان هوش و استعداد تحصیلی را نیز نام برد.

۱۶- گزینه «3» روایی ملاکی میزان ارتباط بین نمرات حاصل از یک آزمون با نمرات حاصل از یک آزمون یا وسیله ی اندازه گیری دیگر است. این نوع روایی خود به دو نوع روایی همزمان و روایی پیش بینی تقسیم می شود که روش تعیین آن ها با هم شباهت دارد. تفاوت آن ها در این است که روایی پیش بینی برای اس پیش بینی موفقیت افراد در امور تحصیلی یا شغل آینده به کار می رود و در روایی همزمان، هدف جانشین ساختن آزمون پیش بینی با آزمون ملاک است.

۱۷- گزینه «3» برای بررسی روایی یک مجموعه سؤال باید نمرات خرده آزمون ها با نمره کل آزمون همبستگی زیادی داشته باشند؛ زیرا فرض بر این است که همه ی آزمون ها روی هم حوزه ی رفتاری واحدی را اندازه گیری می کنند. در ضمن خرده آزمون ها نباید با یکدیگر همبستگی زیادی داشته باشند؛ زیرا در چنین حالتی فرض می شود که همه ی آن ها یک چیز را اندازه می گیرند و لذا تکراری هستند.

۱۸ گزینه «۱» صحیح می باشد. روایی محتوایی به این مطلب اشاره می کند که نمونه سؤال های مورد استفاده در یک آزمون تا چه حد معرف کل جامعه ی ممکن است که می توان از محتوا با موضوع مورد نظر تهیه کرد. بنابراین اگر یک معلم بخواهد برای درس خود یک آزمون پیشرفت تحصیلی زوا بسازد، آزمون او باید در برگیرنده نمونه ای درست و دقیق از مطالب آن درس باشد که به دانش آموزان داده است. در ضمن این آزمون نباید چیزی خارج از محتوا و هدف های آن درس را اندازه گیری بکند. جدول مشخصات برای انجام چنین کاری است.

۱۹- گزینه «۴» در یک آزمون روا به این نکته توجه می کنیم که این آزمون به محتوا و هدف های آن درس توجه کند؛ بنابراین، مربوط بودن محتوا و اهداف، نکته کلیدی در روایی یک آزمون می باشد.

۲۰- گزینه «3» روایی به هدفی اشاره دارد که آزمون برای تحقق بخشیدن به آن درست شده است. این سؤال به روایی واگرا اشاره دارد. در روایی واگرا، دو آزمون که صفت های متفاوتی را می سنجند، باید با هم همبستگی پایینی داشته باشند. وقتی همبستگی بین این دو آزمون بالا رود، روایی پایین می آید

۲۱- گزینه «۴» روایی صوری به این مطلب اشاره می کند که سوال های آزمون تا چه حد در ظاهر شبیه به موضوعی هستند که برای اندازه گیری آن تهیه شده اند. روایی ملاکی نیز به میزان ارتباط بین نمرات حاصل از یک آزمون با نمرات حاصل از یک آزمون یا وسیله دیگر اشاره دارد. روایی سازه نیز برای سازه های مربوط به نظریه های علمی به کار می رود. یک آزمون را در صورتی می توان دارای روایی سازه دانست که نمرات حاصل از اجرای آن به مفاهیم یا سازه های نظریه مورد نظر مربوط باشند. روایی محتوایی نیم به این مطلب اشاره دارد که آزمون تا چه حد می تواند رفتار، خصیصه و یا شایستگی مورد نظر را اندازه گیری کند.

۲۲- گزینه « یک آزمون از مجموعه سؤالاتی تشکیل شده است که به منظور سنجیدن یکسری خصوصیات در آزمون شونده طراحی شده و در کنار یکدیگر قرار گرفته اند. بنام این می توان گفت وقتی یک سؤال در آزمون، همان هدفی را می سنجد که مورد نظر آزمون ساز است، به هدف آزمون دست یافته است. این نکته یادآور روایی آزمون نیز هست. وقتی می گوییم آزمونی روا است که بتواند محتوا و اهداف آن درس را به خوبی اندازه گیری کند، به همین امر اشاره می نماید.

۲۳- گزینه «۱» سؤالات تشریحی شامل آزمونی می شود که آن سؤال از سوی معلم و پاسخ از سوی آزمون شنونده ارائه می شود. در این آزمون های توانایی پاسخ یادگیرنده و میزان درک و فهم او مورد آزمون قرار می گیرد. اگر در آزمون تشریحی، سؤالات معرف نباشند، اجرای آزمون های متعدد می تواند برطرف کننده مشکل باشند.

۲۴- گزینه «3» روایی به هدف آزمون اشاره دارد. به عنوان مثال آزمون هوشی روا، آزمونی است که تنها هوش افراد را بررسی کند و نه چیز دیگری را.

۲۵- گزینه «۲» روایی سازه جنبه نظری دارد. یک آزمون در صورتی دارای روایی سازه است که تمرات حاصل از اجرای آن به مفاهیم یا سازه های نظریه مورد نظر مربوط باشد. مثلا یک آزمون یا پرسشنامه اضطراب زمانی دارای روایی سازه است که نمرات حاصل از آن به سازه هایی که در نظریه اضطراب آمده اند، ارتباط داشته باشد و نقص آن در واقع به تفسیر پذیری نمره آسیب بیشتری می زند. به طور مثال متخصص هوش سنجی پس از این که تعریف از هوش ارائه داد برای آن آزمون می سازد و رابطه بین این نمرات را با سرعت یادگیری، میزان پیشرفت تحصیلی و تعیین می کند و بنقص در این جریان، به تفسیر پذیری نمره آسیب می رساند. (چون مثلا تفسیر می شود: افراد باهوش در یادگیری سریع ترند).

۲۶- گزینه ۴» در تحقیقات و پژوهش ها متغیر وابسته و مستقل داریم. این عنوان در تحقیقات همبستگی به متغیرهای پیش بین و ملاک معروف می باشند. چنانچه نمرات آزمودنی ها در متغیر ملاک به هم شبیه تر باشد، باعث کمتر شدن روایی آزمون می شود.

۲۷- گزینه «۲» روایی سازه بیشتر جنبه نظری دارد. به عنوان مثال آزمون اضطراب در صورتی دارای روایی سازه است که بتواند نمرات افراد در گروه مورد آزمون را به نظریه های اضطراب ارتباط دهد یا آزمون هوش، در صورتی دارای روایی سازه می باشد که بتواند نمرات افراد در گروه مورد آزمون را به نظریه های هوش ارتباط دهد. برای تعیین روایی سازه می توان از تفاوت های گروهی استفاده کرد. معنی سازه ← نزدیک به معنی مفهوم می باشد.

اما سازه یک معنی اضافه بر مفهوم هم دارد و آن این است که سازه برای مقاصد ویژه علمی و به طور عمد و با آگاهی ابداع می شود.

۲۸- گزینه «3» تحلیل عامل ((factor analysis برای تعیین روایی سازه کاربرد دارد. تحلیل عامل با تحلیل عوامل یک روش آماری است که از طریق آن تعداد و ماهیت متغیرهایی را که یک آزمون اندازه می گیرد، مشخص می کنند. در این روش روابط درونی میان داده های مورد تحلیل، تعیین می شود. هم راه های دیگری که برای تعیین روایی سازه کاربرد دارند، عبارت اند از تعیین همبستگی آزمون با سایر آزمون های روا، تمایز سنی و همسانی درونی که به همراه تحلیل عاملی، چهار دسته مختلف در تعیین روایی سازه را تشکیل می دهند.

۲۹- گزینه «3» می دانیم انواع روایی عبارت اند از: روایی محتوا، روایی ملاکی و روایی سازه.

روایی ملاکی شامل روایی پیش بین و روایی هم زمان است.

هنگامی که نمرات حاصل از متغیر ملاک پس از گذشت یک فاصله زمانی گردآوری شود، با روایی پیش بین و هنگامی که نمرات حاصل از متغیر ملاک بدون فاصله زمانی و هم زمان با متغیر پیش بین به کار برده شوند، با روایی هم زمان سروکار داریم.

۳۰- گزینه «۲» روایی صوری که یکی از مشتقات روایی محتوایی است، به این مطلب اشاره دارد که سؤال های آزمون تا چه حد در ظاهر شبیه به موضوعی هستند که برای اندازه گیری آن ها تهیه شده اند. همچنین می توان گفت، روایی صوری به درجه ای از منطقی بودن یک آزمون از نظر کسانی که به آن پاسخ می دهند، گفته می شود.

۳۱- گزینه «۱» کاپلان و ساکوز گفته اند: روایی را می توان به صورت توافق بین نمره آزمون و کیفیتی که قرار است اندازه بگیرند، تعریف کرد. بنابر آنچه گفته شد، آزمونی دارای روایی است. که برای اندازه گیری آنچه مورد نظر است، کافی و مناسب باشد. برای مثال یک آزمون پیشرفت تحصیلی ریاضی کلاس پنجم دبستان در صورتی یک آزمون (روا) است که محتوا و هدف های آن درس را به خوبی اندازه گیری کند و جز یادگیری دانش آموزان از آن درس، چیز دیگری را شامل نباشد. مثال دیگر اینکه، یک آزمون هوشی روا، آزمونی است که تنها هوش افراد را اندازه گیری کند نه چیز دیگری را.

۳۲- گزینه «3» مفهوم سازه، جنبه نظری دارد. کلاریجز سازه را این گونه تعریف می کند: «سازه یک مفهوم است، اما یک معنی اضافی بر مفهوم دارد و آن این است کار سازه برای مقاصد ویژه عملی به طور عمد و از روی آگاهی ابداع می شود.» سازه مفهومی انتزاعی است که مستقیما" قابل مشاهده نیست ولی قابل استنباط از نظریات است. بنابراین آزمون هایی که برای سنجش سازه ها درست می شوند به طور مستقیم سازه ها را نمی سنجند، بلکه عملکرد یا رفتارهایی را می سنجند که گمان می رود منعکس کنده سازه ها هستند.

۳۳- گزینه «۲» هیچ گونه روش آماری برای تعیین ضریب روایی محتوایی وجود ندارد. در عوض برای تعیین روایی محتوایی یک آزمون از قضاوت متخصصان در این باره که سؤال های آزمون تا چه میزانی معرف محتوا و هدفهای برنامه یا حوزه مختوایی هستند استفاده میشود. از آنجا که روایی صوری نوع خاصی از روایی محتوایی است، همان روشی که برای تعیین روایی محتوایی (استفاده از نظر متخصصان) استفاده می شود، در اینجا نیز کاربرد دارد.

۳۴- گزینه «3» چون هر چه ضریب تمییز به ۱ نزدیکتر باشد، به این معناست که آزمون هدفی را که برای سنجیدن آن درست شده است به خوبی می سنجد؛ در واقع ضریب تمییز همان همبستگی سؤال با کل آزمون است که هر چه بالاتر باشد، آزمون هماهنگی درونی بیشتری دارد.

۳۵- گزینه «۴» ثرندایک و هیگن چهار ویژگی مطلوب را برای داده های ملاکی ذکر می کنند. این ویژگی های بکر ترتیب اهمیت عبارت اند از: ربط داشتن، بیطرفی، پایایی و در دسترس بودن. ربط داشتن به قابلیت انطباق داده های یک آزمون ملاکی با موقعیت واقعی اشاره دارد.

۳۶- گزینه «۴» با توجه به فرمول ضریب رولیی، گزینه (۴) درست است و مقدار بیشینه ضریب روایی معمولا بین 9/0 – 8/0 است.

۳۷- گزینه «۱» از جمله عوامل درونی مؤثر بر روایی آزمون عبارت اند از:

1- ترتیب قرار گرفتن سؤالات آزمون ۲- کاربرد واژگان و جمله بندی پیچیده در سؤال 3- طول آزمون و تعداد سؤالات.

۳۸- گزینه «۴» در تعیین روایی سازه، تکنیک بررسی همبستگی ها بیشترین کاربرد را دارد.

۳۹- گزینه «۴» ربط داشتن داده های یک آزمون ملاکی، با توجه به قابلیت انطباق آن ها، با موقعیت واقعی شخص مورد قضاوت قرار می گیرد. برای این کار هیچ گونه شواهد تجربی وجود ندارد و باید براساس ملاک منطقی تصمیم گرفت، در نتیجه ربط داشتن، مهم ترین ویژگی متغیر ملاک در مطالعه روایی آزمون است.

۴۰- گزینه «۲» آزمونی دارای روایی است که برای اندازه گیری آنچه مورد نظر است، کافی و مناسب باشد. سه دسته کلی اعتبار وجود دارد: اعتبار محتوا (اعتبار صوری و منطقی)، اعتبار ملاکی (اعتبار هم زمان و پیش بین) و اعتبار سازه. اعتبار سازه، جدیدترین شکل اعتبار است. اعتبار سازه یک آزمون عبارت ای است از: میزان صحت آن در اندازه گیری ساخت نظری یا ویژگی مورد اندازه گیری. استقرار اعتبار سازه، فرایندی مداوم است. روش های مختلف بررسی اعتبار سازه شامل تفاوت های گروهی، تغییرات، همبستگی ها و فرایندهاست.

۴۱- گزینه «۴» یکی از موارد استفاده از روایی آزمون، استفاده از آن برای تصمیم گیری در مورد افراد است. در چنین موردی می توان برای آزمون و ملاک، نقطه ی برش تعیین کرد. برای تصمیم گیری، اغلب از روایی ملاکی استفاده می شود. مثلا تصمیم گیری می شود که با توجه به نمرات یک فرد در آزمون استعداد ریاضی، بهتر است او در رشته ریاضی درس بخواند یا نه اعتبار، همسانی درونی و تکرارپذیری، مفاهیمی مربوط به هم هستند. اعتبار (پایایی) به ثبات و دقت نتایج حاصل از آن اشاره می کند و انواع مختلفی دارد که هر کدام بر یکی از ابعاد و مفاهیم پایایی تأکید دارند. انواع آن عبارت است از: روش بازآزمایی (ثبات نتایج آزمون)، روش فرم های موازی (ترازی نتایج آزمون) و روش های همسانی درونی (همسانی یا یکنواختی ماده ها با اجزای تشکیل دهنده ی یک آزمون).

۴۲- گزینه «۱» روایی وابسته به ملاک، برای پیش بینی رفتار آینده و دقت این پیش بینی به کار می رود. روایی وابسته به ملاک با مطالعه همبستگی بین آزمون و یک ملاک آماری یا تجربی حاصل می شود.

۴۳- گزینه «1» برای تحلیل روایی تشخیصی هرم سؤال یک آزمون، به محاسبه همبستگی هر سؤال با عضویت گروهی نیاز است؛ زیرا در بحث تعیین روایی یک سؤال می خواهیم دقت و کارایی آن مؤلفه را بسنجیم و ناگزیر باید آن سؤال را با گروه مقایسه کنیم.

۴۴- گزینه «3» روایی یعنی اینکه یک آزمون، همان چیزی را بسنجنم که برای سنجش آن ساخته شده است. روایی به تفسیرپذیری نیز اشاره دارد؛ زیرا سؤالات باید قدرت بررسی و تفسیر را داشته باشند.

۴۵- گزینه «۲» روایی سازه یا مفهومی به همسویی سؤالات با تعاریف و نظریات موجود در مورد خصیصه مورد نظر اشاره دارد و دو نوع معتبر همگرا و واگرا دارد. بدین ترتیب، اعتبار سازه یک ابزار اندازه گیری، نمایانگر آن است که ابزار اندازه گیری تا چه اندازه سازه یا خصیصه ای را که مبنای نظری دارد، می سنجد. در بررسی اعتبار سازه باید به تدوین فرضیه هایی درباره مفاهیم اندازه گیری شده، آزمودن این فرضیه ها و محاسبه همبستگی نتایج با اندازه گیری اولیه پرداخت. اگر ضریب همبستگی حاصله بالا باشد، اعتبار سازه افزایش می یابد، ولی اگر همبستگی معنی دار نباشد، علت را می توان به عوامل زیر نسبت داد:

جمع آوری داده ها با اشکال همراه بوده است، پیش بینی و فرضیه ها غلط بوده است و ابزار، خصیصه مورد نظر را اندازه گیری نمی کند.

اگرچه فرآیند برقراری اعتبار سازه امر بی انتهایی است، ولی محقق یا سازنده آزمون می تواند اعتبار سازه یک آزمون را در موقعیت هـای خـاص نشان دهـد.

از جمله شیوه هایی که برای تعیین اعتبار سازه به کار می رود موارد زیر است:

تفاوت های گروهی: اگر نظریه ای تفاوتی را بین گروه ها (یا هیچ تفاوتی را) در نمرات آزمون پیش بینی کند، آزمونی که این تفاوت را نشان دهـد (یـا نـدهـد) دارای اعتبار سازه است. برای مثال، پیش بینی می شود که کودکان و بزرگسالان در بلوغ اجتماعی بـا هـم تـفـاوت دارند. بنابراین اگر آزمـونـی کـه بـرای اندازه گیری خصیصه بلوغ اجتماعی ساخته شده است تفاوت معنی داری را بین کودکان و بزرگسالان نشان دهد، دارای اعتبار سازه است.

تغییرات: چنانچه نظریه ای پیش بینی کند که خصیصه مورد اندازه گیری تحت تأثیر زمان یا مداخله آزمایشی قرار می گیرد، اگر نمرات آزمـون ایـن تـأثیر را نشان دهد، دارای اعتبار سازه است. برای مثال، اگر آزمونی مهارت های کلامی را می سنجد، باید با افزایش سن آزمودنی ها نمرات افزایش یابد یا اگر مداخلـه اجتماعی برای پرورش این مهارت ها صورت گرفته باشد، نمرات پیش آزمون و پس آزمون باید تغییرات معنی داری را نشان دهد.

اعتبار همگرا و واگرا (تشخیصی): هرگاه یک یا چند خصیصه از طریق دو یا چند روش اندازه گیری شوند، همبستگی بین این اندازه گیری ها دو شاخص مهـم اعتبار را فراهم می کند، اگر همبستگی بین نمرات آزمون هایی که خصیصه واحدی را اندازه گیری می کنند بالا باشد، آزمون ها دارای اعتبار همگرا هستند. چنانچه همبستگی بین آزمون هایی که خصیصه های متفاوتی را اندازه گیری می کنند پایین باشد، آزمون ها دارای اعتبار تشخیصی یا واگرا هستند.

۴۶- گزینه «۲» روایی ملاکی آزمون خواسته شده برابر است با:

آزمون فصل شانزدهم

۱- در تعیین کدام یک از روایی های زیر از روش های آماری استفاده نمی شود؟

۱) همزمان

۲) سازه

۳) محتوایی

۴) پیش بینی

2- کدام یک از روایی های زیر، هم مستلزم تحلیل منطقی است و هم تحلیل آماری؟

۱) محتوایی

2) صوری

3) هم زمان

۴) سازه

۳- منظور از کدام روایی، جانشین کردن یک آزمون یا وسیله اندازه گیری به جای آزمون یا وسیله اندازه گیری دیگر است؟

1) سازه

۲) محتوایی

۳) همزمان

۴) پیش بینی

۴- کدام ویژگی آزمون تعیین می کند که آزمون همان چیزی را که باید بسنجد، می سنجد؟

۱) روایی

۲) پایایی

3) عینیت

۴) سهولت نمره گذاری

۵- روایی ملاک و سازه به روش های ................. معروف اند.

۱) قضاوتی

2) تجربی

۳) عینی

۴) قضاوتی- تجربی

6- آزمون های پیشرفت تحصیلی بیشتر با کدام روایی سروکار دارند؟

1) روایی پیش بینی

۲) روایی سازه

3) روایی محتوایی

۴) روایی همزمان

۷- اگر از نمرات فرد در دوره راهنمایی برای پیش بینی موفقیت او در دبیرستان بهره بگیریم، از تمام روایی استفاده کرده ایم؟

۱) محتوایی

2) سازه

۳) همزمان

۴) پیش بینی

۸- مهم ترین ویژگی مطلوب برای داده های ملاکی چیست؟

1) پایایی

2) در دسترس بودن

۳) ربط داشتن

۴) بیطرفی

۹- کدام یک از موارد زیر جزء عوامل درونی مؤثر بر روایی آزمون نیست ؟

۱) ترکیب آزمون شوندگان از لحاظ توانایی موردنظر اندازه گیری

۲) کیفیت سؤال های زمون

۳) ترتیب قرار گرفتن سؤالات آزمون به دنبال یکدیگر

۴) طول یک آزمون با تعداد سؤال های آن

۱۰- وقتی برای تهیه سؤالات آزمون از جدول مشخصات استفاده می کنیم کدام ویژگی آزمون افزایش

می یابد؟

۱) پایایی

2) روایی

۳) عملی بودن

4) عینیت

۱۱- اجرای مجدد یک تست با فاصله زمانی به منظور محاسبه کدام مشخصه اعتبار صورت می گیرد؟

۱) همگونی و تجانس سؤال ها

۲) هماهنگی درونی سؤال ها

۳) معادل بودن تست ها

۴) پایایی نمره ها

۱۲- برای افزایش اعتبار آزمون و کاهش نسبی اشتباه باید ....................

1) بر طول آزمون افزود.

۲) درصد دشواری را ۴۶ درصد کرد.

۳) جای سؤال را عوض کرد.

۴) درصد دشواری سوالات را ۷۵ درصد کرد.

۱۳- هنگامی که در تهیه سؤالات آزمون اهمیت نسبی هر یک از هدف های رفتاری را تعیین می کنیم (جدول مشخصات را تهیه کنیم) این کار موجب می شود تا ......................

1) روایی افزایش یابد.

۲) پایایی افزایش یابد.

۳) محاسبه ضریب دشواری امکان پذیر است.

۴) محاسبه ضریب تمییز به سادگی میسر می شود.

۱۴- از طریق سؤال های چند گزینه ای در هر جلسه آزمون می توان تعداد زیادی سؤال را بررسی کرد. این ویژگی به ...........

1) بالا رفتن روایی محتوا کمک می کند.

۲) تعیین وقت مناسب برای آزمون کمک می کند.

۳) بالا رفتن سطح دشواری کمک می کند.

۴) عینی شدن هر چه بیشتر آزمون کمک می کند.

۱۵- عوامل مؤثر بر روایی درونی عبارت است از .......................

1) دستورالعمل یا راهنمایی آزمون

۲) کیفیت سؤال های آزمون

3) طول آزمون

۴) همه موارد

۱۶- .................. یک آزمون باعث می شود نمونه بهتری از محتوای مورد اندازه گیری را انتخاب کنیم.

1) راهنما

۲) کیفیت سؤال ها

۳) ترتیب قرار گرفتن سؤالات

۴) طول سؤالات

17- عواملی که بر کاهش روایی تأثیر می گذارند عبارت اند از .................. .

۱) درجه دشواری متنوع نباشد.

۲) اسؤال خیلی دشوار

۳) سؤالات خیلی ساده

۴) همه گزینه ها

۱۸- عواملی مثل کمبود زمان، فضای نامناسب، مکان نامناسب جزء عوامل مؤثر بر روایی ................ می باشد.

1) درونی

۲) ملاکی

۳) بیرونی

۴) سازه

۱۹- درباره این ویژگی ملاک مطلوب، توجه به زمان، هزینه و امکانات مورد نیاز ضروری می باشد؟

1) ربط داشتن

2) بی طرفی

۳) پایایی

۴) در دسترس بودن

۲۰- « همه افراد در کسب نمره خوب فرصت مساوی داشته باشند» مربوط به کدام ویژگی ملاک مطلوب می باشد؟

1) بی طرفی

2) پایاپی

3) در دسترس بودن

۴) ربط داشتن

فصل هفدهم

«پایایی آزمون»

همان طور که گفتیم، روایی به هدفی که آزمون برای آن ایجاد شده است اشاره دارد، یعنی آزمون در صورتی رواست که هدف مورد نظر را بسنجد؛ اما پایایی به دقت، اعتمادپذیری، ثبات یا تکرارپذیری نتایج آزمون اشاره دارد. یک آزمون وقتی دارای پایایی است که در هر بار اجرا روی تعدادی از دانش آموزان نتایج یکسانی را به دست دهد. اگر آزمونی را روی تعدادی دانش آموز اجرا کنیم و نتایجی را به دست آوریم، اگر دوباره آزمون را اجرا کنیم، باید مهمان نتایج را به دست آوریم و اگر نتایج مختلفی را به دست آوریم، آزمون پایا نیست.

برای این که آزمونی روا باشد، باید نخست پایا باشد؛ یعنی پایایی شرط روایی است، اما روایی برای پایایی ضروری نیست.

مثال 1: برای این که آزمونی روا باشد باید .................... .

1) خطای معیار اندازه گیری متوسطی داشته باشد.

۲) ابتدا پایا باشد.

3) میانگین نمرات در حد متوسط باشد.

۴) همه موارد

پاسخ: گزینه «۲» برای این که آزمونی روا باشد باید ابتدا پابا باشد، یعنی این که پایایی شرط روایی است.

مثال ۲: ثبات به کدام یک از ویژگی های یک آزمون اشاره دارد؟

۱) روایی

۲) پایایی

۳) عینیت

۴) عملی بودن

پاسخ: گزینه «۲» ثبات به پایایی آزمون مربوط می شود.

مثال 3: کدام شاخص میزان دقت وسیله اندازه گیری را تعیین می کند؟

1) اعتبار

۲) روایی

۳) عملی بودن

4) نرم

پاسخ: گزینه «۱» پایایی وسیله اندازه گیری، عمدتا به دقت نتایج حاصل از آن اشاره دارد. پایایی به دقت، اعتماد، ثبات و تکرارپذیری نتایج آزمون اشاره دارد.

روش های تعیین پایایی

1- روش پایایی مصححان

برای تحسین پایایی آزمون های غیرعینی که نمرات آن ها تحت تأثیر قضاوت تصحیح کنندگان قرار می گیرد، باید از دو یا چند مصحح خواست که پاسخ های آزمون شوندگان را به طور مستقل تصحیح کنند. همبستگی بین نمرات مصححان، پایایی مصححان نام دارد. پایایی مصححان یک آزمون تشریحی را به صورت میزان توافق مصححان، مختلف در نمره گذاری یک مجموعه برگه آزمون تعریف می کنند؛ پس هر چه ضریب همبستگی بین نمرات مصححان بزرگ تر باشد، توافق نظر مصححان هم بیش تر است.

پایایی مصححان مربوط به شیوه کار تصحیح کنندگان است و با پایایی آزمون که مربوط به سؤالات آزمون است، متفاوت است.ضریب پایایی مصححان برای ضریب پایایی نمرات آزمون یک سقف به حساب می آید. به عنوان مثال اگر ضریب توافق بین مصححان 60/0 باشد، ضریب پایایی نمرات آزمون احتمالا" کتر از 60/0 یا حداکثر 60/0 است.

مثال ۴: روش پایایی مصححان برای کدام یک از آزمون های زیر به کار برده می شود؟

1) آزمون های عینی

۲) آزمون های غیر عینی

۳) آزمون های شخصیت

۴) آزمون های نگرش سنج

پاسخ: گزینه «۲» روش پایایی مصححان برای آزمون های غیر عینی می باشد.

2- روش باز آزمایی

در این روش، یک آزمون را دوبار روی گروه واحدی از آزمون شوندگان اجرا می کنیم و همبستگی بین نمرات حاصل از دو بار اجرای آزمون را محاسبه می کنیم. این ضریب همبستگی، ضریب پایایی آزمون است که به آن پایانی بازآزمایی می گوییم. از این روش فقط زمانی می توانیم استفاده کنیم که ویژگی در مورد اندازه گیری در طول زمان، تغییر نکند. به عنوان مثال چنین تصور می شود که هوش یک ویژگی یا صفت نسبتا" پایدار است که در فاصله اجرایی یک در آزمون هوشی در دو نوبت متفاوت تغییر نمی کند. این روش، ساده ترین روش تعیین پایایی است، اما متخصصان اندازه گیری و ارزشیابی از آن انتقاد کرده اند.

چند انتقاد به این روش وارد است:

1- وقتی آزمون را بار اول اجرا می کنیم، آزمون شوندگان تجربه ای به دست می آورند که باعث می شود در اجرای مجدد آزمون بهتر عمل کنند.

۲- چون در هر دو آزمون سؤالات یکسانی به کار رفته است ، نمرات حاصل از اجرای مجدد آزمون به ما نشان نمی دهند که اگر سؤالات تغییر کنند چه تغییراتی در نتایج حاصل می شود.

٣- اگر فاصله زمانی بین دو بار اجرای آزمون زیاد باشد، ممکن است خطای اندازه گیری با تغییراتی که واقعا در یادگیری دانش آموزان ایجاد شده است ، با هم مخلوط شوند.

۴- معلمان و دانش آموزان علاقه ای به اجرای مجدد یک آزمون ندارند و عدم علاقه ممکن است باعث شود دانش آموزان در آزمون دوم نمره کمتری به دست آورند. باید به خاطر داشت که روش باز آزمایی، بیشتر برای آزمون های هوش و شخصیت که در طول زمان چندان تغییر نمی کنند، مفید است.

مثال 5: اثر تجربه در کدام یک از روش های برآورد پایایی وارد عمل می شود؟

1) اسپیرمن

۲) بازآزمایی

۳) دو نیمه کردن

۴) کودر - ریچاردسون

پاسخ: گزینه «۲» در روش باز آزمایی، آزمون در دو نوبت با گروه واحدی از آزمون شوندگان برگزار می شود و نمرات حاصل با هم مقایسه می شوند. در در روش بازآزمایی، به نظر متخصصان اندازه گیری، آزمون شوندگان با سؤالات آزمون آشنا می شوند و این امر در پاسخ دهی آن ها بر اثر تجربه، تأثیر دارد.

مثال ۶: اجرای مجدد یک تست با فاصله زمانی به منظور محاسبه کدام مشخصه اعتبار صورت می گیرد؟

1) همگونی و تجانس سؤال ها

2) هماهنگی درونی سؤال ها

3) معادل بودن تست ها

4) پایایی نمره ها

پاسخ: گزینه «۴» اجرای یک آزمون در فاصله زمانی معین به پایایی نمرات اشاره دارد، همان گونه که در سؤال قبل توضیح داده شده ویژگی تکرار پذیری، مشخصه ی اعتبار است. در روش بازآزمایی، یک آزمون در فاصله ی زمانی مشخص دوبار انجام میشود.

مثال ۷: اثر « اجرای آزمون» در کدام نوع پایایی وجود دارد؟

۱) اجرای دو آزمون موازی

2) اجرای مجدد آزمون

۳) همسانی درونی

۴) دو نیمه کردن آزمون

پاسخ: گزینه «3» روش های تعیین پایایی عبارت اند از: ۱) روش پایایی مصححان ۲) روش بازآزمایی ۳) روش فرم های موازی با هم ارز

٣- روش فرم های موازی با هم ارز

در روش فرم های موازی با هم ارز دو آزمون موازی یا هم ارز، برای یک مطلب تهیه می شود و با فاصله زمانی کوتاهی، این دو آزمون روی یک گروه واحد از آزمون شوندگان اجرا می شود. ضریب همبستگی بین نمرات دو آزمون، ضریب پایایی آن آزمون ها به حساب می آید.

مثال8: محققی می خواهد در آزمون معادل بسازد تا به عنوان پیش آزمون و پس آزمون در یک تجربه که فقط دو روز به طول می انجامد، استفاده کند. بهترین روش برای محاسبه ضریب پایایی این آزمون ها کدام است؟

1) آلفای کرونباخ

۲) اجرای مجدد آزمون

۳) اجرای آزمون های موازی

۴) دو نیمه کردن آزمون

پاسخ: گزینه «3» روش فرم های موازی با هم ارز، یکی از روش های تعیین پایایی است. در این روش در آزمون معادل یا موازی برای یک موضوع تهیه می کنند و آن ها را در فاصله ی کوتاهی (مثلا دو هفته) روی یک گروه واحد از آزمون شوندگان اجرا می کنند. دو فرم آزمون در صورتی موازی هستند که میانگین و واریانس آنها برابر باشند.

گزینه ی اجرای مجدد آزمون بسیار نزدیک به پاسخ صحیح است اما باید به این نکته توجه داشت که در متن سؤال به پیش آزمون - پس آزمون اشاره می کند که به اجرای آزمون موازی نیاز دارد که سؤالات با یکدیگر متفاوت اند، ولی در اجرای مجدد یک آزمون سؤالات آزمون تکراری هستند.

مثال ۹: تفاوت در آزمون موازی بیشتر در چیست؟

1) درجه دشواری

2) شکل ظاهری

۳) قدرت تمییز

۴) محتوای درونی

پاسخ: گزینه «۲» آزمون های موازی با معادل برای یک موضوع تهیه می شوند. دو فرم آزمون در صورتی موازی هستند که میانگین و واریانس آنها برابر باشند. بنابراین شکل ظاهری سؤالات تفاوت دارند، اما برای رسیدن به یک هدف مشترک تلاش می کنند و از قدرت تمییز و ضریب دشواری یکسانی برخوردارند؛ چون دارای میانگین و واریانس برابر هستند و متن و موضوع آن ها محتوای یکسانی دارد، پس در ظاهر و شکل سؤالات با همدیگر تفاوت دارند.

نکته 1: دو آزمون وقتی هم ارز یا معادل هستند که میانگین و واریانس آنها برابر باشد.

روش فرم های موازی برای رفع مشکلات روش بازآزمایی به کار می رود، اما خود این روش هم مشکلاتی دارد که مهمترین مشکل آن، این است که تهیه فرم های هم ارز برای آزمون های پیشرفت تحصیلی کار آسانی نیست. در این روش هم مشکل بی علاقگی دانش آموزان وجود دارد. مشکل دیگر این است که ممکن است دانش آموزان در اجرای فرم اول آزمون، با اصولی آشنا شونگ که عملکرد آنها در فرم دوم تحت تأثیر قرار دهد. برای رفع مشکل روش فرم های هم ارز از روش های دو نیمه کردن استفاده می شود.

مثال ۱۰: ضریب همبستگی بین نمرات آزمودنی ها از دو بار آزمایش، نشان دهنده میزان .................. آزمون است. 1) روایی محتوا

۲) روایی همزمان

۳) ضریب دشواری

۴) ضریب پایایی

پاسخ: گزینه «۴» ضریب پایایی به دقت و تکرار در اندازه گیری اشاره دارد.

۴- روش های همسانی درونی

روش بازآزمایی بر ثبات نتایج آزمون و روش فرم های هم ارز بر ترازی نتایج آزمون تأکید دارند، اما روش های همسانی درونی بر همسانی ماده های تشکیل دهنده یک آزمون تاکید دارند.

در روش های همسانی درونی، آزمون فقط یک بار روی گروه واحدی از آزمون شوندگان اجرا می شود. بعد با یکی از روش های زیر به تعیین پایایی آن می پردازند.

الف - روش دو نیمه کردن: (یا روش تنصیف)

در این روش آزمون را یک بار روی یک گروه از آن شوندگان اجرا می کنیم، سپس آزمون (یعنی سوالات آزمون) را به 2 نیمه تقسیم می کنیم. بهترین راه برای دو نیم کردن آزمون این است که سوالات فرد را در یک نیمه قرار می دهیم و سوالات زوج را در نیمه دیگر. سپس ضریب همبستگی بین 2 نیمه آزمون را محاسبه می کنیم. ضریب حاصل پایایی هر یک از 2 نیمه را نشان می دهد.

وقتی ضریب همبستگی بین دو نیمه را محاسبه کردیم آن را در فرمول اسپیرمن - براون قرار می دهیم تا پایایی کل آزمون را به دست آوریم.

فرمول اسپیرمن براون:

= ضریب پایایی کل آزمون

= ضریب پایایی بین دو نیمه های آزمون

مثال ۱۱: در روش تنصیف بهترین راه برای دو نیمه کردن سؤالات آزمون این است ؟

۱) سؤالات را به دو نیمه مساوی تقسیم کنیم.

۲) سؤالات فرد را در یک نیمه و سؤالات زوج را در نیمه ی دیگر قرار دهیم.

۳) به صورت تصادفی آن ها را به دو نیمه تقسیم کنیم.

۴) از روش تصادفی منظم استفاده کنیم.

پاسخ: گزینه «۲» در روش تنصیف بهترین راه برای نیمه کردن سؤالات آزمون این است که، سؤالات فرد را در یک نیمـه و سـؤالات زوج را در نیمه مدیگر قرار دهیم.

مثال ۱۲: ضریب پایایی آزمونی با روش تنصیف (ضریب همبستگی مابین دو نیمه) ۸۰/0 شده است. ضریب پایایی کل آزمون، کدام است؟

1) 80/0

2) 84/0

3) 88/0

4) 91/0

طبق مفروضات سؤال ← ۸/0 =

پاسخ: گزینه «۳» در روش تنصیف (دو نیمه کردن)، آزمون مورد نظر یک بار با گروه واحدی از آزمون شوندگان اجرا و سپس آزمـون بـه دو نیمـه تقسیم می شود. مثلاً سؤالات فرد آزمون در یک گروه و سؤالات زوج آزمون در گروه دیگر قرار می گیرد. برای محاسبه پایایی کـل آزمـون نـیـز از روش اسپیرمن - براون استفاده می شود.

مثال ۱۳: اگر ضریب همبستگی بین دو نیمه زوج و فرد آزمونی 6/0 باشد، ضریب پایایی کل آزمون چه قدر خواهد بود؟

1) 6/0

2) 75/0

3) 8/0

4) 9/0

پاسخ: گزینه «۲» برای تعیین ضریب همبستگی از فرمول اسپیرمن - براون استفاده می کنیم. داریم:

ضریب پایایی

مثال ۱۴: اگر ضریب همبستگی بین نیمه های یک آزمون 40/0 باشد، ضریب پایایی کل آزمون چقدر است؟

1) 42/0

2) 66/0

3) 57/0

4) 45/0

پاسخ: گزینه «3»

اگر فرمول اسپیرمن ـ براون را کمی تغییر دهیم، می توانیم از آن برای محاسبه ضریب پایایی یک آزمون، پس از آن که تعدادی سؤال به آن افزوده شده استفاده کنیم.

= ضریب پایایی آزمون جدید

= ضریب پایایی آزمون اولیه

= k نسبت سوال های آزمون جدید به تعداد سوالات آزمون اولیه

دقت کنید که k نسبت طول دو آزمون است، نه تعدادی سؤال های جدید که به آزمون اولیه افزوده شده است.

مثال ۱۵: اگر به یک آزمون ۲۰ سؤالی که ضریب پایایی آن 70/0 است ۱۰ سؤال اضافه کنیم، پایایی آزمون جدید چند خواهد شد؟

1) 65/0

2) 77/0

3) 70/0

4) 85/0

پاسخ: گزینه «2»

اکنون فرض کنید، سازنده یک آزمون می خواهد بداند که چند سؤال دیگر باید به سؤالات آزمونی که قبلا ساخته است، اضافه کند تا ضریب اعتبـار آزمـون جدید به اندازه ی مورد نظر برسد. او برای این کار می تواند از رابطه زیر استفاده کند:

در این فرمول n به معنای ضریب افزایش تعداد سوال های آزمون، rd ضریب اعتبار مطلوب و ro ضریب اعتبار مشاهده شده آزمون موجود می باشد.

مثال ۱۶: اگر همبستگی میان دو نیمه آزمون برابر با ۴۵/0 باشد، پایایی آن آزمون چه قدر است؟

1) 55/0

2) 58/0

3) 60/0

4) 62/0

پاسخ: گزینه «۴» برای تعیین همبستگی میان دو نیمه یک آزمون از فرمول اسپیرمن براون استفاده می کنیم:

مثال ۱۷: یک آزمون ۱۰ سؤالی، دارای ضریب ۸۰٪ است. در صورتی که بخواهیم ضریب پایایی این آزمون را به ۹۰ ٪ برسانیم، طول این آزمون را چه مقدار باید افزایش دهیم؟

1) 25/2

2) 9/3

3) 8/5

4) 7/9

پاسخ: گزینه «۱) با استفاده از فرمول بالا، فقط ضریب افزایش تعداد سؤال های آزمون را داریم؛ اگر بخواهیم دقیقا بدانیم که چه تعداد سؤال جدید باید به آزمون اضافه کنیم باید ضریب به دست آمده را در تعداد سؤال های قبلی ضرب کنیم تا تعداد کل سؤال های جدید به دست آید. اگر این مقدار را از تعداد سؤال های اولیه کم کنیم، تعداد سؤال هایی که باید به آزمون اضافه شود، به دست می آید.

مثال ۱۸: پایایی یک آزمون ۳۰ سؤالی محاسبه شده است. برای رسیدن به پایایی تعداد سؤال ها چه قدر باید باشد؟

1) 60

2) 120

3) 150

4) 180

پاسخ: گزینه «۴» برای افزایش پایایی مفروضات سؤال را در فرمول جایگزین می کنیم:

ضریب پایایی خواسته شده

ضریب پایایی مشاهده شده

= n چند برابر شدن سؤال

مثال ۱۹: فرض کنید آزمونی با ۲۵ سؤال دارای ضریب اعتبار 80/0 است، اگر بخواهیم ضریب اعتبار آن را تا ۹۵/0 افزایش دهیم چه تعداد سوال جدید باید به آزمون اضافه کنیم تا این ضریب پایایی به دست آید؟

1) 82

2) 87

3) 94

4) 96

پاسخ: گزینه «3»

ب - روش کودو- ریچاردسون

در این روش هم ، آزمون فقط یک بار اجرا می شود با این تفاوت که همه سوالات (یا ماده های) آزمون مورد تحلیل قرار می گیرند. در این روش از دو فرمول20 KR و21 KR استفاده می شود.

ضریب پایایی

تعداد سوالات آزمون

واریانس کل سوالات آزمون

نسبت دانش آموزانی که به سوال پاسخ درست داده اند.

نسبت دانش آموزانی که به سوال پاسخ غلط داده اند.

مثال ۲۰: اگر در آزمونی تعداد سؤالات ۱۰ تا باشد و واریانس کل سؤالات 75/2 باشد و باشد، ضریب پایایی (20KR) چقدر است؟

1) 42/0

2) 47/0

3) 49/0

4) 50/0

پاسخ: گزینه «۱»

اگر تمامی سوالات آزمون دارای ضریب دشواری متوسط و شبیه به هم باشند، می توان به جای فرمول20KR از فرمول 21KR به شرح زیر استفاده کرد.

21KR

که در این فرمول، باید میانگین pها و q ها را برای تمام سوالات حساب کرد و آن را در n یعنی تمام سؤال ها ضرب کرد و در فرمول گذاشت. ضریب حاصل از فرمول 21KR کم تر از ضریب حاصل از فرمول 20KR است. در فرمول 21KR با گذاشتن مقدار به جای و فرمول زیر به دست می آید که فقط با داشتن سه جزء اطلاعات یعنی:

n: تعداد سؤالات، : میانگین نمرات آزمون و : واریانس کل آزمون

21KR

می توانیم ضریب پایایی را محاسبه کنیم.

مثال ۲۱: در کدام یک از روش های تعیین پایایی، به کار کرد افراد در تک تک سؤال های آزمون توجه می شود؟

1) کودر - ریچاردسون

۲) تنصیف

۳) باز آزمایی

۴) فرم های موازی (همتا)

پاسخ: گزینه «1» در روش کودر - ریچاردسون یک آزمون اجرا می شود و همه ماده های آزمون تحلیل می شوند؛ به عنوان مثال نمرات افراد درباره تک تک سؤالات، بررسی می شوند. برای استفاده از فرمول کودر- ریچاردسون 20KR و 21KR، فرض بر این است که سؤال های آزمون را می توان به صورت 1 برای پاسخ درست و صفر برای پاسخ غلط، تصحیح کرد و پرسشنامه های نگرش سنج کار از یک طیف ۱ تا ۵ یا ۷ برخوردارند، در این چارچوب قرار نمی گیرند. در فرمول های موازی اساسا" سؤالات دو دسته اند. در تنصیف نیز یک آزمون به دو دسته جداگانه تقسیم می شود و در روش باز آزمایی یک آزمون در دو نوبت انجام می گیرد.

مثال ۲۲: ضریب محاسبه شده بر اساس فرمول کودر - ریچاردسون چه نوع شاخصی است؟

۱) همسانی درونی

۲) رواییValidity))

3) روایی محتوایی

۴) خطای معیار اندازه گیری

پاسخ: گزینه «۱» روش کودر - ریچاردسون یکی از انواع روش های برآورد پایایی از طریق همسانی درونی است که در آن آزمون تنها یک بار اجرا و همه ماده های آزمون تحلیل می شود.

مثال ۲۳: آزمونی دارای ۴۰ سؤال است. میانگین این آزمون ۱۸ و انحراف معیار آن ۴ است. ضریب پایایی آزمون چند است؟

پاسخ: گزینه «۴»

وقتی می توان از فرمول 21KR استفاده کرد که فرض کنیم درجه دشواری سؤالات آزمون شبیه به هم هستند و هر چقدر که سؤالات از نظر درجه دشواری متفاوت باشند، از ضریب پایایی 21KR کاسته می شود.

فرض اساسی در فرمول های کودر - ریچاردسون (21KR و 20KR) این است که سؤالات آزمون را می توان به صورت ۱ برای پاسخ های درست و صفر برای اما پاسخ های غلط تصحیح کرد. بنابراین نمی توان این فرمول ها را برای محاسبه ضریب پایایی پرسشنامه هایی که پاسخ های آنها از ۱ تا ۴ یا ۵ را شامل می شوند استفاده کرد.

