اصول طراحی سوالات امتحانی

تهیه و تنظیم : نونا سادات آریاشکوه

کارشناس واحد EDO دانشکده بهداشت دانشگاه علوم پزشکی قزوین

مقدمه

یکی از عوامل بسیار مهم در فرآیند آموزش و یادگیری موثر و همچنین حفظ و تداوم آن ارزشیابی از آموخته های دیگران است و این ارزشیابی زمانی اثر بخش و پایدار می باشدکه مجموعه سوالات هر آزمون بر اساس معیارهای صحیح علمی و استاندارد تهیه شود و با توجه به ناهمگونی های متعدد و دخالت سلایق فردی و احتمالا عدم آشنایی بعضی از طراحان سوالات امتحانی با ابعاد فنی مسئله و عدم رعایت حیطه های شناختی و استفاده از تمام فصول کتاب در طراحی ، پیامدهای منفی حاصل از آن بر امر ارزشیابی آثار سوئی به همراه داشته است که مورد توجه متخصصین فن و اساتید علوم تربیتی و سنجش و اندازه گیری قرار گرفته است . لذا در این مقاله سعی شده اصول طراحی سوال در سه محور ۱- ملاک های فنی سوالات ۲- ادبیات سوالات ۳- مشخصات ظاهری سوالات به تفصیل شرح داده شود و برای یادآوری همکاران عزیز حیطه های شناختی در جدولی(جدول ۱/۱) ضمیمه گردد.

ملاک های فنی سوالات {لازمه اینکار ، مشخص بودن نوع سوال و سطح یادگیری مورد سنجش سوال و ... است. یعنی رعایت سطوح حیطه های شناختی ، عاطفی ، روانی – حرکتی و در بخش حیطه شناسی بر رعایت طبقات دانش – درک و فهم – کاربرد –تجزیه و تحلیل- ترکیب و ارزشیابی توجه شود }

۱- سوالات از ساده به مشکل تنظیم شود .

۲- از انواع سوالات استفاده شود .(حداقل ۵ نوع)

از جمله سوالات امتحانی می توان عینی و انشایی را بر شمرد.

در سوالات عینی ارزشگذاری یا تصحیح آنها ، به دور از حب و بغض قضاوت و نظارت شخص مصحح انجام می گیرد.

مانند سوالات چند گزینه ای – صحیح و غلط- کامل کردنی – جور کردنی – تلفیقی صحیح و غلط تصحیح کردنی

در سوالات کامل کردنی باید :

۱- کلمه یا عبارت حذف شده از مفاهیم با ارزش جمله باشد.

۲- در سوال بیش از یک جای خالی گنجانده نشود و جای خالی در وسط یا آخر جمله باشد.

۳- فقط یک پاسخ معین و مشخص داشته باشد.

در سوالات جور کردنی باید:

۱- تعداد پاسخ ها بیش از سوالات باشد.

۲- سوال و پاسخ بصورت مختصر نوشته شود .

۳- به منظور اینکه پاسخ سوال ها به راحتی پیدا نشود ، سوالات باید با یکدیگر متجانس باشند .

تلفیقی : در این نوع سوال از دانش آموزان خواسته می شود در صورت غلط بودن عبارت صحیح را در قسمت مربوطه بنویسد.

سوالات انشایی نیز به دو دسته کوته پاسخ و گسترده پاسخ تقسیم می شوند.

در سوالات کوتاه پاسخ سوال تا حدی محدود است و بر مبنای الگوی تعیین شده مشخص و تهیه می شود.

در سوالات گسترده پاسخ آزمایش شونده به پیروی از الگوی مشخصی موظف نیست و در نحوه تنظیم و ارائه پاسخ ، آزادی کامل دارد .

در اینگونه سوالات طول پاسخ و زمان لازم برای آن از طرف طراح تعین می شود.

۳- سوالات همنوع (مانند چند گزینه ای و جور کردنی و...) پشت سرهم قرار گیرند و دستور کار لازم را داشته باشند .

۴- هر سوال به طور مستقل فقط یک هدف آموزشی خاص را مورد سنجش قرار دهد. (و در نوشتن سوالات پایانی ، باید اهداف مهم تر آورده شود) یعنی پاسخ به سوال و یا بخشی از سوال مشروط بر پایان سوالات دیگر نباشد.

۵- هر سوال در سنجش یک هدف آموزشی بدون تکرار طرح شود (یعنی هر سوالی باید برای یک هدف خاص طراحی گردد)

۶- بارم هر سوال با جواب مورد انتظار متناسب باشد.

باید نمره داده شده به سوالات با پاسخ های مورد پیش بینی ، هم سنگ باشد . بعبارت دیگر بارم سوال با توجه به اهمیت هدف مورد سنجش و میزان پاسخ پیش بینی شده هم ارزش باشد . بنابر این به سوالی که هدف آموزشی مهمی را می سنجد ومیزان پاسخ مورد نظر هم نسبتا زیاد است باید نمره بیشتری تعلق بگیرد و بالعکس.

نمره گذاری و بارم بندی بر دو نوع است.

۱- نمره گذاری تحلیلی که بر نکات یا اجزای خاص پاسخ تکیه می کند.

۲- نمره کذاری کل نگر که تمامیت پاسخ را بصورت یک کل در نظر می گیرد.

۷- در تدوین سوالات بودجه بندی کتاب رعایت شود .