متال ۲۴: با کدام یک از روش های زیر، متوسط ضریب بابایی (reliability) حاصل از دو نیمه کردن سؤال های آزمون در تمام حالت های ممکن به دست می آید؟

۱) اجرای فرم های موازی با همتا

۲) آلفای کرونباخ

3) روش بازآزمایی

۴) همبستگی آزمون با یک ملاک خارجی

پاسخ: گزینه «۲» روش دو نیمه کردن آزمون، نوعی روش ضریب آلفاست، در واقع ضریب آلفا معرف میانگین ضرایب پایایی حاصل از همه دو نیمه های ممکن یک آزمون است تفاوت بین روش ضریب آلفه و روش دو نیمه کردن در واحد تحلیل است، در روش دو نیمه کردن، یک نیمه آزمون با نیمه دیگر آزمون مقایسه می شود، ولی در روش ضریب آلفا هر سؤال با تک تک سؤال های دیگر مقایسه می شود.

مثال ۲۵: در کدام یک از روش های زیر، ضریب پایایی (Reliability) آزمون با یک بار اجرای آزمون محاسبه می شود؟

1) بازآزمایی و دو نیمه کردن

۲) کودر - ریچاردسون و فرم های همتا

۳) دو نیمه کردن و کودر - ریچاردسون

۴) فرم های همتا و دو نیمه کردن

پاسخ: گزینه «3» برای کنترل عوامل درونی (همسانی درونی) در محاسبه پایایی، روش های زیر وجود دارند: دو نیمه کردن آزمون، آلفای کرونباخ، کودر - ریچاردسون ۲۰ و کودر- ریچاردسون ۲۱

مثال ۲۶: عامل p در فرمول کودر - ریچاردسون با کدام شاخص زیر ارتباط بیشتری دارد؟

۱) ضریب روایی

۲) ضریب دشواری

3) شاخص تمییز

۴) خطای معیار اندازه گیری

پاسخ: گزینه «۲» می دانیم که p ضریب دشواری سؤال است و واریانس سؤال طبق این محاسبه به دست می آید:

= P(1-P) واریانس سؤال

بنابراین در فرمول کودر- ریچاردسونه ۲0 = p داریم که به ضریب دشواری اشاره دارد.

ج - روش ضریب آلفای کرونباخ

وقتی آزمون از بخش ها و قسمت های مختلف تشکیل شده باشد، مثل زمانی که یک آزمون از تعدادی خرده آزمون تشکیل شده است، برای محاسبه ضریب پایایی کل آزمون، از روش آلفای کرونباخ استفاده می کنیم که کار این روش از اجزاء یا قسمت های آزمون برای محاسبه پایایی استفاده می شود. زمانی که قسمت های آزمون همان سؤال های آزمون باشند و تصحیح سؤلات به صورت 0 و 1 هوا باشد، ضریب آلفا با 20KR برابر خواهد شود.

ضریب پایایی کل آزمون، =j تعداد بخش های آزمون، واریانس نمرات هر بخش آزمون،

واریانس نمرات کل آزمون

ضریب آلفا را زمانی که نمره گذاری آزمون به صورت دو ارزشی (صفر و ۱) نیست، بلکه به صورت چند ارزشی است، هم می توان استفاده کرد؛ مثل پرسش نامه های شخصیت یا نگرش سنج که پاسخ دهنده به آن ها به صورت طیفی از کاملا موافق تا کاملا مخالف پاسخ می دهد.

به عنوان مثال: یک آزمون از ۴ خرده آزمون تشکیل شده است که واریانس این خرده آزمون ها برابر است با:

ضریب پایایی آزمون را محاسبه کنید. (j= 4)

روش های همسانی درونی، ضریب پایایی را با توجه به تعداد سؤال های آزمون (k) و میانگین همبستگی میان سؤال ها برآورد می کنند. اگر این دو عامل با هم ترکیب شوند، فرمول زیر به دست می آید که به این فرمول هم ضریب آلفا می گویند.

ضریب همبستگی هر یک از سؤال های آزمون با سؤال دیگر همان آزمون

= میانگین همه ضرایب همبستگی بین سؤال ها

روش دو نیمه کردن آزمون، نوعی روش ضریب آلفاست. ضریب آلفا معرف میانگین ضرایب پایایی حاصل از همه دو نیمه های ممکن، یک آزمون است. تفاوت بین روش ضریب آلفا و روش دو نیمه کردن در واحد تحلیل است. در روش دو نیمه کردن، یک نیمه آزمون با نیمه دیگر آزمون مقایسه می شود، ولی در روش ضریب آلفا هر سوال با تک تک سؤال های دیگر مقایسه می شود.

مثال ۲۷: در کدام یک از روشهای محاسبه ی ضریب پایایی (Reliability) اجرای آزمون فقط در یک دفعه کافی است؟

۱) فرم های همتا ا و دو نیمه کردن

2) بازآزمایی و دو نیمه کردن

۳) فرم های همتا و کودر - ریچاردسون

۴) دو نیمه کردن و کودر - ریچاردسون کردن

پاسخ: گزینه «۴» در روش کودر-ریچاردسون آزمون یک بار اجرا می شود، اما در این روش همه ی ماده های آزمون تحلیل می شوند؛ همچنین در روش دو نیمه کردن آزمون را اجرا (با گروه واحدی این کار را انجام می دهیم) و سپس بعد از اجرا آن آزمون را به دو نیم ته ون این دو نیم تقسیم می کنیم و ضریب همبستگی حاصل از نمرات دو نیمه ی آزمون، ضریب پایایی هر یک از دو نیمه ها خواهد بود.

مثال ۲۸: اگر بخواهیم ضریب پایایی یک آزمون نگرش سنج ۵ درجه ای را اندازه بگیریم بهتر است از کدام روش زیر استفاده کنیم؟

۱) ضریب آلفای کرونباخ

۲) کودر ریچاردسون ۲۰

۳) کودر- ریچاردسون ۲۱

4) دو نیمه کردن

پاسخ: گزینه «۱» چنانچه بخواهیم پایایی سؤال های آزمون را که به صورت اجزاء و قسمت های مختلف هستند محاسبه کنیم، باید همه بخـش هـای و قسمت ها را که خرده آزمون می باشند، اندازه گیری نماییم که برای این محاسبه از ضریب آلفا استفاده روش آلفای کرونباخ نمرات سؤالات به صورت یک طیف از کاملا موافق تا کاملاً مخالف است و هیچ یک از جواب ها درست یا غلط محسوب نمی شوند. با آلفای کرونباخ که به روش لیکـرت نـمـره گذاری می شود، آزمون صفتی را می سنجد که دارای نوسان است.

روش آلفای کرونباخ یکی از روش های همسانی درونی است که در آن بر یکنواختی ماده ها یا اجزای یک آزمون تأکید می شود.

روش تعیین پایایی آزمون های سرعت

آزمون ها را می توان به ۲ دسته آزمون های قدرت و آزمون های سرعت تقسیم کرد. آزمون های قدرت آزمون هایی هستند که در آن هـا بـه آزمـون شـونـده هـا فرصت کافی داده می شود را به سؤالات پاسخ دهند. آزمون های پیشرفت تحصیلی معمولاً از نوع آزمـون هـای قـدرت هستند. این گونـه آزمـون هـا شـامل سؤال های چالش انگیزند و هدف اصلی آن ها این است که سطح عملکرد یا میزان دانش و توانایی آزمـون شوندگان را بسنجد، اما آزمـون هـای سـرعت – آزمون هایی هستند که برای پاسخ دادن به سؤال ها محدودیت زمانی وجود دارد و اکثر آزمون شوندگان در محدوده معین زمانی نمی توانند به همه سـؤال هـا ر پاسخ دهند. هدف از آزمون های سرعت، سنجیدن سرعت عمل آزمون شوندگان است. هیچ یک از روش های تعیین پایایی دو نیمه کردن و کودر ریچاردسون برای آزمون های سرعت مناسب نیستند؛ زیرا در آزمون های سرعت، آزمون شوندگان تقریبا همه سؤال هایی را که وقت رسیدگی به آن ها را دارند، درست جواب می دهند و تنها تفاوت بین آنها در تعداد سؤال هایی است که جواب می دهند. بنابراین اگر نمرات نیمه اول را با نیمه دوم آزمون همبسته سازیم ضریب حاصل یک ضریب بسیار کوچک خواهد بود؛ زیرا همه آزمون شوندگان سؤالات نیمه اول را درست جواب می دهند، در حالی که در نیمه دوم آزمـون تفاوت های زیادی میان نمرات مشاهده خواهد شد. گرچه می توان روش بازآزمایی را به کار برد، اما بهتر است برای تعیین پایایی آزمـون هـای سرعت از روش فرم های هم ارز استفاده کنیم.

مثال ۲۹: برای محاسبه پایایی (Reliability) در آزمونهای سرعت از کدام شیوه استفاده می شود؟

۱) تنصیف زمانی

۲) فرم های هم ارز

۳) تنصیف با زوج و فرد کردن سؤالات

۴) تنصیف یا دو نیمه کردن بالایی و پایینی سؤالات

پاسخ: گزینه «۲» روش های تعیین پایایی عبارت اند از: ۱- روش پایایی مصححان ۲- روش بازآزمایی ۳- روش فرم های هم ارز ۴- روش های همسانی درونی شامل: روش دو نیمه کردن آزمون - روش کودر - ریچاردسون و روش ضریب آلفای کرونباخ.

آزمون های سرعت به آزمون هایی گفته می شود که برای جواب دادن به آن ها محدودیت های زمانی وجود دارد و اکثر آزمـون شـوندگان در محدوده زمانی تعیین شده نمی توانند آن را به پایان برسانند. سؤال های آزمون سرعت بسیار ساده اند و هدف، سنجیدن سرعت عمل آزمون شوندگان می باشد. در ارتباط با روش های سنجش پایایی، هیچ یک از روش های دو نیمه کردن یا کودر - ریچاردسون بـرای تعیین پایایی آزمـون هـای سرعت مناسب نیست، چون در آزمون های سرعت آزمون شوندگان تقریباً همه ی سؤال هایی را که وقت رسیدگی به آن ها را دارند، درست جواب می دهند و تنهـا تفـاوت آنها در تعداد سؤال هایی است که جواب می دهند؛ بنابراین اگر نمرات نیمه ی اول آزمون را با نمرات نیمه ی دوم آزمون همبسته سازیم، ضریب حاصل یک ضریب بسیار کوچک خواهد بود؛ زیرا همه ی آزمون شوندگان سؤالات نیمه اول را درست جواب می دهند در حالی که در نیمه ی دوم آزمـون تفاوت هـای زیـادی مـیـان نمرات مشاهده خواهد شد. اگر به جای نیمه ی اول و دوم آزمون، سؤال های زوج و فرد را جدا کنیم و همبستگی آن ها را به دست بیاوریم، این ضریب خیلی بزرگ خواهد بود؛ زیرا برای هر یک از آزمون شوندگان نمرات فرم و زوج تقریباً یکسان است. روش دیگر بازآزمایی است؛ از آنجایی که یادآوری مطالب و تمرین روی مطلب سبب افزایش تفاوت بین عملکرد میشود، بهترین روش برآورد پایایی آزمون های سرعت این است که از روش فرم های هم ارز استفاده کنیم.

مثال۳۰: کدام روش، برآورد جامع تری از پایایی آزمون به دست می دهد؟

۱) بازآزمایی

۲) دو نیمه کردن

۳) کودر - ریچاردسون

۴) فرم های همتا

پاسخ: گزینه «۲» روش های همسانی درونی عبارتند از روش دو نیمه کردن آزمون، روش کودر ریچاردسون و روش ضریب آلفا که به ترتیب از اهمیت برخوردارند. در روش دو نیمه کردن، آزمون یک بار با یک گروه واحد آزمون شوندگان اجرا می شود و پس از اجرا به دونیمه تقسیم میشود. ضریب همبستگی حاصل از نمرات دو نیمه آزمون، ضریب پایایی هریک از دو نیمه خواهد بود که به فرمول اسپیرمن - براون مشهور است. گزینه ی سه به دلیل قرار گرفتن در گروه روش های همسانی درونی می تواند داوطلب را به خطا بیندازد.

پایایی نمرات اختلاف

نمره اختلاف می تواند اختلاف بین نمره های دانش آموزان در یک آزمون، در دو فرم موازی یک آزمون، نمره اختلاف بین عملکرد دو گروه دانش آموز در یک آزمون یا اختلاف نمرات یک گروه دانش آموز در آزمون های مختلف باشد. مثلاً زمانی که می خواهیم اختلاف بین پیش آزمون و پس آزمون را مشخص کنیم.

پایایی نمرات اختلاف از طریق فرمول زیر به دست می آید:

= rdپایایی نمره اختلاف

= پایایی آزمون اول

پایایی آزمون دوم

همبستگی دو آزمون

مثال ۳۱: اگر پایایی نمرات پیش آزمون 70/0 و پایایی نمرات پس از آزمون 80/0 باشد، در صورتی که ضریب همبستگی بین دو آزمون 35/0 باشد، پایایی نمره ی اختلاف چقدر می باشد؟

1) 59/0

2) 65/0

3) 61/0

4) 57/0

پاسخ: گزینه «۳»

نکته۲: هرچه همبستگی بین دو آزمون بیش تر باشد، پایانی نمره اختلاف آن دو آزمون کم تر است و بالعکس.

عوامل مؤثر در پایایی آزمون های وابسته به هنجار

۱ـ هرچه تعداد سؤالات آزمون بیشتر باشد، پایایی هم افزایش می یابد.

۲- هرچه سؤالات آزمون متجانس تر و همگون تر باشند، پایایی هم افزایش می یابد.

۳ـ سؤالاتی که ضریب تمییز بالایی دارند، پایایی آزمون را افزایش می دهند.

۴- سؤالات دارای ضریب دشواری متوسط، پایایی آزمون را افزایش می دهند.

5ـ هرچه آزمون شوندگان تجانس داشته باشند، پایایی افزایش می یابد.

۶ـ ضریب پایایی آزمون های سرعت بیشتر از آزمون های قدرت است.

مثال ۳۲: یک روش مطمئن برای افزایش ضریب پایایی آزمون چیست؟

۱) تغییر نوع سؤال های آزمون

۲) افزایش سطح دشواری سؤال ها

۳) افزایش تعداد سؤال های آزمون

۴) طرح سؤال های مرتبط از نظر محتوا

پاسخ: گزینه «3» تعدادی از عوامل مؤثر بر افزایش پایایی آزمون عبارت اند از:

1- با افزودن بر تعداد سؤالات آزمون، پایایی آن نیز افزایش می یابد.

۲- با متجانس تر و همگون تر کردن سؤالات یک آزمون، پایایی افزایش می یابد.

۳- با افزایش سؤالاتی که قدرت تمییز زیاد دارند، پایایی آزمون افزایش می یابد.

۴- با افزودن تعدادی سؤال دارای ضریب دشواری متوسط به یک آزمون پایایی آن نیز افزایش می یابد.

5- اگر آزمون را برای گروهی از افرادی اجرا کنیم که تجانس کمتری با یکدیگر دارند، این آزمون پایایی بیشتری دارد.

6- آزمون های سرعت بیشتر از آزمون های قدرت، پایایی دارند.

گزینه ی ۲، می تواند یکی از گزینه هایی باشد که ذهن داوطلب را به سمت و سوی نادرست هدایت کند، اما با توجه به این که برای بالا بردن پایایی آزمون باید ضریب دشواری متوسط برای سؤالات در نظر گرفت، بنابراین گزینه ی درست نمی باشد.

مثال ۳۳: کدام یک از روش های محاسبه پایایی قریب پایایی را بیشتر از سایر روش ها برآورد می کند؟

1) آلفای کرونباخ

۲) کودر - ریچاردسون۲۰

۳) کودر - ریچاردسون ۲۱

۴) دو نیمه کردن اسپیرمن - براون

پاسخ: گزینه «1» با توجه به این که روش دو نیمه کردن نوعی روش ضریب آلفا است، بنابراین ضریب آلفا معرف، میانگین ضریب های پایایی حاصل از همه ی دو نیمه های ممکن یک آزمون است؛ همچنین توجه به این نکتان ضروری است که در روش ضریب آلفا، آزمون از بخش ها یا قسمت هایی تحت عنوان خرده آزمون، تشکیل شده است و نسبت به حالتی که آزمون به صورت صحیح (۱) و غلط (0) در روش کودر - ریچاردسون است، قوی تر عمل می کند.

مثال ۳۴: از جمله روش های افزایش پایایی (Reliability) آزمون کدام است؟

۱) افزایش تعداد گزینه ها

۲) افزایش تعداد سؤال های آزمون

۳) افزایش سطح دشواری سؤال ها

۴) کاهش سطح دشواری سؤال ها

پاسخ: گزینه «۲» عوامل مؤثر در افزایش پایایی شامل:

۱- با افزودن تعداد سؤالات، پایایی افزایش می یاید.

۲- تجانس و همگونی سؤالات باعث افزایش پایایی می شود.

۳- با افزودن تعداگر سؤال با تمییز بالا، پایایی افزایش می یابد.

۴- با افزودن تعداد سؤال با دشواری متوسط، پایایی آزمون افزایش می یابد.

5- با اجرای آزمون برای افرادی که تجانس کمتری دارند، پایایی افزایش می یابد.

6- آزمون های سرعت بیشتر از آزمون های قدرت، ضریب پایایی را نشان میدهند.

مثال ۳۵: برای افزایش اعتبار آزمون و کاهش مقدار نسبی اشتباه باید ............................

1) بر طول آزمون افزود.

۲) درصد دشواری را ۴۰ درصد کرد.

۳) جای سؤالات را عوض کرد.

۴) درصد دشواری سؤالات را ۷۵ درصد کرد.

پاسخ: گزینه «۱» یکی از راه های افزایش پایایی یک آزمون است. به طور کلی راه های افزایش پایایی آزمون عبارتند از:

1) افزایش طول آزمون با سؤالات همگن

۲) حذف سؤالات ناهماهنگ با آزمون

۳) افزایش سؤالاتی با درجه دشواری متوسط

۴) افزایش سؤالاتی با ضریب تمییز بالا

برآورد نمره حقیقی با استفاده از ضریب پایایی آزمون

مفهوم پایایی به این نکته اشاره دارد که هیچ آزمونی کاملا پایا نیست و نمره های هر آزمون همیشه با مقداری خطا همراه است.

این خطا که بخشی از نمره به دست آمده یا نمره مشاهده شده (نمره خام) است، جزء خطای نمره محسوب می شود. عوامل مختلفی در پدید آمدن این دم جزء خطا دخیل هستند. بنابراین واریانس خطا یک واریانس تصادفی یا غیر منظم است.

هر نمره مشاهده شده از ۲ جزء تشکیل شده است: ۱- نمره حقیقی ۲- نمره خطا

نمره حقیقی

نمره خطا نمره مشاهده شده

یک واریانس خطا وجود دارد که تصادفی و نامنظم است و یک واریانس حقیقی وجود دارد که غیرتصادفی و منظم است. واریانس حقیقی نشان دهنده توانایی مورد سنجش است ولی واریانس خطا ناشی از یک عامل تصادفی است.

در تفسیر نمره های یک آزمون همواره از این مفاهیم استفاده می شود: نمره مشاهده شده که همان نمره خام است، نمره حقیقی و نمره خطا. نمره مشاهده شده یا نمره به دست آمده نمره ای است که یک آزمون شونده در آزمون به دست می آورد. نمره حقیفی به میانگین فرضی تعداد بیشماری آزمون گفته می شود که در مورد یک آزمون شونده اجرا می شوند و فرض بر این است که تمرین بر نمرات پی در پی کسب شده تأثبر ندارد. تفاوت بین نمره مشاهده شده و نمره حقیقی نمره خطا یا خطای اندازه گیری نام دارد.

واریانس نمره حقیقی واریانس نمره مشاهده شده

با توجه به اطلاعات بالا پایایی آزمون را از طریق یکی از فرمول های زیر می توان محاسبه کرد.

نکته 3: ضریب پایایی، ارقام بین 0 تا ۱+ را شامل می شود و ضریب پایایی ۱ نشان می دهد که صد در صد واریانس مشاهده شده واریانس حقیقی است و ضریب پایایی صفر، یعنی تمام واریانس مشاهده شده واریانس خطاست.

مثال ۳۶: اگر واریانس نمرات مشاهده شده در یک آزمون برابر ۵۰ و واریانس خطا معادل ۱۰ باشد، مقدار ضریب پایایی Reliablity coefficient چه قدر خواهد بود؟

1) 20/0

2) 45/0

3) 80/0

4) 90/0

پاسخ: گزینه «3» اگر واریانس نمرات مشاهده شده آزمون را با ، واریانس فرد حقیقی را با و واریانس نمرات خطای آن را با نشان دهیم، رابطه بین این واریانس ها به شرح زیر خواهد بود:

که با توجه به متن سؤال که ما فقط واریانس نمرات مشاهده شده و واریانس خطا را در اختیار داریم، از فرمول زیر استفاده می کنیم.

مثال ۲۷: اگر واریانس نمره حقیقی آزمونی ۷۵/0 و واریانس نمرات مشاهده شده آن ۸۰/0 باشد، پایایی این آزمون چقدر می باشد؟

1) 88/0

2) 93/0

3) 85/0

4) 97/0

پاسخ: گزینه «۲»

خطای معیار اندازه گیری

با استفاده از ضریب پایایی می توانیم خطای معیار اندازه گیری را محاسبه کنیم خطای معیار اندازه گیری با انحراف معیار توزیع نمرات خطا برابر است.

=SEMخطای معیار اندازه گیری

=SDtانحراف معیار آزمون

= پایایی

فرض کنید یک آزمون دارای ضریب پایایی 83/0 و انحراف استاندارد ۷ باشد. طبق فرمول بالا مقدار خطای معیار اندازه گیری برابر است با:

مثال ۳۸: اگر انحراف معیار یک آزمون ۳ و ضریب پایایی آن ۶۴/0 باشد، خطای معیار اندازه گیری آن چقدر است؟

1) 8/1

2) 9/1

3) 4/2

4) 2/3

پاسخ: گزینه «۱» فرمول خطای معیار اندازه گیری با استفاده از ضریب پایایی به این شرح است:

=SEMخطای معیار اندازه گیری

=انحراف معیار آزمون = 3

=ضریب پایایی آزمون = 64/0

مثال ۳۹: اگر انحراف معیار یک آزمون ۴ و ضریب پایایی آن ۶۴/0 باشد، خطای معیار اندازه گیری کل آزمون چه قدر خواهد بود؟

1) 8/12

2) 6/9

3) 2/3

4) 4/2

پاسخ: گزینه «۴» فرمول خطای معیار اندازه گیری را با اعداد متن سؤال جایگزین می کنیم.

استفاده از خطای معیار اندازه گیری برای برآورد نمره حقیقی

عملکرد دانش آموز در نیمرخ نمره های او، به صورت نوارهایی در حول نمره خام یا رتبه درصدی او نشان داده می شود و هر نوار به اندازه یک خطای معیار اندازه گیری بالا و یک خطای معیار اندازه گیری پایین هر یک از نمره ها رسم می شود. این نوار همان فاصله اطمینان است. منظور از فاصله اطمینان دامنه ای از نمرات است که با احتمال معین در برگیرنده نمره حقیقی است.

خطاهای معیار اندازه گیری عکس پایایی است؛ یعنی هرچه خطای معیار اندازه گیری کم باشد، پایایی زیاد و خطا کم است و هر چقدر خطای معیار اندازه گیری بزرگتر باشد، پایایی کاهش می یابد. در واقع خطای معیار اندازه گیری آشکارا نشان می دهد که به فرد در بازآزمایی یک آزمون معین چه مقدار به تغییر خواهد کرد. نمره هر فردی در یک آزمون دارای مقداری خطاست. غالبا خطاها اندک اند، اما ممکن است بعضی از آنها خیلی بزرگ باشند. در ضمن نیمی از خطاها مثبت اند و سبب بزرگ تر شدن نمره مشاهده شده می شوند و نیم دیگر منفی اند و نمره مشاهده شده را کم نشان می دهند، ولی ما تنها نمره مشاهده شده را در اختیار داریم و نمی دانیم که نمره خطای یک فرد چقدر در چه جهتی است.

البته بهتر است در اطراف نمره مشاهده شده، یک محدوده درست کنید متناسب با نمره مشاهده شده به اضافه و منهای یک خطای معیار اندازه گیری آزمون به این محدوده نوار اطمینان می گویند. آشکار است هرچه خطای معیار اندازه گیری آزمون بزرگتر باشد، این نوار پهن تر و هرچه خطای معیار اندازه گیری کوچک تر باشد، این نوار باریک تر است.

نکته ۴: خطای معیار اندازه گیری با واریانس رابطه مستقیم و با پایاایی رابطه عکس دارد.

مثال ۴۰: افزایش خطای معیار اندازه گیری یک آزمون با تغییرات عوامل، بستگی دارد به ................... واریانس و .................. پایایی.

۱) افزایش و کاهش

۲) افزایش و افزایش

۳) کاهش و کاهش

۴) کاهش و افزایش

پاسخ: گزینه «1» بین واریانس آزمون و خطای معیار اندازه گیری رابطه مستقیم وجود دارد؛ یعنی با افزایش واریانس در یک آزمون، خطای معیار نیز افزایش می یابد، اما بین ضریب پایایی و خطای معیار اندازه گیری رابطه معکوس وجود دارد؛ یعنی هرچقدرپایایی یک آزمون افزایش یابد، خطای معیار نیز کاهش می یابد و آزمون از اعتبار و دقت بیشتری برخوردار است.

مثال ۴۱: اگر پایایی آزمونی افزایش یابد، کدام نمره کاهش پیدا خواهد کرد؟

1) استاندارد

۲) خطا

۳) واقعی

۴) مشاهده شده

پاسخ: گزینه «۲» خطای معیار اندازه گیری با پایایی نسبت عکس دارد.

با استفاده از خطای معیار اندازه گیری می توان حدود نمره حقیقی را برآورد کرد که برای این کار از فرمول زیر استفاده می کنیم:

نمره دانش آموز = X، میانگین =

در سطح اطمینان ۹۹% برابر است با ← 58/2

در سطح اطمینان ۹۵% برابر است با به ←96/1

در سطح اطمینان ۶۸% برابر است با ←1

مثال ۴۲: اگر ضریب پاپایی یک آزمون برابر با ۷۲/0 باشد و انحراف معیار آن آزمون ۵ باشد، خطای معیار اندازه گیری را محاسبه کنید.

1) 64/2

2) 80/1

3) 3

4) 50/1

پاسخ: گزینه «۱»

مثال ۴۳: نمره دانش آموزی در آزمونی ۵۰ می باشد. اگر خطای معیار اندازه گیری برابر با ۲ باشد، با ۹۵% اطمینان نمره او بین کدام دو حد قرار دارد؟

1) ۴۲و۵۲

۲) ۴۶ و۵۰

۳) ۴۶و۵۴

۴) ۴۶و۵۲

پاسخ: گزینه «3»

56

مثال ۴۴: نمره دانش آموزی در یک آزمون ۷۰ می باشد. اگر انحراف استاندارد آزمون ۵ و ضریب پایایی آن ۸۴/0 باشد، با ۹۵% اطمینان حدود نمره واقعی را بر آورد کنید. ۱

1) ۷۰ و ۷2

2) ۷۲ و۷۴

3) 66 و 70

۴) ۶۶ و ۷۴

پاسخ: گزینه «۴»

76

مثال ۴۵: یک آزمون تحت کدام شرایط دارای پایایی است؟

1) همبستگی نمره خطا و مشاهده شده معنادار باشد.

۲) نمره های مشاهده شده با نمره های واقعی همبسته باشند.

۳) بین نمره های خطا و واقعی همبستگی بالا باشد.

۴) نمره های مشاهده شده با نمره های خطا همبسته باشند.

پاسخ: گزینه «۲» عوامل متعددی در افزایش پایایی یک آزمون تأثیر گذار هستند که به ایجاد همبستگی بین نمره های واقعی با نمره های مشاهده شده منجر می شوند. این امر در صورتی رخ می دهد که سؤالات یک آزمون با یکدیگر متجانس تر و همگون تر باشند، به این ترتیب برای خطا کمتر می شود و نمره حقیقی و نمره ی مشاهده شده به هم نزدیک تر می شوند. نمره های خطا ناشی از عوامل تصادفی می باشند، نمره خطا + نمره حقیقی = نمره مشاهده شده

مثال ۳۴۶: واریانس واقعی یک آزمون ۵۰ سؤالی ۶۴ است. اگر تعداد سؤال های این آزمون از طریق افزودن سؤال های همگون با آزمون اصلی به ۱۰0 افزایش داده شود، واریانس واقعی آن چه قدر می شود؟

1) 128

2) 200

3) 256

4) 400

پاسخ: گزینه «3»

2 C =

مثال ۴۷: نمره دانش آموزی در یک آزمون ۶۰ می باشد. اگر انحراف استاندارد این آزمون ۵ و ضریب پایایی آن ۸۴/0 باشد، با ۹۵ درصد احتمال حدود نمره واقعی او را برآورد کنید.

۱) بین ۵۷ و ۶۳

۲) بین ۵۶ و ۶۴

۳) بین ۸۵ و ۶۲

۴) بین ۵۵ ور۶۵

پاسخ: گزینه «۲» با توجه به این که 05/0 a = و است بنابراین داریم:

56

مثال ۴۸: واریانس واقعی یک آزمون ۱۴۴ است (۲۴۴ ). اگر تعداد سؤال های این آزمون از طریق افزودن سؤال های همگون با آزمون اصلی به دو برابر افزایش داده شود، واریانس واقعی آن چقدر می شود؟

1) 576

2) 288

3) 144

4) 72

پاسخ: گزینه «۱» با توجه به این که سؤالات آزمون دو برابر افزایش یافته اند، با این فرمول محاسبه می شود:

2 C =

رابطه بین روایی و پایایی

پایایی شرط روایی است؛ یعنی این که یک آزمون باید پایا باشد تا بتوان روا باشد. هر چقدر پایایی افزایش یابد، روایی هم افزایش می یابد و هر چقدر از مقدار پایایی کاسته شود از روایی هم کاسته می شود.

خطای اندازه گیری، همبستگی بین دو آزمون x و y یا روایی پیش بینی را کاهش می دهد. همبستگی بین اندازه های ناپای x و y از همبستگی واقعی بین آن ها کمتر خواهد بود. اسپیرمن برای رفع این مشکل و برآورد رابطه واقعی بین x و y فرمول اصلاح کاهش را ارائه کرد:

همبستگی بین x و y بعد از اصلاح،

پایایی x

پایاییy

به عنوان مثال: اگر ضریب همبستگی بین ۲ آزمون x و y ۴۰/0 باشد، ضریب پایایی آزمون x ، یعنی

باشد و ضریب پایایی آزمون y یعنی باشد، ضریب همبستگی اصلاح شده بین x و y چقدر است؟

در استفاده عملی از فرمول اصلاح کاهش می توان تعیین کرد که اگر پایایی یکی از دو آزمون یا هر دو آنها را به مقدار معینی افزایش دهیم، همبستگی بین x و y یعنی چه مقدار افزایش خواهد داشت. از فرمول رو به رو برای پاسخ دادن به این سوال استفاده می کنیم:

ضریب همبستگی بین x و y پس از اصلاح ضریب پایایی x و یا y

ضریب پایایی آزمون x پیش از تغییر

ضریب پایایی آزمون x و پیش از تغییر

ضریب پایایی آزمون y پس از تغییر

ضریب پایایی آزمون yو پس از تغییر

مثال ۴۹: بین دو آزمون که ضرایب پایایی آن ها 60/0 و 75/0 است، همبستگی 30/0 را به دست آورده ایم، یعنی . اگر پایایی دو آزمون را به 85/0 افزایش دهیم، چقدر بر مقدار افزوده می شود؟

1) 14/0

2) 38/0

3) 18/0

4) 20/0

پاسخ: گزینه «۲»

پایایی روش های درجه بندی با رتبه بندی

معروف ترین روش برای تعیین پایایی روش های درجه بندی، روش گیلفورد است. در این روش تعدادی داور تعدادی از افراد را از بالا به پایین یا از زیاد به کم درجه بندی با رتبه بندی می کنند بعد نتایج حاصل از این رتبه بندی را در فرمول زیر قرار می دهیم و ضریب پایایی درجه بندی را به دست می آوریم:

میانگین همبستگی های بین داوران با پایایی روش درجه بندی یک داور

=M تعداد داوران

=N تعداد افراد مورد داوری

=S مجموع رتبه ها برای هر نفر

رقم حاصل از این فرمول، پایایی یک داور است. برای به دست آوردن پایایی همه داوران باید از فرمول اسپیرمن - براون استفاده کنیم.

= ضریب پایایی همه داوران

= ضریب پایایی یک داور

= k نسبت تعداد کل داوران به یک داور

مثال ۵۰: اگر تعداداوران ۵ نفر باشد و ضریب پایایی یک داور ۴۸/0 باشد، ضریب پایایی ۵ داور را محاسبه کنید.

1) 82/0

2) 78/0

3) 86/0

4) 89/0

پاسخ: گزینه «۱»

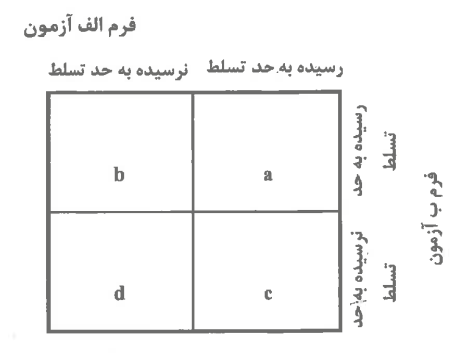
پایایی آزمون های وابسته به ملاک

در تعیین پایایی آزمون های وابسته به ملاک، هم می توانیم از روش بازآزمایی استفاده کنیم و هم از روش فرم های هم ارز که روش فرم های هم ارز ترجیح داده می شود. پایایی آزمون های وابسته به ملاک تعیین می کند که وضعیت اکثر آزمون شوندگان از لحاظ دستیابی یا عدم دستیابی به حد تسلط در دو موقعیت مختلف تا چه اندازه از ثبات برخوردار است.

زمانی که اکثر آزمول شوندگان از لحاظ دستیابی یا عدم دستیابی به حد تسلط در دو موقعیت آزمون (دو بار اجرای یک آزمون یا اجرا در فرم موازی یک آزمون) یکسان باشند، آزمون موردنظر از پایایی بالایی برخوردار است.

روش های تعیین ضریب پایایی آزمون های وابسته به ملاک

1- روش درصد توافق: به ضریب پایایی به دست آمده از این طریق، ضریب ثبات تصمیم گیری هم می گویند. برای محاسبه این ضریب، یک چهارخانه تشکیل می دهیم و بعد اطلاعات لازم را در فرمول قرار می دهیم و ضریب پایایی را محاسبه می کنیم.



N = a + b + c + d

a: افرادی که در فرم الف و ب آزمون به حد تسلط نرسیده اند.

b: افرادی که در فرم الف آزمون به حد تسلط رسیده اند ولی در فرم ب آزمون به حد تسلط نرسیده اند.

c:کسانی که در فرم الف آزمون به حد تسلط نرسیده اند ولی در فرم ب به حد تسلط رسیده اند.

:dافرادی که در فرم الف آزمون و فرم ب آزمون به حد تسلط رسیده اند.

مثال ۵۱: دو فرم موازی از یک آزمون ۳۰ سؤالی را روی یک گروه ۴۰ نفری از دانش آموزان اجرا کرده ایم. نمره رسیدن به حد تسلط برای این افراد نمره ۲۴، یعنی کسب دست کم ۸۰ درصد نمره آزمون است. با استفاده از اطلاعات جدول مقابل ضریب پایایی را محاسبه کنید؟

1) 60

2) 77

3) 67

4) 69

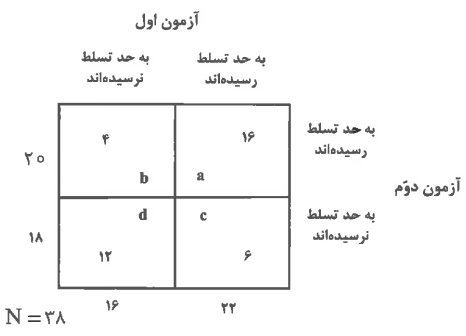
|  |  |  |
| --- | --- | --- |
| 25 | 4 | 29  11 |
| 5 | 6 |
| 30 10 40 | | |

در این روش هم از چهار خانه مانند چهار خانه مقابل استفاده می کنیم. اطلاعات لازم را در فرمول kappa قرار داده، ضریب پایایی را محاسبه می کنیم.

سه و چهار عدن



به عنوان مثال: برای چهار خانه نیم ضریب پایایی کاپا را محاسبه کنید.



3-روش لیوینگ استون: هالیوینگ استون این طور استدلال می کند که تفاوت بین نمره حد تسلط و میانگین آزمون حد تسلط را می توان واریانس حقیقی تصور کرد؛ حتی زمانی که بین نمره های مشاهده شده تفاوت اندکی وجود دارد یا بین آن ها هیچ گونه تفاوتی وجود ندارد، این به اصطلاح نمره حقیقی می تواند چشم گیر باشد. او برای تعیین پایایی فرمول زیر را ارائه داد:

پایایی آزمون وابسته به ملاک

پایایی آزمون با روش وابسته به هنجار (20KR ، دو نیمه کردن و...)

واریانس آزمون براساس تفاوت نمره های مشاهده شده و نمره حد تسلط

براساس این فرمول، اگر میانگین نمره های آزمون و نمره حد تسلط برابر باشند ضریب پایانی لیوینگ استون با ضریب پایایی محاسبه شده با روش های مورد استفاده در آزمون های وابسته به هنجار برابر است، اما وقتی که بین میانگین آزمون و نمره حد تسلط تفاوت وجود دارد، ضریب لیوینگ استون بزرگ تر از ضرایب پایایی محاسبه شده با روش های مورد استفاده در آزمون های وابستگر به هنجار یا سنتی خواهد بود.

عوامل مؤثر در افزایش پایایی آزمون (وابسته به ملاک)

1- 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 سؤال های زیادتری بنویسید.

۳- هدف های خود را تا حد ممکن دقیق و مشخص بنویسید.

۴- از روش عینی نمره گذاری استفاده کنید.

5- سؤال های نمونه شبیه به سؤال های اصلی آزمون را در اختیار آزمون شنوندگان قرار دهید.

6- سؤال های آزمون را سازمان یافته و روشن در اختیار آزمون شوندگان قرار دهید.

پایایی روش های سنجش مستقیم رفتار

روش های سنجش مستقیم رفتار مشکل روایی ندارند، اما مسأله پایایی در آنها مطرح است؛ چون ممکن است این روش ها دقت کافی نداشته باشند. برای تعیین پایایی سنجش محصولات ماندنی رفتار از فرمول زیر استفاده می کنیم:

برای تعیین پایایی روش ثبت فراوانی رفتار از فرمول زیر استفاده می کنیم:

از بزرگترین فراوانی ثبت شده محاسبه ضریب پایایر روش سنجش طول مدت نهفتگی رفتار، مانند روش تعیین پایایی روش سنجش فراوانی رفتار است با این تفاوت که در این روش طول مدت رفتار در صورت کسر به جای کوچک ترین فراوانی ثبت شده، کوچکترین طول مدت ثبت شده برای رفتار و در مخرج کسر، به جای بزرگترین فراوانی ثبت شده، بزرگترین طول مدت ثبت شده را قرار می دهند. برای تعیین پایایی روش سنجش طول مدت و دوره نهفتگی رفتار از فرمول زیر استفاده می کنیم:

برای تعیین ضریب پایایی روش های ثبت فاصله ای و ثبت نمونه گیری زمان از فرمول زیر استفاده می کنیم:

اگر دو مشاهده گر مستقل به ثبت رفتار واحدی بپردازند، آنچه باید معین شود این است که فاصله زمانی متوالی ثبت شده از سوی آنها از لحاظ علامت های مربوط به وقوع پاسخ یا عدم وقوع پاسخ در هر یک از فاصله های زمانی با هم مطابقت دارد یا نه.

مثال ۵۲: برای کدام یک از موارد زیر مشکلم روایی مطرح نیست بلکه مسأله پایایی مطرح است؟

۱) روش های سنجش آزمون های سرعت

۲) روش های سنجش نمرات اختلاف

3) روش های سنجش مستقیم

۴) همه موارد

پاسخ: گزینه «٣» در روش سنجش مستقیم پایایی مطرح است و مشکل روایی اهمیتی ندارد.

تست های طبقه بندی شده فصل هفدهم

۱- انحراف معیار توزیع نمرارخطا در نهایت آزمایش از یک فرد را اصطلاحا خطای معیار ................. می گویند.

(سراسری ۸۱)

1) برآورد

۲) اندازه گیری

۳) نمونه گیری

۴) پیش بینی

۲- در کدام یک از روش های برآورد پایایی آزمون، عامل رشد و یادگیری ممکن است در برآورد پایایی تأثیر بگذارد؟

(سراسری ۸۱)

1) بازآزمایی

۲) دو نیمه کردن

۳) اجرای همزمان فرم های همتا

۴) کودر - ریچاردسون

۳- آزمون های سرعت در مقایسه با آزمون های غیر سرعت معمولا ضریب .................... را نشان می دهند.

(سراسری ۸۱)

1) روایی بیشتری

۲) روایی کمتری

3) پایایی بیشتری

۴) پایایی کمتری

۴- ضریب پایایی یک آزمون ۲۰ سؤالی ۶/0 است. اگر ۲۰ سؤال به آن اضافه شود، ضریب پایایی چقدر می شود

(سراسری ۸۱)

1) 65/0

2) 70/0

3) 75/0

4) 80/0

5- برای محاسبه همسانی درونی آزمون هایی که پاسخ سؤالاتشان به صورت چند ارزشی نمره گذاری می شود، کدام شیوه مطلوب تر است؟

(سراسری ۸۱)

۱) تنصیف

۲) ضریب آلفا

3) کودر- ریچاردسون ۲۰

۴) کودر - ریچاردسون ۲۱

6- اگر در آزمونی، تعداد سؤال ها کاهش یابد، معمولا:

(سراسری ۸۱)

1) پایایی (Reliability) افزایش می یابد.

۲) روایی (Validity) کاهش می یاید.

3) روایی (Validity) افزایش می یابد.

۴) پایایی (Reliability) کاهش می یابد.

۷- از هر یک از اوراق تشریحی فلسفه و زبان، ۱۰۰ ورقه را در اختیار تعدادی معلم قرار داده ایم تا جداگانه تصحیح کنند. نتایج به دست آمده مطابق جدول زیر است. درصد توافق معلمان در مورد اوراق فلسفه چقدر است؟

(سراسری ۸۲)

|  |  |  |  |
| --- | --- | --- | --- |
| نام درس | قبول از نظر همه معلمان | مردود از نظر همه معلمان | قبول از نظر یک عده و مردود از نظر یک عده دیگر |
| فلسفه | 27 | 33 | 40 |
| زبان | 15 | 25 | 60 |

1) 27

2) 33

3) 40

4) 60

8- به منظور از پایایی (Reliability) آزمون کدام است؟

(سراسری 82)

۱) همسانی در اندازه گیری

۲) کاهش خطای اندازه گیری

۳) بررسی هدف مورد اندازه گیری

۴) یکسانی محتوای آزمون

۹- اعتبار یک آزمون از طریق دو نیمه کردن معادل ۷۰/0 است. اعتبار کل آزمون چقدر است؟

(سراسری ۸۲)

1) 35/0

2) 75/0

3) 82/0

4) 4/1

۱۰- در محاسبه ی کدام شاخص، از خطای استاندارد اندازه گیری استفاده می شود؟

(سراسری ۸۲)

1) محاسبه ی ضریب تمیز کل آزمون

۲) محاسبه ی ضریب دشواری آلفای کرونباخ

۳) روایی آزمون

۴) ضریب پایایی آزمون

1۱- تعیین پایایی به شیوه های مرسوم در پژوهش، بر اساس کدام یک از نظریه های زیر است؟

(آزاد ۸۲)

1) کلاسیک

2) منحنی خم سؤال

۳) منحنی ویژگی آزمون

۴) خطای نمونه گیری و اندازه گیری

۱۲- کدام یک از روش های نامبرده در زیر پایایی بیش از اندازه واقعی برآورد می کند؟

(آزاد ۸۲)

1) کودر - ریچاردسون ۲1

۲) کودر - ریچاردسون ۲۰

۳) کرونباخ

۴) بازآزمایی

۱۳- کدام یک از آزمون های زیر دارای بالاترین ضریب پایایی است؟

(آزاد ۸۲)

۱) همگن

۲) ناهمگن

۳) کوتاه

۴) بلند ولی نامتجانس

۱۴- انحراف معیار آزمونی ۱۵ و ضریب پایایی آن ۶۴/0 می باشد. خطای معیار اندازه گیری کدام است؟

( آزاد ۸۲)

1) 3

2) 4

3) 9

4) 12

۱۵- یک آزمون ۱۸۰ سؤالی را به ۶ آزمون ۳۰ سؤالی تقسیم کرده ایم. اگر بخواهیم ضریب پایایی همسانی درونی این آزمون را حساب کنیم، بهترین روش کدام است؟

(سراسری ۸۳)

1) 20KR

2) 21KR

3) آلفای کرونباخ

۴) فرمول اسپیرمن - براون

۱۶- ضریب پایایی برای یک آزمون ۱۰۰ سؤالی برابر ۹۰/0 است. اگر طول آزمون را به تقلیل دهیم، ضریب پایایی چقدر خواهد بود؟

(سراسری ۸۳)

1) 73/0

2) 78/0

3) 85/0

4) 87/0

۱۷- فرمول کودر - ریچاردسون ۲۰ زمانی مناسب ترین کارایی را دارد که ......