۸- متن سوال راهنمای پاسخ آن سوال یا دیگر سوالات نباشد .(هر پرسش باید از پرسش های دیگر مستقل باشد و پاسخ هیچ پرسشی نباید از روی متن یا گزینه های پرسش های دیگر حدس زده شود .)

۹- سوالات با عبارتی دقیق گویا و دور از ابهام طرح شوند.

اصول ساده نویس و روشنی در بیان سوال اهمیت خاصی دارد . طرز بیان سوال باید چنان باشد که آزمایش شوندگان کاری راکه باید انجام دهند را بفهمند . به بیان دیگر عدم پاسخ گویی دانش آموزان باید صرفا به دلیل نداشتن دانش و مهارت لازم در مورد فرآیند یادگیری مربوط به سوال باشد نه عدم توانایی در درک مفهوم سوال

۱۰- در طراحی سوالات به متن کتاب و محتوای اصلی درس تاکید شود . تا میزان پاسخگویی دانش آموزان بیانگر درجه توان و عملکرد تحصیلی وی در آن باشد و دانش آموزان را از حفظ طوطی وار محصور شدن در قالب مطالبی که فهم و یادگیری او را در حد ضعیفی نگه دارد خودداری شود . چنانچه نیاز به انعکاس سوال یا سوالاتی که تمرین کتاب یا .... باشد ، تغییراتی جزیی در آن ضرورت دارد و هیچ سوالی نباید خارج از کتاب باشد.

۱۱- برای سوالات انشایی دامنه و حد در نظر گرفته شود .

زمان و رشد سنی دانش آموزان را باید در نظر گرفت و حدود و اندازه تقریبی و طول هر پاسخ سوال انشایی از طرف طراح مشخص و تعیین شود .

۱۲- در توزیع بارم هر سوال به قسمت های مختلف پاسخ آن دقت شود .

نمره اختصاص داده شده به هر سوال می بایست با اجزای تقسیم شده اش متناسب باشد . مثلا برای یک سوال ۳ جزیی نمیتوان ۱ نمره در نظر گرفت.

۱۳- سوالات از نظر علمی صحیح طرح شوند.

سوالی از نظر علمی صحیح است که در مورد صحت و سقم پاسخ آن بین مصححان اتفاق نظر وجود داشته باشد و سوالی که بدنه (صورت مسئله) آن غلط باشد ، یک سوال علمی نیست.

۱۴- سوالات از قوه تمییز و مناسب برخوردار باشند .

حساسیت سوالات طراحی شده باید به گونه ای باشد که دانش آموزان موفق را از دیگر داننش آموزان جدا سازد.

۱۵- سوالات باید از اعتبار لازم برخوردار باشد.

اعتبار یعنی اگر یک آزمون دو بار در مورد یک آزمودنی بکار رود نتایج تقریبا یکسانی به دست می آید می توان گفت که اعتبار یکآزمون ثبات و پایایی نتایج آن در اجراهای متعدد است.

۱۶- نوع سوالات با سطح یادگیری مورد سنجش و اهداف امتحان ، هم خوانی داشته باشد.

۱۷- سوالات از روایی لازم برخوردار باشند .

روایی یعنی آزمون چیزی را اندازه بگیرد که برای آن ساخته شده است . مثلا آزمونی که برای اندازه گیریمعلومات دانش آموزان در درس ریاضی ساخته شده نباید معلومات آنان را در ادبیات فارسی اندازه بگیرد.

انواع روایی : ۱- روایی ظاهری ۲- روایی محتوایی ۳- روایی هم زمان ۴- روایی پیش بین ۵- روایی سازه (روایی مفهومی) شاید توضیح بیشتری بخواهد..

۱۸- در سوالات چند قسمتی باید سوالات به گونه ای طرح شوند که هر قسمت دارای پاسخ مجزا باشد و عدم پاسخگویی به یک قسمت دانش آموزان را از پاسخ به قسمت های دیگر محروم نکند.

۱۹- زمان اختصاص داده شده به پاسخگویی به سوالات یکامتحان توسط طراح بر اساس سطح دشواری سوالات و بار در نظر گیری تمام عوامل فیزیکی و شرایط برگزاری امتحان اختصاص داده شود .بدیهی است که زمان منظور شده در یک امتحان از مجموع زمان های محاسبه شده تک تک سوالات بر اساس پیش بینی زمان لازم شده توسط دانش آموزان برای پاسخ به هر سوال طبق راهنمای تصحیح سوالات انجام می شود .این زمان شامل: زمان لازم برای خواندن سوال + زمان لازم برای تفکر و یادآوری پاسخ + زمان لازم برای نوشتن پاسخ +بازنگری آن خواهد بود .(متناسب با توانایی دانش آموز متوسط است)

۲۰- از دادن سوالات اختیاری خودداری شود.

۲۱- چنانچه پاسخگویی به سوال یا سوالاتی نیاز به دستورالعمل جداگانه یا استفاده از ابزار خاصی دارد ، طراح نسبت به درج نکات لازم قبل از سوال یا سوالات مورد نظر ترجیحا در ابتدای سوالات امتحانی بعد از مندرجات سر برگ با رعایت قوانین مربوطه در این زمینه اقدام نماید.

۲۲- چنانچه برای پاسخگویی به سوالی نیاز به مقیاس – درجه – تعداد و یا بیان ویپگی خاصی است حتما در دنباله سوال آورده شود .

۲۳- چنانچه پاسخ سوالی از منبع خاصی مد نظر است در متن سوال بهمنبع مورد نظر اشاره شود .

۲۴- سقف حداقل و حداکثر تعداد سوالات امتحانی با توجه بهنوع درس – حجم و اهمیت محتوا- سطوح یادگیری مورد سنجش نوع سوال- زمان مناسب پاسخ به سوالات و شرایط سنی دانش آموزان تعین گردد.

بدیهی است سقف حداقل و حداکثر سوال در امتحانات دروس مختلف یکسان نخواهد بود ولی حداقل تعداد ۱۲ و حداکثر ۲۵ سوال پیشنهاد می شود . (کم بودن تعداد سوالات باعث بالا رفتن بارم و از دست دادن نمره و زیاد بودن تعداد باعث خستگی دانش آموزان و تصحیح مشکل اوراق امتحانی می شود)

۲۵- فهم نهایی سوالات امتحانی قبل از ارسال توسط طراحبازنگری و چنانچه نیاز به اصلاحات داشت ، به موقع اقدام شود .

۲۶- ارزش نمره ای هر سوال با توجه به سطوح یادگیری مورد سنجش حداقل ۲۵/. برای سوال آسان و حداکثر ۴ نمره برای سوالات سطوح بالاتر منظور گردد . در مواردی که سوال چند قسمت دارد نباید ارزش نمره ای هر قسمت از ۲۵/. کمتر تعین شود.

۲۷- تقسیم بندی سطح دشواری سوالات بر اساس ملاک های مهم از جنبه سطوح یادگیری مورد سنجش در حیطه شناختی صورت گیرد . بعنوان مثال در یک امتحان برای تعین سطح دشواری امتحان ممکن است پیشنهاد شود :

۱- سوالات طبقه دانش (آسان)۲- سوالات طبقه درک و فهم (متوسط پائین)۳- سوالات طبقه کاربرد و تجزیه و تحلیل(سطح بالا) ۴- سوالات طبقه ترکیب و ارزشیابی (دشوار) بدیهی است این تقسیم بندی بنابرماهیت محتوا در دروسی مختلف تغییر می کند.

سطوح ارزشیابی / سطوح حیطه شناخته / توزیع درصد سوالات دامنه نمره

آسان	دانش	./۲۵	./۴۰	./۳۰	./۳۵	./۴۰	./۳۵	۵-۸
متوسط	درک فهم- کاربرد	./۵۵	./۴۰	./۵۰	./۴۵	./۵۰	./۵۰	۸-۱۱
دشوار	تحلیل و ترکیب و ارزشیابی	./۲۰	./۲۰	./۲۰	./۲۰	./۱۰	./۱۵	۲-۴
		./۱۰۰	./۱۰۰	./۱۰۰	./۱۰۰	./۱۰۰	./۱۰۰	

نکته: در امتحان ۱۵ نمره ای دامنه نمره را در ۳ ضرب می کنیم

۴

۲۸- سطح دشواری سوالات امتحانی ، از قبل توسط طراح سوال بصورت نظری و مطابق با اهداف امتحان و با توجه به موارد زیر تعین می گردد. ۱- سوالی دشوار است که تا ۲۵/. دانش آموزان بتوانند به آن پاسخ صحیح دهند .۲- سوالی متوسط محسوب می شود که ۲۶ تا ۵۰ درصد دانش آموزان بتوانند به آن پاسخ صحیح بدهند. ۳- سوالی آسانمحسوب می گردد که بیش از ۷۵/. دانش آموزان بتوانند به آن پاسخ صحیح بدهند.

۲۹- فرم کلی سطح دشواری سوالات امتحانی بر اساس ۲۰/. آسان ۳۰/. متوسط پائین و ۳۰/. متوسط بالا و ۲۰/. دشوار طراحی می شود بدیهی استکه این فرم می تواند بر اساس اهداف امتحان و .... تغییر یابد.

سطح دشواری سوالات را می توان بصورت زیر تعین کرد:

$$P = \frac{\text{عامل حدس} + ۱}{۲} \quad \text{یا} \quad P = \frac{\text{درصد احتمالی پاسخ تصادفی} + ۱۰۰/.\text{پاسخ های دست به سوال}}{۲}$$

نوع سوال	۲ گزینه ای	۳ گزینه ای	۴ گزینه ای	۵ گزینه ای	کوتاه پاسخ و کامل کردنی (تشریحی)
درصد احتمالی پاسخ درست	./۵۰	۳۳	۲۵	۲۰	۰

مثلا برای سوال صحیح و غلط احتمال اینکه جواب درست یا غلط بدهند ./۵۰

پس ضریب سختی $\frac{\text{./۵۰} + ۱}{} = \text{./۷۵}$

ادبیات سوالات

۱- متن سوالات از غلط املایی محفوظ باشد.

یک ورقه امتحانی ضمن اینکه وسیله ای برای برآورد میزان یادگیری دانش آموزان می باشد ، بطور غیر مستقیم خود نیز یک وسیله آموزشی محسوب می شود . وجود غلط املایی در متن سوالات ارزش غیر مستقیم ورقه امتحان را کم خواهد کرد.

۲- از نشانه ها و نقطه گذاری ها و رسم الخط صحیح استفاده شود :

برای بیان دقیق طرح لازم است از نشانه ها مثل علامت سوال ویرگول ، تعجب ، پرانتز،خط تیره و غیره استفاده کرد.