(سراسری ۸۳)

۱) سطح دشواری آزمون مناسب باشد.

۲) سؤال های آزمون در ارزشی باشند.

۳) اختلاف بین دشواری سؤال ها زیاد باشد.

۴) سؤال ها حداقل چهار ارزش داشته باشد.

18- از واریانس های هر یک از ۴ سؤال یک آزمون انشایی به ترتیب ۹، 8/4، 2/10و ۱۶ محاسبه شده است. اگر واریانس کل سؤالات برابر ۱۰۰ باشد، ضریب پایایی آزمون چقدر خواهد شد؟

(سراسری ۸۳)

۱) ۶۵/0

2) 75/0

3) 80/0

4) 85/0

۱۹- اگر خطای معیار یک آزمون ۳ باشد، فراگیری که در آن نمره ۱۰ گرفته است، نمره واقعی وی به احتمال ۶۸ درصد بین کدام دو نمره قرار خواهد گرفت؟

(سراسری ۸۳)

1) ۱۰ و ۳

۲) ۱۳ و ۷

۳) ۱۳ و ۳

۴) ۱۳ و ۱۵

۲۰- در کدام یک از روش های تعیین پایایی، واریانس خطای ناشی از حافظه وجود ندارد؟

(سراسری ۸۳)

1) دو نیمه کردن آزمون

۲) اجرای مجدد آزمون

۳) فرم های همتا با فاصله زمانی

۴) فرم های همتا بدون فاصله زمانی

۲۱- نمره ای که در اثر اجرای یک آزمون به دست می آید، نمره .................. نامیده می شود.

(آزاد ۸۲)

1) واقعی

۲) حقیقی

۲) مشاهده شده

۴) خطا

۲۲- همبستگی بین دو آزمون موازی چه نامیده می شود؟

(آزاد ۸۳)

۱) اعتبار

۲) پایایی

۳) رابطه

۲۳- در کدام یک از شرایط زیر می توان پایایی آزمون را با استفاده از روش بازآزمایی تعیین کرد؟

(آزاد ۸۳)

1) رفتار مورد اندازه گیری فرد باشد.

۲) ویژگی مورد اندازه گیری تحت تأثیر زمان قرار داشته باشد.

3) صفت مورد نظر با گذشت زمان تغییر نکند.

۴) آزمون فقط پیشرفت تحصیلی باشد.

۲۴- برای افزایش پایایی یک آزمون کدام یک از فعالیت های زیر مناسب است؟

(آزاد ۸۳)

۱) کاهش تعداد سؤالات

۲) افزایش طول آزمون

3) بالا بردن اعتبار آزمون

۴) کاهش تعداد گزینه های غلط

۲۵- در صورتی که پایایی آزمون مساوی یک باشد، آزمون کدام یک از عوامل زیر را اندازه گیری می کند؟

(آزاد ۸۳)

۱) رفتار نهفته

۲) خصیصه مکنون

۳) ویژگی مورد مطالعه

۴) ویژگی آشکار

۲۶- در کدام یک از حالت های آزمون x و موازی هستند؟

(آزاد ۸۳)

1)

2)

3)

4)

۲۷- چنانچه باشد، انتظار می رود که رابطه ........... برقرار باشد.

(سراسری ۸۴)

1)

2)

3)

4)

۲۸- هر گاه در یک آزمون هوشی، انحراف معیار هوشبهرها ۱۰، ضریب پایایی (Reliability) ۷۵/0 و نمره احمد ۹۰ باشد، کدام یک از احکام زیر در تفسیر نمره صادق خواهد بود؟

(سراسری ۸۴)

1) به احتمال 3۴% نمره واقعی احمد بین 85 و ۹۵ است.

۲) به احتمال 68% نمره واقعی احمد بین ۸۰ و ۱۰۰ است.

3) به احتمال 95% نمره واقعی احمد بین ۷۵ و ۱۰۵ است.

۴) به احتمال ۹۹٪ نمره واقعی احمد بین ۷۵ و ۱۰۵ است.

۲۹- در یک آزمون تشریحی چهار سؤالی، واریانس کل آزمون ۸۱ و واریانس سؤال ها به ترتیب 8،4،9و16 است. ضریب آلفای آزمون کدام است؟

(سراسری ۸۴)

۱) ۷۲/0

2) 80/0

3) 85/0

4) 90/0

۲۰- در کدام یک از روش های تعیین پایایی (Reliability)، تمرین، بیشترین تأثیر را می تواند در پراش خطای آزمون داشته باشد؟

(سراسری ۸۴)

۱) تنصیف

۲) فرم های همتا

۳) بازآزمایی

۴) کودر - ریچاردسون

۳۱- اگر ضریب (Reliability) یک آزمون یا توزیع نرمال برابر با 84/0 باشد، خطای معیار اندازه گیری بر حسب نمرات Z چقدر خواهد بود؟

(سراسری ۸۴)

1) 16/0

2) 36/0

3)40/0

4) 68/0

۳۲- برآورد پایایی (Reliability) به کمک ثبات درونی در وضعیتی بهتر برآورد می شود که آزمون:

(سراسری ۸۴)

1) دارای سؤال متجانس و همگون باشد.

۲) به قسمت های موازی قابل تقسیم نباشد.

۳) دارای سؤال قابل تفکیک نباشد.

۴) دارای سؤال های مستقل و قابل تفکیک باشد.

۳۳- مقادیر پایایی حاصل از دو روش 20KR و 21KR در حالتی برابر خواهند بود که دشواری همه سؤال های آزمون ................... باشند، در غیر این صورت مقدار 21KR ................. از مقدار 20KR خواهد شد.

(سراسری ۸۴)

1) یکسان - کوچک تر

۲) یکسان - بزرگتر

۳) برابر یک - کوچکتر

۴) برابر یک - بزرگتر

۳۴- برای افزایش پایایی آزمون و کاهش مقدار نسبی خطا باید:

(آزاد ۸۴)

۱) درصد دشواری را ۷۵ درصد کرد.

۲) بر تعداد سؤال های آزمون افزود.

۳ آزمودنی ها را بیشتر همگن کرد.

۴) بر طول هر سؤال افزود.

۳۵- یک اشتباه نظام دارد که معمولا از طریق عملکرد متفاوت دو یا چند گروه از افراد در آزمون مشخص می شود........................ گفته می شود.

(آزاد ۸۴)

1) ضریب عدم تعین

2) جنبه های قانونی آزمون

۳) سوگیری آزمون

۴) سوگیری فرهنگی

۳۶- در برآورد هشدار خطا در نمره آزمون کدام اندازه پراکندگی استفاده می شود؟

(آزاد ۸۴)

1) انحراف متوسط

۲) انحراف چارکی

۳) دامنه میان چارکی

۴) انحراف معیار

۳۷- اگر ارزش خطای معیار اندازه گیری نزدیک به صفر باشد به این معنی است که:

(آزاد ۸۴)

1) با تعداد سؤال ها ناهمبسته است.

۲) اندازه گیری زیاد اشتباه دارد.

3) هم اندازه گیری بسیار پایایی دارد.

۴) آزمون فاقد روایی است.

۳۸- اگر انحراف معیار یک آزمون ۱۵ و ضریب پایایی آن ۸۴/0 باشد، خطای معیار اندازه گیری کدام است؟

(آزاد 8۴)

1) 5/7

2) 5/4

3) 6

4) 9

۳۹- اگر تعداد سؤال های یک آزمون را دوبرابر کنیم، خطای معیار اندازه گیری:

(آزاد ۸۴)

1) دوبرابر می شود.

2) نصف نمی شود.

۳) نصف می شود.

۴) تغییر نمی کند.

۴۰- چنانچه باشد، آنگاه کدام یک از رابطه های زیر صحیح است؟

(سراسری ۸۵)

1)

2)

3)

4)

۴۱- خطای استاندارد اندازه گیری آزمونی برابر با 5/2 است. دانش آموزی در این آزمون نمره ی ۸۴ کسب کرده است، نمره حقیقی این دانش آموز به احتمال ۹۵ درصد :

(سراسری 85)

۱) همان ۸۴ است.

۲) بین ۷۹ و ۸۹ است.

۳) بین 5/81 و 5/86 است.

۴) بین 5/76 و 5/91 است.

۴۲- پایایی (Reliability) آزمونی را که برای سنجش تسلط دانش آموزان دوره ی دبیرستان به زبان انگلیسی تهیه شده است با استفاده از ۱۲۰ دانش آموز سال های اول دوم و سوم متوسطه محاسبه کردیم و مقدار 80/0+ به دست آمده است. اگر نمرات دانش آموزان سال دوم را حذف و ضریب پایایی را مجددا با استفاده از نمرات دانش آموزان سال اول و سوم محاسبه نماییم، ضریب پایایی به احتمال زیاد:

(سراسری ۸۵)

۱) تغییری نخواهد کرد.

۲) افزایش خواهد یافت.

3) کاهش خواهد یافت.

۴) تغییر خواهد کرد، اما معلوم نیست در کدام جهت

۴۳- هرگاه یک آزمون، موضوعات متفاوتی را اندازه گیری کند، استفاده از کدام روش احراز پایایی منجر به ضریبی کمتر از روش های دیگر خواهد شد؟

(سراسری ۸۵)

۱) تنصیفی ( Split- half )

۲) بازآزمایی ( Test - retest)

۳) کودر - ریچاردسون ( Kuder- Richardson ) .

۴) فرم های معادل (Parallel forms)

۴۴- در کدام یک از گروه های زیر، ضریب پایایی یک آزمون واحد بیشتر است؟

(سراسری ۸۵)

1) جمعیت پسران ۱۲ ساله

۲) جمعیت دختران بالای ۱۵ سال

۳) جمعیت مردان بالای ۲۵ سال

۴) جمعیت مردان و زنان بالای ۲۰ سال

۴۵- اگر ضریب همبستگی دو نیمه کردن یک آزمون نزدیک به صفر باشد، ولی سؤالات هر نیمه با هم رابطه داشته باشند، نتایج تحلیل عامل سؤالات چگونه خواهد بود؟

(سراسری ۸۵)

1) دو عامل گروهی

۲) دو عامل اختصاصی

۳) یک عامل مهم عمومی

۴) یک عامل گروهی و یک عامل اختصاصی

۴۶- دلیل اینکه معلم میخواهد نمره امتحان دانش آموز کمتر تحت تأثیر خطاهای تصادفی قرار گیرد، کدام است؟

(سراسری 85)

1) تعیین اثر دقیق آموزش

۲) به حد مطلوب رساندن ضریب دشواری

۳) تنظیم ضریب تمییز، متناسب با توانایی شاگردان

۴) نزدیک کردن نمره مشاهده شده به نمره واقعی شاگرد

۴۷- رابطه مفاهیم پایایی (Reliability) و روایی (Validity) را کدام جمله زیر به درستی بیان می کند؟

(آزاد ۸۵)

۱) هر گاه آزمون پایایی داشته باشدف روایی هم دارد.

۲) هر گاه آزمونی روا نباشد نمی تواند پایایی نشان دهد.

۳) روایی و پایایی دو مفهوم کاملا مستقل از هم هستند.

۴) برای این که آزمون روا باشد حتما باید پایا هم باشد.

۴۸- کدام روش تأمین پایایی از نظر استفاده بهتر از نمونه ها برتری دارد؟

(آزاد ۸۵)

۱) آزمون - آزمون مجدد

۲) دو فرم همتا

۳) دو نیمه کردن

۴) کودر - ریچاردسون

۴۹- اگر خطای معیار اندازه گیری آزمونی برابر با ۱ باشد، با ۹۵ درصد اطمینان نمره واقعی فردی که در این آزمون ۱۵ را به دست آورده بین کدام دو حد قرار می گیرد؟

(سراسری ۸۶)

1) 16-14

2) 96/16-13

3) 28/16-72/13

4) 58/17-42/12

50- چنانچه تعداد سؤالات یک آزمون را دو برابر کنیم، واریانس خطای آزمون چه تغییری می کند؟

(سراسری 86)

1) نصف می شود.

۲) دو برابر می شود.

۳) چهار برابر می شود.

۴) تغییر نمی کند.

51- به انحراف معیار آزمونی برابر با 8 ضریب اعتبار ((Reliability آن برابر با 75/0می باشد. خطای معیار اندازه گیری این آزمون چقدر است؟

(سراسری ۸۶)

1) 8

2) 6

3) 4

4) 2

۵۲- در کدام یک از روش های تعیین اعتبار (Reliability) آزمون ها به کار کرد افراد در تک تک سؤالات آزمون توجه می شود؟

(سراسری ۸۶)

۱) بازآزمایی

۲) تنصیف

۳) فرم های همتا

۴) کودر – ریچاردسون

۵۳- میانگین و انحراف معیار نمره های دانش آموزان یک کلاس ۳۱ نفری در آزمونی ۵۰ سؤالی به ترتیب برابر با ۳۰ و ۶ شده است. ضریب اعتبار (Reliability) این آزمون چقدر است؟

(سراسری ۸۶)

1) 33/0

2) 68/0

3) 90/0

4) 04/1

۵۴- چنانچه ضریب همبستگی بین دو نیمه یک آزمون ۵/0باشد، ضریب اعتبار (Reliability) کل آزمون چقدر است؟

(سراسری 86)

1) 1

2) 5/0

3) 67/0

4) 75/0

55- اگر ضریب اعتبار (Reliability) آزمونی برابر با یک باشد، کدام یک از گزینه های ذیل درست است؟

(سراسری 86)

۱) نمره مشاهده شده برابر با نمره خطاست.

۲) نمره مشاهده شده بزرگتر از نمره واقعی است.

۳) نمره مشاهده شده برابر با نمره واقعی است.

۴) نمره واقعی برابر با نمره خطاست.

56- در کدام یک از روش های تعیین اعتبار (Reliability)، میزان تغییر در طول زمان به عنوان منبعی از خطا در نظر گرفته می شود؟

(سراسری ۸۶)

۱) بازآزمایی

۲) تنصیف

3) کودر - ریچاردسون

۴) فرم های همتا بدون فاصله زمانی

۵۷- چه رابطه ای بین ضریب اعتبار (Reliability) و روایی (Validity) آزمون وجود دارد؟

(سراسری ۸۶)

1) اعتبار آزمون همواره بزرگ تر یا مساوی روایی می باشد.

۲) اعتبار آزمون همواره کوچک تر از ضریب روایی می باشد.

۳) اعتبار آزمون همواره بزرگتر از ضریب روایی آزمون می باشد.

۴) اعتبار آزمون کوچکتر یا مساوی روایی می باشد.

5۸- عبارت « نسبت واریانس نمرات واقعی آزمودنی ها به واریانس مشاهده شده آنان» معرف کدام ضریب زیر است؟

(آزاد ۸۶)

۱) پایایی

۲) تمییز

۳) دشواری

۴) همبستگی

۵۹- ضریب پایایی شامل چه مقادیری است؟

(آزاد ۸۶)

1) از 1- تا ۱+

۲) از صفر تا ۱-

۳) محدودیتی وجود ندارد.

۴) از صفر تا ۱+

۶۰- ضریب پایایی که از فرمول 20KR به دست می آید، تابع کدام یک از موارد زیر است؟

(آزاد ۸۶)

1) درجه همبستگی بین سؤال ها

2) ضریب تمییز سؤال ها

۳) ضریب دشواری آزمون

۴) ضریب پیش بینی آزمون

۶۱- کدام عبارت منعکس کننده این است که نمره واقعی برابر با امید ریاضی نمره مشاهده شده می باشد؟

(سراسری ۸۷)

1) E(X) = T

2) E(T) = X

3) E(T) = E

4) E(E) = T

۶۲- اگر بخواهیم ضریب اعتبار آزمون را بالا ببریم بهتر است به چه کاری دست بزنیم؟

(سراسری ۸۷)

۱) آزمون را به دو فرم موازی تقسیم کنیم.

۲) با سؤالات مشابه تعداد سؤالات آزمون را افزایش دهیم

۳) سؤالات آسان تری را در آزمون بگنجانیم.

۴) سؤالات دشوارتری را در آزمون بگنجانیم.

۶3- اگر همه سؤالات آزمون دارای ضریب دشواری یکسان باشند، کدام یک از روش های برآورد اعتبار را پیشنهاد می کنید؟

(سراسری ۸۷)

1) آلفا کرونباخ

۲) 21KR

3) 20KR

۴) کپا

۶۴- در کدام یکی از روش های تعیین پایایی، پراش (واریانس) خطای مربوط به حافظه بیشتر وجود دارد؟

(سراسری ۸۷)

1) کودر - ریچاردسون

۲) فرم های همتا بدون فاصله زمانی

۳) فرم های همتا با فاصله زمانی

۴) تنصیف

۶۵- آزمونی را با فاصله ی ۱۰ روز در یک گروه ۸ نفری دوباره اجرا کرده ایم و نتایج زیر به دست آمده است. ضریب اعتبار (Reliability) آزمون چقدر است؟

(سراسری ۸۷)

12-11-10-9-8-7-6-5 : نوبت اول

14-13-12-11-10-9-7 : نوبت دوم

1) 1-

2) 0

3) 1+

۴) بزرگتر از صفر ولی کوچکتر از یک

۶۶- اجرای دو فرم همتای یک آزمون در یک زمان مربوط به کدام یک از ویژگی های ضریب اعتبار است؟

(سراسری ۸۷)

۱) ثبات درونی

2) ضریب ثبات

3) ضریب معادل بودن

۴) معادل منطقی که

67- در آزمونی که ضریب اعتبار آن برابر با 84/0 و واریانس آن برابر با ۴ می باشد، فردی نمره ۱۱ را به دست آورده است. با ۹۵٪ اطمینان نمره واقعی این فرد در دامنه .................. قرار می گیرد.

(سراسری ۸۷)

1) 57/12- 43/9

2) 32/11-68/10

3) 64//11- 36/10

4) 14/14- 86/7

68- یک آزمون ۴۰ سؤالی دارای اعتبار (Reliability)۶۰/0 است. اگر تعداد سؤالات این آزمون را به ۶۰ سؤال افزایش دهیم، اعتبار آزمون چقدر خواهد شد؟

(سراسری ۸۸)

1) 63/0

2) 69/0

3) 73/0

4) 79/0

69- چنانچه آزمونی دارای ۱۰ سوال و اعتبار أن مساوی 80/0باشد، طول این آزمون چه مقدار باید افزایش یابد تا اعتبار آن به ۹۰/0 برسد؟

(سراسری 88)

1) 2

2) 25/2

3) 5/2

4) 52/2

۷۰- در کدام نوع آزمون، جمع کل نمرات کمتر اهمیت دارد، بلکه نمرات قسمت های مختلف یا درصد پاسخ های درست به گروه های سؤالات خاص اهمیت دارد؟

(سراسری ۸۸)

1) قدرت

۲) سرعت

3) تشخیصی

۴) بازده

71- از جمله مفروضات استفاده از فرمول کودر- ریچاردسون ۲۱ (21KR) کدام است؟

(سراسری ۸۸)

1) نمره گذاری سؤال، صفر و یک باشد.

۲) ضریب تمییز سؤالات پایین باشد.

۳) سطح دشواری سوالات متفاوت باشد.

۴) سطح دشواری سؤالات نزدیک به یک باشد.

۷۲- هنگامی که از ضریب آلفا (Coefficient of Alfa) استفاده می شود که:

(سراسری ۸۸)

1) انحراف معیار نمرات کمتر از ۳ باشد.

۲) آزمون از چندین بخش مجزا تشکیل شده باشد.

۳) ضریب دشواری سؤالات مشابه باشد.

۴) نمونه های آزمون شونده ناهمگن باشد.

۷۳- کدام یک از انواع اعتبار (Reliability) بیشتر بیانگر دقت آزمون است؟

(سراسری ۸۸)

۱) بازآزمایی

2) فرم های همتا

۳) دو نیمه کردن

۴) بین ارزیابان

۷۴- « خطای معیار اندازه گیری» آزمون به کدام عوامل به ترتیب بستگی دارد؟

(سراسری ۸۸)

۱) با انحراف استاندارد و پایایی آزمون ارتباط مستقیم دارد.

۲) با انحراف استاندارد و پایایی آزمون ارتباط معکوس دارد.

۳) با انحراف استاندارد رابطه معکوس و با پایایی آزمون ارتباط مستقیم دارد.

۴) با انحراف استاندارد ارتباط مستقیم و با پایایی آزمون ارتباط معکوس دارد.

۷۵- هنگامی که ضریب اعتبار (Reliability) آزمونی مساوی یک باشد، ............... خواهد بود.

(سراسری ۸۸)

1)

2)

3)

4)

76- کدام عبارت در مورد رابطه روایی و اعتبار صحیح است؟

(سراسری ۸۸)

1) اعتبار شرط لازم و کافی برای روایی است.

۲) روایی آزمون معمولا بیشتر از اعتبار آن است.

۳) اعتبار شرط لازم ولی ناکافی برای روایی است.

۴) آزمونی که روایی ندارد، اعتبار هم ندارد.

۷۷- کدام یک از روش های زیر روش برآورد پایایی نیست؟

(آزاد ۸۸)

۱) روش اجرای مجدد

2) روش جانسون

۳) روش دو فرم موازی

۴) روش ریچاردسون

۷۸- آزمون توانایی با قدرت آن نوع از آزمون ها است که:

(آزاد ۸۸)

۱) وقت محدودی دارد.

۲) آزمودنی ها می توانند به تمام سؤال ها پاسخ دهند.

3) مستلزم استعداد و اطلاعات ویژهای است.

۴) جواب غلط در آن وجود ندارد.

۷۹- کدام مورد رابطه احتمالی بین سؤال و ویژگی مورد اندازه گیری را دقیق تر نشان می دهد؟

(سراسری ۸۹)

۱) تحلیل عاملی

۲) پایایی سؤال

۳) منحنی ویژه سؤال

۴) اعتبار وابسته به ملاک

۸۰- برای محاسبه اعتبار (Reliability) آزمون های فرافکن، کدام روش مناسب تر است؟ /

(سراسری ۸۹)

۱) نمره گذاران

۲) تنصف

۳) آلفای کرونباخ

۴) فرم های همتا

۸۱- برای تعیین اعتبار (Reliability) نمرات سؤالات تشریحی، کدام روش مناسب تر است؟

(سراسری ۸۹)

1) 21KR

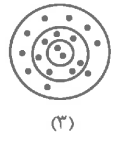
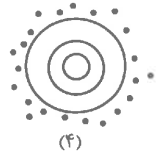
2) 20KR

3) دو نیمه کردن

۴) آلفای کرونباخ

۸۲- در کدام شکل، اعتبار (Reliability) وجود دارد اما روایی (Validity) وجود ندارد؟

(سراسری ۸۹)

۸۳- در شرایط یکسان، ضریب اعتبار بالای حاصل از روش باز آزمایی در کدام مورد مطلوب تر است؟

(سراسری 89)

1) نگرش به درس ریاضی

۲) پیشرفت تحصیلی در درس ریاضی

3) اضطراب امتحان ریاضی

۴) استعداد ادراک فضایی در ریاضی

۸۴- اگر واریانس نمرات برابر با ۹ و ضریب اعتبار (Reliability) برابر با ۷۵/0 باشد، خطای استاندارد اندازه گیری چقدر است؟

(سراسری ۸۹)

1) 5/4

2) 5/1

3) 25/2

4) 75/0

۸۵- کدام مورد تعریف اعتبار (Reliability) آزمون را نشان نمی دهد؟

(سراسری 89)

1) همبستگی بین نمره های فرم های موازی

۲) همبستگی بین نمره های واقعی و نمره های مشاهده شده

۳) نسبت واریانس واقعی به واریانس مشاهده شده

۴) مجذور همبستگی بین نمره های واقعی و نمره های مشاهده شده

۸۶- کدام ضریب برای تعیین اعتبار (Reliability) آزمون های ملاک - مرجع مناسب تر است؟

(سراسری ۹۰)

۱) ضریب تعمیم پذیری

۲) ضریب همبستگی دو نیمه آزمون

۳) ضریب همبستگی نمره مشاهده شده با نمره واقعی در آزمون

۴) ضریب همبستگی ناشی از اجرای آزمون در دو زمان مختلف

۸۷- مناسب ترین روش تعیین اعتبار (Reliability) آزمون های سرعت کدام است؟

(سراسری ۹۰)

۲) بازآزمایی

۳) آلفای کرونباخ

۴) دو نیمه کردن

۸۸- در تحلیل اعتبار (Reliability) سؤالات، ضریب همبستگی بین نمرات ....................... در یک گروه محاسبه می شود.

(سراسری ۹۰)

۱) هر جفت سؤال

۲) دو نیمه سؤالات

3) هر سؤال و نمره کل

۴) هر سؤال و نمره ملاک

۸۹- در تحلیل اعتبار (Reliability) آزمونی مشتمل بر سؤالات صحیح - غلط، استفاده از کدام روش مناسب تر است؟

(سراسری ۹۰)

1) پیرسون

۲) کودر - ریچاردسون

۳) اسپیرمن - براون

۴) ضریب رولون

۹۰ - روایی (Validity) یک آزمون برابر با ۸۰/0 و نمرات آزمون روی مقیاس آ گزارش شده است، خطای معیار برآورد کدام است؟

(سراسری ۹۰)

1) 8/0

2) 6

3) 8

4) 4/6

۹۱- چنانچه برابر با ۱۵/0 واریانس مشاهده شده باشد، مقدار اعتبار (Reliability ( کدام است؟

(سراسری 90)

1) 85/0

2)

3)

4)

۹۲ - کدام روش تعیین اعتبار (Reliability) نوسانات ناشی از گذشت زمان را بهتر کنترل می کند؟

(سراسری 90)

1) بازآزمایی

۲) فرم های معادل

3) همبستگی نمرات ارزیابان در زمان های مختلف

۴) هماهنگی درونی

۹۳– اگر باشد و طول آزمون دو برابر شود، ضریب اعتبار (Riliability) آزمون جدید چه مقدار خواهد شد؟

(سراسری ۹۰)

1) 5/0

2) 6/0

3) 8/0

4) 7/0

۹۴- کدام مورد همبستگی نمرات خطا و نمرات مشاهده شده را نشان می دهد؟

(سراسری ۹۱)

1)

2)

3)

4)

۹۵- کدام گزینه رابطه فرمول KR20 و KR21 را بهتر توصیف می کند؟

(سراسری ۹۱)

KR21 < KR20 (1

KR21≥ KR20 (2

KR21≤ KR20 (3

KR21 > KR 20 (4

۹۶- آیا می توان بدون مطمئن شدن از روایی تست، فقط با در نظر گرفتن اعتبار (Reliability) آزمون را به کار برد؟

(سراسری 91)

1) بلی، زیرا اهمیت اعتبار بیشتر از روایی است.

2) بلی، زیرا وجوگر اعتبار بالا دلیل روایی است.

۳) خیر، زیرا ممکن است با وجود اعتبار، آزمون با خصیصه مورد نظر مرتبط نباشد.

۴) خیر، زیرا وقتی اعتبار بالاست روایی پایین است

۹۷- کدام مورد همزاد مفهوم اعتبار (Reliability) در نظریه پرسش - پاسخ است؟

(سراسری ۹۱)

۱) تابع آگاهی

۲) تابع درست نمایی

3) تابع نسبت پاسخ صحیح

۴) تابع اجایو نرمال

۹۸- کدام یک از ویژگی های مربوط به اعتبار (Reliability) آزمون است؟

(سراسری ۹۱)

۱) عدم تغییر معنادار نتایج در آزمایش های مکرر

۲) قدرت پیش بینی عملکرد افراد در آینده

۳) سنجش همان چیزی که برای آن ساخته شده است.

۴) وجود همبستگی بین نمره آزمون و نمره های ملاک

۹۹- کدام عامل در افزایش ضریب اعتبار (Reliability) مؤثرتر است؟

(سراسری ۹۱)

۱) همگونی گروه

۲) ناهمگونی گروه

۳) دشواری آزمون

۴) سادگی آزمون

۱۰0- اگر طول آزمون با سؤالات مشابه دو برابر شود، اعتبار آزمون چه تغییری می کند؟

(سراسری 96)

1) دو برابر می شود.

۲) تغییری نمی کند.

۳) نصف می شود.

۴) تا حدودی اضافه می شود.

۱۰۱- اگر واریانس واقعی دو برابر مقدار واریانس خطا باشد، ضریب اعتبار (Reliability کدام است؟

(سراسری ۹۱)

1) ۶۷/0

2) 33/0

3) 5/0

4) قابل محاسبه نیست و به اطلاعات بیشتری نیاز است.

۱۰۲- کدام گزینه رابطه میان روایی و اعتبار (Reliablity) را بهتر بیان می کند؟

(سراسری 92)

۱) اگر آزمون روا باشد، اعتبار نیز دارد.

۲) اعتبار آزمون نسبت به روایی آن اهمیت بیشتری دارد.

۳) آزمون می تواند على رغم نداشتن اعتبار، روایی داشته باشد.

۴) اگر آزمون اعتبار داشته باشد، روایی نیز دارد.

۱۰۳- خطای استاندارد برآورد بیشتر تحت تأثیر کدام عامل قرار می گیرد؟

(سراسری 92)

۱) انحراف استاندار متغیر ملاک

۲) روایی آزمون

۳) اعتبار آزمون

۴) انحراف معیار خطای اندازه گیری

۱۰۴- در کدام یک از حالات از اعتبار (Reliability) یک آزمون کاسته می شود؟

(سراسری ۹۲)

1) اجرای آزمون در زمان محدود

۲) کم کردن تعداد گزینه ها .

3) اجرای آزمون روی آزمودنی های غیر متجانس

۴) طولانی کردن آزمون

۱۰۵- در صورتی که واریانس نمرات خطا مساوی با صفر باشد، کدام یک از موارد درست است؟

(سراسری ۹۲)

۱)

2)

3)

4)

۱۰۶- در صورتی که ضریب همبستگی دو نیمه آزمونی مساوی 80/0 باشد، ضریب اعتبار آن چقدر است؟

(سراسری ۹۲)

1) 89/0

2) 44/0

3) 8/0

4) 7/0

۱۰۷- در آزمونی ۲۰ سؤالی که اعتبار آن مساوی 80/0 است، طول آزمون باید چند برابر افزایش یابد تا ضریب اعتبار به ۹۰/0 برسد؟

(سراسری ۹۲)

1) 25/2

2) 7/24

3) 5/25

4) 28

۱۰۸- در آزمونی که واریانس آن برابر با ۲۵ و ضریب اعتبار (Reliability) آن برابر با 84/0 است، با ۶۸% اطمینان نمره واقعی فردی که نمره ۷۵ به دست آورده، در کدام محدوده قرار می گیرد؟

(سراسری ۹۲)

1) 79- 73

2) 95- 55

3) 85- 65

4) 77- 73

۱۰۹- اگر واریانس نمرات واقعی یک آزمون ۲ و واریانس نمرات خطا ۱ باشد، در صورتی که طول آزمون ۲ برابر شود، ضریب اعتبار آزمون چقدر خواهد شد؟

(سراسری ۹۲)

۱) ۵/0

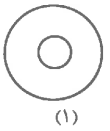
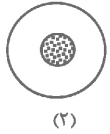
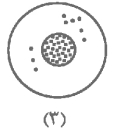
2) 6/0

3) 8/0

4) 7/0

۱۱۰- کدام شکل بیانگر آزمونی است که اعتبار (Reliability) دارد ولی روایی (Validity) ندارد؟

(سراسری ۹۳)

۱۱۱ـ کدام عمل باعث کاهش اعتبار (Reliability) در آزمون می شود؟

(سراسری 93)

۱) همگون کردن (تجانس) سؤالات

2) همگون کردن (تجانس) آزمودنی ها

۳) کاهش زمان پاسخ دهی به سؤالات

4) افزایش تعداد گزینه های سوالات

۱۱۲ـ اگر ضریب اعتبار (Reliability) یک آزمون برابر با ۷۵/0 و واریانس آن برابر ۹ باشد، خطای استاندارد اندازه گیری چقدر است؟

(سراسری ۹۳)

1) 75/0

2) 5/1

3) 25/2

4) 75/6

۱۱۴- کدام یک از آزمون های روان شناختی، به طور کلی دارای پایایی بیشتری هستند؟

(سراسری 94)

۱) توانایی

۲) شخصیت

۳) نگرش

4) علاقه

۱۱۳- اعتبار (Reliability) برآوردشده یک آزمون ۸۰ سؤالی،90/0 است. چنانچه طول این آزمون به نصف برسد، اعتبار آن چقدر خواهد شد؟

(سراسری ۹۳)

1) 81/0

2) 84/0

3) 88/0

4) 91/0

۱۱۵ـ مهم ترین ویژگی یک آزمون که باید موردنظر قرار گیرد، کدام است؟

(سراسری 94)

1) پایایی (reliability)

۲) عینیت (objectivity)

۳) عملی بودن (practically)

۴) روایی (validity)

۱۱۶ـ کدام مورد، می تواند ضریب پایایی (reliability) یک آزمون استاندارد باشد؟

(سراسری 94)

1) 86/0

2) 48/0

3) 30/1

4) 93/0-

۱۱۷- پایایی (reliability) یک آزمون ۲۰ سؤالی عزت نفس برابر با 90/0 است. کدام مورد در خصوص سؤالات این آزمون درست است؟

(سراسری 94)

۱) نمونه معرفی از رفتارهای بیانگر عزت نفس را اندازه می گیرند.

۲) در یک نمونه معرف از جامعه هنجاریابی شده اند.

۳) با نمره کل آزمون همبستگی پایینی دارند.

4) عزت نفس را به طور همسانی اندازه می گیرد.

۱۱۸ـ خطای معیار اندازه گیری یک آزمون برابر با 2/1 است. در صورتی که پایایی آزمون برابر با 64/0 باشد، واریانس نمرات آزمون چقدر است؟

(سراسری ۹۴)

1) 2

2) 4

3) 6

4) 8

پاسخنامه تست های طبقه بندی شده فصل هفدهم

1-گزینه «۲» یکی از فواید ضریب پایایی آزمون این است که با استفاده از آن می توان خطای معیار اندازه گیری آزمون را تعیین کرد. خظای معیار اندازه گیری برابر است با انحراف معیار توزیع نمرات خطا. فرض کنید از شخصی بخواهیم تا توپی را ۲۰۰ بار پرتاب کند و فرض کنید که انجام دادن این کار بدون تأثیر پذیرفتن شخص از خستگی یا تمرین ممکن باشد. نتیجه این عمل یک توزیع فراوانی از فواصلی خواهد بود که از این ۲۰۰ بار پرتاب شدن توپ حاصل می شود. این توزیع فراوانی مربوط به نمرات مشاهده شده است که دارای یک میانگین است که ما می توانیم آن را به طور تقریبی نمره حقیقی آن شخص در پرتاب کردن توپ بدانیم. این توزیع همچنین یگر انحراف معیار دارد که نحوه ی پراکندگی نمرات را نشان می دهد؛ از آنجا که این انحراف معیار خطای پرتاب توپ به وسیله ی آن شخص است، به آن خطای معیار اندازه گیری می گویند.

۲- گزینه «۱» روش های تعیین پایایی متنوع می باشند و به ۴ دسته تقسیم می شوند.

١- روش پایایی مصطحان، ۲ روش بازآزمایی ۳- روش فرم های موازی، ۴- روش های همسانی درونی: که این روش خود شامل این زیر گروه ها می شود:

الف: روش دو نیمه کردن آزمون

ب: روش کودر - ریچاردسون

ج: روش ضریب آلفای کرونباخ

بازآزمایی یکی از ساده ترین روش های تعیین پایایی است که در این روش یک آزمون در نوبت برای یک گروه آزمون شونده اجرا می شود و نمرات حاصل با همدیگر مقایسه می شوند. ضریب همبستگی به دست آمده بین نمرات حاصل از دو بار اجرا، ضریب پایایی آزمون نام دارند. در روش بازآزمایی چون یک گروه در در نظر گرفته می شوند، فاصله زمانی ممکن است منجر به رشد کم یادگیری آنان شود و تغییرات حاصل از آزمون را تحت تأثیر قرار دهد. گزینه ی سوم می تواند به گزینه ی صحیح نزدیک باشد که دو بار اجرا آزمون را دارد، اما در این فرمهای همتا، دو نوع سؤال برای یک آزمون تهیه می شود که با روش بازآزمایی تفاوت دارد.

۳- گزینه «3» مؤلفه های شش گانه ای در تعیین و افزایش پایایی آزمون وابسته با هنجار نقش دارند که یکی از آنها به این نکته اشاره دارد که آزمون های سرعت دارای ضریب پایایی بیشتری از آزمون های قدرت دارند. با توجه به گزینه ها اگر داوطلب نتواند تفاوت روایی و پایایی را تشخیص دهد، گزینه اول به عنوان نزدیک ترین گزینه انتخاب می شود که این انتخاب نادرست می باشد.

۴- گزینه «3» با استفاده از فرمول اسپیرمن - براون می توان چنین محاسبه ای انجام داد:

ضریب پایایی آزمون جدید

ضریب پایایی آزمون اولیه = 6/0

=K نسبت سؤال های جدید به سؤال های آزمون اولیه = 2،

۵- گزینه «۲» روش آلفای کرونباخ یکی از زیر گروه های روش همسانی درونی است که برای تعیین پایایی کاربرد دارد. در این روش، اجزا یا قسمت های آزمون برای سنجش ضریب پایایی آزمون به کار می روند. زمانی که یک آزمون از خرده آزمون تشکیل شده است، برای سنجش ضریب پایایی آزمون از روش آلفای استفاده می کنیم. گزینه ۳ کودر- ریچاردسون ۲۰ یا 20KR نزدیک ترین گزینه به گزینه ی صحیح می باشد. اگر قسمت های آزمون (خرده آزمون) همان سؤال های آزمون باشند و سؤال ها به صورت دو ارزشی درست (۱) و غلط (5) تصحیح بشوند، ضریب آلفا با 20KR برابر می شود که این گزینه داوطلب را به اشتباه می اندازد.

گزینه «۴» عوامل و مولفه های شش گانه ای وجود دارند که می توان از آن ها برای افزایش پایایی سؤال و آزمون استفاده کرد (که این پایایی مستقیما در روایی آزمون تأثیر دارد و آن را افزایش می دهد). یکی از این عوامل به این نکته اشاره دارد که اگر تعداد سؤالات آزمون افزایش یابد در نتیجه، پایایی آزمون نیز افزایش می یابد. با توجه به این که در متن سؤال این قاعده به صورت برعکس بیان شده است، گزینه ۴ صحیح می باشد. همان طور که گفتیم افزایش و کاهش پایایی مستقیما در افزایش و کاهش روایی تأثیر می گذارد؛ بنابراین می توان گفت که گزینه ۲ نیز به عبارتی درست است، اما چون ابتدا باید پایایی آزمون افزایش و کاهش پیدا کند تا روایی آن نیز افزایش و کاهش پیدا کند گزینه ۴ صحیح تر می باشد. ذکر این نکته لازم است که پایایی به دقت و اعتمادپذیری یا تکرارپذیری نتایج یک آزمون اشاره دارد.

۷- گزینه «۴» توافق معلمان شامل: قبول از نظر همه معلمان و مردود از نظر همه معلمان می شود که به این موضوع تنها در گزینه ی ۴ اشاره شده است. گزینه نزدیک می تواند گزینه ی شماره ۳ باشد که داوطلب اگر تعداد افراد قبول و مردود از نظر معلمان را جمع نمایم و از تعداد کل کسر نماید به این گزینه می رسد و دچار خطا می شود، زیرا از مسیر درست جهت رسیدن به جواب استفاده نکرده است.

۶۰ = ۲۷ + ۳۳ = جمع قبول و مردود از نظر همه

۴۰ = ۶۰ – ۱۰۰ ⟹ عدد اشتباه = جمع حاصل - تعداد کل

8- گزینه «۱» پایایی یک وسیله ی اندازه گیری آزمون است که عمدتا به دقت نتایج اشاره دارد. پایایی به دقت، اعتماد، ثبات و تکرارپذیری نتایج آزمون تأکید دارد. اگر یک آزمون بتواند در دفعات متفاوت اجرا شود و نتایج یکسان ارائه کند، این امر ثابت می کند که از پایایی برخوردار است. گزینه ی ۲، کاهش خطای اندازه گیری، خود یکی از اجزای تعیین پایایی است و گزینه ۴ هم به روایی اشاره دارد نه پایایی. پایایی به دنبال کسب اطلاعات مفید از دقت حاصل از نتایج آزمون است.

۹- گزینه «3» روش دو نیمه کردن، زیر مجموعه روش های همسانی درونی در تعیین پایایی آزمون است. در روش دو نیمه کردن یک آزمون را با یک گروه آزمون شونده اجرا می کنیم و پس از اجرا آن را دو نیمه می کنیم. مثلا سؤالات فرد را در یک گروه و سوالات زوج را در گروه دیگر قرار می دهیم. برای محاسبه پایایی از فرمول اسپیرمن - براون استفاده می کنیم.

70/0 = ضریب همبستگی بین نیمه های آزمون

۱۰- گزینه «۴» پایایی یک وسیله ی اندازه گیری عمدتا به دقت نتیجه آزمون اشاره دارد: خطای استاندارد اندازه گیری نیز برای محاسبه ی دقیق تر شاخص از پایایی کاربرد دارد. یکی از فواید پایایی آزمون این است که با استفاده از آن می توان خطای معیار اندازه گیری آزمون را تعیین نمود. اگر آزمونی دارای دقت کمتری در نمرات باشد، این موضوع که خطای استاندارد اندازه گیری منعکس می شود. گزینه های ۱ و ۲ به مراحل تحلیل یک آزمون اشاره دارد.

۱۱- گزینه «1»

۱۲- گزینه «۴» سه ضریب پایایی اول به روش های احراز پایایی همسانی درونی مربوط هستند که فقط آزمون یک بار اجرا می شود، اما در روش بازآزمایی دو بار آزمون اجرا می شود. آزمون شوندگان بار اول با آزمودن آشنا می شوند و یا بر اثر تمرین خود را تقویت می کنند و بار دوم بهتر به سؤالات پاسخ می دهند.

۱۳- گزینه «۴» به طور کلی می توان با افزودن بر تعداد سؤالات و متجانس و همگون تر کردن سؤالات یک آزمون، پایایی آن را افزایش داد علاوه بر این دو عامل می توان سؤالاتی با ضریب دشواری متوسط و ضریب تمییز بالا نیز برای افزایش پایایی آزمون به کار برد.

۱۴- گزینه «3» با توجه به فرمول زیر جواب صحیح که همان ۹ است به دست می آید:

ضریب پایایی انحراف معیار

۱۵- گزینه «3» زمانی که یک آزمون از قسمت های مختلف تشکیل شود، مثل زمانی که یک آزمون از تعدادی خرده آزمون تشکیل یافته است، از روش آلفای کرونباخ استفاده می کنیم.

۱۶- گزینه «۴»

ضریب پایایی آزمون جدید

ضریب پایایی آزمون اولیه،

=k نسبت سؤال آزمون جدید به آزمون اولیه.

۱۷- گزینه «3» فرض اساسی فرمول های کودر - ریچاردسون این است که سؤال های آزمون را می توان به صورت ا برای پاسخ صحیح و صفر برای پاسخ غلط تصحیح کرد. بنابراین اگر سؤال های آزمون در ارزشی باشند، مناسب ترین زمان برای استفاده از فرمول کودر- ریچاردسون می باشد.

۱۸- گزینه «3» طبق روش ضریب آلفا:

برای سوال این مفروضات را داریم

۱۹- گزینه «۲» با توجه به مفروضات سؤال، خطای معیار این آزمون ۳ می باشد و نمره نیز ۱۰ است. بنابراین با احتمال ۶۸% داریم:

حداکثر ۱۳

= 1× 3± 10 نمره = نمره واقعی

حداقل 7

۲۰- گزینه «۱» در روش دو نیمه کردن، آزمون را یکبار با گروه واحدی از آزمون شوندگان اجرا می کنیم و پس از اجرا، آن را به دو نیمه تقسیم می کنیم. بهترین راه تقسیم، براساس سؤالات زوج و فرد می باشد. بنابراین در این آزمون، خطای ناشی از حافظه نقشی ندارد؛ زیرا فقط یک بار آزمون برگزار شده است.

۲۱- گزینه «3» نمره حقیقی و واقعی هر دو یک مفهوم هستند و به میانگین فرضی تعداد بیشماری آزمون گفته می شود که در مورد یک آزمون شونده اجرا می شوند و فرض می شود که تمرین بر نمرات پی در پی کسب شده تأثیر ندارد. نمره خطا به تفاوت بین نمره مشاهده شده (نمره خام) یا همان نمره ای که فرد در انر اجرای آزمون به دست می آورد و نمره حقیقی گفته می شود. بنابراین گزینه ۳ صحیح می باشد. اگر معرف نمره مشاهده شده ، معرف نمره حقیقی و معرف نمره خطا باشد، تساوی مقابل را در مورد آنها خواهیم داشت:

۲۲- گزینه «۲» یکی از راه های تعیین پایایی این است که ضریب همبستگی بین نمرات حاصل از اجرای فرم های موازی را به دست آوریم. از دیگر روش های تعیین پایایی می توان به روش پایایی مصححان، روش بازآزمایی و روش های همسانی درونی اشاره کرد.

۲۳- گزینه «3» ساده ترین روش تعیین پایایی یک آزمون، بازآزمایی آن است که این روش تنها زمانی قابل استفاده است که ویژگی مورد اندازه گیری در طول زمان تغییر نکند. به عنوان مثال چنین تصور می شود که هوش یک ویژگی یا صفت نسبتا پایدار است که در فاصله اجرای یک آزمون هوشی در دو نوبت متفاوت تغییر نمی کند.

۲۴- گزینه «۲» یکی از راه هایی که می توان برای افزایش پایایی یک آزمون از آن استفاده کرد افزایش طول آزمون است. از دیگر مواردی که می توان در این رابطه استفاه کرد، موارد زیر است:

١- متجانس و همگون کردن سؤالات یک آزمون

۲- افزودن تعدادی سؤال با ضریب تمییز بالا به یک آزمون

٣- افزودن تعدادی سؤال با ضریب دشواری متوسط به یک آزمون

۴- اجرای آزمون در مورد گروهی که تجانس کمتری با هم دارند

5- استفاده از آزمون های سرعت به جای آزمون های قدرت

کاهش تعداد سؤالات یک آزمون باعث کاهش پایایی می شود، بنابراین گزینه اصحیح نمی باشد. بالا بردن اعتبار آزمون در افزایش پایایی مؤثر نیست، د بلکه این افزایش پایایی است که باعث افزایش روایی آزمون می شود، بنابراین گزینه م نیز صحیح نمی باشد. با کاهش تعداد گزینه های غلط شانس افراد برای پاسخگویی صحیح به سؤال افزایش می یابد و گروه بیشتر متجانس می شود. در صورتی که گروه دارای تجانس کمتر باعث افزایش پایایی می شود، بنابراین این گزینه نیز اشتباه می باشد.

۲۵- گزینه «۴» زمانی که پایایی مساوی یک باشد، به این معنا است که نمره ی مشاهده شده با نمره ی حقیقی برابر است و هیچگونه خطایی نداشته است؛ پس آزمون ویژگی آشکار آزمون شونده را سنجیده است.

۲۶- گزینه «۲» دو فرم یک آزمون در صورتی موازی هستند که میانگین واریانس آن ها برابر باشند. بعضی از صاحب نظران معتقدند که برای داشتن , آزمون های موازی واقعی علاوه بر میانگین و واریانس، پایایی آزمون ها و رابطه میان نمرات آزمون ها و متغیرهای دیگر نیز باید برابر باشد. ضریب همبستگی بین نمرات حاصل از اجرای فرم های موازی، ضریب پایایی آن آزمون ها به حساب می آید. از بین گزینه های موجود گزینه ۲ که تساوی بدون ضریب همبستگی در آزمون را نشان می دهد، از بقیه گزینه ها به جواب سؤال نزدیک تر می باشد.

۲۷ـ گزینه «۲» اگر ضریب پایایی آزمون صفر باشد، خطای معیار اندازه گیری در حداکثر قرار دارد و واریانس نمره مشاهده شده با واریانس نمره خطا برابر است.

۲۸- گزینه «۴»

75= 5- 10- 90 = حداقل نمره

انحراف معیار = 10

ضریب پایایی = 75/0

۲۹- گزینه «۱» طبق فرمول روش ضریب آلفا:

*4*j = *بخش های آزمون*

۳۰- گزینه «۳» در روش بازآزمایی، آزمون را به گروه واحـدی از آزمـون شـوندگان در دو نوبت می دهند و نمرات حاصل را با هم مقایسه می کنند. روش بازآزمایی از نظر متخصصان اندازه گیری و ارزشیابی مورد انتقاد است. به نظر آنان تجربه نوبت اول آزمون سبب می شود که آزمون شوندگان با سؤال های آزمون آشنا شوند. این آشنایی سبب می شود که آنان در آزمون نوبت دوم بهتر عمل کنند.

۳۱- گزینه «۳» بر حسب نمرات Z انحراف معیار برابر ۱ است.

1= انحراف معیار Z

۳۲- گزینه «۴» زمانی که آزمون دارای سؤال های مستقل و قابل تفکیک باشد، برآورد پایایی به کمک ثبات درونی بهتر انجام می شود. در روش همسانی درونی، بر همسانی یا یکنواختی ماده ها یا اجزای تشکیل دهنده یک آزمون تأکید می شود.

۳۳- گزینه «۱» طبق تعریف، زمانی که همه سؤال های آزمون از لحاظ دشواری یکسان باشند، پایایی دو روش 20KR و 21KR برابر می شود. هر قدر از لحاظ درجه دشواری میان سؤال های آزمون تفاوت وجود داشته باشد، به همان نسبت از ضریب پایایی حاصل از فرمول 21KR می شود. به طور کلی می توان گفت استفاده از فرمول 21KR، آسان تر از استفاده از فرمول 21KR است.

SEM=SD+ Itt = 0/AF

۳۴- گزینه «۲» به طور کلی برای افزایش پایایی می توان فعالیت های زیر را انجام داد:

الف) افزودن به تعداد سؤالات یک آزمون

ب) متجانس تر و همگون تر کردن سؤالات یک آزمون

ج) افزودن تعدادی سؤال با ضریب تمییز بالا و تمییز دشواری متوسط به آزمون

چ) اجرای آزمون در بین گروه های متجانس

ح) اجرای آزمون های سرعت به جای آزمون های قدرت

۳۵- گزینه «۳»

۳۶- گزینه «۴» خطای معیار اندازه گیری برابر است با انحراف معیار توزیع نمرات خطا؛ بنابراین در برآورد تعداد خطا در تمـردی آزمـون از انحراف معیار استفاده می شود. فرض کنید از شخصی بخواهیم تا توپی را ۲۰۰ بار پرتاب کند، طوری که که انجام دادن این کار بدون تات تأثیر پذیرفتن شخص از خستگی یا تمرین ممکن باشد. نتیجه این عمل توزیعی است با میانگین که ما می توانیم آن را به طور تقریبی نمره حقیقی آن شخص در پرتاب کردن توپ بدانیم. ایـن توزیع همچنین یک انحراف معیار دارد که نحوه ی پراکندگی نمرات را نشان می دهد. از آنجا که این انحراف معیار خطای پرتاب توپ به وسیله آن شخص است، به آن خطای معیار اندازه گیری می گویند.

۳۷- گزینه «۳» بین پایایی و خطائی معیار اندازه گیری رابطه ی معکوس وجود دارد، یعنی با افزایش خطا مقدار پایایی کاهش و با کاهش آن مقدار پایایی افزایش می یابد. بنابراین هرچه مقدار خطا کم باشد (به صفر نزدیک باشد) پایایی آزمون افزایش می یابد.

۳۸- گزینه «3» با توجه به فرمول خطای معیار اندازه گیری، باسخ صحیح که همان 6 است به دست می آید.

ضریب پایایی آزمون انحراف معیار آزمون

۳۹- گزینه «3» هر چه تعداد ماده ها یا سؤال های یک آزمون بیشتر باشد، پایایی آن آزمون بزرگتر و در نتیجه خطای معیار اندازه گیری آن هم کوچکتر است. بنابراین اگر تعداد سؤال های یک آزمون را ۲ برابر کنیم خطای معیاز نصف می شود.

۴۰- گزینه «3» طبق رابطه ی 1> pxx´ > 0 در نتیجه:

وقتی ضریب پایایی بین صفر و یک باشد

۴۱- گزینه «۲» خطای استاندارد = 5/2

۴۲- گزینه «۲» با کاهش تعداد آزمودنی ها، پایایی افزایش می یابد. یکی از مواردی که موجب افزایش پایایی می شود این است که اگر یک آزمون را در مورد گروهی از افراد اجرا کنیم که از نظر توانایی موردنظر تجانس کمتری با یکدیگر دارند، آزمون پایایی بیشتری را نشان خواهد داد.

۴۳- گزینه «3» هرگاه آزمون، موضوعات متفاوتی را بسنجد، ضریب دشواری سؤال ها نیز متفاوت می شود که در این صورت استفاده از روش کودر - ریچاردسون، ضریب پایایی را کمتر از آنچه هست، برآورد می کند.

۴۴- گزینه «۴» با توجه به اینکه گروه زن ها و گروه مرد ها دارای تجانس کمتری هستند، پایایی آزمون به دلیل همین عدم تجانس افزایش می یابد.

۴۵- گزینه «۱» ضریب همبستگی نزدیک به صفر یعنی ارتباطی بین سؤالات وجود ندارد (اندک است) اما چون سؤالات هر نیمه با هم رابطه دارند، نتیجه تحلیل عاملی سؤالات، دو عامل گروهی خواهد بود؛ چون در تحلیل عاملی به تعداد و ماهیت متغیرهای مورد آزمون توجه می شود و با توجه به عامل دو دسته بودن سؤالات، گزینه ۱ صحیح است.

46- گزینه «۴» همان طور که طبق فرمول می دانیم: نمره مشاهده شده = نمره واقعی + نمره خطا

اگر معلم قصد دارد که نمره شاگرد مورد آزمون، کمتر تحت تأثیر خطای تصادفی قرار گیرد، در واقع میخواهد نمره واقعی مساوی با نمره مشاهده شده باشد. برای تحلیل این موضوع می توان چنین استنباط کرد: زمانی که نمره واقعی بدون خطا یا به همراه خطای کمتری است، معلم می تواند موقعیت واقعی شاگردان را بررسی کند و میزان یادگیری واقعی (بدون حدس و گمان) را در آزمون بسنجد.

۴۷- گزینه «۴» رابطه بین روایی و پایایی به این قرار است که یک آزمون باید پایا باشد تا بتواند روا باشد؛ برای مثال یک آزمون حساب برای اندازه گیری محتوا و هدف های درس تاریخ روا (مناسب) نیست، اما همین آزمون می تواند مطالب حساب را که اندازه می گیرد با دقت (به طور پایا) اندازه گیری کند.

۴۸- گزینه «۴» در روش های آزمون - آزمون مجدد و دو فرم همتا آزمودنی ها دوبار به پاسخ گویی می پردازند و تجربه آزمون اول به آن ها در پاسخ گویی به آزمون دوم کمک می کند. روش های دو نیمه کردن و کودر- ریچاردسون تنها یک بار اجرا می شوند، اما در روش کورد - ریچاردسون که شامل فرمول های 20KR و 21KR می باشد، چون همه ی ماده های آزمون تحلیل می شوند، نسبت به بقیه گزینه ها برتری دارد.

۴۹- گزینه «۲» برای پاسخگویی به این تست ابتدا توجه می کنیم که 95% اطمینان برابر با 96/1می باشد؛ سپس به تست ها مراجعه می کنیم و فاصله نمرات در تست را در نظر میگیریم.

۵۰ - گزینه «۲» چنانچه تعداد سؤالات یک آزمون را دو برابر کنیم، واریانس خطای آزمون نیز دو برابر می شود.

51 - گزینه «3»

SEM = خطای معیار اندازه گیری برحسب نمرات خام، SDT = انحراف معیار نمرات مشاهده شده،

ضریب پایایی آزمون

52 - گزینه «۴» در روش کودر - ریچاردسون، آزمون یکبار اجرا می شود؛ اما همه موارد آزمون مورد تحلیل قرار می گیرند. روش کودر - ریچاردسون برای بررسی همسانی درونی آزمون و تعیین پایایی آزمون استفاده می شوم و به 20KR و ۲۱ KR شهرت دارد. در این آزمون به کار کرد افراد در تک تک سؤالات آزمون توجه می شود.

۵۳ - گزینه «۲»

مفروضات سؤال:

54 - گزینه «3» با توجه به فرمول محاسبه ضریب پایایی آزمون برای آزمونی که دو نیمه شده است، عدد ۶۷/0 به دست می آید.

۵۵ - گزینه «۴» مفهوم پایایی بر این باور استوار است که هیچ آزمونی کاملا پایا نیست و لذا با نمره های هر آزمون قدری خطا همراه است. این خطا که بخشی از نمره به دست آمده یا نمره مشاهده شده (نمره خام) است، جزء خطای نمره نیز محسوب می شود.

در هر نموة مشاهده شده ۲ جزء داریم: نمره خطا (مشاهده شده ی خطا) و نمره حقیقی (جزء حقیقی)

پس علاوه بر واریانس خطاء واریانس حقیقی داریم. بنابراین ضریب پایایی صفر نشان دهندهی این است که تمامی واریانس مشاهده شده واریانس خطا است و تا در مقابل، طبعا" ضریب پایایی ۱ معرف این است کار صد در صد واریانس مشاهده شده واریانس حقیقی انست؛ یعنی نمره واقعی برابر با نمره خطاست.

56- گزینه «1» ساده ترین روش تعیین آزمون، بازآزمایی است. این روش در صورتی قابل اجرا است که ویژگی مورد اندازه گیری در طول زمان تغییری نکند.

به عنوان مثال، چنین تصور می شود که فوت یک صفت یا ویژگی نسبتا پایدار است و در فاصله اجرای یک آزمون در دو نوبت تغییر نمی کند.

57-گزینه «1» اعتبار (پایایی) یعنی همبستگی آزمون با خودش ولی روایی یعنی همبستگی با یک ملاک یا آزمون دیگر. به این ترتیب، اعتبار همواره بزرگتر از روایی و یا در موارد نادر مساوی آن است.

۵۸- گزینه «1» در تغییر نمره های یک آزمون همواره با مفاهیم نمره مشاهده شده یا نمره خام، نمره حقیقی (میانگین فرضی تعداد بیشماری آزمون که در مورد یک آزمون شونده اجرا می شود و فرض بر این است که تمرین بر نمرات کسب شده پی در پی تأثیر ندارد) و نمردم خطا که ناشی از یک عامل تصادفی مانند اشتباهات نمره گذاری است مواجه هستیم که بین آن ها تساوی زیر برقرار است.

نمره خطاء + نمره حقیقی = نمره مشاهده شده

حال اگر واریانس نمرات مشاهده شده را با ، واریانس نمرات حقیقی را با و واریانس نمرات خطا را با او نشان دهیم، رابطه بین آنها را می توان به صورت زیر نشان داد:

با توجه به مطالب بالا می توان پایایی آزمون را به صورت نسبت واریانس نمره حقیقی به واریانس نمره مشاهده شده تعریف کرد که به این صورت بیان می شود: همچنین با توجه به تساوی پایایی را می توان به این در صورت نیز بیان کرد:

۵۹ـ گزینه «۴» در یک تعریف از ضریب پایایی، این ضریب مساوی است با همبستگی بین نمره های مشاهده شده در آزمون های موازی؛ چنانچـه نمـره هـر یک از آزمودنی ها در زمانی که با آزمون های مواری آزمون می شوند، یکسان باشد. نمره های مشاهده شده هر آزمـون مـقـداری پراکندگی داشـته باشـد، آزمون دارای ضریب پایی کامل است (). در صورتی که نمره مشاهده شده هر آزمودنی در یک آزمون با نمره همان آزمودنی در آزمونی دیگر همبسته نباشد، آن آزمون ها به طور کلی پایایی ندارند (). بنابراین می توان گفت که پایایی شامل مقادیری بین ۰ تا ۱ میباشد.

۶۰- گزینه «۳» ضریب پایایی که از طریق فرمول به دست می آید به صورت مقابل بیان می شود:

در این فرمول واریانس هر سؤال بر حسب pq به محسوب می شود که در آن p همان ضریب دشواری است و q نیز مساوی ۱ - p است بنابراین می تـوان گفت که ضریب پایایی که از طریق این فرمول بدست می آید، تابع ضریب دشواری آزمون می باشد. در صورتی که تمام سؤال های آزمون دارای ضریب دشواری متوسط باشند، می توان به جای فرمول از فرمول استفاده کرد که به شرح زیر می باشد:

۶۱ـ گزینه «۱» ضریب پایایی ۱ معرف این است که صد در صد واریانس مشاهده شده، واریانس حقیقی یا واقعی است.

۶۲- گزینه «۲» برای بالا بردن ضریب اعتبار آزمون باید در یک آزمون به تعداد سؤالات مشابه بیفزایم. اعتبار آزمون زمانی بالا می رود که قابلیت تکرارپذیری داشته باشد.

۶۳- گزینه «2» برای برای استفاده از فرمول 21KR، فرض بر این است که همه سؤال ها و یا ماده های آزمون از لحاظ دشواری متوسط و شبیه بـه هـم هستند. بنابراین زمانی که همه سؤالات ضریب دشواری یکسانی داشته باشند، از روش 21KR استفاده می شود. 21KR و 20KR برای بررسی و تعیین پایایی کاربرد دارند که روش کودر - ریچاردسون نامیده می شوند.

۶۴- گزینه «۳» فرمهای موازی (همتا) با فاصله زمانی گزینه صحیح می باشد. در روش فرم های موازی، دو آرمون معادل با یک موضوع تهیه می شود و آن ها را در فاصله کوتاهی (زمانی در حدود ۲ هفته) به یک گروه واحد از آزمون شوندگان می دهند. یکی از مشکلات عمده این روش این است که تهیـه فـرم هـای موازی خصوصاً در آزمون های پیشرفت تحصیلی آسان نمی باشد؛ ضمن آنکه بعضی از آزمون شوندگان در جواب دادن به سؤال های فرم اول آزمون با اصـولی آشنا می شوند که عملکرد آنها را در فرم دوم تحت تأثیر قرار این موضوع خطای حضور حافظه را شامل می شود که زمانی که فرم های همتا بـدون فاصله زمانی اجرا شوند، کم می شود.

۶۵- گزینه «۳» ضریب پایایی بین ارقام صفر تا ۱+ است. ضریب پایایی صفر نشان میدهد که تمامی واریانس های مشاهده شده واریانس خطا است. در مقابل، ضریب پایایی مشاهده شده نشانگر واریانس حقیقی است. با دقت در نمرات آزمودنی ها در نوبت اول و دوم متوجه می شویم که به نمره هر فرد در بـار دوم اجرای آزمون ۲ نمره اضافه شده است؛ بنابراین همبستگی بین آن ها کامل و مثبت است (۱+).

۶۶ـ گزینه «۳» در فرم های موازی همتا، دو آزمون به صورت معادل یا موازی برای یک موضوع تهیه شوند. دو فـرم مـوازی (همتا) دارای میانگین و واریانس های برابر می باشند.

۶۷- گزینه «۴» برای پاسخ به این سؤال ابتدا خطای معیار اندازه گیری را محاسبه می کنیم، سپس پاسخ را در فرمول زیر قرار می دهیم تا پاسخ به دست آید: در این سؤال ابتدا باید واریانس را تبدیل به انحراف معیار کنیم.

(خطای معیار اندازه گیری) (96/1) ± نمره حقیقی

۶۸- گزینه «۲» برای محاسبه ضریب پایایی کل آزمون از این فرمول استفاده :

60/0 × 5/0

۶۹- گزینه «۲»

۷۰- گزینه «3» آزمون های تشخیصی به دلیل این برگزار می شوند که میزان یادگیری افراد را بسنجند. در واقع اگر هر کدام از قسمت های آزمون را در نظر بگیریم، میزان درک یادگیرنده مهم می باشد. در واقع تشخیص اینکه هر آزمون شونده چه مقدار از مطلب مورد نظر را یاد گرفته است، از ارزش برخوردار است.

۷۱- گزینه «1» فرض استفاده از فرمول 21KR بر این پایه استوار است که همه سؤالها یا ماده های آزمون از لحاظ درجه دشواری شبیه به هم هستند. فرض اساسی فرمول های کودر - ریچاردسون این است که سؤال های آزمون را می توان به صورت1 برای پاسخ درست و صفر برای پاسخ غلط تصحیح کرد.

۷۲- گزینه «3» ضریب آلفا راهی برای تعیین پایایی در آزمون است. این روش زمانی کاربرد دارد که آزمون از بخش های مختلف تشکیل شده باشد؛ مثل زمانی که یک آزمون از تعدادی خرده آزمون تشکیل شود. روش آلفای کرونباخ در شرایطی که نمرات سؤال در طیف کاملا موافقم تا ... کاملا مخالفم قرار گیرد، بسیار مناسب است؛ زیرا هیچ یک از جواب های سؤالات درست یا غلط نیست و فقط نظر یادگیرنده و آزمون شونده را بررسی می کند.

۷۳- گزینه «۲» در روش فرم های همتا، میزان همبستگی بین دو فرم همتا از یک آزمون، بررسی می شود. این نوع از پایایی (اعتبار) بیشتر دقت آزمون را نشان می دهد.

۷۴- گزینه «۴» خطای معیار اندازه گیری از نتایج و قواید پایایی و اعتبار آزمون است. برای درک بهتر به مثال زیر توجه کنید.

اگر از شخصی بخواهیم تا توپی را ۲۰۰ بار پرتاب کند، او این کار را انجام می دهد. فرض می کنیم انجام دادن این کار بدون تأثیرپذیری شخص از خستگی یا تمرین باشد. حالا ما یک توزیع فراوانی حاصل از ۲۰۰ بار پرتاب شدن توپ داریم. این توزیع دارای یک میانگین است که ما این میانگین را به طور تقریب و نمره حقیقی فرد در پرتاب توپ می دانیم. همچنین، این توزیع یک انحراف معیار دارد. این انحراف معیار خطاهای پرتاب توپ به وسیله آن شخص است که ا به آن خطای معیار اندازه گیری می گویند؛ بنابراین فقدان دقت نمرات، یعنی فقدان دقتی که از ناپایایی آزمون ناشی می شود، در خطای معیار اندازه گیری منعکس می شود.

۷۵- گزینه «۲» هنگامی ضریب اعتبار آزمونی مساوی یک است که نمره های واقعی (T) و نمره های مشاهده شده (X) کاملا به هم مربوط باشند. هنگامی که ضریب اعتبار عدد خاصی باشد، میزان پایایی همان عدد است.

۷۶- گزینه «3» آزمونی دارای روایی است که برای اندازه گیری آنچه مورد نظر است، کافی و مناسب باشد و بتواند محتوا و هدف آزمون را به خوبی اندازه بگیرد. پایایی نیز یعنی اینکه تکرار آزمون در دفعات مختلف به ما نتایج مساوی بدهد. پایا بودن شرط روا بودن هر آزمون است، اما برای روا بودن هر آزمون شروط دیگری نیز لازم است؛ یعنی برای تعیین هدف و محتوا باید به روایی آزمون هم دقت داشت.

۷۷- گزینه «۲» به طور کلی روش های برآورد پایایی عبارت اند از: روش پایایی مصححان - روش بازآزمایی - روش فرم های هم ارز – روش های همسانی درونی (روش دو نیمه کردن آزمون، روش کودر - ریچاردسون و روش ضریب آلفای کرونباخ).

۷۸- گزینه «۲» به طور کلی می توان آزمون ها را به دو دسته: قدرت و سرعت تقسیم نمود. آزمون های قدرت آزمون هایی هستند که در اجرای آن به آزمون شوندگان وقت کافی داده می شود تا به همه سؤال های آن جواب دهند. این گونه آزمون ها شامل سؤال های چالش انگیزند و هدف اصلی آنها این است که و سطح عملکرد یا میزان دانش و توانایی آزمون شوندگان را بسنجند. در مقابل آزمون های سرعت به آزمون هایی گفته میشود که برای جواب دادن به و سؤال های آنها محدودیت زمانی وجود دارد و اکثر آزمون شوندگان در محدوده زمانی تعیین شده نمی توانند آن را به پایان برسانند. سؤال های این نوع آزمون ها بسیار ساده اند و هدف آنها سنجیدن سرعت عمل آزمون شوندگان است.

۷۹- گزینه «3» با استفاده از نظریه سؤال - پاسخ IRT میتوان برخی از ویژگیهای سؤال یا ماده آزمون را تعیین کرد. این روش یکی از روش های دقیق برای تحلیل سؤال آزمون خصوصأ در آزمون های وابسته به ملاک است. این ویژگی های سؤال برای بسیاری از مقاصد آزمون سازی از جمله تحلیل سوال های در آزمون مفیدند. منحنی ویژه سؤال ICC که برای هر یک از سؤال های آزمون رستم می شود، احتمال پاسخ درست دادن به هر سؤال را به توانایی آزمون شونده ربط می دهد. به سخن دیگر، منحنی ویژه سؤال یک بازنمایی نموداری از رابطه بین احتمال پاسخ درست دادن به یک سؤال و موقعیت آزمون شونده در صفت مورد اندازه گیری توسط آزمون است.

۸۰- گزینه «۲» آزمون های فرافکن با هدف شناخت شخصی طراحی شده اند. این آزمون ها تخیل شخص را بر می انگیزند و فرض بر این است که شخص با برآوردهای نخیلی اش اطلاعاتی درباره شخصیت خود ارائه می دهد. آزمون «رورشاخ» و آزمون اندریافت موضوع، آزمون هایی رایج در این زمینه هستند که برای بزرگسالان طراحی شده اند. برای کودکان و سالمندان نیز آزمون های فرافکن خاصی وجود دارد. هدف اصلی در پرسشنامه های شخصیت، عینیت است. به این معنا که هرچه را می بینند، تفسیر و نظر خود را اعلام می کنند. این پرسشنامه ها را به راحتی می توان نمره گذاری کرد و پایایی و روایی آن ها را سنجید.

۸۱- گزینه «۴» روش های تعیین پایایی متنوع اند. این روش ها طبق تعریف مرفی و دیوید شوفر (۱۹۹۴) طبقه بندی می شوند: روش پایایی مصححان، روش بازآزمایی، روش فرم های هم ارز و روش های همسانی درونی که روش همسانی درونی خود شامل روش دو نیمه کردن آزمون روش کودر - ریچاردسون (20KR، 21KR) و روش ضریب آلفای کرونباخ است.

اگر به صورت کلی بخواهیم روش ضریب آلفای کرونباخ و دو نیمه کردن آزمون را با هم مقایسه کنیم، خواهیم دید که روش دو نیمه کردن آزمون، نوعی روش ضریب آلفای کرونباخ است. به عبارت دیگر، ضریب آلفا معرف میانگین ضریب های پایایی حاصل از همه دو نیمه های ممکن یک آزمون است. بنابراین برای تعیین اعتبار (Reliability) نمرات سؤالات تشریحی، از روش آلفای کرونباخ استفاده می کنیم.

۸۲- گزینه «۲» این شکل ها نتایج تیراندازی با سه تفنگ مختلف را نشان می دهد. چنانکه در متن سؤال می بینیم، نتایج تیراندازی با تفنگ در شکل اول، هم پایا است و هم روا؛ هم اعتبار دارد و هم روایی. نتایج تیراندازی با تفنگ دوم پایا هست اما روا نیست؛ زیرا تیرها تقریبا به نقطه ای اصابت شده اند که نقطهی موردنظر تیرانداز نیست. نتایج تیراندازی با تفنگ در شکل سوم نه پایا است و نه روا؛ نه اعتبار دارد و نه روایی. نتایج تیراندازی گزینه و شکل چهارم، بی اعتبار است.

۳- گزینه «۱» روش بازآزمایی، یکی از روش های تعیین پایایی (اعتبار) است. این روش، ساده ترین نوع روش تعیین پایایی است که در آن، آزمون را در دو نوبت به گروه واحدی از آزمون شوندگان می دهند و نمرات حاصل را با هم مقایسه می کنند. این روش تنها زمانی قابل استفاده است که ویژگی مورد اندازه گیری در طول زمان تغییر نکند. به عنوان مثال، چنین تصور می شود که هوش یک ویژگی یا صفت نسبتا پایدار است که در فاصله اجرای یک آزمون هوشی در دو نوبت متفاوت تغییر نمی کند. با این توضیحات، نگرش به درس ریاضی نیز می تواند از روش بازآزمایی ارزشیابی شود. گزینه ۳ به اضطراب امتحان ریاضی اشاره دارد. این گزینه می تواند به عنوان گزینه نزدیک به صحیح مطرح شود، اما باید به متن سؤال دقت کرد که اشاره به روش بازآزمایی دارد و روش بازآزمایی برای تعیین پایایی استفاده می شود، اما اضطراب امتحان ریاضی و بررسی آن به روایی آزمون اشاره می کند و با توجه به مفهوم نظری گزینه ، روایی سازه مطرح می گردد.

۸۴- گزینه «۲» فرمول برآورد خطای معیار اندازه گیری به این شرح است:

خطای معیار اندازه گیری = SEM ، انحراف معیار = SDT، ضریب پایایی آزمون

مفروضات گزینه به این شرح است:

مفروضات را در فرمول قرار می دهیم :

۸۵- گزینه «۲» پایایی (اعتبار) یک وسیله اندازه گیری است که عمدتا به دقت نتایج حاصل از آزمون اشاره دارد. کاپلان و ساکوز گفته اند: «پایایی به دقت، اعتمادپذیری یا تکرارپذیری نتایج آزمون اشاره دارد.»

پایا بودن آزمون، شرط روا بودن آن است. در یک آزمون از نمره حقیقی هیچ آزمون شونده ای آگاهی نداریم، بنابراین باید آن را برآورد کنیم، به این منظور از نمره مشاهده شده استفاده می کنیم.

۸۶- گزینه «۴» پایایی آزمون های وابسته به ملاک تعیین می کند که وضعیت تعدادی دانش آموز از لحاظ دستیابی یا عدم دستیابی به حد تسلط به دو موقعیت آزمون (دو بار اجرای یک آزمون با اجرای دو فرم موازی از یک آزمون) یکسان باشد، در این صورت نتیجه می گیریم که آزمون از پایایی برخوردار است. البته باید در نظر داشت که زمان بین اجرای دو آزمون باید بسیار کوتاه باشد، به گونه ای که در آن فاصله، هیچ عامل مشخصی بر بهبود عملکرد آزمون شوندگان در دو بار اجرای آزمون تأثیر نگذارد.

۸۷- گزینه «۲» روش های دو نیم کردن با کودر- ریچاردسون برای تعیین پایایی آزمون های سرعت مناسب نیستند؛ چون در آزمون های سرعت آزمون شوندگان تقریبا همه سؤال هایی را که وقت رسیدگی به آنها را دارند، درست جواب می دهند و تنها تفاوت بین آنها در تعداد سؤال هایی است که جواب می دهند. بنابراین اگر نمرات نیمه اول را با نمرات نیمه دوم آزمون همبسته بسازیم، ضریب حاصل بسیار کوچک خواهد بود و اگر سؤال هارا زوج و فرد و جدا کنیم و همبستگی بین آن ها را بدست آوریم، این ضریب خیلی بزرگ خواهد بوگیر ضریب آلفای کرونباخ نیز مخصوص آزمون هایی است که به صورت چند اززشی تعیین می شوند، بنابراین بهترین روش برای محاسبه آزمون های سرعت روش بازآزمایی است. ساده ترین روش تعیین پایایی یک آزمون، روش بازآزمایی اشت البته متخصصان معتقدند که آزمون اول، به آزمون شوندگان تجربه ای می آموزد تا در نوبت دوم بهتر با سؤالات پاسخ گویند. بنابراین برای رفع مشکلات این روش، روش فرم های موازی با هم ارز وجود دارد که در آن، دو آزمون معادل برای یک موضوع تهیه می شود و در فاصله ی زمانی کوتاهی به یک گروه آزمون شونده ارائه می شود.

۸۸- گزینه «۲» یکی از راه های تعیین اعتبار این اسنت کلر آزمونی را که یکبار با گروه واحدی از آزمون شوندگان اجرا کرده ایم، پس از اجرا به دو گروه تقسیم کنیم و ضریب همبستگی بین دو نیمه را محاسبه کنیم.

۸۹- گزینه «۲» فرض اساسی فرمول های کودر - ریچاردسون این است که سؤال های آزمون را می توان به صورت یک برای پاسخ های درست و صفر برای پاسخ های غلط تصحیح کرد. از آنجا که سؤالات آزمونی را که به صورت صحیح و غلط طراحی نشده است، می توان به این صورت (0و1) تصحیح کرد، روش کودر - ریچاردسون برای تحلیل اعتبار آن ها روش مناسبی است.

۹۰- گزینه «۲» برای به دست آوردن خطای معیار اندازه گیری باید از فرمولی استفاده کرد که در مقابل می آید:

در بدنه سؤال ، ۸۰/0 بیان شده است. از طرفی مقیاس T دارای میانگین ۵۰ و انحراف معیار ۱۰ است.

پاسخ به گزینه (۲) نزدیکتر است.

۹۱- گزینه «1» اگر واریانس نمرات مشاهده شده آزمونی را با ،واریانس نمرات حقیقی را با و واریانس نمرات خطا را با نشان دهیم رابطه بین این واریانس ها به شرح زیر خواهد بود:

۹۲- گزینه «۴» گزینه های ۱، ۲ و ۳ به روش هایی اشاره دارند که در آن ها آزمون دو بار اجرا می شود و همین امر باعث ایجاد تغییراتی در نمرار افراد بر اثر گذشت زمان می شود، اما در روش های همسانی درونی (گزینه ۴) یک بار آزمون اجرا می شود و با یکی از روش های دو نیمه کردن کودر - ریچاردسون و ضریب آلفای کرونباخ، ضریب پایایی آن محاسبه می شود. همین اجرای یکباره آزمون باعث می شود که نوسانات ناشی از گذشت زمان به طور کامل کنترل شود.

۹۳- گزینه «٣» بر استفاده از فرمول زیر که رابطه بین ضریب اعتبار آزمون و افزایش تعداد سؤال های آزمون را نشان می دهد، پاسخ صحیح (80/0) به دست می آید.

ضریب اعتبار آزمون جدید که تعداد سؤال های آن برابر شده است

ضریب افزایش تعدام سؤال ها

ضریب اعتبار آزمون پیش از آنکه تعداد سؤال های آن افزایش یابد

۹۴- گزینه «3» همبستگی نمرات خطا و نمرات مشاهده شده را با فرمول نشان می دهند که از این همبستگی برای به دست آوردن خطای معیار اندازه گیری استفاده می کنند.

در خطای معیار اندازه گیری بر حسب نمرات خام :SEM، انحراف معیار آزمون (نمرات مشاهده شده) : ، ضریب پایایی آزمون

۹۵- گزینه «3» به فرض اینکه تمام سؤال های آزمون دارای ضریب دشواری متوسط باشند، می توان به جای 20KR از فرمول 21KR استفاده کرد. پس: KR۲۰ ≥ 21KR

۹۶- گزینه «3» فقط با در نظر گرفتن پایایی (اعتبار) یک آزمون نمی توانی از روایی آن اطمینان پیدا کرد؛ زیرا ممکن است با وجود پایایی، آزمون با خصیصه مورد نظر مرتبط نباشد.

۹۷- گزینه «۱» مفهوم اعتبار در نظریه کلاسیک با تابع آگاهی در نظریه سؤال - پاسخ همزاد است.

۹۸- گزینه «۱» پایایی به دقت، اعتمادپذیری، ثبات یا تکرارپذیری نتایج آزمون اشاره می کند. اگر آزمونی در هر بار اجرا روی تعدادی دانش آموز، نتایج مختلفی به دست دهد، آن آزمون پایا نیست.

۹۹- گزینه «۲» اگر یک آزمون را در مورد گروهی از افراد اجرا کنیم، که از نظر توانایی مورد سنجش، تجانس کمتری با یکدیگر دارند، آزمون پایایی بیشتری نشان خواهد داد.

۱۰۰- گزینه «۴» اگر طول آزمون با سؤالات مشابه دو برابر شود، اعتبار (پایایی) آزمون تا حدودی اضافه می شود.

۱۰۱- گزینه «3»

۱۰۲- گزینه «۱» رابطه بین روایی و پایایی از این قرار است که یک آزمون باید پایا باشد تا بتواند روا باشد. اگر آزمونی در هر بار اجرا روی تعدادی دانش آموز , نتایج مختلفی به دست بدهد، یک آزمون پایا نخواهد بود و در واقع هیچ چیز را به درستی اندازه گیری نخواهد کرد و اگر یک آزمون چیزی را به درستی , اندازه گیری نکند، هیچ اطلاع مفیدی به ما نخواهد داد. برای اینکه یک آزمون روا باشد، نخست باید پایا باشد؛ یعنی پایایی، شرط روایی است، اما روایی برای پایایی ضروری نیست. سه حالت در رابطه پایایی و روایی می توان داشت:

1) یک آزمون نه پا است نه روا. ۲) یک آزمون پایاست ولی روا نیست. ۳) یک آزمون هم پایاست و هم روا

پس اعتبار (پایایی)، شرط لازم روایی است، اما اعتبار به تنهایی روایی آزمون را تضمین نمی کند. پایایی یک آزمون بر اعتبار آن آزمون تأثیر دارد؛ زیرا از لحاظ نظری همبستگی نمره های یک آزمون با نمره ی دیگر نمی تواند بیش از همبستگی نمره های همین آزمون با نمره های واقعی خود باشد. یعنی یا

به بیان دیگر، همبستگی بین نمره های واقعی و نمره مشاهده شده، برابر است با ریشه دوم ضریب پایایی که معمولا شاخص پایایی نامیده می شود.

پایایی

در نتیجه، ضریب اعتبار نمی تواند بزرگتر از جذر پایایی باشد، ولی می تواند بزرگتر از ضریب پایایی خودش باشد.

مثال: اگر باشد (ضریب اعتماد یا روایی) است:

۱۰۳- گزینه «۲» خطای معیار برآورد هم متأثر از (انحراف استاندارد متغیر ملاک) و هم متأثر از (روایی) است.

خطای معیار برآورد عبارت است از:

مقدار خطای معیار برآورد، از طرفی تابع مقدار انحراف معیار توزیع متغیر مورد پیش بینی (متغیر ملاک) است و از طرف دیگر تابع مقدار همبستگی بین نمره های این متغیر با نمره های متغیر پیش بین است.

هرگاه مایل باشیم که نمره دقیق فرد را از روی متغیر ملاک پیش بینی کنیم، باید ضریب اعتبار را بر حسب خطای معیار برآورد، که شاخصی شبیه خطای معیار اندازه گیری در مبحث پایایی است، تفسیر کنیم. خطای معیار اندازه گیری، میزان خطای احتمالی فرد را که ناشی از ناپایایی آزمون است، نشان می دهد و عبارت است از:

به همین نحوه، خطای معیار برآورد نیز میزان خطای احتمالی را که به علت کم اعتباری (عدم روایی آزمون در نمره پیش بینی شده ی ملاک رخ خواهد داده نشان می دهد و عبارت است از:

مجذور ضریب اعتبار

و

انحراف معیار نمره های ملاک

هر گاه آزمون، اعتبار کامل داشته باشد در این صورت خطای برآورد، برابر صفر خواهد بود. از سوی دیگر در آزمونی که ضریب اعتبار آن برابر صفر است، خطای معیار برآورد، برابر با انحراف معیار توزیع دیگر و در آزمونی که ضریب اعتبار آن برابر صفر است، خطای معیار برآورد، برابر با انحراف معیار توزیع ملاک خواهد بود:

۱۰۴- گزینه «۲» اعتبار یا پایایی، یک وسیله اندازه گیری است که عمدتا به ثبات و دقت نتایج حاصل از آن اشاره می کند.پایایی به دقت، اعتماد پذیری، ثبات یا تکرارپذیری نتایج آزمون اشاره می کند.

عوامل زیادی بر افزایش پایایی اثر دارند که یکی از آنها، طولانی کردن آزمون است. با افزودن به تعداد سؤالات یک آزمون، پایایی آن نیز افزایش می یابد. عامل دیگر، عدم تجانس گروه آزمون شوندگان است. اگر یک آزمون را در مورد گروهی از افراد اجرا کنیم که از نظر توانایی مورد سنجش، تجانس کمتری بایکدیگر دارند، آزمون، پایایی بیشتری خواهد داشت.

عامل دیگر، اجرای آزمون در زمان محدود یا آزمون های سرعت است. آزمون های سرعت معمولا از آزمون هایی که سرعت عملکرد در آنها عامل مهمی به حساب نمی آید (آزمون های قدرت) ضریب پایایی بیشتری را نشان می دهند، ولی کم کردن تعداد گزینه ها باعث کاهش اعتبار می شود. بهتر است برای هر سؤال بین ۳ تا ۵ گزینه در نظر گرفته شود؛ زیرا میزان حدس پذیری افزایش می یابد و خطا زیاد می شود و در نتیجه، پایایی کاهش می یابد.

۱۰۵- گزینه «۲» هنگامی که است، می توان دریافت:

۱- اندازه گیری بدون خطا انجام شده است.

۲- برای همه ی آزمودنی ها است.

۳- کل واریانس نمره مشاهده شده، منعکس کننده واریانس نمره واقعی است .

۴- تمام اختلاف های بین نمره های مشاهده شده، منعکس کننده اختلاف های نمره واقعی هستند.

5- همبستگی بین نمره های مشاهده شده و نمره های واقعی مساوی یک است

6- همبستگی بین نمره های مشاهده شده و نمره های خطا، مساوی صفر است

۱۰۶- گزینه «1» فرمول اسپیرمن - پراون که پایایی کل آزمون از آن به دست می آید، عبارت است از:

۱۰۷- گزینه «۱» در وضعیتی که ما را می دانیم و قصد داریم را به مقدار مطلوب برسانیم، می توانیم از فرمول بازنویسی شده اسپیرمن - براون به صورت مقابل، جهت تعیین N مورد لزوم برای دستیابی به خواسته شده، استفاده کنیم:

۱۰۸- گزینه «۴» دامنه اطمینان برای نمره واقعی یک آزمون عبارت است از ( Z در نقطه ۶۸% برابر ۱ است.)

109- گزینه «3»

۱۱۰- گزینه «4»

آزمونی که نه روایی دارد و نه پایایی 

آزمونی که روایی دارد، اما پایایی ندارد. 

آزمونی که پایایی دارد، اما روایی ندارد. 

آزمونی که هم روایی دارد و هم پایایی . 

۱۱۱- گزینه «۲» پایایی به معنای ثبات و کارایی ید آزمون در زمانهای بعدی است. یکی از عواملی که باعث افزایش پایایی می شود، افزایش تعداد گزینه های سؤالات است و در مقابل، یکی از عواملی که باعث کاهش پایایی می شود، همگونی و تجانس آزمودنی هاست.

۱۱۲- گزینه «۲» برای بررسی خطای استاندارد اندازه گیری به روش زیر عمل می کنیم:

پایایی

انحراف معیار،

=SEMخطای استاندارد اندازه گیری،

۱۱۳- گزینه «1» برای بررسی اعتبار آزمون هنگامی که طول آزمون نصف می شود، به روش زیر عمل می کنیم:

۱۱۴- گزینه «۱» آزمون یکی از ابزارهای مهم پژوهش در روان شناسی معاصر است. بعضی اطلاعات روانشناختی درباره ماهیت مشکلات و ناتوانایی های فرد، از طریق آزمون روانی به دست می آید. با استفاده از آزمون ها می توان انواع توانایی ها، رغبت ها، نگرش ها و دستاورد ها را اندازه گیری کرد. آزمون ها به روانشناسان کمک می کنند تا داده های وسیعی درباره افراد به دست آورند، بی آنکه نظم عادی زندگی آن ها به هم بخورد یا نیازی به وسایل پیچیده آزمایشگاهی باشد. به طور کلی آزمون های توانایی، میزان پایایی بیشتری نسبت به سایر آزمون ها دارند؛ یعنی نتایج این آزمون ها در آینده ثبات بیشتری دارد.

۱۱۵- گزینه «۴» شرط لازم و کافی آزمون های روان سنجی، روایی (validity) است، یعنی آزمونی که توانایی سنجش چیزی را داشته باشد که برای سنجش آن ساخته شده است.

مفهوم اعتبار (روایی) به این سؤال پاسخ می دهد که ابزار اندازه گیری تا چه حد خصیصه مورد نظر را می سنجد. بدون آگاهی از اعتبار ابزار اندازه گیری نمی توان به دقت داده های حاصل از آن اطمینان داشت. ابزار اندازه گیری ممکن است برای اندازه گیری یک خصیصه ویژه دارای اعتبار باشد، در حالی که سر برای سنجش همان خصیصه روی جامعه دیگر از هیچ گونه اعتباری نباشد. برای مثال یک آزمون ریاضی ممکن است برای سنجش توانایی ریاضی دانش آموزان پایه پنجم ابتدایی از اعتبار لازم برخوردار باشد، اما برای سنجش توانایی ریاضی دانش آموزان پایه سوم راهنمایی فاقد اعتبار باشد.

۱۱۶- گزینه «1» هر چقدر که ضریب پایایی عدد بزرگتر باشد، قدرت ثبات آزمون بالاتر است.