۳- از جملات کوتاه و روان استفاده شود . به طوری که دانش آموزان آگاه بهمحتوای درس با صرف حداقل زمان ممکن به منظور و هدف سوالی پی ببرند متن سوال با بیانی کاملا ساده و روشن و دور از ابهام نوشته شود و از بکار بردن کلمات و عبارات دشوار در سوال و طولانی نمودن سوال که باعث می شود دانش آموزان نتوانند مهارت و توانایی خود را ظاهر سازند خودداری شود.

۴- نکات دستوری در تهیه و تنظیم سوالات رعایت شود .

رعایت نکات دستوری به فهم بهتر سوال کمک خواهد کرد . برای درک معنا و مفهوم و منظور طرح علاوه بر رعایت نشانه گذاری نکات دستوری مثل بکار بردن ارکان صحیح جمله (فاعل + مفعول+ فعل) (مسندالیه + مسند + فعل ربطی) و غیره ..... مد نظر قرار گیرد.

مشخصات ظاهری سوالات

۱- مشخصات سربرگ رعایت شود .

۱- نام و نام خانوادگی ۲- شماره دانش آموزی ۳- نام درس ۴- پایه ۵- آموزشگاه ۶- استان ۷- ناحیه ۸- نوبت امتحانی ۹- مدت امتحان ۱۰- تعداد صفحات سوالات ۱۱- تاریخ امتحان ۱۲- نام طراح

موارد ۷-۱ موجب راهنمایی و تسریع در امور می گردد.

۱۰-۸ باعث سهولت در رسیدگیبه اعتراضات دانش آموزان می شود

مورد ۱۰- باعث می شود دانش آموزان با اطلاع از زمان مربوط طوری وقت خود را تنظیم کنند که حداقل دچار کمبود زمان پاسخگویی نشوند.

مورد ۱۱ – صفحه بندی مرتب ، توجهدادن آزمودنی در پائین صفحه به صفحه بعد و درج عبارتی مانند پایان یا موفق باشید در انتهای صفحه آخر – تعداد صفحات سوالات دقیقا مشخص شود تا اگر صفحه ای کم یا تکراری قرار داده شده باشد ، دانش آموز یا مسئولین برگزاری مطلع و نسبت به رفع آن بکوشند.

۲- فضای کافی و مناسب برای پاسخ سوالات در نظر گرفته شود .(در صورت نداشتن پاسخنامه) باید فضای لازم متناسب با پاسخ هر سوال در نظر گرفته شود تا مشکلاتی برای دانش آموزان و مصححان ایجاد نشود . چنانچه پاسخ سوالی می بایست به صفحه دیگر انتقال یابد حتما دستور کار مربوط به انتقال پاسخ سوالات در زیر هر سوال باید در سر برگ سوالات آزمون قید گردیده باشد.

۳- سوالات خوانا باشد. کیفیت تایپ ، خوانایی خط و شیوایی سوالات مطلوب باشد.

۴- سوالات باکیفیت بالا تکثیر شود . کم رنگ بودن سوالات ، چاپ نشدن ابتدا و انتهای سوالات اثرات منفی در پاسخگویی دانش آموزان دارد(لذا ضرورت دارد طراح نظارت و کنترل لازم را بر کیفیت تکثیر سوالات قبل از برگزاری امتحان داشته باشد)

۵- بارم سوالات در سمت چپ ورقه قید گردد. بارم هر سوال فقط در سمت چپ ورقه امتحان نوشته شود . چنانچه در زیر سوال یا بالای سوال درج گردد مشکلاتی در آزمون پدیدار خواهد کرد خصوصا درس ریاضی.

۶- مجموع نمرات سوالات با نمره اختصاص یافته به ورقه مطابقت داشته باشد. باید بارم تک تک سوالات با جمع بارم کلیه سوالات هم خوانی و مطابقت داشته باشد(۱).

جدول ۱/۱ ، سطوح حیطه شناختی (۱).