قابلیت اعتماد که واژه هایی مانند پایایی، ثبات و اعتبار برای آن به کار برده می شود و معادل انگلیسی آن Reliability است، یکی از ویژگی های ابزار , اندازه گیری (پرسشنامه یا مصاحبه با سایر آزمون های روانشناسی) است. مفهوم یاد شده با این امر سروکار دارد که ابزار اندازه گیری در شرایط یکسان تا مچه به اندازه نتایج یکسانی به دست می دهد. از جمله تعریفهایی که برای قابلیت اعتماد ارائه شده است می توان به تعریف ارائه شده توسط ایبل و فریسبی (۱۹۸۹) اشاره کرد: «همبستگی میان یک مجموعه از نمرات و مجموعه دیگری از نمرات در یک آزمون معادل که به صورت مستقل بر یک گروه آزمودنی به دست آمده است.»

۱۱۷- گزینه «۴» طبق نمره به دست آمده که نمره بالایی نیز محسوب می شود، آزمون همسانی، ثبات و پایایی قدرتمند و بالایی دارد.

۱۱۸- گزینه «۲» واریانس نمرات آزمون با استفاده از فرمول زیر محاسبه می شود:

آزمون فصل هفدهم

۱- کدام یک از عوامل زیر پایایی آزمون را افزایش می دهد؟

1) افزودن تعدادی سؤال با ضریب دشواری بالا

۲) افزودن تعدادی سؤال با ضریب دشواری پایین

۳) افزودن تعدادی سؤال با ضریب دشواری متوسط

۴) افزودن تعدادی سؤال با ضریب تمیز پایین

۲- زمانی که نمرات سؤال به صورت چندارزشی تعیین می شود کدام شیوه پایایی مناسب است؟

۱) آلفای کرونباخ

2) 20KR

3) 21KR

4) تنصیف

۳- ضریب پایایی یک آزمون ۳۰ سؤالی 60/0 است. اگر ۱۵ سؤال به آزمون اضافه کنیم، ضریب پایایی چقدر است؟

1) 80/0

2) 78/0

3) 69/0

4) 65/0

۴- دانش آموزی در یک آزمون نمره ۵۵ را کسب کرده است. اگر خطای استاندارد اندازه گیری این آزمون 5/1 باشد، نمره حقیقی این دانش آموز بـه احتمال ۹۵ درصد بین کدام دو حد قرار می گیرد؟

1) 57- 53

2)94/57- 06/52

3) 58-53

4) 58- 96/51

۵- برای تعیین پایایی آزمون های سرعت کدام شیوه مناسب تر است؟

۱) دو فرم هم ارز

2) 20KR

3) تنصیف

4) 21KR

۶- میانگین و انحراف معیار دانش آموزان یک کلاس ۴۰ نفری در یک آزمون که ۶۰ سؤال دارد، به ترتیب ۲۵ و ۳ است. ضریب پایایی چقدر است؟

1) 50/0

2) 63/0

3) 86/0

4) 34/0

۷- اگر پایایی یک آزمون ۱ باشد، خطای معیار اندازه گیری آن چقدر است؟

1) یک

۲) بزرگ تر از یک

3) صفر

4) هیچکدام

۸- اگر ضریب همبستگی بین دو نیمه آزمون 75/0باشد، ضریب پایایی کل آزمون چقدر است؟

1) 74/0

2) 57/1

3) 86/0

4) 63/0

۹- در کدام روش تعیین پایایی، همه سؤالات یا ماده های آزمون تحلیل می شوند؟

1) تنصیف

2) بازآزمایی

۳) دو فرم هم ارز

۴) کودر- ریچاردسون

۱۰ـ اگر واریانس نمرات مشاهده شده در یک آزمون برابر با ۳۰ و واریانس خطا معادل ۶ باشد، ضریب پایایی چقدر خواهد بود؟

1) 6/0

2) 8/0

3) 7/0

4) 75/0

۱۱ـ رابطه بین روایی و پایایی چگونه است؟

1) روایی شرط پایایی است.

۲) این دو با هم رابطه عکس دارند.

۳) پایایی شرط روایی است.

۴) این دو از هم مستقل اند.

۱۲- واریانس آزمونی برابر با ۱۶ و ضریب اعتبار آن 60/0 است. خطای معیار اندازه گیری آزمون چقدر است؟

1) 5/3

2) 4

3) 5/1

4) 5/2

۱۳- یک آزمون ۲۷۰ سوالی را به 9 آزمون ۳۰ سؤالی تقسیم کرده ایم. بهترین روش برای محسبه ضریب پایایی کدام است؟

۱) آلفای کرونباخ

2) کودر- ریچاردسون

3) تنصیف

۴) بازآزمایی

۱۴ـ برای تعیین پایایی روش های سنجش مستقیم رفتار از کدام روش استفاده می شود؟

۱) کودر - ریچاردسون

۲) بازآزمایی

۳) فرم های هم ارز

4) میان مشاهده گران

۱۵ـ ضریب پایایی شامل چه مقادیری می شود؟

1) از صفر تا ۱+

2) از ۱- تا ۱+

3) از صفر تا 1

۴) هیچ یک از موارد

۱۶- کدام عبارت زیر در مورد روش باز آزمابی صحیح نمی باشد؟

۱) زمانی که آزمون را بار اول اجرا می کنیم، آزمون شوندگان برای اجرای مجدد آزمون تجربه به دست می آورند.

2) اگر فاصله زمانی بین دو بار اجرای آزمون زیاد نباشد، ممکن است خطای اندازه گیری یا تغییراتی که واقعا در یادگیری دانش آموزان ایجاد شده است، با هم مخلوط شوند.

3) معلمان و دانش آموزان علاقه ای به اجرای مجدد یک آزمون ندارند.

۴) چون در هر دو آزمون سؤالات یکسانی به کار رفته است نمرات حاصل در اجرای مجدد آزمون به ما نشان نمیدهد که اگر سؤالات تغییر کنند چه تغییراتی در نتایج حاصل می شود.

۱۷- دو آزمون وقتی هم ارز هستند که ......................

۱) میانگین آن ها برابر باشد.

۲) میانگین و واریانس آن ها برابر باشد.

۳) واریانس آن ها برابر باشد.

۴) انحراف معیار آن ها برابر باشد.

۱۸- مهم ترین مشکل روش پایایی بازآزمایی کدام است؟

۱) اجرای دو فرم هم ارز مورد علاقه دانش آموزان نیست.

۲) برآورد مناسبی به دست نمی دهد.

۳) اگر فاصله بین دو بار اجزا زیاد باشد، خطای اندازه گیری زیاد می شود.

۴) تهیه فرمهای هم ارز برای آزمون های پیشرفت تحصیلی کار آسانی نیست.

۱۹- روش باز آزمایی بر ...................... و روش فرم های هم ارز بر ..................... تأکید دارند اما روشهای همسانی درونی بر ........................ تأکید دارند.

۱) همسانی ماده های تشکیل دهنده آزمون - ترازی نتایج - ثبات نتایج

۲) ثبات نتایج - ترازی نتایج - همسانی ماده های تشکیل دهنده ی آزمون

۳) ترازی نتایج - ثبات نتایج - همسانی ماده های تشکیل دهنده ی آزمون

۴) ترازی نتایج - همسانی ماده های تشکیل دهنده ی آزمون - ثبات نتایج

۲۰- در روش تنصیف از طریق کدام فرمول، ضریب پایایی کل آزمون را به دست می آوریم؟

1) 20KR

۲) پیرسون

۳) اسپیرمن – براون

4) 21KR

۲۱- اگر ضریب همبستگی بین نیمه های یک آزمون 67/0 باشد، ضریب پایایی کل آزمون چقدر است؟

1) 75/0

2) 80/0

3) 83/0

4) 87/0

۲۲- آزمونی ۳۰ سؤالی دارای ضریب پایایی 80/0 است. اگر ۱۵ سؤال به آن اضافه کنیم، پایایی آزمون جدید چند خواهد شد؟

1) 70/0

2) 75/0

3) 80/0

4) 86/0

۲۳- در کدام یک از روش های برآورد پایایی، همه ی سؤالات با ماده های آزمون مورد تحلیل قرار می گیرند؟ ۱) فرم های موازی

۲) همسانی درونی

3) کودر- ریچاردسون

4) آلفای کرونباخ

۲۴- چه موقع می توان به جای فرمول 20KR از فرمول 21KR استفاده کرد؟

۱) زمانی که سؤالات آزمون دارای ضریب دشواری پایین و شبیه به هم باشند.

۲) زمانی که واریانس سؤالات شبیه به هم باشد.

۳) زمانی که سؤالات آزمون دارای ضریب دشواری متوسط و شبیه به هم باشند.

۴) زمانی که انحراف معیار سؤالات شبیه به هم باشد.

۲۵- اگر واریانس کلی آزمونی برابر با که باشد و این آزمون از ۳ خرده مقیاس تشکیل شده باشد که واریانس آنها به ترتیب زیر باشد، پایایی آن را محاسبه کنید.

1) 85/0

2) 96/0

3) 90/0

4) 78/0

۲۶- اگر واریانس نمره ی خطای یک آزمون ۱۵/0 و واریانس نمرات مشاهده شده آن ۸۵/0 باشد، پایایی آن چقدر خواهد بود؟

1) 78/0

2) 82/0

3) 87/0

4) 90/0

۲۷- اگر میانگین نمره های آزمون و نمره ی حد تسلط برابر باشند، ضریب پایایی لیوینگ استون با کدام یک از روش های محاسبه و پایایی برابر است؟

1) روش محاسبه پایایی آزمون های سرعت

۲) روش های محاسبه پایایی آزمون های قدرت

۳) روش محاسبه پایایی نمرات اختلاف

۴) روش های محاسبه پایایی آزمون های وابسته به هنجار

۲۸- اگر پایایی نمرات پیش آزمون ۶۸/0و پایایی نمرات پس آزمون ۷۵/0 باشد، در صورتی که ضریب همبستگی بین این آزمون ۳۷/0 باشد، پایایی نمرات اختلاف چقدر می باشد؟

۱) 53/0

2) 60/0

3) 64/0

4) 67/0

۲۹- کدام یک از واریانس های زیر در شمار اجزاء واریانس خطا در روش تنصیف است؟

1) عامل بلوغ

۲) عامل روش

۳) عامل حدس زدن

4) عامل اجرا

۳۰- انتقال اثر و طول فاصله بین دو اجرای آزمون در کدام یک از روش های محاسبه پایایی تأثیر دارد؟

۱) دو نیمه کردن

۲) بازآزمایی

۳) کود - ریچاردسون 20

۴) کودر- ریچاردسون ۲۱

آزمون 1

1-کدام یک از گزینه های زیر در مورد قواعد اندازه گیری درست نیست؟

1) قواعد اندازه گیری خط مشی هایی برای نشان دادن مقدار شیء مورد اندازه گیری هستند.

۲) قواعد اندازه گیری معمول ترین و متداول ترین وسیله اندازه گیری نمونه ای از رفتار است.

۳) منظور از کاربرد قاعده در اندازه گیری این است که روش های نسبت دادن اعداد باید به طور روشن بیان شوند.

۴) بیان قواعد اندازه گیری به عینی کردن تصحیح آزمون کمک می کند.

۲- در ارزشیابی .................. اگر بخواهیم پیشرفت تحصیلی گروهی دانش آموز را ارزشیابی کنیم، عملکرد دانش آموزان را با یکدیگر مقایسه می کنیم.

1) وابسته به ملاک

۲) وابسته به هنجار

۳) تکوینی

۴) تراکمی

٣- مقیاس مورد استفاده در آزمون های چند گزینه ای چیست؟

۱) اسمی (Nominal)

2) تربیتیOrdinal)

۳) فاصله ای (Interval)

۴) نسبی (Ratio)

4- پرسش های شفاهی برای کدام یک از ارزشیابی ها، بهترین استفاده را دارند؟

۱) تکوینی

۲) تراکمی

۳) تشخیصی

۴) رفتار ورودی

5- در کدام رویکرد ارزشیابی، اختلاف نظرهای ارزشیابان مختلف (موافق و مخالف) مورد تأکید قرار می گیرد؟

1) رویکرد مبتنی بر هدف

۲) رویکرد مبتنی بر نظر متخصصان

3) رویکرد مبتنی بر مدیریت

۴) رویکرد مبتنی بر اختلاف نظر متخصصان

6- در کدام یک از گزینه های زیر، بیشترین اصالت را به اهداف از پیش تعیین شده می دهند؟

۱) ارزیابی

۲) ارزشیابی

۳) اندازه گیری

4) سنجش

7- کدام یک از گزینه های زیر در مورد ضریب تمییز، درست می باشد؟

۱) هر چقدر ضریب تمییز بزرگتر باشد قوة تمییز أن سؤال بیشتر است.

2) هر چقدر ضریب تمییز کوچکتر باشد قوه تمییز آن سؤال بیشتر است.

3) هر چقدر ضریب تمییز به ۵/0 نزدیکتر باشد قوه تمییز آن سوال بیشتر است.

۴) سؤالی که ضریب تمییزش، به ۵/0 نزدیکتر باشد بهتر است.

8- همه ی موارد زیر از محاسن آزمون های صحیح و غلط است بجز؟

1) حدس زدن

۲) سادگی تهیه

۳) عینی بودن تصحیح جواب ها

۴) نمونه گیری خوب

۹- در یک آزمون ۶۰ سؤالی پنج گزینه ای، دانشجویی به ۴۰ سؤال پاسخ درست و به ۱۶ سؤال پاسخ غلط داده است و بقیه سوالات را بدون پاسخ گذاشته است. نمره اصلاح شده او با احتساب نمره منفی چقدر است؟

1) 46

2) 38

3) 34

4) 36

۱۰- در جدول زیر پاسخ های داده شده توسط گروه های قوی و ضعیف یک کلاس به یک سؤال چهارگزینه ای ارائه شده است. در این جدول گزینه ی (ج) گزینه ی درست می باشد. ضریب دشواری سؤال چقدر است؟

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| گزینه ها | الف | ب | ج | د |
| 27% بالا | 3 | 2 | 7 | 3 |
| 27% پایین | 7 | 2 | 3 | 3 |

1) 34/0

2) 35/0

3) 33/0

4) 36/0

۱۱- کدام یک از موارد زیر، جزو قواعد سؤال های چند گزینه ای نیست؟

1) گنجاندن حداکتر داوطلب در هر سؤال

۲) متناسب بودن همه ی گزینه های یک سؤال

۳) گزینه های انحرافی باید توجه همه ی آزمون شوندگان را جلب کند.

۴) موارد ۱ و ۳

۱۲- اگر متغیر اول دو ارزشی ساختگی باشد و متغیر دوم نیز دو ارزشی ساختگی باشد، کدام یک از انواع همبستگی است؟

1) گشتاوری پیرسون

۲) دورشت های

3) تتراکوریک

۴) فی (فای)

۱۳- اگر اندازه انحراف استاندارد توزیع گروهی از نمرات افزایش یابد، به شرط یکسانی سایر شرایط چه تغییری در مربوط به هر نفره ایجاد می شود؟

۱) مثبت می شود

۲) منفی می شود

۳) کاهش می یابد

۴) افزایش می یابد.

۱۴- هدف از روایی .......................... مشخص کردن این است که آیا میتوان یکی از دو آزمون را به جای هم استفاده کرد؟

1) پیش بینی

۲) همزمان

3) ملاکی

4) صوری

15- در چه صورتی، میانگین و میانه و نما روی هم قرار می گیرند؟

۱) منحنی بهنجار باشد.

۲) منحنی نرمال باشد.

۳) انحراف استاندارد بالای 5 باشد.

۴) موارد ۱ و ۳

۱۶- هر قدر از لحاظ درجه دشواری میان سؤال های آزمون تفاوت وجود داشته باشد، ضریب پایایی (Reliability) حاصل از کدام نوع، کاسته می شود؟

1) پیرسون

۲) کودر - ریچاردسون

۳) آلفای کرانباخ

۴) فرم های موازی

۱۷- «آیا نمرات آزمون به مفاهیم نظری که آزمون برای سنجش آن ها درست شده است، مربوط اند؟» سؤال فوق معرف کدام نوع روایی است؟

1) سازه

۲) محتوایی

3) ملاکی

4) صوری

۱۸- بهترین روش برای تعیین پایایی (Reliability) آزمون های سرعت چیست؟

1) دو فرم موازی

۲) آلفای کرانباخ

3) کودر - ریچاردسون

۴) دو نیمه کردن

۱۹- یک آزمون ۳۰ سؤالی دارای اعتبار (Reliability) 60/0 است. اگر تعداد سؤالات این آزمون را به ۵۰ سؤال افزایش دهیم، اعتبار آزمون چقدر می شود؟

1) 61/0

2) 71/0

3) 67/0

4) 74/0

۲۰- نمره ۷۵ در مقیاس T معادل چه نمره ای در مقیاس نمرات هوش با (میانگین ۱۰۰ و انحراف معیار برابر ۱۶) می باشد؟

1) 125

2) 130

3) 135

4) 140

آزمون (۲)

1- هر کدام ویژگی آزمون های استاندارد شده تضمین می کند که اگر چند نفر از آزمون واحدی استفاده کنند، نتایج یکسانی به دست خواهند داد؟

۱) قواعد اندازه گیری

۲) نمرات استاندارد

۳) هنجارهای آزمون

۴) همه ی موارد

۲- اگر باشد، آنگاه چقدر است؟

1) 2/0-

2) 0

3)

4) 1

٣- کدام یک از گزینه های زیر در مورد علاقه و نگرش درست است؟

۱) علاقه و نگرش دو نام متفاوت برای موضوع واحدی هستند.

۲) علاقه به یک نهاد ولی نگرش به یک فعالیت مربوط می شود.

۳) علاقه و نگرش دو جنبه از شخصیت هستند

۴) علاقه و نگرش کاملا جدا از هم و نامرتبط اند.

۴- در کدام نوع ارزشیابی، نمی توان از آزمون وابسته به هنجار استفاده کرد؟

۱) جایابی

۲) تکوینی

3) تشخیصی

۴) تراکمی

5- اجرای پیش آزمون و اجرای آزمون رفتار ورودی، جزء کدام مرحله ارزشیابی است؟

۱) طرح ریزی

۲) فرآیندی

۳) فرآوردهای

۴) گزینه های ۲و۳

۶- یک آزمون از نوع صحیح و غلط، شامل دو سؤال است. احتمال این که در این آزمون، شانسی بتوان نمره گرفت چقدر است؟

1)

2)

3)

4)

۷- چنان چه ضریب همبستگی بین دو نیمه یک آزمون ۵/0 باشد، ضریب اعتبار (Reliability) کل آزمون چقدر است؟

1) 1

2) 5/0

3) 67/0

4) 75/0

۸- مهمترین گام در تهیه آزمون های پیشرفت تحصیلی چیست؟

۱) تعیین کردن هدف

۲) سنجش عملکرد

۳) تعیین کردن سؤال

۴) تهیه جدول مشخصات

9- سؤال زیر چه نوع هدف آموزشی را می سنجد؟

در یک پاراگراف به طور خلاصه توضیح دهید که چرا فشارسنج بهترین ابزار پیش بینی وضع هواست؟

1) ارزشیابی

2) دانش

۳) کاربرد

4) فهم

۱۰- کدام یک از ویژگی های زیر از ویژگی های آزمون چند گزینه ای نیست؟

۱) ساختن این آزمون آسان است.

۲) آزمون چند گزینه ای انعطاف پذیرتر هست.

3) آزمون چند گزینه ای شامل اهداف سطوح بالا نیز می شود.

۴) عینی ترین آزمون است.

۱۱- اگر تعداد پاسخ دهندگان به یک سؤال ۸۰ نفر باشند و از این تعداد ۶۵ نفر به آن سؤال جواب درست داده باشند، ضریب دشواری این سؤال چقدر خواهد بود؟

1) 80

2) 81

3) 78

4) 79

۱۲- تنها در صورتی ضریب تمییز به حداکثر می رسد که ضریب دشواری برابر ..................... باشد.

1) 1

2) 5/0

3) 5

4) 1/0

۱۳- اگر متغیر اول، پیوست و متغیر دوم دو ارزشی واقعی باشد، استفاده از کدام نوع همبستگی مناسب است؟ ١) دو رشته ای نقطه ای

2) دو رشته ای

3) فای

۴) تتراکوریک

۱۴- کدام یک از شاخص های آماری زیر، اطلاعاتی کلی درباره توزیع نمرات یک آزمون به ما می دهد؟

۱) ضریب همبستگی بین سؤال ها

۲) ضریب دشواریم سؤال ها

۳) ضریب تمییز سؤال ها

۴) انحراف معیار

۱۵۔ کدام یک از نمرات معیار زیر، دارای میانگین ۵ و انحراف معیار ۲ است؟

1) نمرات ۹ بخشی

۲) نمرات SAT

۳) نمره T

۴) نمره NCE

۱۶- کدام یک از عوامل زیر از نوع عوامل بیرونی مؤثر بر روایی است؟

1) راهنمایی آزمون

۲) کیفیت سؤال آزمون

۳) ترتیب قرار گرفتن سؤال های آزمون

۴) کمبود زمان آزمون

۱۷- ثرندایک و هیگن برای داده های ملاکی، چهار ویژگی مطلوب ذکر کرده اند، این ویژگی ها به ترتیب اهمیت عبارت اند از:

۱) در دسترس بودن - ربط داشتن - پایایی - بی طرفی

۲) در دسترس بودن - بی طرفی – پایایی -ربط داشتن

3) ربط داشتن - در دسترس بودن - بی طرفی - پایایی

۴) ربط داشتن - بی طرفی - پایایی - در دسترس بودن

18- انحراف معیار آزمونی برابر با ۱۵ و ضریب اعتبار ((Reliability آن برابر با ۹۰ می باشد، خطای معیار اندازه گیری این آزمون چقدر است؟

1) 7/4

2) 2/5

3) 5/4

4) 1/5

۱۹- سؤالات یک آزمون دارای درجه دشواری متوسط و ضریب تمییز ۹۸/0 هستند، توزیع نمرات آزمون به چه صورت خواهد بود؟

1) نرمال

۲) U شکل

۳) کجی مثبت

۴) مسطح

۲۰- کدام یک از گزینه های زیر در مورد رابطه ی روایی و پایایی درست است؟

۱) آزمون روا، پایا نیز هست.

2) پایایی شرط لازم برای روایی است.

۳) آزمون پایا حتما روا هست.

۴) همه ی موارد

آزمون (۳)

1- کدام یک از گزینه های زیر، مهم ترین وجه اشتراک بین پژوهش آموزشی و ارزشیابی آموزشی است؟

۱) هدف هر دو تصمیم گیری است.

۲) هر دو به نظریه وابسته است.

۳) هدف هر دو استنتاج است.

۴) هر دو فعالیت علمی محسوب می شوند.

2- از محدودیت های نرم درصدی این است که:

1) فاصله واحدها در طول مقیاس برابر نیست.

۲) دارای اعداد اعشاری است.

۳) برای افراد عادی قابل فهم نیست.

۴) محاسبه و تفسیر رتبه های درصدی دشوار است.

۳- کدام یک از گزینه های زیر از ویژگی های جدول مشخصات آزمون نیست؟

۱) جدول مشخصات به ایجاد تعادل بین آموزش و سنجش کمک می کند.

۲) مهم ترین گام در تهیه آزمون پیشرفت تحصیلی تهیه جدول مشخصات است.

۳) جدول مشخصات فقط برای آزمون های ملاکی ضروری است و برای هنجاری ضرورتی ندارد.

۴) جدول مشخصات به معلم کمک می کند تا از مجموع بزرگی از سوال های ممکن، نمونه ای را انتخاب کند که به خوبی معرف محتوا و هدف ها است.

۴- فعالیت هایی مثل «تحلیل و تفسیر داده ها»، «تهیه گزارش» در کدام یک از مراحل ارزشیابی، قرار دارد؟

۱) مرحله طرح ریزی

۲) مرحله تولیدی

۳) مرحله فرآیندی

۴) موارد ۱ و ۲

5- کدام یک از آزمون های عملکردی یا واقعی زیر، کمترین فاصله را با عملکرد در زندگی واقعی دارد؟

۱) نمونه کار

۲ آزمون شناسایی

۳) آزمون کتبی عملکردی

۴) انجام عملکرد در موقعیتهای شبیه سازی شده

6-کدام یک از گزینه های زیر از معایب آزمون های چند گزینه ای است؟

۱) پاسخ های این آزمون ها با عینیت کافی قابل تصحیح نیست.

2) آزمون های چند گزینه ای فقط برای اهداف در سطح دانش مناسب است.

3) این آزمون ها از سایر آزمون های عینی انعطاف پذیری کمتری دارند.

۴) ساختن این آزمون ها بسیار دشوار است.

7- کدام یک از گزینه های زیر، در مورد ضریب دشواری، صحیح تر است؟

۱) ضریب دشواری یعنی درصد کل آزمون شوندگانی که به یک سوال جواب غلط می دهند.

۲) ضریب دشواری هرچه به یک نزدیک تر باشد، سؤال آسان تر است.

۳) ضریب دشواری هرچه به یک نزدیک تر باشد، سوال سخت تر است.

۴) سوال های خوب، سوال هایی هستند که ضریب دشواری آن ها نزدیک به صفر باشد.

8- در یک آزمون پنج گزینه ای که دارای ۶۰ سوال است، دانش آموزی به ۴۵ سوال پاسخ داده است و از این تعداد ۹ پاسخ غلط و ۳۶ پاسخ درست است. نمره اصلاح شده این دانش آموز برابر است با:

1) 75/33

2) 5/32

3) 33

4) 35

9- وقتی یکی از متغیرها پیوسته و متغیر دوم دو ارزشی واقعی باشد، از کدام نوع همبستگی استفاده می شود؟ 1) دو رشته ای

۲) تتراکوریک

۳) دو رشته ای نقطه ای

۴) فی (فای)

۱۰ـ در کدام یک از هنجارها، موقعیت نسبی یک دانش آموز در گروهی که او خود عضوی از آن است، نشان داده می شود؟

۱) هنجارانی

2) هنجار درصدی

3) نمرات معیار

۴) موارد ۲ و ۳

۱۱- فراهم آوردن بازخورد برای معلم و تشخیص و رفع مشکلات احتمالی یادگیری دانش آموزان و نواقص آموزشی معلم، هـدف کدام یک از انواع ارزشیابی است؟

۱) جایابی

۲) تکوینی

3) تشخیصی

4) تراکمی

۱۲- کدام یک از گزینه های زیر در مورد پایایی آزمون، صحیح تر است؟

1) هر چه تعداد سوال های آزمون کمتر باشد، پایایی آن آزمون بزرگ تر می شود و در نتیجه خطای معیار اندازه گیری آن هم بزرگتر می شود.

۲) هرچه تعداد سوال های آزمون بیشتر باشد، پایایی آزمون بزرگتر می شود و در نتیجه خطای معیار اندازه گیری آن بزرگتر می شود.

۳) هرچه تعداد سوال های آزمون بیشتر باشد، پایایی آن آزمون بزرگتر می شود و در نتیجه خطای معیار اندازه گیری آن کوچک تر می شود.

۴) ضریب پایایی صفر نشان دهنده این است که تمامی واریانس مشاهده شده، واریانس حقیقی است.

۱۳- چنانچه ضریب همبستگی بین دو نیمه یک آزمون 70/0 باشد، ضریب اعتبار (reliability) کل آزمون چقدر است؟

1) 80/0

2) 81/0

3) 82/0

4) 83/0

۱۴ـ کدام مفهوم، اصطلاحی است که به هدفی که آزمون برای تحقق بخشیدن به آن درست شده است. اشاره می کند؟

1) پایایی

2) روایی

۳) همبستگی

۴) عینیت

۱۵ـ هنجار سن ذهنی یا سن عقلی، اولین بار توسط چه کسی مطرح شد؟

۱) پیرسون

۲) اسپیرمن

۳) بینه

4) ثورندایک

۱۶- برای تعیین ضریب پایایی آزمون هوش و شخصیت، کدام یک از روش های زیر مفیدتر است؟

۱) روش پایایی

۲) روش بازآزمایی

۳) روش فرم های موازی

4) روش کودر – ریچاردسون

17- کدام یک از مفاهیم زیر مربوط به پایایی (reliability) نیست؟

1) سازه

2) دقت

3) اعتمادپذیری

4) ثبات

۱۸- کدام یک از انواع روایی، بیشتر برانگیزش تأکید می کند؟

۱) محتوایی

۲) ملاکی

3) صوری

4) پیش بینی

۱۹- اگر نمره علی در مقیاس Z برابر با 5/1- باشد، نمره ی علی در مقیاس T چند خواهد بود؟

1) 35-

2) 35

3) 65

4) 65-

۲۰ـ چنان چه در آزمون های گسترده پاسخ (انشایی) به دانشجویان شانس انتخاب داده شود تا از بین سوالات به چند سوال به دلخواه پاسخ دهنـد، این اقدام چگونه بر تفسیر نمره های آزمون تأثیر می گذارد؟

۱) تفسیر ارزشیابی ملاکی را دقیق تر می سازد.

2) تفسیر ارزشیابی ملاکی نمره ها اثر می گذارد.

۳) بر تفسیر ارزشیابی هنجاری نمره ها اثر می گذارد.

4) بر ارزشیابی های هنجاری و ملاکی اثر می گذارد.

سؤالات آزمون سراسری ۹۵

۱- در الگوی ارزشیابی سیپ (CIPP)، پس از سنجش نیازها، چه عنصری از برنامه مشخص می شود؟

1) منابع گردآوری داده ها

۲) هدف های برنامه

3) ابزارهای گردآوری داده ها

۴) روش های اجرای برنامه

۲- کم توجهی به مطلوبیت اهداف یک برنامه به هنگام ارزیابی آن، از مهم ترین نقاط ضعف کدام الگو به شمار می رود؟

۱) تحقق هدف ها

2) سیپ

۳) مبتنی بر مدافعه

۴) هدف آزاد

۳- اگر اطلاعات دو سؤال i وj به صورت مقابل باشند، جهت و مقدار کوواریانس بین آنها کدام است؟

1) 25/0

2) 75/0-

3) 1+

4) 0

4- این که یادگیرنده با فکر کردن درباره کارهای خود، خودارزیابی را نیز می آموزد، از مزایای اصلی کذام شیوه سنجش است؟

۱) نمونه کار در موقعیت های طبیعی

۲) آزمون عملکردی

۳) انجام عملکرد در موقعیت های شبیه سازی

۴) کارپوشه

۵- «معلمی یک آزمون پیشرفت تحصیلی استانداردشده را به منظور این که آیا دانش و مهارتهای تدریس شده در کلاس درس را پوشش می دهد بررسی کرده است». هدف وی بررسی کدام روایی آزمون بوده است؟

1) واگرا

۲) ملاکی

۳) محتوایی

۴) همگرا

۶- «استنباط هایی که بر مبنای نمرات آزمون صورت می گیرد تا چه اندازه درست، مفید و مناسب است». بیانگر کدام ویژگی آزمون است؟

۱) پایایی

2) هنجار

۳) کارایی

4) روایی

۷- با توجه به منحنی ویژگی سؤال، کدام مورد درست است؟

۱) ضریب تمییز بالای سؤال

۲) دشواری پایین سؤال

۳) بالا بودن همبستگی سؤال با نمره کل

۴) حدس پذیری بالای سؤال



۸- نمرات خام در آزمودنی در یک آزمون استاندارد به ترتیب ۱۵ و ۲۴ است. در صورتی که نمره استاندارد Z معادل با این دو نمره به ترتیب ۲- و۱ باشد، میانگین و انحراف استاندارد توزیع نمرات خام به ترتیب کدام است؟

1) 21- 3

2) 5/19 – 5/7

3) 30- 6

4) 20- 4

9-کدام مورد ویژگی هایی از نظام مورد ارزیابی است که به عنوان جنبه های اصلی نظام برای قضاوت مورد استفاده قرار می گیرد؟

۱) اهداف

۲) معیار

۳) ملاک

۴) نشانگر

۱۰- «پژوهشگری همبستگی بین نمرات یک آزمون عزت نفس را با نمرات چندین آزمون دیگر به منظور تعیین این که آیا واقعا آزمون عزت نفس را اندازه می گیرد، بررسی کرده است»، وی به دنبال کدام ویژگی آزمون بوده است؟

1) روایی محتوایی

2) پایایی

۳) روایی پیش بین

۴) روایی سازه

۱۱- وقتی هدف، جایگزین کردن نمره های یک پرسشنامه به جای مصاحبه بالینی است، از کدام نوع روایی استفاده می شود؟

۱) پیش بین

2) سازه

۳) همزمان

۴ ) واگرا

۱۲- در یک نمونه ۱0۰ نفری میزان کای دو () محاسبه شده برای مقایسه ضریب دشواری دو سؤال یک آزمون برابر با است. ضریب همبستگی فای بین دو سؤال چقدر است؟

1) 6/0

2) 3/0

3) 90/0

4) 09/0

۱۳- در کدام یک از رویدادهای ارزیابی، بیشترین تأکید بر شیوه های عینی است؟

1) اجرای عمل

۲) مدیریت گرا

۳) هدف آزاد

۴) هدف گرا

۱۴- در یک پرسش نامه ۱۵ سؤالی که برای اندازه گیری سازه چند بعدی رضایت شغلی طراحی شده است، ۱۲ سؤال در زمینه بند رضایت از حقوق دریافتی طراحی شده است. کدام ویژگی این پرسشنامه مطلوب نیست؟

1) روایی محتوایی

۲) پایایی

۳) همسانی درونی

۴) روایی ملاکی

۱۵- برای برآورد توانایی درک واژه یک فرد در یک آزمون ۱۰۰ کلمه ای، منطقی ترین نسبت کدام است؟

1) 70/0

2) 30/0

3) 80/0

4) 50/0

۱۶- ضریب تمییز یک سؤال برابر با 25/0D = است. در صورتی که از ۲۰ نفر افراد گروه قوی ۱۴ نفر به سؤال پاسخ درست داده باشند، چند نفر از افراد گروه ضعیف به سؤال پاسخ درست داده اند؟

1) 6

2) 4

3) 9

4) 5

۱۷- آزمونی که فقط یک ویژگی را اندازه گیری کند و نمرات حاصل از اجرای آن هم نرمال باشد، دارای چه مقیاس اندازه گیری است؟

1) نسبتی(ratio)

۲) فاصله ای ((interval

۳) ترتیبی (ordinal)

۴) اسمی (nomial)

۱۸- در یک آزمون که سؤالات به طور تصادفی مرتب شده اند، ضرایب دشواری سؤالاتی که در ابتدای آزمون قرار دارند بالا، اما ضرایب دشواری سؤالات آخر آزمون پایین است. نوع آزمون و بهترین روش محاسبه پایایی آن کدام است؟

1) سرعت - بازآزمایی

۲) توانایی- بازآزمایی

۳) توانایی - همسانی درونی

۴) سرعت - همسانی درونی

۱۹- در یک آزمون ۲۰ سؤالی، انحراف استاندارد نمرات ۵ = S و خطای معیار اندازه گیری نیز ۵ = SEM است. پایایی آزمون کدام است؟

1) 5/0

2) 0

3) 40/0

4) 1

۲۰- اگر پایایی (reliability) یک آزمون ۳۰ سؤالی 80/0 باشد و ۱۵ سؤال آن به صورت تصادفی برای ساخت فرم کوتاه آن انتخاب شود، پایایی آن کدام است؟

1) 60/0

2) 40/0

3) 67/0

4) 50/0

پاسخنامه آزمون سراسری ۹۵

1-گزینه «۲» الگوی ارزشیابی سیپ چهار نوع ارزشیابی را ارائه می کند؛ چون عنوان سیپ (CIPP) از حرف اول کلمات موقعیت یا زمینه (context)، درون داد (input)، فرایند (process) و فراورده (product) به دست آمده است.

ارزشیابی زمینه: این ارزشیابی مهم ترین نوع ارزشیابی های الگوی سیپ است و ارتباط نزدیکی با تصمیمات مربوط به برنامه ریزی دارد.

ارزشیابی درون داد: هدف کلی ارزشیابی درون داد، کمک به تصمیم گیرندگان برای بررسی راه حل های گوناگون، تحقق هدف ها و تکامل برنامه ها برای رسیدن به اهداف است و به مجریان برنامه ها کمک می کند تا از فراهم کردن درون دادهای غیرلازم خودداری کنند؛ زیرا باعث عدم تحقق اهداف می شود.

ارزشیابی فرایند: این نوع ارزشیابی قضاوت درباره ی اجرای یک برنامه است و هدف از آن آشکارسازی نقائص برنامه در هنگام اجراست.

ارزشیابی برون داد یا فراورده: هدف از این ارزشیابی، فراهم سازی اطلاعات لازم برای تصمیم گیری درباره اثربخشی و کارایی برنامه است که به موجب آن درباره ادامه یا قطع اجرای برنامه در مراحل بعدی تصمیم گیری می شود.

با توجه به توضیحات ارائه شده، پس از سنجش نیازها، هدف برنامه مشخص می شود.

۲- گزینه «۱» ارزشیابی به عنوان یکی از مراحل مهم فرایند برنامه ریزی در ابعاد مختلف از جمله برنامه ریزی آموزشی محسوب می شود و تاکنون الگوها و روش های مختلفی برای انجام آن طراحی و پیشنهاد شده که یکی از کامل ترین آنها که به الگوی سیپ (cipp) معروف است، توسط استافیل بیم طراحی و ارائه گردیده است. این الگو از دسته ی الگوهای مدیریت گرا بوده که مدیران و تصمیم گیران با استفاده از بررسی های پیمایشی، پرسشنامه ای، مصاحبه و به منظور اثربخشی و کنترل کیفیت از آن استفاده می کنند. در این الگو، ارزشیابی عبارت است از فرایند تعیین کردن، به دست آوردن و فراهم ساختن اطلاعات توصیفی و قضاوتی در مورد ارزش و مطلوبیت هدف ها، طرح، اجرا و نتایج. هر الگوی ارزشیابی شامل سه مرحله است:

1- طراحی اطلاعاتی که باید جمع آوری شود.

۲- شیوه به دست آوردن اطلاعات

۳- فراهم کردن اطلاعات برای به دست آوردن قسمتهای مورد نظر الگوی سیپ خود مستلزم انجام چهار نوع ارزشیابی است :

۱- ارزشیابی زمینه (Context evaluation)

۲- ارزشیابی درون داد (Input evaluation)

٣- ارزشیابی فرایند (Process evaluation)

۴- ارزشیابی برون داد (Product evaluation)

سر عامل ارزیابی می تواند برای ارزشیابی یک برنامه، با توجه به موفقیت از هر چهار نوع ارزشیابی بالا استفاده کند و یا آنکه فقط یک یا چند قسمت از برنامه را به صورت مستقل مورد ارزشیابی قرار دهد. معیار اساسی برای هر چهار نوع ارزشیابی عبارت اند از ارزش و مطلوبیت. کم توجهی به مطلوبیت اهداف یک برنامه به هنگام ارزیابی آن، از مهم ترین نقاط ضعف تحقق اهداف به شمار می رود.

۳۰-گزینه «1»

۴- گزینه «۴» اینکه یادگیرنده با فکر کردن درباره ی کارهای خود، خودارزیابی را نیز می آموزد. از مزایای اصلی کارپوشه است. کارپوشه نوعأ مجموعه ای از کارهای یک یادگیرنده است که کوشش، پیشرفت و موفقیت تحصیلی او را در یک زمینه خاص نشان می دهد. یادگیرنده از کارپوشه برای درک و ارزشیابی پیشرفت خود استفاده زیادی می کند. یادگیرنده در کارپوشه با اندیشیدن دربارهی نمونه کارهایی که در کارپوشه خود قرار می دهد، ارزشیابی از خود را می آموزد. این عمل موجب گسترش مهارت های تفکر، تحلیل انتقادی و راهبردهای فراشناختی می شود.

5- گزینه «3» وقتی معلمی به منظور بررسی دانش و مهارت های تدریس شده در کلاس، یک آزمون پیشرفت تحصیلی استاندارد شده را اجرا می کند، در واقع روایی محتوایی را بررسی کرده است. باید توجه داشت که از میان انواع روش های بررسی روایی (روایی محتوا، ملاکی و سازه)، روایی محتوایی مربوط به آزمون های پیشرفت تحصیلی است.

6- گزینه «۴» مفهوم روایی به معنای آن است که آزمون مورد نظر تا چه حد خصیصه مورد نظر را می سنجد و به دقت حاصل از اندازه گیری آن خصیصه اشاره دارد. روایی شامل سه نوع روایی محتواء ملاکی و سازه است. و پایایی به یکسانی نتایج در اندازه گیری های متعدد اشاره دارد؛ یعنی ابزار اندازه گیری در شرایط یکسان تا چه اندازه نتایج یکسانی را ارائه می دهد.

هنجارسازی یا استاندارد سازی به تبدیل نمرات خام به مقیاسی قابل مقایسه اشاره دارد؛ یعنی تبدیل نمرات خام آزمون ها به نمرات استاندارد که امکان مقایسه را فراهم می کند؛ بنابراین استنباط های مربوط به درست، مفید و مناسب بودن، بیانگر روایی آزمون است.

۷- گزینه «۴» با توجه به محل تقاطع خط و محورy و بالاتر از ۵/0 بودن آن، حدس پذیری بالای سؤال استنباط می شود. هر چه محل تقاطع منحنی ویژگی سؤال و محور عرضی (yها) بالاتر باشد، به این معناست که احتمال پاسخ شانسی یا حدسی بیشتر است..

۸- گزینه «1»

با توجه به فرمول و جایگزینی اعداد، فقط گزینه (۳) یعنی پاسخ صحیح است.

۹- گزینه «3» ملاک، جنبه ی اصلی برای قضاوت در نظام مورد ارزیابی در آزمون های ملاک مرجع است از لحاظ روش تفسیر نمره های آزمون، می توان آزمون ها را که دو گروه ملاک مرجع (Criterion Referenced) و گروه مرجع تقسیم کرد. در آزمون های ملاک مرجع عملکرد هر آزمودنی با توجه به یک ملای مشخص و بدون توجه به عملکرد سایر افراد گروه ارزیابی می شود؛ مثلا در آزمون های نمونه ای از - کار، آزمودنی کار معینی را در مدت زمان معینی انجام میدهد و کار او با یک ملاک که قبلا تهیه شده مقایسه می شود از طرف دیگر، در آزمون های گروه مرجع عملکرد هر آزمودنی با متوسط عملکرد افراد گروه مورد مقایسه قرار گرفته و تفسیر می شود. مثلا برای تفسیر از هنجارهای گروه همسالان آزمودنی استفاده می شود.

۱۰- گزینه «۴» روایی سازه به شواهد همگرا و انتزاعی برای سازه مورد نظر اشاره دارد؛ یعنی سازه مورد نظر تا چه اندازه یک سازه یا خصیصه را که برای اندازه گیری آن ساخته شده است، می سنجد. یکی از روشهای جمع آوری شواهد همگرا، بررسی همبستگی مقیاس مورد نظر با آزمون هایی است که همان خصیصه را اندازه گیری می کنند؛ بنابراین بررسی همبستگی نمرات یک آزمون اعتماد به نفس با آزمون های دیگر که همان خصیصه را می سنجند. روایی سازه را بررسی می کند.

۱۱- گزینه «3» روایی ملاکی عبارت است از کارآمدی یک ابزار اندازه گیری در پیش بینی رفتار یک فرد در موقعیت های خاص برای این منظور عملکرد هر فرد در آزمون با یک ملاک مقایسه می شود. روایی ملاکی بر دو نوع است: ۱) روایی پیش بین ۲) روایی همزمان

روایی پیش بین عبارت است از بررسی رابطه نمره های آزمونی که برخی ویژگی ها را می سنجد و آنچه ادعای پیش بینی آن را دارد. روایی پیش بین می تواند به وسیله رابطه عملکرد در یک آزمون با آزمون ملاک رفتاری به دست آید. به عبارت دیگر، ضریب همبستگی نمره های حاصل از اجرای آزمون با نمره های متغیر ملاک، نمایانگر روایی پیش بین است.

روایی همزمان در مواردی به کار می رود که داده های حاصل از دو اندازه گیری در یک زمان در دسترس باشد. در این گونه موارد، عملکرد در یک آزمون به عملکرد در آزمون دیگر مرتبط می شود. اعتبار همزمان در مواردی محاسبه می شود که هدف جانشین کردن یک ابزار اندازه گیری به جای ابزار دیگری باشد. بنابراین با توجه به توضیحات ارائه شده، گزینه (۳) پاسخ درست است.

۱۲- گزینه «۳» با توجه به فرمول فی، ضریب همبستگی فی 3/0 است.