نمونه هدف های کلی آموزشی و فعل های رفتاری هر طبقه در حیطه شناختی

طبقات	تعریف هر طبقه	هدف کلی آموزشی (نتایج یادگیری	فعلهای رفتاری (به صورت زمان حال ساده و سوم شخص مفرد)
دانش	در این طبقه از دانش آموز انتظار می رود تا چیزهایی را که آموخته است عینا پاسخ دهد و شامل یادآوری بازشناسی آموخته ها می شود	اصطلاحات متداول را میداند. مفاهیم اصلی را می داند. روشها را می داند	تعریف می کند . نام می برد. بازگو می کند . فهرست می کند . نسخه برداری می کند . بیان می کند. تکرار می کند. فکر می کند. مشخص می کند . جور می کند . توصیف می کند.
درک و فهم (فهمیدن)	آن دسته از هدف ها ، رفتار ها یا پاسخ هایی است که با فهمیدن و درک کردن همراه است و لازمه آن فهم معنی ظاهری یا تحت اللفظی پیام ها یا منظورهای موجود در سوالات است مانند ترجمه کردن ، تفسیر کردن ،استخراج (درک اطلاعات ناشناخته از اطلاعات شناخته شده) خواندن نقشه ، جدول ، نمودار ، فرمول های ریاضی و ...	اصول را می فهمد تئوریها را می فهمد قوانین را می فهمد	مثال می زند .حل می کند .ترجمه می کند .تخمین می زند .توضیح می دهد .خلاصه می کند .تمیز میدهد .برگردان می کند .بسط می دهد مطابقت میکند. پیش بینی می کند. تعمیم می دهد. بازنویسی می کند
کاربرد	. کاربرد یعنی استفاده از آموخته های پیشین در موقعیتهای جدید و حل مسایل جدید	مفاهیم و اصول را در موقعیت جدید به کار می برد . قوانین و نظریه ها را در موقعیتهای عملی به کار می برد.	تغییر می دهد. محاسبه می کند . نمایش می دهد . کشف می کند . اندازه گیری می کند .کنترل می کند . با مهارت انجام می دهد . اصلاح می کند . به کار می برد . تهیه می کند . تولید می کند . نشان می دهد . تولید می کند . آماده می کند .
تجزیه و تحلیل	به معنی تقسیم مطالب کلی به اجزا تشکیل دهنده آن و تجزیه و تحلیل یعنی تعین اجزا و کشف و بیان روابط بین قسمتهای مختلف یک مطلب و نحوه سازمان بندی آنها و شناخت اصول حاکمبر ساخت مطالب	ساختار یک اثر هنری را تجزیه و تحلیل می کند . جداول و نمودارها را تجزیه و تحلیل می کند . یک متن ادبی را تجزیه و تحلیل می کند .	به اجزا تقسیم می کند . با نمودار نشان می دهد . مجزا می کند . تجزیه می کند . تفکیک می کند . هجی می کند . به تصویر در می آورد . استفاده نمی نماید . می شکافد.
ترکیب	توانایی به هم پیوسته عناصر و اجزا به منظور تهیه یک صورت یا کل جدید ترکیب را می توان توانایی آفرینندگی و تولید دانست که با استفاده از دانسته های پیشین انجام می گیرد.	اصول و قواعد را با یکدیگر ترکیب می کند . طرح جدیدی ابداع می کند . مطالب را با یکدیگر ترکیب می کند .	طبقه بندی می کند . درهم می آمیزد . انشا می کند . تجدید نظر می کند . سخنرانی می کند . داستان می نویسد . شعر می سراید. تجدید سازمان می کند . مدون می کند . نمودار تهیه می کند . گروه بندی می کند . ترکیب می کند .
ارزشیابی	بالاترین هدف فرآیند ذهنی است که در آن شخص قضاوتی مبتنی بر هدف و منظوری خاص درباره ارزشها ، عقاید ،کارها،راه حل ها ، روش ها .مواد و غیره	یک اثر هنری را ارزشیابی می کند . روشهای تدریس را ارزشیابی می کند . قوانین یادگیری را ارزشیابی می کند	ارزیابی می کند . مقایسه می کند . نتیجه گیری می کند . مقابله می کند . انتقاد می کند. تفسیر می کند ربط می دهد. قضاوت می کند.داوری می کند.خلاصه می کند

چگونه سوال های آزمون را تحلیل کنیم؟

هدف از تحلیل سئوالهای آزمون و بررسی تک تک سئوالها و تعیین میزان دقت و نارسایهای آنهاست .

در تحلیل سئوالهای آزمون ، نقاط قوت و ضعف آزمون و کیفیت سئوالهای آن تعیین می شود . لازم است مدرسین پس از اجرای هر آزمون سئوالهای آن را تحلیل کنند و با استفاده از نتایج حاصل به تجدید نظر در آزمون و بهبود کیفیت سئوالها برای استفاده های بعدی اقدام نمایند .

برای تهیه یک آزمون خوب تحلیل پاسخهای آزمون شوندگان و بررسی تجربی آنها ضروی است .

مراحل تحلیل سئوال :

اطلاعات مورد نیاز برای تحلیل سئوالهای یک آزمون پاسخهایی هستند که آزمون شوندگان به هر سئوال داده اند بنابراین باید تعیین شود که در هر سئوال چند نفر گزینه درست را انتخاب کرده اند ، هریک از گزینه های انحرافی چند نفر را به خود جلب کرده است و چند نفر آن را بی جواب گذاشته اند . برای این منظور ، باید هر سئوال را در کارتی مانند شکل یک وارد نمود .

عنوان آزمون : روانشناسی تربیتی تاریخ اجرا : ۹۹/۱۱/۲

موضوع سئوال : هدفهای رفتاری

هدف یادگیری : دانشجو بتواند بین هدفهای کلی رفتاری ، هدفهای غیررفتاری و هدفهای رفتاری ناکامل تمیز قائل شود

یادگیرنده بتواند به کمک کلمه هایی که معلم در اختیارش می گذارد جمله های انگلیسی بسازد

هدف بالا :

الف : یک هدف رفتاری است ، ملاک عملکرد دارد ، اما شرایط عملکرد ندارد

*ب : یک هدف رفتاری است ، شرایط عملکرد دارد ، اما ملاک عملکرد ندارد

ج : یک هدف کامل رفتاری است

د : یک هدف غیر رفتاری است

گروهها
الف
ب
ج
د
بدون پاسخ

۲۵٪ بالا
۰
۵
۳
۰
۲
۱۰

۲۵٪ پایین
۰
۲
۳
۰
۰
۱۰

ضریب دشواری : ۳۵

ضریب تمیز یا تشخیص : ۰,۳

این کارت را می توان برای هر سئوال در روی یک صفحه کاغذ یا مقوا درست کرد و نتایج تحلیل آن سئوال را در یک طرف و صورت اصلاح شده سئوال را در طرف دیگر کارت نوشت.

اطلاعات موجود در شکل فوق که برای تحلیل یکی از سئوالهای آزمون و محاسبه شاخصهای آماری آن سئوال ضروری هستند از طریق تعیین دو گروه قوی (بالا) و ضعیف (پایین) آزمون شوندگان و تعیین نحوه پاسخدهی آنان به گزینه های مختلف سئوال بدست آمده اند .