۱۳- گزینه «۴» در ارزیابی هدف گرا، بیشترین تأکید بر شیوه های عینی است. انواع الگوهای ارزیابی: از جمله دسته بندی های ارزیابی می توان به دسته بندی هفت گانه مرکز مطالعات ارزیابی (دانشگاه کالیفرنیا) اشاره کرد. این دسته بندی هفت گانه به شرح زیر است:

۱) الگوی هدف گرا

۲) الگوی تصمیم گرا

3) الگوی پاسخگویانه

۴) الگوهای مبتنی بر طرح های تحقیق آزمایشی (برای تبیین رابطه علّی)

5) الگوی هدف - آزاد

۶) الگوی مبتنی بر مدافعه

۷) الگوهای کاربرد گرا

این دسته بندی چنان است که الگوهای دسته اول بر سنجش پیشرفت تحصیلی بادگیرندگان تأکید دارد. دسته دوم الگوها، بر نیازهای اطلاعاتی سس تصمیم گیرندگان تأکید می کند. سومین دسته الگوها بر فرآیند برنامه تأکید دارد. در حالی که چهارمین دسته الگوها، بر تعیین آثار برنامه و برقراری رابطه علت و معلولی در برنامه مورد ارزیابی تأکید دارد. به عبارت دیگر، دسته چهارم الگوها از روش های تحقیق آزمایشی برای برقراری رابطه على استفاده می کند . این دسته الگو را می توان مثلا در ارزیابی اثربخشی برنامه های آموزشی ضمن خدمت معلمان به کار برد. دسته پنجم، شامل الگوهای هدف - آزاد است. این و دسته از الگوها فارغ از چارچوب آغازین آثار برنامه مورد ارزیابی را می سنجند. به عبارت دیگر، این الگوی ارزیابی بدون توجه به این که چه چیزی هدف اصلی برنامه بوده است، به گردآوری آن دسته از اطلاعات ارزیابی می پردازد که بتواند میزان توان برنامه را در برآوردن نیازهای افراد ذیربط بسنجد. دسته ششم از طریق برقراری و فرآیندی شبه قانونی و شنیدن آرا و نظرهای مدافعان و غیر مدافعان، درک افراد را از برنامه می سنجد. در کاسته هفتم، الگوهای کاریردگرا قرار دارند. در این دسته الگوها، اطلاعات ارزیابی چنان گردآوری می شوند که حداکثر استفاده را برای افراد ذی ربط داشته باشد.

۱۴- گزینه «۱» سازه رضایت شغلی، یک سازه چند بعدی است و شامل چندین عامل است. اینکه از میان ۱۵ سؤال این مقیاس، ۱۲ سؤال مربوط به یک عامل رضایت از حقوق دریافتی باشد، نشانه عدم روایی محتوای آزمون است؛ یعنی محتوای آزمون با آنچه آزمون برای سنجش آن ساخته شده است، همخوانی ندارد.

۱۵- گزینه «۴» توانایی برآورد درک یک فرد در یک آزمون ۱۰۰ کلمه ای آن است که حداقل به ۵۰ درصد سؤالات یا نیمی از سؤالات پاسخ صحیح دهد که با توجه به تعداد سؤالات، ۵۰ درصد ۱۰۰ سؤال 5/0 است.

۱۶- گزینه «3»

۱۷- گزینه «۲» آزمونی که فقط یک ویژگی را اندازه گیری می کند و نمرات حاصل از اجرای آن هم نرمال است، دارای یک مقیاس فاصله ای می باشد. یک متغیر را می توان در سطوح مختلف اندازه گیری کرد. انتخاب سطح مناسب برای اندازه گیری متغیر مورد مطالعه باعث میشود که داده های گردآوری شده، گویای واقعیت مورد مطالعه باشند. به طور کلی، چهار سطح یا مقیاس برای اندازه گیری متغیرها می توان در نظر گرفت: مقیاس اسمی، مقیاس رتبه ای، مقیاس فاصله ای و مقیاس نسبتی.

1- مقیاس اسمی (Nominal Scale): مقیاس اسمی برای اندازه گیری متغیرهای مقوله ای به کار می رود. این مقیاس شامل حداقل دو مقوله متمایز است که هیچ گونه تقدم یا تأخر در آن وجود ندارد. به عبارت دیگر، میان مقوله های مقیاس اسنادی نمی توان ترتیب خاصی در نظر گرفت؛ مثلا برای متغیر جنسیت دو مقوله نمی توان ترتیب ویژه ای منظور داشت.

۲- مقیاس رتبه ای (Ordinal Scale): مقیاس رتبه ای برای اندازه گیری متغیرهایی به کار میرود که پیوسته باشند و تفاوت حالت های مختلف صفت متغیر فقط از نظر سلسله مراتب وضع افراد قابل نمایان باشد. مثلا برای تعیین رتبه های کنکور، می توان مقیاس رتبه ای را به کار برد و افراد را در مراتب رتبه اول، دوم تا رتبه آخر دسته بندی کرد. از جمله متغیرهایی که مقیاس رتبه ای برای آنها به کار می رود، متغیر نگرش می باشد.

٣- مقیاس فاصله ای (Interval Scale): مقیاس فاصله ای مقیاسی است که به وسیله آن می توان متغیرهای کمی را که دارای مبدأ اختیاری هستند، از اندازه گیری کرد. به وسیله این مقیاس نه تنها می توان افراد را رتبه بندی کرد، بلکه تفاوت آنها از نظر صفت متغیر مورد مطالعه را نیز می توان معین کرد، اما این مقیاس دارای صفر مطلق نیست؛ مثلا برای اندازه گیری هوش و ارزیابی عملکرد از مقیاس فاصله ای استفاده می شود؛ زیرا متغیر هوش دارای مقدار صفر نمی باشد.

4- مقیاس نسبتی (Ratis Scale): مقیاس نسبتی، که بالاترین سطح اندازه گیری است، مقیاسی دارای مبدأ صفر مطلق است که از فاصله های مساوی برخوردار است. برای هر دو مقدار این مقیاس می توان نسبتی را تعیین کرد که حاکی از بیشی مقدار صفت متغیر در یک فرد نسبت به فرد دیگر مورد مطالعه باشد. مثلا در اندازه گیری سال های خدمت کارکنان، می توان فردی یافت که تازه استخدام شده باشد و سابقه خدمتش صفر باشد. همچنین، مثلا می توان نسبت سابقه خدمت دو نفر را که یکی دارای سابقه خدمت ۱۰ سال و دیگری ۵ سال است، حساب کرد. در این مثال، سابقه خدمت نقز او دو برابر نفر دوم است و نسبت سابقه شان ۲ می باشد که مساوی نسبت سابقه خدمت دو نفر دیگر از کارکنان است که سابقه خدمت آنان به ترتیب ۶ سال و ۳ سال می باشد. به عبارت دیگر، مقیاس نسبتی علاوه بر دارا بودن ویژگی های مقیاس فاصله ای دارای مبدأ واقعی (صفر مطلق) نیز می باشد.

۱۸- گزینه «۱» در آزمون های سرعت، همهی سؤالات دارای ضرایب دشواری یکسانی هستند، پس آزمون موردنظر نمی تواند از نوع سرعت باشد. با توجه به اینکه بهترین روش بررسی پایایی، بازآزمایی است، گزینه (1) درست است.

۱۹- گزینه «۲» با توجه به مقادیر انحراف استاندارد نمرات (۵) و خطای معیار اندازه گیری (۵)، پایایی آزمون با جایگزینی در فرمول زیر به دست خواهد آمد:

20-گزینه «3» پایایی آزمون به معنای ثبات آزمون در طول زمان است. برای بررسی پایایی در شرایطی که چند سؤال آن به صورت تصادفی انتخاب می شود، به شیوه زیر عمل می کنیم:

سوالات آزمون دکتری 96

1- املاک مورد استفاده در اندازه گیری صنعت مورد بررسی در آزمون های هنجاری، ملاکی، استاندارد شده و معلم ساخته از چه نوعی است؟

١) نسبی- دبی - مطلق - مطلق

۲) نسبی - مطلق - نسبی – مطلق

3) مطلق - نسبی - نسبی - مطلق

۴) مطلق - نسبی - مطلق - نسبی

۲- توزیع نمرات یک آزمون عزت نفس نرمال با میانگین ۷۰ و انحراف استاندارد ۱۲ است. در صورتی که عزت نفس 200 دانش آموز با این آزمون اندازه گرفته شود، عزت نفس چند نفر بین ۵۸ تا ۸۲ قرار می گیرد؟

1) 34

2) 68

3) 136

4) 168

۳- کدام مورد درباره اندازه گیری ملاک و هنجار مرجع درست است؟

1) اندازه گیری ملاک مرجع برای ارزشیابی آموزش های انفرادی مناسب تر است.

۲) اندازه گیری ملاک مرجع برای رتبه بندی و مقایسه آزمودنی ها مناسب تر است.

3) اندازه گیری هنجار مرجع برای داوری در مورد میزان مهارت کسب شده در هدف خاصی مناسب تر است.

۴) اندازه گیری هنجار مرجع برای تعیین آزمودنی هایی که میزان خاصی از توانایی را کسب کرده اند مناسب تر است.

۴- توزیع نمرات یک آزمون پیشرفت تحصیلی نرمال با میانگین ۳۰ و انحراف استاندارد ۶ است. کاهش یک نمره ای کدام یک از نمرات زیر موجب کاهش بیشتر در رتبه درصدی معادل با آن می شود؟

1) 20

2) 24

3) 36

4) 32

5- همبستگی بین وزن با عادات غذایی برابر با 32/0 است. وزن از طریق یک ترازوی دقیق و بدون خطای اندازه گرفته شده است؛ اما عادات غذایی از طریق پرسشنامه خود گزارش دهی با پایایی 64/0اندازه گرفته شده است. همبستگی واقعی بین وزن و عادات غذایی کدام است؟

1) 36/0

2) 4/0

3) 6/0

4) 68/0

6- معلم درس ریاضیات قصد دارد برآوردی از خطای اندازه گیری ناشی از طرح سؤالات از نمونه ای از اهداف و محتوای آموزشی داشته باشد. برای این منظور از کدام روش محاسبه پایایی می تواند استفاده کند؟

1) فرم های موازی

۲) دونیمه کردن

۳) آلفای کرانباخ

۴) بازآزمایی

7- یک آزمون تشخیص اختلال یادگیری در بین ۵۰ دانش آموز سالم و ۵۰ دانش آموز مبتلا به اختلال یادگیری اجرا شد. از بین دانش آموزان سالم نمره ۵ نفر و از بین دانش آموزان مبتلا به اختلال یادگیری نمره ۴۰ نفر بالاتر از نقطه برش آزمون لیست ویژگی آزمون (specifity) کدام است؟

1) 10/0

2) 45/0

3) 80/0

4) 90/0

۸- اصلی ترین هدف از تبدیل نمره های خام (rew score) به نمره های فرمولی (formula score) کدام است؟

۱) اعمال جریمه برای حدس تصادفی

۲) اعمال تشویق برای حدس آگاهانه

۳) محاسبه نمره کل برای محاسبه رتبه

۴) محاسبه نمره کل برای مقایسه پذیری افراد

۹- کدام مورد به توزیع نمرات در آزمون های ملاک - مرجع مربوط است؟

1) یکنواخت

۲) کجی مثبت

3) کجی منفی

۴) نرمال

10-کدام ویژگی سؤالات آزمون برای به کارگیری نظریه های نمره کار آمد واقعی و الگوهای خصیصه مکنون الزامی است؟

۱) وابستگی موضوعی

۲) دشواری متوسط

3) چندبعدی بودن

۴) همگنی

۱1- برای بررسی پایایی ارزشیابی کیفی معلمان دوره ابتدایی از پیشرفت تحصیلی دانش آموزان از کدام نوع پایایی می توان استفاده کرد؟

1) بازآزمایی

۲) بین داوران

3) کودر - ریچاردسون

۴) همسانی درونی

۱۲- در یک آزمون تشریحی ۱۰ سؤالی نمره هر سؤال ۲ است. در صورتی که میانگین نمره سؤال اول آن 6/1 باشد، ضریب دشواری آن کدام است؟

1) 16/0

2) 84/0

3) 80/0

4) 20/0

۱۳- در صورت ثابت ماندن سایر شرایط آزمون همبستگی دورشته ای نقطه ای سؤال با نمره کل در آزمونی با چند سؤال بیشتر خواهد بود؟

1) 10

۲) 15

3) 20

4) 25

۱۴- منحنی ویژگی سؤال در چه صورت دارای برازش خواهد بود؟

۱) پراکندگی داده ها به توزیع نرمال نزدیک باشد.

۲) توزیع نمونه گیری داده ها به توزیع نرمال نزدیک باشد.

۳) مجذور فواصل بین داده های مشاهده شده کمترین باشد.

۴) مجذور فواصل بین داده های مشاهده شده و نمودار کمترین باشد.

۱۵- اگر دو صفت روان شناسی با استفاده از سؤال های (صحیح و غلط) و (چند گزینه ای) اندازه گیری شود، مقادیر موجود بر روی قطر ماتریس همبستگی حاصل کدام مورد را نشان می دهند؟

1) روایی ملاکی

۲) روایی همزمان

3) همسانی درونی

۴) پایایی بازآزمایی

۱۶- اگر نمره های افراد در دو متغیر x و y رتبه بندی و سیس همبستگی بین رتبه ها با استفاده از روش پیرسون محاسبه شود، کدام شاخص همبستگی به دست می آید؟

1) اسپیرمن

2) چهارخانه ای

۳) تائوی کندال

۴) دورشته ای

۱۷- در صورتی که پایایی یک آزمون پیشرفت تحصیلی که سؤالات به صورت دو ارزشی نمره گذاری می شوند را با دو روش آلفای کرانباخ (a) و کودر - ریچاردسون ۲۰ (۲۰-KR) محاسبه کنیم، کدام مورد درست است؟

1) 20 – KR < a

2) 20 – KR a =

3) 20 – KR > a

4) 20 – KR ≥ a

۱۸- کدام مورد درباره استفاده از آزمون های تشریحی در اندازه گیری پیشرفت تحصیلی درس ریاضیات درست است؟

۱) تهیه آن برای معلمان ریاضی دشوار است.

۲) بر روش مطالعه ریاضی دانش آموزان اثر مثبت دارد.

۳) نمونه معرفی از محتوای درس ریاضیات را اندازه می گیرد.

۴) از معلمان سایر دروس برای نمره گذاری آن می توان استفاده کرد.

۱۹- در صورتی که ضریب همبستگی دورشته ای نقطه ای سؤال با نمره کل آزمون برابر با 80/0 باشد، ضریب دشواری سؤال کدام است؟

1) 80/0

2) 50/0

3) 20/0

4) 10/0

۲۰- در صورتی که پایایی یک آزمون با استفاده از روش دونیمه کردن ۷۵/0 باشد، همبستگی بین دو نیمه آزمون کدام است؟

1) 25/0

2) 50/0

3)40/0

4) 60/0

پاسخنامه آزمون دکتری ۹۶

1- گزینه «۲» یک تقسیم بندی که برای آزمون ها وجود دارد، ملاک مورد استفاده برای قضاوت درباره نمره هر فرد است. در صورتی که از یک ملاک مطلق (مشخص و بدون توجه به عملکرد سایر افراد گروه) برای قضاوت درباره نمره فرد استفاده شود «آزمون ملاک مرجع» نامیده می شود. آزمون های معلم ساخته و به طور کلی هر و آزمون ملاکی در این طبقه قرار می گیرند. دسته دیگر «آزمون های گروه مرجع» هستند، در این نوع آزمون ها عملکرد هر آزمودنی براساس یک ملاک نسبی قضاوت میشود. این ملاک نسبی از عملکرد سایر افراد گروه به دست می آید. آزمون های هنجاری و استاندارد شده از آزمون های گروه مرجع محسوب می شوند.

۲- گزینه «3» برای به دست آوردن تعداد افرادی که نمره آنها بین ۵۸ تا ۸۲ است و با توجه به نرمال بودن توزیع، ابتدا نمرات Z معادل ۵۸ و ۷۲ را طبق فرمول مقابل محاسبه می کنیم:

با توجه به نمرات Z که برابر با ۱ و ۱- است و با توجه به منحنی توزیع نرمال ۶۸% افراد در این فاصله قرار می گیرند. حال برای به دست آوردن اینکه چه تعداد از این ۲۰۰ نفر نمره ای بین ۵۸ تا ۷۲ دارند، ۶۸% را در ۲۰۰ ضرب می کنیم:

۱۳۶ نفر نمره ای بین ۵۸ تا ۷۲ کسب کرده اند.

3- گزینه «1» هدف از بررسی هنجارها برای یک آزمون، تعیین چگونگی مقایسه یک آزمودنی با دیگران است. یک آزمون وابسته به هنجار هر فردی را با یک ناهنجار مقایسه می کند. اما در سال های اخیر گرایش به سمت آزمون هایی بوده است که تعیین کنند دانش آموزان از اطلاعات خاص چه چیزهایی می دانند.

این آزمون ها دانش آموزان را با دیگران مقایسه نمی کنند، بلکه عمل هر دانش آموز را با یک ملاک یا سطح قابل انتظار عمل مقایسه می کنند. با توجه به توضیحات ارائه شده در رابطه با آزمون های «ملاک مرجع» و «هنجار مرجع» و کاربرد هریک از آن ها تنها گزینه (۱) صحیح است.

۴- گزینه «۴» یکی از ویژگی های رتبه درصدی این است که اگر اعداد زیادی نزدیک به این نمره (نمره ای که قصد محاسبه رتبه درصدی آن را داریم) باشند، یک تغییر جزئی در نمره های خام موجب افزایش چشمگیری در رتبه درصدی می شود. با توجه به اینکه در صورت سؤال به نرمال بودن توزیع نمرات اشاره شده است و در توزیع نرمال بیشتر نمرات در اطراف میانگین قرار دارند؛ نمره ۳۲ نزدیک ترین نمره به میانگین می باشد. بنابراین یک کاهش نمره در نمره ۳۲ بیشتر از سایر نمرات موجب تغییر رتبه درصدی آن می شود.

5- گزینه «۲» همبستگی واقعی بین دو مقیاس اندازه گیری از طریق تقسیم همبستگی مشاهده شده بین دو مقیاس بر ریشه دوم حاصل ضرب پایایی هر یک از دو مقیاس به دست می آید. این فرمول که به «تصحیح کاهش» معروف است به صورت زیر قابل محاسبه است:

: همبستگی واقعی بین دو مقیاس

: همبستگی مشاهده شده بین دو مقیاس

: پایایی هر یک از دو مقیاس (در اینجا چون وزن با ترازوی بدون خطا اندازه گیری شده، پایابی آن معادل یک است )

۶- گزینه «1» برای مطالعه پایایی مربوط به آزمون های ملاک مرجع نمرات آزمون را با نمرات مقیاس های دیگر که همبستگی آن را با آزمون پذیرفته اند مقایسه می کنند. از آنجایی که آزمون معلم ساخته یک آزمون ملاک مرجع محسوب می شود، بهترین روش محاسبه پایایی آن استفاده از فرم موازی است.

۷- گزینه «4» ویژگی آزمون به توانایی آزمون برای تشخیص درست افراد بدون بیماری اشاره دارد که از فرمول زیر محاسبه می شود:

X : تعداد افرادی که به اشتباه بیمار تشخیص داده شده اند

Y : تعداد افراد سالمی که به درستی سالم تشخیص داده شده اند.

8- گزینه «۴» مقایسه نمرات خام دشوار است اما تبدیل نمرات خام به نمرات فرمولی امکان مقایسه وضعیت فرد را در دو آزمون مختلف فراهم می کند. همچنین تبدیل نمره خام به نمرام فرمولی به معلمان در فعالیتهای آموزشی کمک می کند. غالبا معلمان در طول یک دوره آموزشی آزمون های مختلفی را اجرا می کنند و در آخر دوره، نمره های این آزمون ها را با هم جمع با معدل نهایی آن ها را به منظور نمره نهایی هر دانش آموز مورد استفاده قرار می دهند. این روش از و نظر آماری درست نیست، مگر اینکه کلیه آزمون ها دارای انحراف استاندارد یکسانی باشند. بهترین کار تبدیل نمره های خام در آزمون به نمره فرمولی است که همه از آن ها دارای میانگین و انحراف استاندارد مساوی باشند. سازمان سنجش گزینه (۱) را صحیح اعلام کرده است که با توجه به توضیحات داده شده نادرست است.

نکته: سازمان سنجش گزینه (۱) (اعمال جریمه برای حدس تصادفی) را به عنوان گزینه صحیح معرفی کرده است، در صورتی که نمره خام با توجه به میانگین و انحراف استاندارد جامعه تبدیل به نمره فرمولی (شامل نمره z نمره T و..) می شود و در این تبدیل هیچ سهمی برای حدس در نظر گرفته نمی شود. به عبارت دیگر، در نمرات فرمولی فرض بر این است که نمره فرد، نمره واقعی او بدون حدس و شانس است.

9- گزینه «3» کچی منفی زمانی اتفاق می افتد که آزمون آسان باشد و فراوانی پاسخ های صحیح زیاد باشد، که این حالت معمولا در آزمون های ملاکی از جمله آزمون های معلم ساخته رخ می دهد.

کجی (چولگی یعنی انحراف از حالت تقارن، بنابراین زمانی که منحنی متقارن باشد کجی برابر با صفر است. اگر توزیع داده ها منقارن باشد، نه شاخص میانگین، میانه و مد در یک نقطه متمرکز و با هم برابرند. اما اگر تجمع داده ها در سمت چپ زیاد باشد و دم نمودار به سمت راست کشیده شده باشد، می گوییم توزیع دادهها دارای کجی (چولگی) مثبت است. در این حالت همیشه میانگین بزرگتر از میانه و نما کوچکتر از میانه است. آزمونهای دشوار دارلی مروجی مثبت هستند؛ زیرا نمرات بیشتر افراد در پایین توزیع قرار دارد. از طرف دیگر در صورتی که تجمع داده ها در سمت راست زیاد باشد و دم نمودار به رسمت چپ کشیده شده باشد، می گوییم توزیع داده ها دارای کجی (چولگی) منفی است. در این حالت همیشه میانگین کوچکتر از میانه و نما بزرگتر از میانه است. کجی منفی زمانی اتفاق می افتد که آزمون آسان باشد و فراوانی پاسخ های صحیح زیاد باشد، یعنی نمرات افراد در بالای توزیع قرار دارد که این حالت معمولا در آزمون های ملاک از جمله آزمون های معلم ساخته رخ می دهد.

10- گزینه «۴» همگنی شاخصی است که درباره میزان صحت دو فرض تک بعدی بودن. آزمون مو بی خطا بودن اندازه گیری آگاهی میدهد و برقراری این دو فرض برای به کارگیری نظریه های نمره کارآمد و الگوی خصیصه مکنون الزامی است.

۱۱- گزینه «۲» نظام های مشاهده رفتاری غالیا به دلیل اختلاف بین نمرات واقعی و نمرات ثبت شده غیر قابل اعتماد هستند. برای ارزشیابی این مسائل فرد نیاز دارد تا پایایی مشاهده گران را برآورد کند. این برآوردهای پایایی نام های مختلفی دارد که شامل پایایی بین ارزیابان، بین نمره دهندگان، بین مشاهده گران یا قضاوت کنندگان می شود.

۱۲- گزینه «3» دشواری در نظریه کلاسیک برابر است با تعداد افرادی که به سؤال پاسخ درست داده اند و دارای دامنه ای از صفر تا یک می باشد. در اینجا نمره هر سؤال در دامنه ای بین صفر تا دو قرار دارد و میانگین سؤال اول 6/1 است، بنابراین با تقسیم 6/1 بر ۲، دشواری سؤال به دست می آید. دشواری سؤال موردنظر برابر با

8/0 است.

۱۳- گزینه «1» هرچه تعداد سؤالات آزمونی کمتر باشد، واریانس کل آزمون نیز کوچک تر است. براساس فرمول ضریب همبستگی دورشته ای نقطه ای نیز هر قدر واریانس آزمون کوچکتر باشد، میزان ضریب همبستگی دورشتهای نقطه ای بزرگ تر خواهد بود. بنابراین آزمون ده سؤالی در مقایسه با سایر آزمون ها همبستگی دورشته ای نقطه ای بیشتری دارد.

۱۴- گزینه «۴» با ارزش ترین شواهد نیکویی برازش به وسیله باقی مانده ها فراهم شده و به بهترین وجه به کمک نمودارها تفسیر می شوند. وقتی که و باقی مانده ها کوچک هستند و به طور تصادفی در اطراف منحنی ویژگی سؤال توزیع شده اند، می توانیم بگوییم که با عملکرد در سؤال برازش دارد. در واقع بهترین برازش سؤال با منحنی ویژگی سؤال زمانی است که مجذور فواصل بین داده های مشاهده شده و نمودار کمترین میزان باشد.

۱۵- هیچ کدام از گزینه ها صحیح نیست. (چون صورت سؤال و گزینه های آن مبهم است، گزینه خاصی را نمی توان به عنوان گزینه درست تعیین کرد. به صورت کلی به سؤال پاسخ می دهیم).

برای پاسخ به این سؤال، نخست تعریف مختصری در مورد هر یک از گزینه ها ارائه می دهیم روایی ملاکی: منظور از روایی ملاکی، میزان ارتباطی است که بین نمرات حاصل از یک آزمون یا وسیله ی اندازه گیری دیگر وجود دارد. هرگاه از نمرات یک آزمون برای پیش بینی عملکرد افراد در آزمونی دیگر استفاده کنیم، یا روایی ملاکی سروکار داریم.

روایی ملاکی: به دو نوع روایی پیش بینی و روایی همزمان تفسیم می شود. اگر دو آزمون مزبور پس از گذشت یک فاصله زمانی بین آنها اجرا شوند و نمرات یک آزمون برای پیش بینی عملکرد افراد در آزمون دیگر به کار گرفته شود، روایی پیش بینی مدنظر است. اما اگر هر دو آرمون، همزمان اجرا شوند و رابطه بین آن ها به طور همزمان تعیین شود، روایی همزمان مدنظر است.

پایایی و بازآزمایی: در این روش، یک آزمون واحد را دو بار روی گروهی از آزمون شوندگان اجرا می کنیم و همبستگی بین نمرات حاصل از دو بار اجرای آزمون را محاسبه می کنیم. به این ضریب همبستگی، پایایی و بازآزمایی می گویند.

روش بازآزمایی، بر ثبات نتایج آزمون تأکید دارد، اما روش های همسانی درونی بر همسانی ماده های تشکیل دهنده یک آزمون تأکید دارند. در روش های همسانی درونی، آزمون فقط یک بار بر روی گروه واحدی از آزمون شوندگان اجرا می شود. در این روش، ساخت درونی آزمون بررسی میشود و ملاک مورداستفاده برای بررسی همسانی درونی، نمره کل آزمون است. برای این کار باید همبستگی بین نمرات آزمون های فرعی یا خرده آزمون ها با نمره کل آزمون محاسبه شود و هر یک از خرده آزمون ها که با نمره کل آزمون، همبستگی اندکی را نشان داد، حذف شود. نمرات خرده آزمون ها باید با نمره کل آزمون، همبستگی زیادی داشته باشد و خرده آزمون ها نباید با هم همبستگی زیادی داشته باشند. زیرا در این صورت فرض می شود که همه آن ها یک چیز را اندازه می گیرند و لذا تکراری هستند.

صورت این سؤال مبهم است. اشاره نشده است که این دو سری سؤال ها، در قالب یک آزمون اجرا شده اند یا دو آزمون. اگر دو سری سؤالات (صحیح و غلط و چندگزینه ای) در قالب یک آزمون اجرا شده باشند که کلا گزینه های (۱)، (۲) و (۳) حذف می شوند. اما در عین حال، چون همبستگی دو صفت با هم مطرح است (اشاره شده که همبستگی سؤالات مربوط به دو صفت روانشناسی با یکدیگر)، نمی توان گفت که پاسخ صحیح، گزینه (۳) است. چون در ر همسانی درونی، باید همبستگی بین خرده آزمون ها با نمره کل آزمون محاسبه شود.

اگر هم منظور سؤال این باشد که سؤالات صحیح و غلط، مربوط به یک آزمون و سؤالات چندگزینه ای، مربوط به آزمون دیگری هستند، در این صورت گزینه (۳) و (۴) حذف می شوند. اما چون مشخص نکرده و آزمون همزمان اجرا می شوند یا پس از گذشت یک فاصله زمانی، باز هم نمی توان گزینه صحیح را تشخیص داد. . چون روایی همزمان، بخشی از روایی ملاکی است که در آن، هر دو آزمون همزمان اجرا می شوند.

۱۶- گزینه «۱» ضریب همبستگی اسپیرمن صورتی از ضریب همب پیرسون است و زمانی به کار می رود که نمره ها رتبه بندی شده باشند یا اینکه به جای اعداد، رتبه های آن ها در دست باشد.

۱۷ـ گزینه «۲» زمانی که سؤال ها به صورت دوارزشی نمره گذاری می شوند، پایایی به دست آمده از هر دو روش آلفای کرونباخ و کودر ریچاردسون ۲۰ برابر است.

pq : واریانس سؤال های دوارزشی (زمانی که نمره گذاری دوارزشی باشد مانند واریانس سؤال این گونه محاسبه می شود.)

: واریانس نمره های خام آزمون

K : تعداد سوال ها

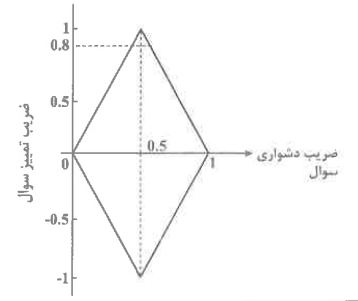
۱۸- گزینه«۳» یکی از ویژگی های آزمون های تشریحی تقویت عادات صحیح مطالعه در همه زمینه ها به ویژه در زمینه ترکیب و ارزشیابی است. ولی سه گزینه دیگر از ویژگی های آزمون های غیرتشریحی است.

۱۹ـ گزینه «۲» اطلاعات داده شده در صورت سؤال برای محاسبه ضریب دشواری کافی نیست، بنابراین برای رسیدن به پاسخ صحیح باید از رد گزینه های اشتباه استفاده کرد.

ضریب همبستگی دورشته ای نقطه ای سؤال با نمره کل آزمون، بیانگر ضریب تمییز آن سؤال است. درواقع، هرچه این ضریب همبستگی بالاتر باشد، یعنی رآن سؤال با سایر سؤالات آزمون هم خوانی بیشتری دارد و چیزی را می سنجد که در سؤال های دیگر آزمون می سنجند.

از طرفی، هرچه ضریب تمییز بزرگ تر باشد، آن سؤال بهتر می تواند آزمون شوندگان قوی کند عیف را از هم جدا کند. ضریب تمیز بین محدوده 1± تغییر می کند و زمانی به حداکثر 1± می رسد که ، ضریب دشواری سؤال، 5/0 یعنی درحد وسط باشد.

با استفاده از شکل زیر می توان رابطه بین ضریب دشواری و ضریب تمییز را نشان داد. از آنجا که ضریب تمییز گفته شده در صورت سؤال، یعنی 8/0، ضریب تمییز نسبتاً بالایی است، می توان گفت که ضریب دشواری سؤال در محدودهی نزدیک به متوسط یعنی نزدیک به 5/0 است، از آنجا که گزینه های دیگر در فاصله های زیادی از دشواری متوسط قرار دارند، می توان با تقریب، گزینه (۲) را به عنوان پاسخ سؤال برگزید.



۲۰- گزینه «۴» برای پاسخ به این سؤال باید از فرمول اسپیرمن - براون استفاده کرد. بنابراین:

سؤالات آزمون سراسری ۹۶

۱- اگر همه سؤال های یک آزمون پیشرفت تحصیلی روان شناسی عمومی به هدف های این درس مربوط باشند و مشخص باشد که هر سؤالی از کدام هدف درس طرح شده است؛ این آزمون پیشرفت تحصیلی دارای چه نوع روایی است؟

۱) پیش بین

۲) سازه

۳) محتوایی

۴) همزمان

۲- از کدام آزمون آماری برای محاسبه ضریب همبستگی فای () بین دو سؤال دو ارزشی استفاده می شود؟

1) اسپیرمن

2) خی دو

3) دو رشته ای

۴) دو رشته ای نقطه ای

۳- اگر نمره T دانش آموزی در درس ریاضی معادل ۷۰ باشد، بدین معنی است که نمره او:

۱) دو انحراف معیار بالاتر از میانگین است

۲) یک انحراف معیار بالاتر از میانگین است.

3) یک انحراف معیار پایین تر از میانگین است.

۴) دو انحراف معیار پایین تر از میانگین است.

۴- اگر واریانس مشاهده شده نمرات یک آزمون ۱۵۰ و واریانس نمرات خطای آزمون ۳۰ باشد، پایایی (reliability) آزمون کدام است؟

1) 20/0

2) 30/0

3) 70/0

4) 80/0

5- اطلاعات زیر مربوط به پاسخ های گروه های قوی و ضعیف به یک سؤال است، پاسخ درست سوال گزینه «ج» است. ضریب تعییز سؤال کدام است؟

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| گروه ها/ گزینه ها | الف | ب | ج | د | بدون پاسخ | تعداد کل |
| قوی | 0 | 2 | 20 | 3 | 0 | 25 |
| ضعیف | 4 | 6 | 10 | 4 | 1 | 25 |

1) 2/0

2) 3/0

3) 4/0

4) 6/0

6- اگر یادگیرنده بتواند اطلاعات مرتبط و غیرمرتبط داده شده در یک مسئله را از هم تشخیص دهد، در یادگیری آن مطلب به کدام سطح از فرایندهای شناختی دست یافته است؟

۱) ارزشیابی کردن

۲) به یاد آوردن

3) فهمیدن

۴) تحلیل کردن

۷- کدام آزمون از عینیت و دقت بیشتری برخوردار است؟

1) شناختی

2) شخصیت

3) رغبت

4) نگرش

۸- کدام روایی، هم مستلزم تحلیل منطقی هم تحلیل آماری - همبستگی و تفاوت های گروهی است؟

1) همزمان

2) سازه

3) صوری

۴) پیش بین

۹- سؤال زیر چه نوع هدف آموزشی را می سنجد؟

«در تعیین رفتار انسان ها طبیعت بیشتر نقش دارد یا تربیت؟ نظر خود را با ذکر دلیل توضیح دهید.»

۱) ارزشیابی - دانش مفهومی

۲) تحلیل کردن دانش مفهومی

۳) فهمیدن - دانش مفهومی

۴) کاربستن - دانش روندی

۱۰- عامل اصلی تمایز بین آزمون های ملاکی و هنجاری کدام است؟

1) تعداد سؤال هایی که در آزمون گنجانده می شود.

۲) نوع سؤال هایی که در آزمون ها طرح می شوند.

۳) مهارت های مورد سنجش که معرف اهداف آموزشی هستند.

۴) معیاری که به واسطه آن عملکرد فرد در آزمون تفسیر میشود.

۱۱- پایایی یک آزمون پیشرفت تحصیلی 6/0 است. اضافه کردن دو سؤال با کدام ضریب دشواری موجب افزایش بیشتر در پایایی آزمون می شود؟

1) 10/0

2) 25/0

3) 45/0

4) 85/0

۱۲- ضریب تمییز (D) یک سؤال 45/0 است. در صورتی که از ۲۰ نفر گروه قوی ۱۳ نفر به سؤال پاسخ صحیح داده باشند، چند نفر از گروه ضعیف به سؤال پاسخ درست داده اند؟

1) 3

2) 4

3) 6

4) 8

۱۳- بر اساس فرمول کودر - ریچاردسون، پایایی یک آزمون در چه صورت کامل می شود؟

1) مجموع واریانس سؤالات بزرگتر از واریانس نمره کل باشد.

۲) نیمی از افراد نمره کامل و نیمی دیگر نمره صفر بگیرند.

3) تمام افراد به تمام سؤال ها پاسخ درست بدهند.

۴) همبستگی بین سؤالات آزمون متوسط باشد.

۱۴- با فرض اینکه j نماد سؤال jام ، x نماد نمره آزمون و y نماد متغیر دیگر باشند، عبارت های و به ترتیب به کدام مورد اشاره دارد؟

1) روایی سؤال - واریانس آزمون

۲) پایایی سؤال - روایی سؤال

۳) پایایی سؤال - واریانس آزمون

۴) روایی آزمون - خطای اندازه گیری سؤال

۱۵ـ در صورتی که نمرات خام یک آزمون را به رتبه های درصدی تبدیل کنیم، کدام مورد درست است؟

۱) فاصله بین نمرات خام و رتبه های درصدی مساوی است.

۲) همبستگی بین نمرات خام و رتبه های درصدی کامل است.

۳) ترتیب افراد در نمرات خام و رتبه های درصدی یکسان است.

۴) شکل توزیع نمرات خام و رتبه های درصدی همانند هم است.

۱۶- کدام یک از روش های تعیین پایایی برای ویژگی های ناپایدار مانند شادمانی مناسب نیستند؟

۱) کودر - ریچاردسون

2) دونیمه کردن

۳) آلفای کرانباخ

4) بازآزمایی

۱۷ـ بسیاری از نقطه ضعف های نظریه کلاسیک اندازه گیری در کدام پیش فرض آن ریشه دارند؟

۱) نبود همبستگی بین نمره خطا و واقعی

2) تصادفی بودن تغییرات خطا در تمام نمره ها

۳) در نظر گرفتن هر نوع تغییر تصادفی به عنوان خطا

۴) تعریف پایایی براساس نسبت واریانس خطا به واریانس واقعی

۱۸- یک سازمان دولتی براساس نتایج یک آزمون استخدامی ۳۰۰ کارمند جدید استخدام کرده است، بعد از یک سال همبستگی بین نمرات ایـن افراد در آزمون استخدامی و عملکرد شغلی آنها برابر با ۰/۰۸ = r میباشد، کدام ویژگی روان سنجی این آزمون استخدامی مناسب نیست؟

۱) روایی ملاکی

۲) روایی محتوایی

۳) همسانی درونی

۴) پایایی بازآزمایی

۱۹ـ کدام مورد از ویژگی های آزمون های معیار شده نیست؟

۱) نمره گذاری آن ها عینی است.

۲) با دقت طراحی، تحلیل و بازبینی می شوند.

۳) نتایج آن ها از طریق مقایسه با جداول هنجار تفسیر می شود.

۴) دستورالعمل های اجرا، نمره گذاری و تفسیر انعطاف پذیر است.

۲۰- پایایی یک آزمون هوشی که نمرات آن به صورت نمره های استاندارد IQ گزارش می شود، برابر 75/0 است، خطای استاندارد آن کدام است؟

1) 5/3

2) 15

3) 5/7

4) 25

پاسخنامه آزمون سراسری 96

1- گزینه «۳» روایی آزمون عبارت است از میزان کارآیی آن برای اندازه گیری خصیصه ای که به منظور اندازه گیری آن خصیصه ساخته شده است. برحسب هدف های خاص آزمون و روش های برآورد روایی آن، انواع مختلف روایی وجود دارد. در واقع روایی یعنی توافق بین نمره آزمون با صفت یا خصیصه ای که آزمون برای اندازه گیری آن ساخته شده است. روایی شواهدی است برای تفسیر و استنباط هایی درباره نمره آزمون و این شواهد بر سه نوع اند: شواهد وابسته به سازه، شواهد وابسته به ملاک و شواهد وابسته به محتوا.

روایی محتوا یکی از انواع روایی است که غالباً در آزمونهای پیشرفت تحصیلی مورد استفاده قرار می گیرد. برای تحلیل روایی محتوای آزمون پیشرفت تحصیلی باید معلوم کرد که محتوای آزمون تا چه اندازه با هدف های آموزشی رابطه دارد و سؤال های مطرح شده در آزمون تا چه اندازه معرف محتوا و هدف های آموزشی است. به منظور تحلیل روایی محتوای آزمون پیشرفت تحصیلی باید ابتدا یک جدول مشخصات شامل محتوا و هدف های آموزشی تهیه شود.

۲- گزینه «۲» ضریب کا از شاخص هایی است که در بررسی همبستگی هر سؤال با کل آزمون و برای بررسی ثبات و همسانی درونی بین یک سؤال با سؤال های دیگر آزمون به کار می رود. گیلفورد (۱۹۴۵) برای محاسبه ضریب فای فرمول ساده زیر را ارائه کرده است که براساس آزمون مجذور خی استوار است. شرط استفاده از ضریب فای این است که تعداد افراد گروه بالا و پایین برابر باشند.

: نسبت افراد گروه پایین که به سؤال پاسخ صحیح داده اند.

: نسبت افراد گروه بالا که به سؤال پاسخ صحیح داده اند.

p : نسبت پاسخ های درست آزمودنی ها به سؤال موردنظر.

۳- گزینه «۱» برای پاسخ به این سؤال با استفاده از فرمول نمره T نمره Z را به صورت زیر محاسبه می کنیم:

نمره Z برابر با ۲ شده است، بنابراین نمره این دانش آموز دو انحراف معیار بالاتر از میانگین قرار دارد.

۴- گزینه «۴» آزمون در صورتی دارای پایایی است که عاری از خطای اندازه گیری غیرنظامدار باشد. خطاهای اندازه گیری غیرنظامدار که بر نمرات آزمون تأثیر می گذارند، غیرقابل پیش بینی هستند و پایایی را کاهش می دهند. از سوی دیگر خطاهای اندازه گیری نظامدار نمرات آزمون را تحت تأثیر قرار می دهند چون اثر آن ها نسبتاً ثابت و قابل پیش بینی است لذا پایایی را کاهش نمی دهند.

پایایی از طریق تقسیم واریانس نمرات واقعی بر واریانس نمرات مشاهده شده قابل محاسبه است. ابتدا واریانس نمرات واقعی را به صورت زیر محاسبه می کنیم:

سپس با استفاده از فرمول مقابل پایایی را محاسبه می کنیم:

پایایی برابر با 8/0 است.

5- گزینه «3» ضریب تمیز یک سؤال برابر است با نسبت افراد گروه قوی که به سؤال پاسخ صحیح داده اند، منهای نسبت افراد گروه ضغیف که به سؤال پاسخ صحیح داده اند؛ لذا ابتدا نسبت افراد گروه قوی و نسبت افراد گروه ضعیف که به سؤال پاسخ صحیح دادهاند (انتخاب گزینه ج) را محاسبه و این دو مقدار را از هم کم می کنیم:

6- گزینه «۴» قلمرو شناختی شش سطح مختلف دارد که از آن ها به ترتیب با عنوان های دانش، درک و فراگیری، کاربستن، تحلیل کردن، ترکیب یا آفرینندگی و ارزشیابی نام می برند. تحلیل کردن چهارمین سطح قلمرو شناختی است و مستلزم فعالیت های ذهنی ای است که دانش آموز بتواند یک مطلب را به اجزای تشکیل دهنده آن تقسیم کند، روابط بین اجزا را کشف و درک کند و نحوه سازمان یافتن اجزا برای به وجود آوردن ساخت کلی مطلب و اصول حاکم بر آن را دریابد. هدف رفتاری زیر نمونه ای از رفتار آموخته شده در سطح تحلیل کردن است :

عقاید مهم و نتایج حاصل از خواندن یک متن را استخراج و تفسیر کنید.

۷- گزینه «1» آزمون های هوش، استعدادهای خاص و آزمون های پیشرفت عمدتأ توانایی های شناختی آدمی، را مورد سنجش قرار می دهند. آزمونهای رغبت، نگرش و ارزش ها بیشتر به سنجش بعد عاطفی معطوف اند. پرسشنامه های خودسنجی که در سنجش رغبت، نگرش و ارزش ها به کار می روند در مقایسه با آزمون های شناختی از دقت و عینیت کمتری برخوردارند.

8-گزینه «۲» روایی آزمون عبارت است از میزان کارایی آن برای اندازه گیری خصیصه ای که به منظور اندازه گیری آن خصیصه ساخته شده است. روایی بر سه نوع است: روایی محتوا، روایی ملاکی و روایی سازه.

روایی سازه از طریق مجموعه ای از فعالیت ها که در آن محقق همزمان سازه را تعریف و ابزاری را برای اندازه گیری آن تولید می کند، تعیین می شود. روایی زمانه شواهکم جمع آوری شدهای درباره معنای یک آزمون را در برمی گیرد. این امر با نشان دادن رابطه بین یک آزمون و آزمون ها و ابزارهای دیگر انجام می شود. لذا روایی سازه مستلزم تحلیل منطقی همراه با تحلیل آماری - همبستگی و تفاوتهای گروهی است. استقرار روایی سازه برای آزمون یک فرایند مداوم است و به فرایند ساختن نظریه های پیچیده علمی براساس مجموعه شواهد علمی شباهت دارد. شواهد قطعی روایی سازه هرگز براساس مشاهدات و منحصر به فرد به دست نمی آید. کمپل و فیسک (۱۹۵۹) برای تهیه آزمون هایی که روایی سازه دارند، اصولی را مطرح کردهاند، دو نوع شواهد برای با معنا کردن آزمون دارای روایی سازم لازم است. هر اندازه که امکان دارد شواهدی در مورد روایی همگرا و روایی افتراقی آزمون فراهم می شود.

9- گزینه «1» قلمرو شناختی شش سطح مختلف دارد که از آن ها به ترتیب با عنوان های دانش، درک و فراگیری، کازیستن، تحلیل کردن، ترکیب با آفرینندگی و ارزشیابی نام می برند. ارزشیابی آخرین سطح قلمرو شناختی و عالی ترین سطح آن است و بدون رسیدن به سطوح پایین تر، دست یابی به این سطح غیر ممکن است. داوری آگاهانه درباره چیزی براساس ملاک های معین که برای یک منظور مشخص انجام می گیرد ارزشیابی است. در این مثال آزمودنی باید نظر خود را با ذکر دلیل بیان کند. بنابر این هدف آموزشی در سطح ارزشیابی است و ارزشیابی به دانش مفهومی تعلق دارد.