گروه بالا از ۲۵٪ تا ۳۳٪ آزمون شوندگان دارای نمره های بالا و گروه پایین از ۲۵٪ تا ۳۳٪ آزمون شوندگان دارای نمره پایین تشکیل می یابد . این اطلاعات برای تک تک سئوالهای هر آزمون ضروری هستند .

برای تعیین گروههای بالا و پایین ، برگه های آزمون را به ترتیب نمره ای که گرفته اند از کوچک به بزرگ مرتب کنید . بعد از بالاترین نمره شروع کنید و تعداد برگه هایی را که می خواهید در گروه بالا قرار دهید انتخاب کنید و به همین ترتیب از پایین ترین نمره شروع کنید و تعداد برگه هایی را که می خواهید در گروه پایین قرار دهید نیز بر گزیند .

متخصصان آزمون سازی پیشنهاد کرده اند که اگر تعداد کل دانش آموزان که در آزمون شرکت کرده اند بیشتر از ۲۰ و تا ۴۰ نفر هستند ۱۰ برگه بالا و ۱۰ برگه پایین را انتخاب کنید و کاری به بقیه برگه ها نداشته باشید اگر تعداد دانش آموزان ۲۰ نفر یا کمتر است برگه های آزمون را کاملاً به دو دسته بالا و پایین تقسیم بندی کنید .

در شرایطی که تعداد کل آزمون شوندگان بیش از ۴۰ نفر است بهترین رقم برای گروه بالا و پایین ۲۷٪ کل برگه هاست این رقم برای شرایطی که نمرات دارای توزیع بهنجار باشد پیشنهاد شده است .

اما برای آزمونهای کلاسی هر یک از ارقام ۲۵ تا ۳۳ درصد به عنوان تعداد برگه های گروه بالا یا پایین مناسب است بعد از تهیه گروههای بالا و پایین لازم است برای هر سئوال اطلاعات زیر را تهیه کنید:

۱-تعداد افراد گروه بالا که هر یک از گزینه های سئوال را انتخاب کرده یا آن را بی جواب گذاشته اند

۲-تعداد افراد گروه پایین که هرک از گزینه های سئوال را انتخاب کرده یا آن را بی جواب گذاشته اند .

محاسبه ضریب دشواری سئوال

درصد کل آزمون شوندگانی که به یک سئوال جواب درست می دهند ضریب دشواری آن سئوال است که با حرف P نشان داده می شود . اگر در تحلیل یک سئوال کلیه افراد یا کلیه برگه های امتحانی دخالت داشته باشند برای محاسبه ضریب دشواری سئوال آن کافی است که تعداد کل افرادی را که به آن سئوال جواب درست داده اند بر تعداد کل آزمون شوندگان تقسیم کنیم و نتیجه را در ۱۰۰ ضرب نماییم رقم حاصل ضریب دشواری سئوال است مطابق فرمول $P=100(T/R)$

در این فرمول R معرف تعداد کسانی است که به سئوال جواب درست داده و T نشان دهنده تعداد کل آزمون شوندگان است . به عنوان مثال اگر تعداد پاسخ دهندگان به یک سئوال ۸۰ نفر باشند و از این تعداد ۶۵ نفر به آن سئوال جواب درست داده باشند ضریب دشوا ری این سئوال برابر خواهد بود با ۸۱=۶۵/۸۰ *۱۰۰ = P

در مواردی که تعداد آزمون شوندگان (تعداد برگه های آزمون) زیاد است و اطلاعات ما به نحوی به پاسخدهی افراد گروه بالا و گروه پایین محدود می شود لازم است از فرمول زیر استفاده کنیم .

۱۰۰*انتخاب های درست گروه پایین + انتخابهای درست گروه بالا = ضریب دشواری سئوال = P

تعداد افراد گروه بالا + تعداد افراد گروه پایین

برای محاسبه ضریب دشواری سئوال شکل یک به نحو زیر عمل می نمائیم .

۳۵٪ = ۱۰۰ * ۲+۵ = ضریب دشواری سئوال = P

۲۰

رقم ۳۵ نشان می دهد که درصد نسبتاً کمی از دو گروه بالا و پایین پاسخ درست سئوال را برگزیده اند .

براساس اطلاعات مربوط به سئوال می توان نتیجه گرفت که این سئوال نسبتاً دشوار است زیرا حتی نصف گروه انتخابی (بالا و پایین) نتوانسته اند به آن جواب درست بدهد براساس توضیحات بالا ، هر اندازه ضریب دشواری یک سئوال بزرگ تر (به ۱۰۰ نزدیکتر) باشد آن سئوال آسان تر است و هر اندازه که این ضریب کوچک تر (به صفر نزدیکتر) باشد سئوال دشوارتر است . بنابراین به جای ضریب دشواری می توان از ضریب آسانی یا سهولت نام برد اما معمول این است که به آن ضریب دشواری بگویند .

چنان که ملاحظه شد فرمول بالا ضریب دشواری را برحسب یکعدد صحیح دورقمی نشان می دهد.

برخی از متخصصان اندازه گیری و ارزشیابی ترجیح می دهند که این ضریب مانند ضریب تمیز که در قسمت بعد توضیح داده خواهد شد نیز برحسب اعداد اعشاری باشد که تفسیر آن با تفسیر ضریب تمیز همانند صورت گیرد . در این حالت فرمول محاسبه ضریب دشواری به صورت زیر است

نسبت افراد گروه پایین که به سئوال جواب درست داده اند + نسب افراد گروه بالا که به ۲سئوال جواب درست داده اند= ضریب دشورای

برای محاسبه ضریب دشواری سئوال شکل یک بار این فرمول به نحو زیر عمل می کنیم

۰,۵+. / ۲ = . /۳۵ = ضریب دشواری

۲

ملاحظه می کنید که ارقام حاصل از دو فرمول برابرند اما در فرمول اول ضریب حاصل به صورت عدد صحیح نشان داده می شود در حالیکه در فرمول دوم به صورت عدد اعشاری است

تفسیر ضریب دشواری

یکی از راههای قضاوت درباره مفید بودن سئوالهای یک آزمون آن است که ببینیم سئوال تا چه اندازه به پراکندگی یا واربانس نمرات کمک می کند .

واربانس نمرات گروهی پاسخ دهنده به یک آزمون از دو قسمت تشکیل می یابد:

۱- واریانس سئوالها

۲- همبستگیهای بین سئوالها

هر سئوالی که دارای واریانس بزرگتری است و با سایر سئوالهای آزمون همبستگی بیشتری دارد به واریانس کل آزمون کمک بیشتری می کند در مقابل سئوالی که با سایر سئوالها همبستگی ندارد و دارای واریانس کوچکی است کمک زیادی به واریانس کل آزمون نمی کند اگر در تحلیل آزمون به انتخاب سئوالهایی بپردازیم که همبستگی زیادی با یکدیگر دارند و در ضمن دارای واریانسهای بزرگتری هستند در مجموع آزمون بهتری تدارک دیده ایم .

واریانس سئوال = $(P - 1)^* P$

در فرمول بالا ، P برابر است با درصد کسانیکه به سئوال جواب درست داده اند یعنی همان ضریب دشواری ، یک سئوال زمانی دارای حداقل واریانس است که ضریب دشواری آن ۰ یا ۱ باشد یعنی هیچ یک از آزمون شوندگان به آن سئوال جواب درست نداده باشد یا اینکه همه آزمون شوندگان به آن جواب درست داده باشند

واریانس سئوال= $(P- 1)^* P$

واریانس سئوال = ۰*(۰-۱)=۰

واریانس سئوال= ۱*(۱-۱)=۰

اما زمانی که ضریب دشواری سئوال P=۰,۵ باشد واریانس سئوال حداکثر خواهد بود

(P- 1) * واریانس سئوال

۰=(۰,۵-۱)*۰,۵ = واریانس سئوال

۰,۲۵=۰,۵*۰,۵= واریانس سئوال

بهتر است سئوال هایی برای آزمون انتخاب شوند که ضریب دشواری آنها از ۱ کمتر ، و از صفر بیشتر و به ۰,۵ نزدیک باشد . البته انتخاب سئوالهایی با ضرایب دشواری مناسب به نوع سئوالهای آزمون مربوط است در آزمونهای صحیح – غلط ضریب دشواری مناسب به نوع سئوالهای آزمون مربوط است در آزمونهای صحیح – غلط ضریب دشواری p = ۰,۵ زمانی بدست می آید که همه آزمودنیها را با حدس زدن به سئوال جواب بدهند .

آلن وین گفته اند برای آزمونهای چند گزینه ای سطح بهینه دشواری اندکی کمتر از وسط فاصله بین ۱ و سطح موفقیت با حدس زدن است . منظور از سطح موفقیت با حدس زدن عبارت است از نسبت پاسخهای درستی که با حدس زدن به دست می آید بنابراین برای سئوالهای چهارگزینه ای سطح موفقیت با حدس زدن حدود ۲۵۰,۰است و لذا سطح بهینه دشواری در وسط ۰,۲۵ و ۱ یعنی در حدود ۰,۶ خواهد بود.

برای تعیین رقم دقیق مصرف ضریب دشواری بهینه سئوال ابتدا سطح موفقیت ۱۰۰٪ را از سطح عملکرد و با حدس زدن (مثلاً ۲۵٪) کم می کنیم و نتیجه را به ۲ تقسیم می نماییم و نقطه میانی را مشخص می کنیم.

۰,۳۷۵ = ۰,۷۵ = ۰,۲۵ – ۱,۰۰ = نقطه میانی

۲ ۲

پس از آن رقم حاصل از محاسبات بالا (یعنی نقطه میانی) را با سطح مصرف عملکرد از طریق حدس زدن جمع می کنیم.

ضریب دشورای بهینه = عملکرد با حدس زدن + نقطه میانی

۰,۶۲۵ = ۰,۲۵ + ۰,۳۷۵

راه ساده تر برای بدست آوردن همان نتیجه این است که عدد ۱ را با عملکرد مبتنی بر حدس زدن جمع کنیم و نتیجه را بر ۲ تقسیم نمایید.

۱+۰,۲۵ = ۰,۶۲۵

۲

ضریب های دشواری بین ۰,۳ تا۰,۷ حداکثر اطلاع را درباره تفاوت بین آزمودنیها بدست می دهد.

محاسبه ضریب تمیز سئوال

ضریب تمیز را با d نشان داده می دهند . قدرت سئوال را در تمایز گذاری یا تشخیص بین گروه قوی و گروه ضعیف آزمون شوندگان مشخص می کند و یعنی معلوم می نمایند که سئوال تا چه اندازه می تواند گروه قوی را از گروه ضعیف جدا سازد . برای محاسبه ضریب تمیز یک سئوال از فرمول زیر استفاده می شود .