۱۰- گزینه «۴» آزمون ها از جنبه های مختلف قابل طبقه بندی هستند؛ یکی از این جنبه ها ملاک داوری درباره نمرات آزمودنی ها است که بر این اساس آزمون ها به دو دسته ملاکی و هنجاری دسته بندی می شوند. تفاوت آزمون ملابل و هنجاری در معیاری است که برای قضاوت درباره آزمودنی ها در نظر گرفته می شود. در آزمون ملاکی، عملکرد هر آزمودنی با یک ملاکر مشخص و بدون توجه به عملکرد سایر افراد گروه مقایسه و درباره او قضاوت می شود؛ ولی در آزمون های هنجاری عملکرد هر آزمودنی با متوسط عملکرد افراد گروه نرم مثلا گروه همسالان مقایسه و تفسیر می شود.

۱۱- گزینه «3» یکی از راه های افزایش پایایی آزمون، بررسی همبستگی هر سؤال با سایر سؤال ها و با کل آزمون، یعنی تحلیل ضرایب تشخیص سؤال هاست. میدی هیچنین سؤال خیلی آسان و یا خیلی دشوار نیز ممکن است فاقد ضریب تشخیص بوده موجب پایین آمدن ضریب پایایی آزمون شود. لذا سؤالی که دارای ضریب دشواری متوسط باشد به افزایش پایایی بیشتر کمک می کند. به عبارت دیگر در صورتی که سؤالی بسیار آسان یا بسیار سخت باشد، توانایی تمیز کمتری بین افراد قوی و ضعیف خواهد داشت؛ بنابراین سؤال با دشواری متوسط توانایی تمیز بالا و در نتیجه دارای پایایی بالاتری است.

۱۲- گزینه «۱» ابتدا با استفاده از فرمول ضریب تشخیص میزان را محاسبه و سپس تعداد افرادی را که پاسخ صحیح داده اند محاسبه می کنیم.

۱۳- گزینه «3»

:p نسبت افرادی که به هر سؤال پاسخ درست داده اند.

:q نسبت افرادی که پاسخ غلط داده اند.

: واریانس کل آزمون

با توجه به فرمول کودر - ریچاردسون در صورتی که برابر با صفر شود، بیشترین میزان پایایی حاصل می شود و در صورتی برابر صفر می شود که تمام افراد به تمام سؤال ها پاسخ صحیح بدهند یا اینکه تمام افراد به تمام سؤال ها پاسخ اشتباه بدهند. در گزینه (۲) گفته شده که نیمی از افراد نمره کامل و نیمی دیگر نمره صفر گرفته اند، در حالی که شرط پایایی این است که عدد صفر باشد. پس گزینه (۲) سازمان سنجش نمی تواند صحیح باشد.

۱۴- گزینه «۲» عبارت ، نشان دهنده همبستگی بین سؤال با کل آزمون ضرب در انحراف استاندارد سؤال است؛ بنابراین همبستگی بین سؤال و کل آزمون گویای همسانی درونی و در نتیجه پایایی سؤال است.

عبارت ، نشان دهنده همبستگی بین سؤال با متغیر دیگری ضرب در انحراف استاندارد سؤال است؛ با توجه به اینکه روایی ملاکی از همبستگی بین سؤال و ملاک خارجی به دست می آید، این عبارت روایی سؤال را نشان می دهد.

۱۵- گزینه «3» رتبه درصدی، رتبه نسبی یک نمره از توزیع نمره ها را براساس مقیاس ۱۰۰ تعیین می کند و به ما می گوید، چند درصد نمرات در توزیع در زیر زن نمره قرار دارند. نمرات خام غالبا به طور یکنواخت توزیع نمی شوند؛ لذا رتبه های درصدی نشان دهنده واحدهای مساوی در مقیاس اندازه گیری نیستند. با توجه زبه اینکه رتبه درصدی رتبه افراد را براساس نمره خام آن ها تعیین می کند؛ بنابراین ترتیب افراد در نمرات خام و رتبه های درصدی یکسان است.

۱۶- گزینه «4» در روش بازآزمایی، آزمون را درباره گروه نمونه واحدی از آزمودنی ها در دو زمان متفاوت اجرا و ضریب همبستگی بین نمرات حاصل از دو بار اجرای آزمون محاسبه می شود. برآورد پایایی بازآزمایی برای ارزشیابی خطای وابسته به اجرای یک آزمون در دو زمان متفاوت به کار می رود. به کارگیری این روش تنها در مورد صفات و خصایصی مناسب است که معمولا در طول زمان تغییر نمی کنند. بنابر این پایایی بازآزمایی در مورد آزمون هایی که صفات پایدار شخصیت آدمی را اندازه گیری می کنند، مناسب است و برای آزمون هایی که ویژگی های در حال تغییر را اندازه می گیرند ارزشیابی پایایی با روش بازآزمایی مناسب نیست.

۱۷- گزینه «۲» خطای معیار اندازه گیری نوسان های مورد انتظار نمرات را که از خطر ناشی می شود توصیف می کند. در نظریه کلاسیک خطای معیار راندازه گیری درباره همه نمرات یک جامعه صدق می کند. اما این خطاهای معیار تنها در شرایط محدودی همشان خواهند بود.

۱۸- گزینه «۱» روایی ملاکی نشان می دهد که یک آزمون با یک ملاک مشخص مطابقت دارد. چنین شواهدی با همبستگی بالا بین آزمون و اندازه ملاک فراهم می شود. روایی ملاکی بر دو نوع روایی پیش بین و روایی همزمان است. در این سؤال روایی پیش بین مدنظر است و ملاک، عملکرد شغلی است که همبستگی آن با آزمون استخدامی پایین است؛ لذا روایی ملاکی این آزمون نامناسب می باشد.

۱۹- گزینه «۴» آزمون های معیار شده براساس دستورالعمل های خاص و به شیوه یکسان اجراء نمره گذاری و تفسیر می شوند. برای تفسیر نمره آزمودنی در آزمون معیار شده، نرم ها یا هنجارهایی براساس عملکرد گروه نمونه وسیعی از آزمودنی ها فراهم می شود تا نمره آزمودنی با آن مقایسه و موقعیت او نسبت به گروه نرم تعیین شود.

۲۰- گزینه «3» رابطه بین خطای استاندارد اندازه گیری با پایایی به صورت زیر است:

با توجه به اینکه آزمون موردنظر، آزمون IQ است، بنابراین انحراف استاندارد برابر با ۱۵ میباشد و خطای استاندارد اندازه گیری به صورت زیر محاسبه می شود:

سؤالات آزمون دکتری ۹۷

۱- کدام مورد درستی یا نادرستی عبارت های زیر را به ترتیب نشان می دهد؟

- درصد پاسخ های درست را می توان به عنوان مقیاسی از درجه دشواری با واحدهای برابر در نظر گرفت.

- حدس تصادفی در واریانس خطا سهیم نیست پس بر پایایی آزمون تأثیر ندارد.

- درصد پاسخ های درست، دشواری مفهومی سؤال را منعکس می کند.

- سفید گذاشتن سؤال ها به معنی مجازات فرد توسط خود وی است.

۱) نادرست - درست - درست - نادرست

۲) درست - نادرست - درست – درست

3) نادرست - نادرست - نادرست - درست

۴) درست - درست - نادرست - درست

۲- براساس فرمول اسپیرمن - براون پایایی (reliability) یک آزمون در صورت دو برابر شدن مساوی 8/0

است، پایایی اولیه آزمون برابر چند است؟

1) 37/0

2) 50/0

۳) 57/0

۴) 67/0

۳- اگر نمره اصلاح شده (با احتساب نمره منفی) فردی در یک آزمون ۴ گزینه ای برابر با ۶ باشد و تعداد پاسخ های نادرست او برابر با ۱۸ باشد، به چند سؤال آزمون پاسخ درست داده است؟

1) 6

2) 12

3) 18

4) 24

۴- در یک آزمون ۸ سؤالی به چند شیوه می توان پایایی دو نیمه کردن را محاسبه کرد؟

1) 16

2) 32

3) 35

4) 70

5- در صورتی که برای اندازه گیری پیشرفت تحصیلی در درس فیزیک در بین دانش آموزان فارسی زبان از آزمون فیزیک به زبان انگلیسی استفاده کنیم، کدام ویژگی آزمون کاهش می یابد؟

1) پایایی

۲) هنجار

۳) عینیت

۴) روایی

۶- پایایی یک آزمون هوش در بین کودکان ۶ تا ۱۶ ساله برابر با 82/0 است. پایایی این آزمون در یک نمونه از کودکان ۱۰ ساله کدام است؟

1)

2)

3)

4)

7- میانگین و واریانس توزیع نمرات تصادفی (حدس های کورکورانه) در یک آزمون ۵ گزینه ای ۵۰ سؤالی کدام است؟

1) 5، 2

2) 25، 5

3) 10، 8

4) 10، 16

۸- کدام مورد درستی یا نادرستی عبارت های زیر را به ترتیب نشان می دهد؟

- پایایی پایین به معنی کیفیت بد آزمون است.

- تحت برخی شرایط یک آزمون بد ممکن است دارای پایایی خوبی باشد.

- پایایی آزمون های ملاک مرجع قبل و بعد از آموزش مساوی یک است.

1) درست - دوست - درست

۲) نادرست - درست - نادرست

۳) درست - نادرست - نادرست

۴) نادرست - نادرست - درست

9- در نظریه کلاسیک اندازه گیری (CTT) ارزش مورد انتظار کدام نمره با خود آن برابر است؟

1) واقعی (T)

۲) مشاهده شده (X)

3) خطا (E)

۴) واقعی و مشاهده شده

۱۰- با فرض ثابت بودن سایر شرایط، برای ایجاد 5/0 افزایش در پایایی کدام آزمون، نیاز به تعداد سؤالات (n) بیشتری است؟

1)

2)

3)

4)

۱۱- ضریب دشواری ۴ سؤال در زیر ارائه شده است، کدام سؤالات دارای ضریب تمیز (D) کمتری هستند؟

0(I

0/25(II

0/5(III

1(IV

1) II,I

2) IV,I

3) III,II

4) IV,III

۱۲- «در یک آزمون ۲۰ سؤالی که برای اندازه گیری دانش تخصصی اساتید روانشناسی تربیتی تهیه شده بود، از ۱۰ داور متخصص خواسته شد تا هر سؤال را با توجه به مقیاس سه درجه ای ضروری (۲)، مفید(1) و غیرضروری (صفر) درجه بندی کنند، فرض کنید سؤال یک این آزمون را داور ضروری، یک داور مفید و یک داور غیر ضروری درجه بندی کرده اند.» در این صورت ضریب روایی محتوایی سؤال چقدر است؟

1) 87/0

2) 77/0

3) 66/0

4) 65/0

۱۳- در کدام شرایط دادن حق انتخاب چند سؤال از میان تعدادی سؤال بیشتر توجیه پذیر است؟

۱) قصد ایجاد مهارت تصمیم گیری در یادگیرندگان مدنظر باشد.

۲) هدف مقایسه کیفیت و کمیت پاسخ های یادگیرندگان با هم باشد.

۳) مشخص کردن یادگیرندگان قوی تر از یادگیرندگان ضعیف درنظر باشد.

۴) هدف اندازه گیری مهارت یادگیرندگان در نوشتن و پروراندن مطالب باشد.

۱۴- وقتی دانشجویی با شرکت در یک کارگاه مهارت های مقاله نویسی به این باور می رسد که مهارت در مقاله نویسی بسیار ارزشمند است، این بـاور مقاله نویسی به دانشجو در کدام حوزه اهداف آموزش و پرورش از نظر بلوم طبقه بندی می شود؟

۱) شناختی

۲) عاطفی

3) شناختی و عاطفی

4) روانی - حرکتی

۱۵ـ کدام یک از نمرات هنجاری زیر، دارای دامنه یکسانی هستند؟

I) رتبه های درصدی

II) نه بخشی

NCE (III

SAT (IV

1) II,I

2) III,I

3) IV,II

4) IV,III

۱۶- اگر در کل فرایند آزمون هیچ خطایی وجود نداشته باشد، نمره ای که آزمودنی کسب می کند برابر با کدام است؟

۱) نمره واقعی (T)

2) نمره خام

3) نمره خطا

4) نمره مشاهده شده (X)

۱۷ـ در یک سؤال متعلق به یک آزمون ملاکی، از ۴۰ نفر دانش آموز پاسخ دهنده به این آزمون ۱۰ نفر در پیش آزمون و ۴۰ نفر در پس آزمـون بـه سؤال مذکور پاسخ درست داده اند. شاخص تأثیر آموزش (ضریب پیش و پس از آموزش) سؤال چند است؟

1) 55/0

2) 65/0

3) 70/0

4) 75/0

۱۸ـ کدام مورد درباره «CTT» نادرست است؟

۱) پایایی آزمون های بلند کمتر از پایایی آزمون های کوتاه تر است.

۲) خطای استاندارد اندازه گیری درباره همه نمره های یک جامعه خاص صدق می کند.

۳) نمره های آزمون زمانی معنا پیدا می کند که موقعیت آنها با گروه نرم مقایسه شود.

۴) مقایسه نمره های آزمون ها در نرم های چندگانه زمانی بهینه است که آزمون ها موازی یا همتا باشند.

۱۹ـ ملاک های قبولی در چهار امتحان در زیر ارائه شده است، با توجه به این ملاک ها، کدام موارد از نوع هنجار مرجع هستند؟

(I 15 نفری که دارای بالاترین نمره هستند.

(II افرادی که نمره آن ها یک انحراف استاندارد بالاتر از میانگین است.

(III افرادی که در امتحان نمره ۲۴ و بالاتر کسب کنند.

(IV افرادی که که به ۷۰ درصد از سؤالات امتحان پاسخ درست داده اند.

1) III,I

2) IV,II

3) IV,I

4) III,II

۲۰- چنانچه پایایی دو آزمون حداکثر 70/0 و همبستگی بین آن ها نیز 50/0 باشد، مقدار روایی آزمون حداکثر چقدر است؟

1) 20/0

2) 35/0

3) 50/0

4) 71/0

پاسخنامه آزمون دکتری ۹۷

۱ـ گزینه «۳» درصد پاسخ های درست در نظریه کلاسیک معادل ضریب دشواری در نظر گرفته شده است؛ اما این مقیاس دارای واحدهای برابر نیست، بنابراین مورد اول نادرست است.

نمرات فرمولی که برای تعدیل اثر حدس یا پاسخ ندادن به سؤال به کار می رود، ممکن است بر پایایی یا روایی آزمون اثر مثبت یا منفی داشته یا اصلا اثری نداشته باشد. بنابراین نمی توان به طور قطع ادعا کرد که حدس تصادفی بر پایایی بی تاثیر است. در نتیجه مورد دوم نیز نادرست است. هرچقدر ضریب دشواری، یعنی درصد پاسخ های درست بیشتر باشد سؤال آسانتر و هرچه این ضریب به صفر نزدیک تر باشد، سؤال دشوارتر است. بنابراین درصد پاسخ های درست منعکس کننده سهولت یا آسانی سؤال است و مفهوم دشواری از آن برداشت نمی شود. درن درنتیجه مود مورد سوم نیز اشتباه است. در صورتی که آزمونی فاقد نمره منفی برای پاسخهای اشتباه باشد ولی بااین حال برخی از پاسخ دهندگان از جواب دادن به برخی از سؤالات امتناع کرده باشند؛ درواقع شانس بدست آوردن نمره بهتر را از خود دریغ کرده و به نوعی فرد، باعث مجازات خود شده است.

۲- گزینه «۴» فرمول اسپیرمن - براون به صورت زیر است:

در این فرمول نشان دهنده ضریب پایایی (اعتبار) آزمون است که در این سؤال برابر با 8/0 می باشد و ضریب همبستگی بین دو نیمه آزمون و در این سؤال مجهول است و باید محاسبه گردد. با جایگذاری ضریب پایایی در فرمول، همبستگی بین دو نیمه به صورت زیر محاسبه می شود:

۳- گزینه «۲» نمره اصلاح شده فرد با استفاده از فرمول زیر قابل محاسبه است:

با جایگذاری اطلاعات سؤال در فرمول بالا تعداد پاسخ های صحیح فرد موردنظر برابر با ۱۲ به دست آمد.

۴- گزینه «3» برای محاسبه تعداد حالتهای ایجاد دو نیمه های مختلف از ۸ سؤال موجود باید ترکیب ۸ (تعداد کل سؤالات) از ۴ (تعداد سؤالات هر نیمه) را محاسبه و سپس نتیجه را بر ۲ تقسیم نمود.

5- گزینه «۴» روایی یعنی اینکه یک ابزار تا چه حد چیزی را اندازه می گیرد که برای اندازه گیری آن ساخته شده است. به عبارت دیگر منظور از روایی آن تر است که مطمئن شویم ابزار اندازه گیری برای اندازه گیری متغیر تحقیق کارایی لازم را داشته باشد. روایی همیشه مختص هدف ویژه ای است که ابزار و اندازه گیری برای آن مورد استفاده قرار می گیرد. در این سؤال با توجه به اینکه هدف اندازه گیری دانش فیزیک افراد است ولی از سؤالات به زبان انگلیسی واستفاده شده است در صورتی که زبان آزمودنی ها فارسی است؛ این امکان وجود دارد که افرادی با وجود داشتن دانش فیزیک به دلیل تسلط نداشتن بر زبان رانگلیسی از پاسخ صحیح به سؤالات باز بمانند. لذا آزمون برای اندازه گیری هدفی که بدان منظور تهیه شده است . یعنی اندازه گیری دانش فیزیک . آزمودنی ها - مناسب نبوده و نتیجه اینکه دارای روایی نیست.

6- گزینه «1» پایایی یک آزمون برای کودکان ۶ تا ۱۶ سال برابر با ۸۲/0 است، پایایی این آزمون برای کودکان ۱۰ ساله کوچکتر از 82/0 خواهد بود. ضریب پایایی از جنس ضریب همبستگی است و از طریق ضریب همبستگی محاسبه می گردد. از آنجایی که محدودیت در دامنه و استفاده از گروه آزمودنیهای متجانس تر منجر به کاهش ضریب همبستگی می شود؛ بنابراین با محدود کردن آزمودنی ها به یک سن خاص ضریب پایایی کاهش خواهد یافت.

۷- گزینه «3» با توجه به اینکه پاسخ به سوالات براساس شبانس است و سؤالات مورد نظر دارای ۵ گزینه هستند، احتمال پاسخ گویی صحیح به هر سؤال برابر با است. برای بدست آوردن میانگین نمره در این آزمون باید تعداد سوالات آزمون و ۵ در احتمال پاسخ صحیح ضرب شود. همچنین برای محاسبه واریانس این آزمون باید احتمال پاسخ صحیح به هر سؤال در احتمال پاسخ غلط به آن سؤال ضرب و سپس در تعداد سؤالات آزمون ضرب شود.

۸- گزینه «۲» همیشه پایایی پایین آزمون دلیلی بر کیفیت بد آزمون نیست؛ زیرا این امکان وجود دارد که شرایط بیرونی تنیجه آزمون را تحت تاثیر قرار دهد و پایایی کم برآورد شود با اینکه در صورت عدم تجانس درونی بین سؤالات از روش های همسانی درونی برای محاسبه ضریب پایایی استفاده شود یا اینکه در یک آزمون سرعت از روش دو نیمه کردن برای محاسبه ضریب پایایی استفاده شود. بنابراین در صورت وجود چنین شرایطی ضریب پایایی کم دلیل بر کیفیت بد آزمون نیست این امکان وجود دارد که یک آزمون با وجود بد بودن دارای ضریب پایایی خوبی باشد، مثلا ممکن است یک آزمون که فاقد روایی است و در اندازه گیری متغیر مورد نظر نانوان است از نظر پایایی مطلوب باشد. بنابراین ضریب پایایی به تنهایی دلیل بر خوب بودن آرمون نیست. رنمره یک آزمون ملاک مرجع بعد از یک آموزش اثربخش به میزان قابل توجهی نسبت با نمره قبل از آموزش افزایش می یابد و همین امر موجب می شود که بین دو آزمون قبل و بعد از آموزش تفاوت چشمگیر ایجاد شده و ضریب پایایی آزمون کم برآورد شود.

۹- گزینه «1» نظریه کلاسیک اندازه گیری متضمن یک الگوی جمع پذیر است. نمره مشاهده شده یک آزمون مساوی مجموع دو جز است: نمره واقعی T و نمره خطای تصادفی E . نمره های واقعی و خطا ساخت نظری و غیرقابل مشاهده ای دارند. یعنی X=T+E با توجه به این معادله نمره مشاهده شده برحسب خطا تغییر می کند و خطای اندازه گیری نیز تصادفی بوده و میزان آن متغیر است؛ ولی نمره واقعی همیشه ثابت است. البته نمره واقعی و خطا ساخت نظری داشته و قابل مشاهده نیستند.

۱۰- گزینه «۱» برای حل این سؤال باید از فرمول مقابل کمک گرفت:

در این فرمول برابر با ضریب پایایی مطلوب و ضریب پایایی موجود و n به معنای ضریب افزایش تعداد سؤالات آزمون است در این میزان ضریب پایایی مطلوب برابر است با میزان ضریب پایایی موجود به علاوه 05/0. بعد از محاسبه ضریب افزایش سؤالات با توجه به اطلاعات هر گزینه، باید تعداد سؤالات مورد نیاز برای رسیدن به پایایی مطلوب برآورد گردد.

گزینه ۱: ۱۷ سؤال جدید باید اضافه شود.

گزینه ۲: سؤال جدید باید اضافه شود.

گزینه ۳: ۳سؤال جدید باید اضافه شود

گزینه ۴: ۶سؤال جدید، باید اضافه شود.

آزمونی با ویژگی های ذکر شده در گزینه (۱) به تعداد سؤالات بیشتری نیاز دارد.

۱۱- گزینه «۲» سؤالاتی با ضریب دشواری صفر و یک دارای کمترین ضریب تمییز هستند. زیرا در سؤال با ضریب دشواری صفر هیچ کس نتوانسته است به سؤال موردنظر پاسخ دهد و در سؤال با ضریب دشورای یک همه به سؤال پاسخ صحیح داده اند؛ بنابراین سؤالاتی با این ضرایب دشواری قادر به تفکیک آزمودنی ها نبوده و همه افراد در این سؤالات یکسان عمل کرده اند.

۱۲- گزینه «3» فرمول محاسبه ضریب روایی محتوایی به صورت زیر است، که ne در این فرمول برابر با تعداد داورانی است که سوال را ضروری تشخیص داده اند و N برابر با تعداد کل داوران است.

جواب این سوال برابر با 6/0 است که در بین گزینه ها نیست ولی سازمان سنجش گزینه (۳)، یعنی ضریب محتوایی برابر با 66/0 را به عنوان گزینه درست معرفی کرده است.

۱۳- گزینه «۴» به آزمون شوندگان حق انتخاب چند سؤال از میان تعدادی سؤال را ندهید؛ زیرا :

1- تهیه سؤال هایی که ازنظر دشواری همسنگ باشند کار آسانی نیست.

2- آزمون شوندگان، توانایی انتخاب سؤال هایی که جواب آن ها را بهتر می دانند، ندارند.

٣- دادن حق انتخاب، مشکل نمونه گیری را در آزمون تشریحی دشوارتر می کند.

۴- چون آزمون شوندگان قوی تر، بیشتر به سمت سؤال های دشوار می روند. داشتن حق انتخاب به ضرر این افراد تمام می شود. با این حال در مواردی که هدف سؤالات اندازه گیری میزان یادگیری نیست و صرفا جنبه اندازه گیری مهارت پاسخ دهندگان در نوشتن و پروراندن مطالب را دارند، دادن حق انتخاب سؤال به افراد بلامانع است.

۱۴- گزینه «۲» بلوم اهداف آموزشی را به سه حوزه شناختی، عاطفی و روانی - حرکتی تقسیم کرده است، حوزه عاطفی شامل هدف هایی است که تغییرات حاصل در علایق، نگرش ها و ارزش ها و نیز رشد ارج شناسی و سازگاری را نشان می دهد. این حوزه به ۵ مرحله دریافت کردن (توجه کردن)، پاسخ دادن ارزش دادن، سازمان دادن به ارزش ها و شخصیت پذیرفتن تقسیم می گردد. در اینجا فرد موردنظر به این نتیجه می رسد که مهارت مقاله نویسی ارزشمند است؛ بنابراین و به هدف حوزه عاطفی دست یافته است.

۱۵- گزینه «۲» تمره معادل منحنی بهنجار (NCE) در آزمون های میزان شده پیشرفت تحصیل مورد استفاده قرار می گیرد و دارای میانگین ۵۰ و انحراف استاندارد 06/21 است. دلیل انتخاب این میزان از انحراف استاندارد آن است که این نمرات دارای دامنه ای از ۱ تا ۹۹ باشند. در واقع دامنه ی این نمرات با دامنه ی رتبه های درصدی یکسان است با این تفاوت که نمرات NCE در دو انتها بازتر و در میانه توزیع جمع تر هستند.

۱۶- گزینه «۱» براساس نمره کلاسیک آزمون نمره ای که فرد در آزمون می گیرد با مره واقعی یا توانایی واقعی او متفاوت است. تفاوت بین نمره واقعی و نمره مشاهده شده فرد از خطای اندازه گیری ناشی می شود. بنابراین نمره مشاهده شده فرد از جمع نمره واقعی و خطای اندازه گیری ( X=T+E) حاصل می شود. باتوجه به اینکه در این سؤال خطای اندازه گیری برابر با صفر در نظر گرفته شده است؛ نمره مشاهده شده آزمودنی (نمره کسب شده در امتحان) برابر با تمره واقعی او است.

۱۷- گزینه «۴» آزمون های وابسته به ملاک که در سنجش یادگیری در حد تسلط مورد استفاده قرار می گیرند باید سؤالاتی را شامل شوند که پس از آموزش مطالب درسی آسان تر از قبل از آموزش آن مطلب جلوه کنند. فرمول ضریب تمییز پیش و پس آزمون، یا ضریب حساسیت نسبت به آموزش به صورت زیر است، که در این فرمول برابر با تعداد کسانی در پس آزمون و برابر با تعداد کسانی که در پیش آزمون پاسخ درست داده اند و T برابر با تعداد کل افراد است. هرچقدر این ضریب محاسبه شده به ۱+ نزدیک تر باشد سؤال مطلوب تر است.

۱۸- گزینه «1» طبق الگوی نمونه گیری حیطه، هر سؤال آزمون نمونه مستقلی از خصیصه یا توانایی مورد اندازه گیری است. هر اندازه تعداد سؤال ها بیشتر باشد، احتمال اینکه آزمون معرف واقعی خصیصه مورد اندازه گیری باشد افزایش می یابد. بنابراین با به کار بستن الگوی نمونه گیری حیطه با افزایش تعداد سؤال های آزمون اعتبار آن نیز افزایش می یابد.

۱۹- گزینه «3» هدف از بررسی هنجارها برای یک آزمون، تعیین چگونگی مقایسه یک آزمودنی با دیگران است. یک آزمون وابسته به هنجار هر فردی را با یک هنجار مقایسه می کند. این نوع از ارزشیابی به یک ملاک نسبی وابسته است. در این سؤال مورد اول و مورد چهارم به آزمون هنجار مرجع اشاره دارند. در مورد اول ۱۵ نفر با نمره بالاتر مدنظر است یعنی مهم نیست افراد چه نمره ای کسب کرده اند مهم آن است که در مقایسه با سایر افراد جزء ۱۵ نفر اول باشند. در مورد چهارم نیز نمره هر فرد با میانگین گروه مقایسه شده و انتخاب فرد تحت تأثیر میانگین گروه قرار دارد. مورد دوم و سوم به آزمون ملاک مرجع اشاره دارند؛ زیرا ملاک قضاوت، عملکرد خود فرد است.

20- گزینه «۴» همبستگی واقعی بین دو مقیاس اندازه گیری از طریق تقسیم همبستگی مشاهده شده بین دو مقیاس بر ریشه دوم حاصل ضرب پایایی هریک از دو مقیاس به دست می آید. این فرمول که به تصحیح کاهش معروف است به صورت زیر قابل محاسبه است:

: همبستگی واقعی بین دو مقیاس

: همبستگی مشاهده شده بین دو مقیاس

: پایایی هریک از دو مقیاس (در اینجا چون وزن با ترازوی بدون خطا اندازه گیری شده پایایی آن معادل یک است).

سوالات آزمون سراسری 97

۱ـ کدام مورد درباره ارتباط بین خطاهای تصادفی و ثابت با پایایی و روایی درست است؟

۱) افزایش خطای ثابت، روایی اندازه گیری را کاهش می دهد.

۲) افزایش خطای ثابت، پایایی و روایی اندازه گیری را کاهش می دهد.

۳) افزایش خطای تصادفی، فقط پایایی اندازه گیری را کاهش می دهد.

۴) افزایش خطای تصادفی، فقط روایی اندازه گیری را کاهش می دهد.

۲- کدام عبارت در مورد رابطه نمره مشاهده شده (x)، خطا (e) واقعی (t) و پایایی () نادرست است؟

1)

2)

3)

4)

۳ـ درصورتی که در یک آزمون ۴۰ سؤالی همبستگی بین نمره های مشاهده شده و نمره های خطا برابر با باشد، ضریب پایایی آزمـون کدام است؟

1) 25/0

2) 5/0

3) 75/0

4) 80/0

۴- کدام مورد جزء پیش فرض های مربوط به همبستگی دورشته ای () نیست؟

۱) توزیع هر دو متغیر نرمال است.

۲) یکی از متغیرها دو، ارزشی واقعی است.

۳) هر دو متغیر دارای مقیاس پیوسته اند.

۴) یکی از متغیرها براساس آستانه خاصی، افراد را به دو گروه تقسیم می کند.

5- با استفاده از کدام روش، می توان روایی همگرا و واگرا را به طور همزمان بررسی کرد؟

1) بررسی تفاوت های گروهی

۲) تحلیل عاملی

۳) تیلور - راسل

۴) چندروشی - چندصفتی

۶- درصورتی که بین نمرات فرم کوتاه آزمون MMPI با فرم بلند این آزمون همبستگی وجود داشته باشد، بیانگر کدام نوع روایی است؟

۱) هم زمان

۲) محتوایی

۳) پیش بین

۴) تشخیصی

۷ـ معمولاً از بین سؤال های موجود در نمونه اولیه یک آزمون، سؤال هایی برای آزمون نهایی انتخاب می شوند که براساس داده های تجربی دارای ضرایب تمییز بالاتری باشند، با اجرای سؤال های انتخاب شده روی یک نمونه جدید میزان ضریب تمییز چه تغییری میکند؟

۱) به خاطر کاهش تعداد سؤالات میزان ضرایب تمییز بیشتر میشود.

۲) به علت انتخاب سؤال های بهتر ضرایب تمییز کاهش می یابد.

۳) به دلیل اجرای دوباره میزان ضرایب تمییز بیشتر می شود.

۴) به دلیل رگرسیون آماری میزان ضرایب تمییز کمتر می شود.

۸- کدام مورد درباره روایی صوری آزمون، نادرست است؟

1) در مورد رد همه آزمون ها الزامی نیست.

2) معمولاً به صورت ضریب همبستگی بیان می شود.

۳) براساس قضاوت متخصصان تعیین می شود.

۴) گاهی اوقات روایی ذهنی نیز نامیده می شود.

۹- در تحلیل سؤال های آزمون کدام مورد با تغییر نمونه، کمتر دستخوش تغییر می شود؟

1) دشواری سؤال

۲) ضریب تمییز سؤال

۳) منحنی مشخصه سؤال

4) همبستگی سوال با نمره کل

۱۰- بررسی مجدد ویژگی های آماری یک آزمون براساس نمونه جدیدی از آزمودنی ها چه نام دارد؟

1) بررسی انطباقی

۲) تحلیل سازه

۳) تحلیل عاملی

۴) وارسی روایی

۱۱ـ در یک آزمون استاندارد نمرات خام با استفاده از معادله ۴- x۵ = y به نمره جدید تبدیل میشود. درصورتی که نمره تبدیل شـده فـردی ۷۱ باشد، نمره خام وی کدام است؟

1) 15

2) 67

3) 75

4) 351

۱۲- فرض کنید میانگین توزیع نمره های پیشرفت تحصیلی درس فیزیک نرمال با میانگین ۵۰ و انحراف معیار ۲۰ است. به ازای کدام نمـره، رتبـه درصدی و احتمال پاسخ درست به سؤال i برابر ۵۰/۰ خواهد بود؟

۴) همبستگی سؤال با نمره کل

1) 40

2) 50

3) 60

4) 70

۱۳ـ در کدام مقیاس فرض بر این است که اگر آزمودنی به یک سؤال با درجه دشواری معینی پاسخ درست بدهد، همین آزمودنی به کلیه سؤال هـای ساده تر نیز پاسخ درست خواهد داد؟

۱) افتراق معنایی

۲) ترستون

۳) گاتمن

4) لیکرت

۱۴ـ مطابق با نظریه کلاسیک اندازه گیری با افزایش تعداد سؤالات آزمون، کدام مورد درست است؟

1) واریانس نمره واقعی کاهش و واریانس نمره خطا افزایش می بابد.

۲) واریانس نمره واقعی بیشتر از واریانس نمره خطا افزای خطا افزایش می یابد.

۳) واریانس نمره واقعی افزایش و واریانس نمره خطا کاهش می یابد.

۴) واریانس نمره واقعی و واریانس نمره خطا کاهش می یابد.

۱۵ـ کارپوشه هایی که با هدف آموزش به یادگیرندگان مورد استفاده قرار می گیرند، بر چه چیزی تأکید دارند؟

۱) ارزشیابی از خود

۲) انتقال آموخته ها

۳) راهبردهای یادگیری

۴) فرایندهای عالی حافظه

۱۶ـ در آزمون ورودی یک مؤسسه آموزشی افرادی پذیرفته میشوند که حداقل به ۷۰ درصد از سؤالات آزمون با احتساب نمره منفی پاسخ صحیح بدهند، این آزمون از کدام نوع است؟

۱) نسبی

۲) تکوینی

۳) ملاکی

۴) هنجاری

۱۷- کدام شیوه ارزشیابی برای تعیین این که «آیا دانش آموز مهارت های پیش نیاز را می داند؟»، مناسب است؟

۱) تراکمی

۲) تشخیصی

3) تکوینی

۴) ملاکی

۱۸- کدام مورد درباره مقیاس اندازه گیری رتبه ای درست است؟

۱) ویژگی بزرگی، فواصل مساوی صفر مطلق را دارد.

۲) ویژگی بزرگی، فواصل مساوی و صفر مطلق را ندارد.

3) ویژگی بزرگی و فواصل مساوی را دارد ولی صفر مطلق را ندارد.

۴) ویژگی بزرگی را دارد ولی فواصل مساوی و صفر مطلق را ندارد.

۱۹- این هدف که «دانش آموز باید بتواند با آموخته هایش به محاسبه میزان گنجایش یک حوض آب در مقیاس لیتر و مترمکعب بپردازد»، براساس طبقه بندی شناختی بلوم در چه سطحی قرار دارد؟

1) کاربرد

2) فهمیدن

۳) ترکیب

۴) ارزشیابی

۲۰- در تحلیل سوالات چهار گزینه ای، کدام مورد نشان دهنده عملکرد مناسب یک گزینه انحرافی است؟

1) همبستگی مثبت با نمره کلن آزمون

۲) نسبت انتخاب بیشتر توسط گروه قوی

۳) نسبت انتخاب کمتر توسط گروه ضعیف

۴) همبستگی منفی با نمره کل آزمون

پاسخنامه آزمون سراسری 97

1- گزینه «1» روایی آزمون برخلاف پایایی آن، هم تحت تأثیر خطای نظام دار (ثابت) و هم تحت تأثیر خطای غیرنظام دار (تصادفی) است؛ أما پایایی فقط تحت تأثیر خطای غیرنظام دار (تصادفی) است. بنابراین یک آزمون ممکن است در عیر داشتن پایایی فاقد روایی باشد ولی عکس این قضیه صادق نیست؛ یعنی پایایی شرط لازم روایی است. اما پایایی به تنهایی روایی آزمون را تضمین نمی کند.

٢- گزینه «3» گزینه (۱) از نتایج استخراج شده نظریه کلاسیک نمره واقعی است که نشان می دهد در آزمون های موازی واریانس خطا مساوی است با حاصل ضرب واریانس نمرات مشاهده شده در یک منهای همبستگی بین نمرات مشاهده شده (پایایی آزمون).

گزینه (۲) نشان دهنده رابطه بین پایایی یک آزمون با همبستگی نمرات خطا و نمرات مشاهده شده است. معادله نشان داده شده در گزینه (۳) نادرست است؛ در این معادله باید به جای واریانس نمرات مشاهده شده و خطا از انحراف استاندارد آن ها استفاده کرد. گزینه (۴) نشان دهنده این نکته است که مجذور همبستگی بین نمرات مشاهده شده و واقعی برابر نسبت واریانس نمرات واقعی به واریانس نمرات مشاهده شده است. این رابطه از نتایج نظریه کلاسیک نمره واقعی است.

3- گزینه «3» همبستگی بین نمرات خطا و نمرات مشاهده شده مساویاین است. همچنین برابر با پایایی آزمون است.

۴- گزینه «۲» ضریب همبستگی در رشته ای زمانی استفاده می شود که یکی از دو متغیر پیوسته و دیگری دو ارزشی ساختگی باشد. منظور از دو ارزشی ساختگی این است که اعداد اصلی پیوسته هستند نه دو ارزشی، اما موقع محاسبه آن ها به دو ارزشی تبدیل می شوند. شرط دیگر استفاده از این روش بهنجار بودن توزیع نمرات هر دو متغیر است.

5- گزینه «۴» روایی چند روشی - چند صفتی جنبه ای از روایی سازه است. این روش هنگامی به کار برده می شود که دو یا چند صفت با دو یا چندر روش اندازه گیری می شوند. ماتریس چند روشی - چند صفتی یک ماتریس همبستگی است که به جای ضرایب یک بر روی قطر اصلی پایایی های برآورد شده جایگزین شده اند. روایی همگرا از طریق همبستگی های بالا بین نمرات آزمون هایی نشان داده می شود که به وسیله روش های مختلف ویژگی های یکسانی را اندازه گیری می کنند. روایی واگرا نیز از طریق همبستگی های پایین بین نمرات آزمون هایی که ویژگی های مختلفی دارند (به ویژه زمانی که روش های یکسانی به کار برده می شوند) نشان داده میشود.

6 - گزینه «۱» محاسبه همبستگی بین فرم کوتاه و فرم بلند آزمون MMPI نشان دهنده روایی همزمان است. ضرایب روایی همزمان هنگامی که نمرات آزمون برای برآورد ملاک جاری و نه پیش بینی آینده به کار برده شود مناسب است. در اینجا آزمون کوتاه و به عنوان یک پیش بینی کننده و فرم بلند به عنوان یک ملاک دارای روایی همزمان هستند.

۷- گزینه «۴» رگرسیون آماری زمانی اتفاق می افتد که نمرات بسیار بالا با بسیار پایین انتخاب شوند و مورد بررسی دوباره قرار گیرند. در این حالت در اندازه گیری مجدد نمرات به معدل کل گروه بازگشت می کنند. بنابراین اگر در ساخت آزمون سؤالات با ضریب تمییز بالا انتخاب شوند و در گروه جدیدی مورد بررسی قرار گیرند به دلیل رگرسیون آماری ضریب تمییز سؤالات در گروه جدید کمتر برآورد خواهد شد.

۸- گزینه «۲» روایی صوری که گاهی روایی ذهنی نیز نامیده می شود؛ هنگامی برقرار است که فرد آزمون را بررسی کند و نتیجه بگیرد که این آزمون صفت مورد نظر را اندازه گیری می کند. روایی صوری نوعی روایی محتوایی است و روش تعیین آن نیز براساس قضاوت متخصصان می باشد. گاهی مثل تشخیص بیماری های روانی استفاده از روایی صوری مجاز نیست.

۹- گزینه «3» محلی ویرگی سوال نمایش ترسیمی از رابطه بین احتمال پاسخگویی به یک سؤال معین و وضعیت آزمودنی در رفتار مورد اندازه گیری آزمون است. منحنی ویژگی سؤال را می توان برای شناسایی سؤال هایی در گروه های مختلف آزمودنی که دارای عملکرد متفاوتی هستند به کار برد. منحنی ویژگی سؤال با تغییر نمونه نسبت به سایر شاخص ها کمتر دچار تغییر می شود؛ اما سؤال هایی که منحنی ویژگی آنها به صورت قابل توجهی تحت تأثیر اعضای گروه ها قرار می گیرند باید اصلاح یا حذف شوند.

۱۰- گزینه «۴» بررسی مجدد ویژگی های آزمون با نمونه جدید وارسی روایی نام دارد. در واقع روایی و پایایی یک آزمون به شیوه هایی که قابل پیش بینی است تحت تأثیر تغییرات ناهمگنی گروه یا گزینش منظم نمره هاست. در وارسی روایی معادله رگرسیونی که قبلا تعیین شده ایست برای نمونه جدیدی از آزمودنی ها به کار برده می شود و همبستگی چندمتغیری جدیدی برای این نمونه محاسبه می شود.

۱۱- گزینه «۱» در یک آزمون نمره خام هر فرد با کمک معادله به نمره جدید تبدیل می شود. در صورتی که نمره تبدیل شده فردی برابر با ۷۱ (یعنی ۷۱ = y) باشد نمره خام او برابر با ۱۵ بوده است.

۱۲- گزینه «۲» با توجه به این نکته که توزیع نمرات در این آزمون نرمال است منحنی طبیعی در مورد نمرات آن صادق است. لذا با مراجعه به جدول منحنی نرمال نمره استاندارد با احتمال پاسخ درست 5/0 مساوی با صفر است. با جایگذاری این نمره در فرمول زیر نمره با احتمال پاسخ درست 5/0 به دست می آید.

بنابراین به ازای نمره ۵۰ رتبه درصدی و احثمال پاسخ به سؤال i برابر با 5/0است.

گویه ها برحسب دشواری، پبحیدگی،

۱۳- گزینه «3» مقیاس گاتمن شامل مجموعه ای از گویه ها برای سنجش نگرش های تک بعدی است. در این مقیان یا ارزش وزنی مرتب میشوند و موافقت با تأیید یک گویه موافقت با سایر گویه های کم وزن تر را به دنبال دارد.

۱۴- گزینه «۲» براساس نظریه کلاسیک اندازه گیری هرچه تعداد سؤالات آزمون افزایش یابد، پایایی آزمون نیز بیستر می شود. از طرف دیگر پایایی برابر است با واریانس نمرات واقعی بر واریانس نمرات مشاهده شده. با توجه به اینکه واریانس نمرات مشاهده شده از مجموع واریانس نمرات واقعی و نمرات خطا حاصل می شود؛ می توان گفت پایایی برابر است با واریانس نمرات واقعی تقسیم بر مجموع واریانس نمرات واقعی و واریانس نمرا خطا. درنتیجه هرچه واریانس نمرات واقعی با سرعت بیشتری نسبت به واریانس خطا افزایش یابد پایایی بیشتر خواهد شد. بنابر موارد ذکر شده گزینه (۲) پاسخ صحیح است.

۱۵- گزینه «۱» کارپوشه نوعی از آزمون های عملکردی است که شامل مجموعه ای از کارهای یک دانش آموز می شود. در کارپوشه یاناگیرنده برای نشان دادن توانایی ها و پیشرفت خود نمونه هایی از کارهایی را که در طول یک مدت معین انجام داده است با خود به کلاس می آورد و برای ارزشیابی در اختیار معلم قرار می دهد. مزیت مهم کارپوشه آن است که یادگیرندگان با اندیشیدن درباره نمونه کارهای خود، ارزشیابی از خود را می آموزند.

۱۶- گزینه «3» آزمون ها از جنبه های مختلف قابل طبقه بندی هستند؛ یکی از این جنبه ها ملاک داوری درباره نمرات آزمودنی ها است بر این اساس آزمون ها به دو دسته ملاکی و هنجاری دسته بندی می شوند. تفاوت آزمون ملاکی و هنجاری در معیاری است که برای قضاوت درباره آزمودنی ها در نظر گرفته می شود. در آزمون ملاکی، عملکرد هر آزمودنی با یک ملاک مشخص و بدون توجه به عملکرد سایر افراد گروه مقایسه و درباره او قضاوت می شود؟ ولی در آزمون های هنجاری عملکرد هر آزمودنی با متوسط عملکرد افراد گروه نرم مثلا گروه همسالان مقایسه و تفسیر می شود. در این مثال ملاک انتخاب فرد پاسخگویی به ۷۰ درصد سؤالات است و ربطی به پاسخگویی سایر افراد ندارد؛ بنابراین آزمون مورد نظر ملاکی است.

۱۷- گزینه «۲» ارزشیابی تشخیصی با هدف تشخیص مشکلات یادگیری دانش آموزان صورت می گیرد. یک آزمون تشخیصی خوب نه تنها معلم را از نقاط ضعف و قوت دانش آموزان آگاه می کند، بلکه زمینه های ضعف و قوت او را نیز آشکار می کند. همچنین از این ارزشیابی برای اطلاع از توانایی آزمودنی برای شروع نوع خاصی از یادگیری استفاده می شود.