انتخاب های درست گروه پایین – انتخابهای درست گروه بالا= ضریب تمیز سئوال = d

تعداد افراد یک گروه (بالا یا پایین)

ضریب تمیز شکل ۱ به شرح زیر محاسبه می شود.

۵-۲=۰,۳ = ضریب تمیز سئوال = d

۱۰

تفسیر ضریب تمیز

هر قدر ضریب تمیز بزگتر باشد قوه تمیز آن سئوال بیشتر و هر قدر این ضریب کوچک تر باشد قوه تمیز ان کمتر است مثلاً اگر قوه تمیز سئوال ۹۰٪ باشد آن سئوال آزمون شوندگان قوی وضعیف و خیلی خوب از هم جدا خواهد کرد . اما اگر ضریب تمیز سئوال ۰,۱۰ باشد آن سئوال از عهده جداسازی دانش آموزان قوی و ضعیف به خوبی بر نخواهد آمد.

ضریب تمیز صفر حاکی از این است که آن سئوال به هیچ وجه نتوانسته بین گروه قوی ، ضعیف تمایز قائل شود گاه اتفاق می افتد که ضریب تمیز سئوالی منفی است این نوع ضریب تمیز نشان می دهد که در آن سئوال گروه قوی دیتر از گروه ضعیف عمل کرده است . این گونه سئوالها دارای معایب اساسی هستند که یا باید به کلی کنار گذاشته شوند یا در آنها تجدید نظر اساسی صورت پذیرد . اگر سئوال دارای اشکال فنی نباشد دلیل چنین ضریبی می تواند این باشد که افراد گروه بالا یا آن سئوال را به طور کامل یاد نگرفته اند یا آن را به غلط آموخته

اند بنابراین ، تمامی ضریبهای تمیز سئوالهای آزمون باید ابزارها سنجش شخصیت ، علاقه و نگرش ضریب تمیز منفی مشکل زا نیست در این گونه آزمونها ، تمیز دادن بین نسخ ها یا گروههای مختلف کاری مطلوب است و سئوالهای دارای ضریب تمیز بالا در تمایز قائل شدن بین آنها کمک کند (۲).

روش کودر ریچاردسون برای ارزیابی پایایی آزمون

این روش، که به اسم مبدع آن شهرت دارد در سال ۱۹۳۷ ابداع شد برای متغیر های اسمی دو وجهی با کد های صفر و یک طراحی شده است(کارمین و زلر، ۱۹۷۹). روش کودر ریچاردسون میزان همسازی درونی آزمون یعنی میزان تداخل همه سوالات از لحاظ سنجش یک ویژگی مشترک را ارزیابی می کند. استفاده از این آزمون مستلزم آن است که درجه دشواری همه سوالات آزمون یکسان باشد.

این آزمون دارای دو فرمول کودر ریچاردسون ۲۰ و کودر ریچاردسون ۲۱ است. در این روش بر اساس همبستگی درونی سوالات، شاخص کودر ریچاردسون استخراج می شود و چنانچه مقدار این شاخص بیش از 0.7 باشد، میتوان گفت ابزار پایایی دارد.

همانطور که بیان شد کودر و ریچاردسون دو فرمول را برای محاسبه هماهنگی درونی آزمون ها ارائه نموده اند. فرمول اول برآوردی از میانگین ضرایب قابلیت اعتماد برای تمام طرق ممکن تنصیف (دو نیمه کردن) را به دست می‌دهد. این ضریب برابر است با:

$$r_1 = \frac{n}{n-1}\left(1-\frac{\sum pq}{S^2}\right)$$

که در آن :

n تعداد سوال ها – p نسبت پاسخ های صحیح – q نسبت پاسخ های غلط – واریانس نمرات کل

فرمول دوم: برای استفاده از فرمول دوم باید سطح دشواری سوال های آزمون یکسان باشد:

$$r_2 = \frac{n}{n-1}\left(1-\frac{\bar{x}(n-\bar{x})}{nS^2}\right)$$

که در آن :

n تعداد سوال ها – واریانس نمره های کل – میانگین نمرات

برای محاسبه ضریب قابلیت اعتماد با این روش، نیازی به داشتن نسبت پاسخ های صحیح و غلط نیست، بلکه در دست داشتن واریانس نمره های کل و میانگین نمرات کفایت می‌کند.

برای مثال، آزمونی دارای ۴۰ سوال است. این آزمون به گروهی از دانش آموزان داده شده است. میانگین نمره های آزمون ۲۵ و واریانس کل ۲۶ محاسبه شده است. سطح دشواری سوال ها نیز یکسان است. ضریب قابلیت اعتماد این آزمون را محاسبه کنید:

$$r_2 = \frac{40}{39}\left(1-\frac{25(40-25)}{40(26)}\right)= 1.036(1-0.36) = 0.66$$

فرض اساسی رابطه های مطرح شده توسط کودر ریچاردسون این است که سوالات آزمون را می توان به صورت یک برای پاسخ های درست و صفر برای پاسخ های غلط تصحیح کرد. بنابراین نمیتوان این روابط را برای محاسبه ضریب پایایی پرسشنامه هایی که طیف پاسخ های آن ها بیش از دو گزینه دارد مورد استفاده قرار داد(۳).

منابع:

1. https://gamn1k.blogsky.com/1388/11/20/post-88/
2. https://nkums.ac.ir/index.aspx?Catid=3797
3. https://analysisacademy.com/4348