۱۸- گزینه «۴» مقیاس اندازه گیری رتبه ای افراد یا اشیاء را از لحاظ صفت ویژهای رتبه بندی می کند و در آن به تعداد افراد رتبه وجود دارد؛ در واقع مقیاس رتبه ای دارای ویژگی بزرگی است. این مقیاس قادر نیست فاصله تفاوت بین افراد و مشخص کند. در واقع در این مقیاس نمرات و شماره های رتبه بندی فاقد معنای کمی و عددی است. علت این امر نبودن مبدأ خاصی است تا ارزش ها (صفر مطلق) نسبت به آن سنجیده شوند.

۱۹- گزینه «۱» بلوم هدف های حوزه شناختی را به شش طبقه کلی دانش، فهمیدن، کاربرد تجزیه و تحلیل، ترکیب و ارزشیابی تقسیم کرده است. کاربرد به توانایی استفاده از امور انتزاعی، قواعد و قوانین اصول اندیشه ها و روس ها در موقعیت های عینی و عملی اشاره دارد. بنابراین وقتی یک دانش آموز بتواند با آموخته هایش در موقعیت واقعی گنجایش یک حوض را حساب کند در سطح کاربرد است.

۲۰- گزینه «۴» در تحلیل گزینه های انحرافی، قاعده کلی این است که هر گزینه انحرافی باید دست کم یک نفر از افراد گروه ضعیف را به خود جلب کند و اگر هر دو گروه ضعیف و قوی را جلب کرد باید تعداد افراد گروه ضعیف بیشتر از گروه قوی باشد. در واقع باید بین گزینه انحرافی و نمره آزمون همبستگی منفی وجود داشته باشد؛ یعنی هرچه نمره بیشتر است انتخاب گزینه انحرافی کمتر باشد و بالعکس.

سؤالات آزمون دکتری ۹۸

۱- اگر واریانس خطای اندازه گیری در آزمون a و b به ترتیب ۲۵ و ۳۶ باشد، تفاوت بین نمره های یک فرد در این دو آزمون، حداقل چقدر باید باشد تا بتوان آن را در سطح 05/0 معنی دار تلقی کرد؟

1) 81/12

2) 30/15

3) 20/18

4) 15/20

۲- در کدام نوع از سؤال های جور کردنی، تعداد پرسش ها می تواند بیشتر از تعداد پاسخ ها باشد؟

۱) خبری

۲) معمولی

۳) استفهامی

۴) رده بندی

۳- در برآورد اعتبار (Reliability) به روش دو نیمه سازی، اگر واریانس دو نیمه آزمون همگن نباشد، از کدام فرمول باید استفاده کرد؟

۱) رولون

۲) لاندای 2

3) آلفای کرونباخ

۴) کودر - ریچاردسون ۲۰

۴- اگر واریانس یک سؤال دو ارزشی ۲۴/0 باشد، ضریب دشواری آن کدام است؟

1) 20/0

2) 45/0

3) 60/0

4) 70/

5- نمره استاندارد (z) دانش آموزی ۲ است، اگر نمره خام این دانش آموز ۴۴ و میانگین گلاس۴۰ باشد، انحراف استاندارد نمرات کلاس چند است؟

1) 1

2) 5/1

3) 2

4) 5/2

6- اگر در یک آزمون ارزش میانگین برابر با ۳۴، واریانس ۹ و پایایی ۷۵/0 باشد، برای یک آزمودنی با نمره ۳۰، دامنه 5/28 تا 5/31 با چند درصد احتمال محاسبه شده است؟

1) 68

2) 90

3) 95

4) 99

۷- در طبقه بندی بلوم، هدف زیر در کدام سطح فرآیند شناختی قرار دارد؟

«تفسیر خروجی نرم افزار مربوط به پایایی یک آزمون»

1) فهمیدن

۲) کاریستن

۳) ارزشیابی کردن

۴) تجزیه و تحلیل کردن

8- اگر آزمونی طراحی شود که در نحوه اجرا و نمروگذاری دارای انعطاف زیادی باشد، به کدام ملاک آزمون توجه نشده است؟

1) اعتبار

۲) عینیت

۳) شرایط تراز شده

۴) داده های هنجار

۹- اگر مجموع واریانس های سؤال های یک آزمون با واریانس کل آزمون یکسان باشد، ضریب آلفا چقدر است؟

1) 1

2) 5/0

3) 0

۴) قابل محاسبه نیست

۱۰- کدام نوع روایی، بیانگر میزان روشن بودن هدف آزمون هم برای آزمونگر و هم آزمون شونده است؟

1) سازه

۲) محتوایی

3) همزمان

4) صوری

۱۱- کدام مورد در تبدیل خطی نمرات آزمون تغییر می کند؟

1) کجی و کشیدگی

۲) پایایی و روایی

۳) جایگاه افراد روی توزیع

۴) میانگین و انحراف استاندارد

۱۲- شواهد تجربی و روابط نظری که از صحت استنباط ها، تفسیرها یا اعمالی که بر اساس نمره های آزمون صورت می گیرد، حمایت می کند، بیانگر کدام مورد است؟

۱) دقت

۲) روایی

۳) پایایی

۴) همگنی

۱۳- ضریب آلفای کرونباخ یک آزمون ۶ سؤالی برابر با 9/0 است اگر واریانس نمره کل آزمون ۳۰ باشد، مجموع واریانس سؤالات آزمون کدام است؟

1) 5/7

2) 30

3) 5/60

4) 180

۱۴- کدام مورد درباره پایایی بازآزمایی، درست است؟

۱) با کاهش خطای منظم، ضریب پایایی از آزمایی افزایش می یابد.

۲) در این روش، آزمون یک بار بر روی دو گروه از افراد اجرا می شود.

۳) با افزایش فاصله زمانی بین دو بار اجرا، ضریب پایایی بازآزمایی کاهش می یابد.

۴) برای محاسبه ضریب پایایی بازآزمایی، از ضریب اسپیرمن - براون استفاده می شود.

۱۵- شکل توزیع در کدام تبدیل، مستطیلی است؟

۱) نمرات نرمال شده

۲) رتبه درصدی

۳) نمرات Z

4) نمرات بهره هوشی

۱۶- اگر نمره فردی در آزمون یک انحراف معیار پایین تر از میانگین باشد، نمره T او کدام است؟

1) 10

2) 30

3) 40

4) 50

۱۷- در سرشماری سال ۱۳۸۵ از مردم پرسیده اند که «در خانه به چه زبانی (فارسی، ترکی، کردی، لری، سایر زبان ها صحبت می کنند؟»، نتایج به دست آمده در کدام مقیاس قرار می گیرد؟

1) نسبی

3) اسمی

3) ترتیبی

۴) فاصله ای

۱۸- در کدام مورد، فرآیند انجام کار مورد ارزشیابی قرار گرفته است؟

1) از دانش آموز می خواهیم درباره امید انشا بنویسد و آن را ارزشیابی می کنیم

2) از هنرجو می خواهیم یک مدار الکتریکی دو پل بسازد و به آن امتیاز می دهیم.

۳) از هنرجو می خواهیم تابلوی یک منظره را نقاشی کند و به آن امتیاز میدهیم.

۴) از دانشجوی پرستاری می خواهیم فشار خون بیمار را اندازه بگیرد و به آن امتیاز می دهیم.

۱۹- ۲۰۰ نفر دانشجو به یک سؤال چندگزینه ای درس روان سنجی پاسخ داده اند. برای تحلیل سؤال های این آزمون برگه های ۲۷ درصد افراد گروه های قوی و ضعیف برای تحلیل انتخاب شده و بقیه افراد کنار گذاشته شده اند. اگر ۳۷ نفر از افراد گروه قوی و 10 نفر از افراد گروه ضعیف به سؤال مورد نظر پاسخ درست داده باشند، ضریب تمیز سؤال کدام است؟

1) 4/0

2) 45/0

3) 5/0

4) 55/0

۲۰- کدام روش، تأثیر انتخاب سؤال های خاص برای هر نیمه آزمون را در محاسبه پایایی دو نیمه کردن خنثی می کند؟

1) رولن

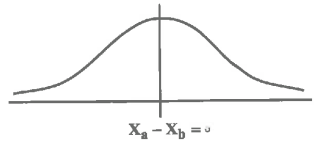
2) 20 - KR

۳) اسپیرمن - براون

۴) آلفای کرونباخ

پاسخنامه آزمون دکتری ۹۸

1- گزینه «۲» فرض می کنیم که با به دست آوردن اختلاف نمره تک تک افراد در این دو آزمون، توزیع نرمالی شکل گرفته است.



یعنی توزیع برای افراد شکل گرفته است:

نمره فرد در آزمون b نمره فرد در آزمون a

(با فرض اینکه میانگین ( ) ها در جامعه است.)

با فرض اینکه نمرات در آزمون a و b از یکدیگر مستقل هستند می توان گفت که:

برای معناداری 05/0(یا همان فاصله اطمینان ۹۵%)، مقدار Z برابر است با ۹۶ / ۱:

2- گزینه «۴» آزمون های جور کردنی، رده بندی با فهرست اصلی از تعدادی سؤال تشکیل شده که هریک از آنها شامل دو ستون است. یک ستون معرف پرسش ها و ستون دیگر نشان دهنده پاسخ هاست. در این نوع از سؤالات تعداد پاسخ ها غالبا" بیشتر از تعداد پرسش هاست.

۳- گزینه «۱» روش رولن (۱۹۳۹) مستلزم تقسیم آزمون کل به دو نیمه آزمون است، ولی متکی بر فرض تساوی واریانس های دو نیمه آزمون نیست. رولون معتقد است که واریانس تفاوت نمره های دو نیمه آزمون به طور کلی ناشی از واریانس خطای دو نیمه آزمون است که روی هم واریانس خطای آزمون کل را تشکیل می دهند.

۴- گزینه «3» واریانس سؤال برابر است با دشواری سؤال ضرب در یک من های دشواری سؤال.

در اینجا با یک معادله درجه دو روبرو هستیم. برای حل این معادله می توان از روش دلتا کمک گرفت.

با توجه به اینکه فقط در بین گزینه ها است، گزینه (۳) پاسخ صحیح می باشد.

5- گزینه «3» با استفاده از فرمول نمره استاندارد، محاسبه انحراف استاندارد به شکل زیر انجام می شود.

6-گزینه «1» برای حل این سؤال ابتدا باید Z را محاسبه و براساس Z احتمال مورد نظر را بدست آورد. برای بدست آوردن تمره Z به ترتیب زیر عمل می کنیم.

مرحله ۱:

مرحله ۲:

d برابر است با Z ضرب در خطای استاندارد و خطای استاندارد برابر است با انحراف استاندارد ضرب در جذر ۱ منهای پایایی آزمون.

مرحله ۲: به جدول Z مراجعه کرده و احتمال متناظر با 2 برابر با یک را پیدا می کنیم.

۷- گزینه «1» بلوم هدف های حوزه شناختی را در شش طبقه کلی دانش، فهمیدن، به کار بستن، تحلیل، ترکیب. و ارزشیایی تقسیم کرده است. فهمیدن یعنی توانایی درک منظور پا مقصود یک مطلب. فهمیدن به پاسخ هایی بیشتر از آنچه قبلا تمرین و آموخته شده اند نیاز دارد. این پاسخ های اضافی عبارت اند از: ترجمه یا برگردان، تفسیر و تعبیر، خلاصه کردن و کشف شباهت ها و تفاوت ها.

۸- گزینه «۲» دسته بندی آزمون ها به چند طریق امکان پذیر است، یک دسته بندی براساس میزان عینیت و چگونگی تهیه آزمون است که بر این اساس آزمون ها به دو دسته آزمون های استاندارد شده و معلم ساخته تقسیم می شود. در آزمون های استانداردشده به وجود روش های یکسان اجرا و نمره گذاری تأکید شده است. در این آزمون ها تحت هر شرایطی باید آزمون طبق روش های مشخص اجرا و نتایج آن نیز طبق ضوابط از پیش تعیین شده تفسیر گردند.

9- گزینه «3» فرمول ضریب آلفا به صورت زیر است:

:n تعداد سؤالات

: واریانس کل آزمون

واریانس هریک از سؤالات آزمون

بنابراین اگر مجموع واریانس سؤالات آزمون با واریانس کلی برابر باشد، ضریب آلفا برابر با صفر خواهد بود.

۱۰- گزینه «۴» روایی محتوا به دو دسته روایی صوری و منطقی قابل تقسیم است. روایی صوری که گاهی اوقات روایی ذهنی نیز نامیده می شود، هنگامی برقرار است که اگر فردی آزمون را بررسی کند نتیجه بگیرد که این آزمون صفت موردنظر را اندازه گیری می کند. فردی که این آزمون را بررسی می کند، می تواند یک آزمودنی یا یک متخصص آزمون باشد؛ آنچه مهم است میزان روشن بودن هدف آزمون برای همه است.

۱۱- گزینه «۴» تبدیل خطی از طریق معادله خطی از نوع Y=aX+b تعریف می شود. در تبدیل خطی، شکل توزیع نمرات تبدیل شده همانند شکل نمرات خام است؛ یعنی کجی و کشیدگی تغییر نمی کند. همچنین با این تبدیل، تغییری در رتبه آزمودنی ها و پایایی و روایی آزمون ایجاد نمی شود. در واقع در تبدیل خطی تنها میانگین و انحراف استاندارد نمرات آزمون تغییر می کند.

۱۲- گزینه «۲» روایی ماهیت مادی ندارد و از جمله ویژگی های ماده هم نیست؛ بلکه روایی صفتی است در خصوص ادعاهایی که مطرح می شود. در واقع روایی از شواهد تجربی و روابط نظری که از صحت استنباط ها، تفسیرها یا اعمالی که براساس نمرات آزمون صورت می گیرد، حمایت می کند.

۱۳- گزینه «۱» فرمول ضریب آلفا به صورت زیر است که با جایگذاری اعداد داده شده در صورت نیؤال، مجموع واریانس سؤالات به صورت زیر قابل محاسبه است:

۱۴- گزینه «3» در پایایی، بازآزمایی یک آزمون در دو نوبت برای آزمودنی های یکسان اجرا می شود و سپس همبستگی نتایج محاسبه می گردد. مسئله ای کم در پایایی با استفاده از روش بازآزمایی وجود دارد، مدت زمان بین دو اجرای آزمون است. اگر مدت زمان بین دو آزمون کوتاه باشد، امکان انتقال اثر وجود دارد، به طوری که آزمودنی پاسخ های آزمون قبلی را به یاد آورده و در این آزمون نیز همان پاسخ ها را بدهد. در این صورت ضریب پایایی به دست آمده بالاتر از ضریب پایایی واقعی خواهد بود. بنابراین با افزایش فاصله زمانی می توان ضریب پایایی برآورد شده را کاهش داد.

۱۵- گزینه «۲» رتبه درصدی مقدار یک ویژگی عبارت است از درصد افراد گروه هنجار که مقدار ویژگی های آن ها مساوی یا کوچکتر از مقدار معینی از ویژگی است. توزیع درصدها در گروه هنجار به صورت مستطیلی است نه به صورت طبیعی.

۱۶- گزینه «3» نمره Z فردی که نمره اش یک انحراف معیار پایین تر از میانگین است، برابر با ۱- است. بنابراین با قرار دادن نمره 2 در فرمول T می توان نمره T او را حساب کرد.

۱۷- گزینه «۲» در مقیاس اسمی از اعداد و ارقام برای طبقه بندی اشیاء و افراد استفاده می شود؛ اما این اعداد و ارقام هیچ گونه معنا و مفهوم ریاضی ندارند و و تنها برای مشخص کردن طبقه های مختلف به کار می روند. با توجه به اینکه در بین طبقات مقیاس اسمی، از نظر رتبه هیچ تفاوتی وجود ندارد، در این سؤال نوع زبانی که افراد با آن صحبت می کنند یک مقیاس اسمی محسوب می شود.

۱۸- گزینه «۴» یکی از رویکردهای ارزشیابی، رویکرد مبتنی بر مدیریت است که در چهار حوزه ارزشیابی از بافت، ارزشیابی از درون داد، ارزشیابی از فرآیند و ارزشیابی از فرآورده قابل بررسی است. در این سؤال، سه مورد اول به دلیل ارزشیابی نتیجه کار به ارزشیابی از فرآورده اشاره دارد؛ اما در مورد آخر چون فرآیند اندازه گیری فشار خون و درستی و دقت انجام این کار مدنظر است، ارزشیابی از فرآیند محسوب می گردد.

۱۹- گزینه «3» برای محاسبه ضریب تمییز از فرمول زیر استفاده می شود، منتها قبل از محاسبه ضریب تمییز باید تعداد افراد هر گروه که برابر با ۲۷درصد کل افراد است مشخص شود.

تعداد افراد در هریک از گروه های بالا و پایین برابر با ۵۴ نفر است.

۲۰- گزینه «۴» روش دو نیمه کردن آزمون، نوعی روش ضریب آلفا است. درواقع ضریب آلفای کرونباخ معرف میانگین ضریب های پایایی حاصل از همه دونیمه های ممکن یک آزمون است. بنابراین با محاسبه ضریب آلفا به جای روش دونیمه کردن می توان تأثیر انتخاب سؤال های خاص برای هر نیمه آزمون را خنثی کرد.

سؤالات آزمون سراسری۹8

1- ضریب پایایی فرم مقدماتی یک آزمون ۲۰ سؤالی، معادل با 75/0 است. اگر آزمون ساز قصد کند ضریب پایایی آزمون نهایی ۹۰/0 باشد، طول آزمون چند برابر افزایش می یابد؟

1) 5/1

2) 2

3) 5/2

4) 3

۲- به کدام علت، پایایی حد بالای اعتبار وابسته به ملاک است؟

1)

2)

3)

4)

۳- اگر نمره خام هوش کسب شده توسط یک آزمودنی ۱۰۵ و نمره خطای وی (خطای تصادفی) ۳ باشد، نمره واقعی این آزمودنی چقدر است؟

1) 99

2) 102

3) 108

4) 111

۴- زمانی که همبستگی یک آزمون استعداد ادبی با یک آزمون استعداد حسابداری را به دست می آوریم، از کدام روش تعیین روایی (validity) استفاده می کنیم؟

۱) هم گرا

2) واگرا

3) تشخیصی

۴) پیش بین

5- کدام مورد، جزو عوامل درونی مؤثر بر روایی آزمون نیست؟

۱) تعداد سؤال ها

۲) ترتیب سؤال ها

۳) کیفیت سؤال ها

۴) انگیزش آزمودنی ها

6- اگر فردی در آزمونی که میانگین و واریانس آن به ترتیب ۳۰ و ۲۵ است، نمره خام ۴۰ بگیرد، نمره ۹ بخشی از چند است؟

1) 8/5

2) 5/6

3) 8

4) 9

۷- به علت اینکه شناسایی دانشجویان مضطرب به صورت بالینی وقت گیر و دشوار است، پژوهشگری برای شناسایی آنها از مقیاس بروز اضطراب تایلور استفاده می کند که همبستگی بالایی با درجه بندی بالینی در دانشجویان دانشکده دارد. در این مثال پژوهشگر از کدام نوع روایی استفاده کرده است؟

1) سازه

۲) محتوی

3) همزمان

4) پیش بین

8- در پایایی به روش دو نیمه کردن، اگر ضریب همبستگی دو نیمه آزمون 50/0 باشد، ضریب پایایی برآورد شده کل آزمون از طریق رابطه اسپیرمن - براون، چقدر افزایش می یابد.

1) 12/0

2) 17/0

3) 22/0

4) 67/0

۹- واریانس کلیک آزمون ۵ سؤالی تشریحی ۱۴۴ می باشد. اگر واریانس سوالات به ترتیب ۱۰، ۵، ۹، ۱۶ و ه باشند، ضریب همسانی درونی این آزمون چقدر است؟

1) 45/0

2) 50/0

3) 90/0

4) 95/0

10- در کدام روش، برآورد پایایی (Reliability)، امکان دارد عامل رشد و یادگیری بر برآورد پایایی اثر بگذارد؟

1) بازآزمایی

۲) فرم های همتا

۳) کودر - ریچاردسون

۴) دو نیمه کردن آزمون

11- در یک امتحان ۱۰۰ سؤالی، اگر میانگین و انحراف معیار نمره های خام آزمودنیها به ترتیب ۶۵ و ۱۰ باشد، ضریب پایایی آزمون کدام است؟

1) 65/0

2) 66/0

3) 70/0

4) 78/0

۱۲- در کیک توزیع بهنجار، با میانگین و واریانس ۱۰۰، نمره خام ۱۲۵، معادل کدام نمره در مقیاس T است؟

1) 65

2) 70

3) 75

4) 80

۱۳- پژوهشگری برای تعیین روایی آزمون خود این فرض را مطرح می کند که اگر آزمون او بتواند بین گروههای خاصی در صفت اندازه گیری شده تفاوت قائل شود، دارای روایی است. این محقق از کدام شیوه برای تعیین روایی آزمون خود استفاده کرده است؟

1) سازه

۲) محتوا

۳) همگرا

۴) پیش بین

۱۴- اگر باشد، حاصل چقدر است؟

1) 3

2) 5

3) 6

4) 7

۱۵- ضریب پایایی آزمونی 91/0 و خطای معیار اندازه گیری برابر ۳ می باشد. انحراف معیار آزمون کدام است؟

1) 9

2) 95/0

3) 10

4) 12

۱۶- یک آزمون عینی برای اندازه گیری پیشرفت تحصیلی در درس ریاضیات طراحی شده است. بعد از اجرای این آزمون در بین گروهی از دانش آموزان بین نفرات افراد با قضاوت معلمان درباره پیشرفت تحصیلی ریاضیات دانش آموزان ارتباطی مشاهده نشد. کدام نوع روایی این آزمون، در خد مناسب نیست؟

1) محتوا

2) واگرا

۳) همزمان

۴) پیش بین

۱۷- در کدام نوع سؤال چند گزینه ای، احتمال آنکه یک یا چند گزینه از نظر دستوری بی معنا باشند، وجود دارد؟

1) منفی

۲) جاخالی

۳) پرسیشی

۴) پاسخ ترکیبی

۱۸- پژوهشگری که تعداد دانشجویان مراجعه کننده به اساتید را به عنوان شاخص محبوبیت اساتید در نظر می گیرد، از کدام شیوه سنجش استفاده کرده است؟

1) نهان یابی

۲) نقش گزاری

۳) عملکردی

۴) قیاس پذیر

۱۹- طبق نمودار زیر، کدام افراد از نظر شدت عاطفه یکسان اند ولی از نظر جهت متفاوت اند؟



1) على و حسین

۲) على و کیوان

۳) ساسان و على

۴) ساسان و حسین

۲۰- یک سؤال تشریحی به صورت ۱ تا ۵ نام گذاری می شود، اگر میانگین آن ۳ باشد، ضریب دشواری آن در بازه صفر تا یک چند است؟

1) 8/0

2) 75/0

3) 6/0

4) 5/0

«مجموعه علوم تربیتی»

۲۱- اگر ضریب دشواری یک سؤال برابر با ۸/0 باشد، ضریب تمیز آن در کدام بازه قرار می گیرد؟

۱) 3/0+ و 3/0-

۲) 4/0+ و 3/0-

۳) 4/0+ و 4/0-

۴) 5/0+ و 5/0-

۲۲- در آزمونی که برای سنجش اضطراب تهیه شده است، میانگین نمره های کودکان مضطرب به نحو معنی داری بیش از کودکان عادی است، این امر، نشان دهندهی کدام نوع روایی (Validity) در آزمون مذکور است؟

1) افتراقی

۲) محتوایی

3) همزمان

۴) ملاکی

۲۳- نمره های یک آزمون مورد استفاده در درس شیمی دوره ی پیش دانشگاهی با نمره های حاصل از آزمون شیمی کنکور سراسری سال پیش، دارای ضریب همبستگی ۹۲/0 است. این ضریب، نشان دهنده ی کدام مورد در آزمون است؟

1) روایی

۲) پایایی

3) همسانی درونی

۴) عینیت

۲۴- ارزش خطای استاندارد اندازه گیری یک آزمون برابر ۲ نمره است. اگر نمرهی یک آزمودنی برابر ۵۰ باشد، نمره ی او با احتمال چند درصد در دامنه ی ۴۳ الی ۵۷ است؟ .

1) 68

2) 95

3) 72/97

4) 100

۲۵- میزان حمایتی که شواهد جمع آوری شده از تفسیر نمره های آزمون برای هدف پیشنهادی به عمل می آورند، بیانگر کدام مورد است؟

1) نرم

۲) پایایی

۳) استاندارد

۴) روایی

پاسخنامه آزمون سراسری ۹۸

1- گزینه «۴» فرمول زیر برای حل این سؤال مناسب است:

در این فرمول برابر با ضریب پایایی مطلوب و ضریب پایایی موجود و به معنای ضریب افزایش تعداد سؤالات آزمون است.

۲- گزینه «۱» پایایی یک آزمون بر روایی آن اثر دارد، زیرا به لحاظ نظری، همبستگی یک آزمون با نمره دیگر نمی تواند بیش از همبستگی نمرات همین آزمون با نمرات واقعی خود باشد؛ یعنی:

3- گزینه «۲» براساس نظریه کلاسیک آزمون، نمره ای که فرد در آزمون می گیرد با نمر واقعی یا توانایی واقعی او متفاوت است. تفاوت بین نمره واقعی و نمره مشاهده شده فرد از خطای اندازه گیری ناشی می شود. بنابراین نمره مشاهده شده فرد از جمع نمره واقعی و خطای اندازه گیری (X=T+E) حاصل می شود. با توجه به اینکه در این سؤال نمره مشاهده شده برابر با ۱۰5 و نمره خطا برابر با ۳ است، نمره واقعی فرد برابر با ۱۰۲ می شود.

۴- گزینه «۲» روایی سازه از طریق مجموعه ای از فعالیت ها که در آن محقق همزمان سازه را تعریف و ابزاری را برای اندازه گیری آن تولید می کند، تعیین می شود. روایی سازه شواهد جمع آوری شدهای درباره معنای یک آزمون را در بر می گیرد. این امر با نشان دادن رابطه بین یک آزمون و آزمون ها و ابزارهای دیگر انجام می شود. لذا روایی سازه مستلزم تحلیل منطقی همراه با تحلیل آماری - همبستگی و تفاوت های گروهی است، در صورتی که همبستگی آزمون موردنظر با آزمون های دیگری که پیازهای متفاوتی را اندازه می گیرند محاسبه شود، روایی واگرا استفاده شده است و در صورت عدم وجود همبستگی روایی واگرا آزمون تأیید می شود.

۵- گزینه «۴» دو دسته از عوامل شامل عوامل درونی و عوامل بیرونی روایی را تحت تأثیر قرار می دهند. عوامل درونی موثر بر روایی آزمون که به ساختار پرسشنامه و ابزار مربوط می شوند عبارتند از: راهنمای آزمون، کیفیت سؤال های آزمون، ترتیب قرار گرفتن سؤال های آزمون به دنبال یکدیگر و طول آزمون یا تعداد سؤال های آزمون در این سؤال گزینه سطح انگیزش آزمودنی ها از جمله موارد مربوط به عوامل بیرونی مؤر بر روایی آزمون محسوب می گردد.

6- گزینه «۴» به منظور پاسخ به این سؤال ابتدا باید نمره آزمون را به نمره Z تبدیل کرده و سپس نمره Z را به نمره نه بخشی تبدیل کرد. اگر نمره این آزمون را با X نشان دهیم نمره Z بصورت مقابل قابل محاسبه است:

بنابراین نمره Z آزمودنی با نمرا خام ۴۰ برابر با یک است. در ادامه باید نمرهZ را به نمره نه بخشی تبدیل کرد. نمرات نه بخشی دارای انحراف معیار ۲ و میانگین ۵ هستند.

۷- گزینه «3» روایی آزمون عبارت است از میزان کارایی آن برای اندازه گیری خصیصه ای که به منظور اندازه گیری آن خصیصه ساخته شده است. روایی بر سم نوع است: روایی محتوا، روایی ملاکی و روایی سازه. روایی ملاکی نیز به دو دسته روایی همزمان و پیش بین تقسیم می شود. ضرایب روایی.همزمان هنگامی که نمرات آزمون برای برآورد ملاک جاری و نه پیش بینی آینده به کار برده شود مناسب است در اینجا نیز روانشناس از روایی همزمان استفاده کرده است.

8- گزینه «۲» فرمول اسپیرمن - براون به صورت مقابل است:

در این فرمول نشان دهنده ضریب پایایی (اعتبار) آزمون و ضریب همبستگی بین دو نیمه آزمون و برابر با ۵/0 است. بنابراین ضریب پایایی کل دوره آزمون به صورت مقابل برآورد می شود:

اما آنچه در این سؤال خواسته شده ضریب پایایی آزمون نیست بلکه میزان افزایش آن است.

ضریب پایایی آزمون 17/0 افزایش داشته است.

9-گزینه «3» فرمول ضریب آلفا بصورت زیر است، در این فرمول معادل تعداد سوالات ، معادل واریانس کل و معادل مجموع واریانس سوالات است.

10- گزینه «1» در پایایی بازآزمایی یک آزمون بر دو نوبت برای آزمودنی های یکسان اجرا می شود و سپس همبستگی نتایج محاسبه می گردد. یک مسأله ای که در پایایی با استفاده از روش بازآزمایی وجود دارد مدت زمان بین دو اجرای آزمون است، رشد و یادگیری دو عاملی هستند که ممکن است در این فاصله زمانی رخ دهد و برآورد پایایی را تحت تأثیر قرار دهند.

11- گزینه «۴» برای برآورد پایایی آزمون با داشتن میانگین و انحراف معیار آزمون باید به فرمول 21K مراجعه کرد. در فرمول زیر N نشان دهنده تعداد سوالات،X نشان دهنده میانگین و نشان دهنده انحراف استاندارد هستند.

۱۲- گزینه «3» به منظور پاسخ به این سؤال ابتدا باید نمره آزمون را به نمره Z تبدیل کرده و سپس نمره Z را به نمره T تبدیل کرد. اگر نمره این آزمون را با X نشان دهیم نمره Z به صورت مقابل قابل محاسبه است:

بنابراین نمره Z آزمودنی با نمره خام ۱۲۵ برابر با 5/2 است. در ادامه باید نمره Z را به نمره T تبدیل کرد. نمرات T دارای انحراف معیار ۱0 و میانگین ۵۰ هستند.

۱۳- گزینه «1» دو دسته شواهد همگرا و شواهد افتراقی برای تعیین روایی تازه لازم است. مطالعه تفاوت بین میانگین نمرات گروه های مختلف یکی از منابع مهم این شواهد است. به عنوان مثال اگر در یک آزمون هوش میانگین نمرات گروه ۶ ساله بیشتر از میانگین گروه ۱۵ ساله و کمتر از میانگین گروه ۷ ساله باشد، این تفاوت نشانگر سازه هوش است.

۱۴- گزینه «۶» واریانس مجموع دو متغیر مساوی است با مجموع واریانس هریک از آن دو متغیر به اضافه دو برابر کوواریانس آن ها، که حل آن در سطر مقابل ارائه شده است.

۱۵- گزینه «۲» رابطه بین خطای معیار اندازه گیری (SEM)، انحراف معیار آزمون (SD) و پایایی () در فرمول زیر نشان داده شده است:

انحراف معیار آزمون موردنظر برابر با 10 است.

16- گزینه «۳» روایی ملاکی زمانی کاربرد دارد که بتوان نمرات آزمون را با ملاکی مرتبط دانست ملاک رفتاری است که نمرات آزمون را برای پیش بینی آن به کار می برد. روایی ملاکی به صورت ضرایب همبستگی بین نمرات آزمون و نمره ملاک بیان می شود. روایی ملاکی به دو شاخه پیش بین و همزمان تقسیم میشود. روایی همزمان برای برآورد ملاک جاری و نه پیش بینی آینده به کار می رود.

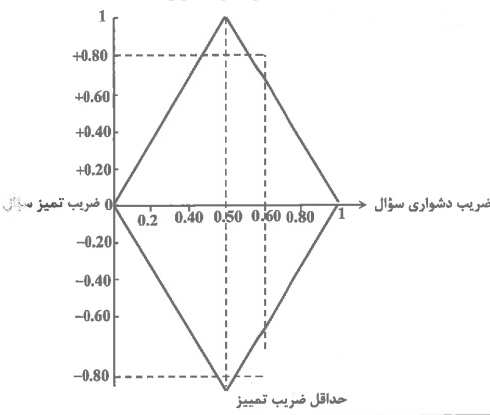
۱۷ـ گزینه «۲» در نوشن سؤال های چندگزینه ای باید بین تنه سؤال و گزینه ها از لحاظ دستوری هماهنگی وجود داشته باشد؛ در غیر این صورت، آزمون شوندگان از این ناهماهنگی به عنوان یک راهنما استفاده می کنند و گزینه درست را حدس می زنند. این اتفاق زمانی رخ میدهد که سؤال چند گزینه ای از نوع جاخالی باشد.

۱۸- گزینه «۱» سنجش نهان یابی (نامرئی) یک از روش های سنجش رفتاری است که مبتنی بر مشاهده است. در این روش آزمودنی یا افراد مورد سنجش نمی دانند که سنجشگر به جمع آوری اطلاعات درباره رفتار آنان می پردازد. فرض زیربنایی این روش آن است که اگر مردم ندانند به وسیله دیگران مشاهده می شوند، رفتارشان طبیعی تر و معمولی تر است.

۱۹ـ گزینه «۱» در این نمودار ساسان به لحاظ عاطفه خنثی است. علی دارای عاطفه مثبت قوی و حسین دارای عاطفه منفی قوی است، لذا علی و حسین از نظر شدت عاطفه یکسان اما به لحاظ جهت متفاوت هستند. نادر و کیوان نیز هر دو به لحاظ شدت عاطفه یکسان و دارای عاطفه متوسط هستند، اما به لحاظ جهت تفاوت دارند. ۲۰- گزینه «۲» روش نیتکو: در این روش ضریب دشواری یک سؤال تشریحی به صورت نمره میانگین گروه آزمون شوندگان برای سؤال تعریف شده است. برای اینکه ضریب دشواری را به مقیاس ضریب دشواری سؤال های دارای ارزش صفر و یک تبدیل کنیم باید نمره میانگین را به دامنه ممکن نمرات سؤال تقسیم کرد:

«مجموعه علوم تربیتی»

۲۱- گزینه «۳» دشواری سؤال بزرگ یا کوچک باشد آن سؤال به خوبی نمی تواند بین افراد مایل قائل شود؛ بلکه سؤالی با ضریب دشواری متوسط توانایی بیشتری در تشخیص افراد قوی و ضعیف خواهد داشت. در شکل زیر رابطه بین ضریب دشواری سؤال و ضریب تشخیص آن نشان داده شده است. براساس این شکل ضریب تشخیص سؤالی با ضریب دشواری 8/0 در بازه 4/0 و 4/0- قرار می گیرد.



۲۲- گزینه «1» کمپل و فسیک (۱۹۵۹) برای تهیه آزمونی که روایی سازه داشته باشد، جمع آوری شواهد همگرا و شواهد افتراقی را ضروری دانستند. یکی از منابع شواهد افتراقی بررسی تفاوت بین میانگین نمرات گروه های مختلف است. با توجه به اینکه در این آزمون بین میانگین نمرات کودکان مضطرب با کودکان عادی تفاوت معنادار نشان داده است، روایی افتراقی آزمون تأیید می شود.

۲۳- گزینه «۱» روایی آزمون عبارت است از میزان کارآیی آن برای اندازه گیری خصیصه ای که به منظور اندازه گیری آن خصیصه ساخته شده است. . هدف های خاص آزمون و روش های برآورد روایی، انواع مختلف روایی وجود دارد. روایی ملاکی یکی از انواع بررسی روایی است. روایی ملاکی نشان میدهد که یک آزمون با یک ملاک مشخص مطابقت دارد. چنین شواهدی با همبستگی بالا بین آزمون و اندازه ملاک فراهم میشود. روایی ملاکی بر دونوع روایی پیش بین و روایی همزمان است. در این سؤال همبستگی بین نمرات شیمی پیش دانشگاهی با نمرات آزمون شیمی کنکور سراسری از نوع روایی ملاکی است.

۲۴- گزینه «۱» برای پاسخ به این سؤال ابتدا باید از طریق فرمول فاصله اطمینان نمره Z را محاسبه و سپس احتمال مربوط به نمره Z را پیدا کرد.

با مراجعه : به جدول Z احتمال اینکه نمره فرد در بازه ۱ و ۱- قرار بگیرد را پیدا می کنیم. احتمال اینکه فردی با نمره خام ۵۰، نمره ای در بازه ۱ تا ۱- بگیرد برابر با ۶۸% است.

۲۵- گزینه «۴» روایی ماهیت مادی ندارد و ازجمله ویژگی های ماده هم نیست؛ بلکه روایی صفتی است در خصوص ادعاهایی که مطرح میشود. در واقع روایی از شواهد تجربی و روابط نظری که از صحت استنباط ها، تفسیرها یا اعمالی که براساس نمرات آزمون صورت می گیرد، حمایت می کند.

پاسخنامه آزمون ها

فصل اول: تعریف اصطلاحات

۱ـ گزینه «۳»

2- گزینه «3»

3ـ گزینه «1»

4- گزینه «4»

5- گزینه «3»

۶- گزینه «4»

7- گزینه «۴»

8ـ گزینه «۴»

9- گزینه «۳»

10ـ گزینه «2»

فصل دوم: مقیاس های اندازه گیری

۱- گزینه «۳»

2-گزینه «1»

3- گزینه «4»

4ـ گزینه «۳»

5- گزینه «3»

۶- گزینه «1»

7ـ گزینه «3»

8- گزینه «2»

9- گزینه «2»

10ـ گزینه «۳»

فصل سوم: دسته بندی آزمون های مورد استفاده در روان شناسی و آموزش و پرورش

1- گزینه «۴»

2- گزینه «2»

3- گزینه «3»

4- گزینه «۳»

5- گزینه «3»

6ـ گزینه «۱»

۷- گزینه «۳»

8- گزینه «۲»

9- گزینه «۲»

10- گزینه «3»

فصل چهارم: دسته بندی ارزشیابی های آموزشی

1- گزینه «4»

۲- گزینه «2»

3- گزینه «1»

4- گزینه «4»

5- گزینه «2»

6- گزینه «4»

7- گزینه «۲»

8- گزینه «3»

۹- گزینه «4»

10- گزینه «1»

فصل پنجم: مراحل و فعالیت های ارزشیابی آموزشی

1- گزینه «۳»

۲ـ گزینه «۲»

3- گزینه «1»

4- گزینه «۳»

5- گزینه «1»

6- گزینه «۳»

۷- گزینه «2»

8- گزینه «2»

9- گزینه «1»

10- گزینه «2»

فصل ششم: تهیه طرح ارزشیابی پیشرفت تحصیلی

1- گزینه «3»

2- گزینه «4»

3ـ گزینه «4»

4- گزینه «3»

5- گزینه «4»

6- گزینه «2»

7- گزینه «۳»

8- گزینه «۳»

۹- گزینه «3»

10- گزینه «۴»

فصل هفتم: آزمون های تشریحی

1- گزینه «1»

2- گزینه «4»

3ـ گزینه «3»

4- گزینه «2»

5- گزینه «۴»

6- گزینه «3»

7- گزینه «۴»

8- گزینه «2»

۹- گزینه «3»

10- گزینه «4»

فصل هشتم: آزمون های کوتاه پاسخ

1- گزینه «2»

2- گزینه «2»

3- گزینه «۴»

4- گزینه «3»

5- گزینه «4»

6- گزینه «4»

۷- گزینه «3»

۸- گزینه «4»

۹- گزینه «3»

10- گزینه «2»

فصل دهم: آزمون های چندگزینه ای

1- گزینه «3»

2- گزینه «3»

3- گزینه «1»

4- گزینه «2»

5- گزینه «4»

6- گزینه «3»

۷- گزینه «2»

۸- گزینه «2»

۹- گزینه «2»

10- گزینه «4»

فصل یازدهم: روش های سنجش فرآیند و فرآورده های یادگیری غیر شناختی

1- گزینه «4»

2- گزینه «1»

3- گزینه «۴»

4- گزینه «2»

5- گزینه «2»

6- گزینه «4»

۷- گزینه «2»

۸- گزینه «1»

۹- گزینه «1»

10- گزینه «3»

فصل دوازدهم: روش های سنجش مستقیم

1- گزینه «1»

2- گزینه «3»

3- گزینه «2»

4- گزینه «3»

5- گزینه «2»

6- گزینه «1»

۷- گزینه «3»

۸- گزینه «3»

۹- گزینه «1»

10- گزینه «2»

فصل سیزدهم: اجرای نمره گذاری و تحلیل آزمون

1- گزینه «4»

2- گزینه «2»

3- گزینه «3»

4- گزینه «4»

5- گزینه «2»

6- گزینه «1»

۷- گزینه «3»

۸- گزینه «1»

۹- گزینه «2»

10- گزینه «3»

11- گزینه «4»

12- گزینه «1»

13- گزینه «2»

14- گزینه «2»

15- گزینه «3»

16- گزینه «2»

1۷- گزینه «3»

1۸- گزینه «4»

1۹- گزینه «1»

20- گزینه «4»

21- گزینه «3»

22- گزینه «2»

23- گزینه «۴»

24- گزینه «1»

25- گزینه «3»

26- گزینه «3»

2۷- گزینه «4»

2۸- گزینه «2»

2۹- گزینه «4»

30- گزینه «4»

فصل چهاردهم: استفاده از روش های آماری برای متغییر نمره های آزمون

1- گزینه «3»

2- گزینه «4»

3- گزینه «2»

4- گزینه «3»

5- گزینه «4»

6- گزینه «1»

۷- گزینه «4»

۸- گزینه «3»

۹- گزینه «1»

10- گزینه «2»

فصل پانزدهم: هنجارها و نیمرخ ها

1- گزینه «4»

2- گزینه «2»

3- گزینه «3»

4- گزینه «4»

5- گزینه «3»

6- گزینه «1»

۷- گزینه «4»

۸- گزینه «3»

۹- گزینه «4»

10- گزینه «4»

11- گزینه «3»

12- گزینه «4»

13- گزینه «1»

14- گزینه «2»

15- گزینه «1»

16- گزینه «2»

1۷- گزینه «3»

1۸- گزینه «3»

1۹- گزینه «4»

20- گزینه «3»

فصل شانزدهم: روایی آزمون

1- گزینه «3»

2- گزینه «4»

3- گزینه «3»

4- گزینه «1»

5- گزینه «2»

6- گزینه «3»

۷- گزینه «4»

۸- گزینه «3»

۹- گزینه «1»

10- گزینه «2»

11- گزینه «4»

12- گزینه «1»

13- گزینه «1»

14- گزینه «1»

15- گزینه «4»

16- گزینه «4»

1۷- گزینه «4»

1۸- گزینه «3»

1۹- گزینه «4»

20- گزینه «1»

فصل هفدهم: پایایی آزمون

1- گزینه «3»

2- گزینه «1»

3- گزینه «3»

4- گزینه «2»

5- گزینه «1»

6- گزینه «2»

۷- گزینه «3»

۸- گزینه «3»

۹- گزینه «4»

10- گزینه «2»

11- گزینه «3»

12- گزینه «4»

13- گزینه «1»

14- گزینه «4»

15- گزینه «1»

16- گزینه «2»

1۷- گزینه «2»

1۸- گزینه «3»

1۹- گزینه «2»

20- گزینه «3»

21- گزینه «2»

22- گزینه «4»

23- گزینه «3»

24- گزینه «3»

25- گزینه «2»

26- گزینه «2»

2۷- گزینه «4»

2۸- گزینه «1»

2۹- گزینه «2»

30- گزینه «2»

منابع و مراجع

منبع اصلی:

سیف، علی اکبر، اندازه گیری، سنجش و ارزشیابی آموزشی، انتشارات رشد، تهران: ۱۳۸۹

منابع فرعی:

١- بلوم، بنجامین و همکاران، طبقه بندی هدف های پرورشی: کتاب اول: حوزه شناختی، ترجمه علی اکبر سیف و خدیجه علی آبادی، انتشارات رشد، تهران: ۱۳۷۴.

2- پاشا شریفی، حسن، اصول روان سنجی و روان آزمایی، انتشارات رشد، تهران، ۱۳۸۱.

٣- دلاور، علی، احتمالات و آمار کاربردی در روانشناسی و علوم تربیتی، انتشارات رشد، تهران: ۱۳۸۹

۴- دلاور، علی، مبانی نظری و عملی پژوهش در علوم انسانی و اجتماعی، انتشارات رشد، تهران: ۱۳۸۰.

۵- مری جی، آلن، وندی، امین، مقدمه ای بر نظریه های اندازه گیری در روان سنجی، ترجمه علی دلاور، انتشارات سمت، تهران: ۱۳۷۴.

6- نادری، عزت اله؛ سیف نراقی، مریم، سنجش و اندازه گیری و بنیادهای تحلیلی ابزارهای آن در علوم تربیتی، نشر میثاق، تهران: ۱۱۳۷۴

۷- هومن، حیدر علی، اندازه گیری های روانی و تربیتی (فن تهیه تست)، انتشارات پیک فرهنگ، تهران: ۱۳۸۹

8- مجموعه سؤالات آزمون های سازمان سنجش تا سال ۹۸