ارزیابی و تحلیل روان سنجی آزمون ها

هدف از تحلیل آزمون بررسی تک تک سوالات و تعیین میزان استاندارد بودن یک آزمون است. ارزیابی هر آزمون سبب می شود که نقاط قوت و ضعف آزمون و کیفیت سوالات مشخص گردد و از نتایج حاصل آن می توان در جهت ارتقاء کیفیت آزمون ها بهره برد.

ارزشیابی آموزشی بر اساس ملاک مورد استفاده به دو دسته زیر تقسیم می شوند :

الف) ارزشیابی وابسته به ملاک :

در این نوع ارزشیابی معیار یا ملاک ارزشیابی از پیش تعیین می شود و عملکرد فراگیر در آزمون با توجه به آن ملاک مورد قضاوت قرار می گیرد . در این نوع ارزشیابی که بیشتر ویژه پیشرفت تحصیلی است نوعی ملاک مطلق که همان هدف های آموزشی از پیش تعیین شده هستند مورد استفاده قرار می گیرند. در این شرایط سنجش برای بررسی دستیابی یا عدم دستیابی به اهداف آموزشی با ابزاری مانند آزمون صورت می گیرد. یعنی کسب موفقیت در آزمون هایی که مبتنی بر ملاک است مستلزم یادگیری هدف های آموزشی در سطحی است که برنامه آموزشی بر آن اساس طراحی شده است .

ب) ارزشیابی وابسته به هنجار :

این نوع ارزشیابی به جای یک ملاک از پیش تعیین شده و مطلق بر اساس نوعی ملاک نسبی طراحی شده است . به این صورت که عملکرد فراگیران را با یکدیگر مقایسه می کنیم . در این روش می توان تعیین کرد که پیشرفت فراگیر نسبت به سایر فراگیران چه وضعی دارد اما نمی توان تعیین کرد که پیشرفت یک دانش آموز نسبت به هدف های آموزشی چگونه است . در پایان این نوع آزمون ها جایگاه فراگیر را بر حسب رتبه ی درصدی یا نوعی هنجار سنی یا کلاسی و در مقایسه با گروه هنجار تعیین می شود. مانند آزمون های دستیاری.

آزمون های وابسته به ملاک دارای یک نمره مرزی برای حد تسلط هستند به زبانی دیگر هدف ما این است که فراگیران ما پس از اتمام دوره آموزشی به سطحی از تسلط در اهداف آموزشی رسیده باشند ، و وظیفه آزمون های وابسته به ملاک نیز دقیقاً تفکیک بین گروه فراگیران مسلط و غیر مسلط است. در همین راستا تحلیل این آزمون ها نیز باید بر همین اساس اجرا گردد.

در تحلیل آزمون های پیشرفت تحصیلی اطلاعات ذیل می بایست در اختیار اساتید قرار گیرد :

۱- ضریب دشواری آزمون :

«تعریف : در کل فراگیران که به یک سوال جواب صحیح می دهند ضریب دشواری آن سوال است . »

اگر تعداد کل فراگیران در یک آزمون ۲۰ نفر یا کمتر باشد برگه های آزمون کلاً به دو دسته بالا و پائین تقسیم می شود. اگر تعداد فراگیران بین ۲۰ تا ۴۰ باشد ۱۰ برگه بالا و ۱۰ برگه پائین انتخاب می گردد. در شرایطی که تعداد برگه های آزمون بیش از ۴۰ نفر است بهترین رقم برای گروه بالا و پائین ٪۲۷ کل برگه هاست . (بین ۲۵ تا ۳۳ درصد به عنوان تعداد برگه های گروه بالا و پائین مناسب است .)

تفسیر ضریب دشواری :

به طور کلی ضریب دشواری بین ۰/۳ تا ۰/۷ حداکثر اطلاعات را درباره تفاوت آزمودنی ها به دست می دهد . برای آزمون های چند گزینه ای سطح بهینه دشواری کمی کمتر از فاصله بین ۱ و سطح موفقیت با حدس زدن است بنابراین در ازمون های چهار گزینه ای سطح بهینه دشواری ۰/۶ خواهد بود .

۱- ضریب تمیز :

« تعریف : قدرت سوال را در تمایز گذاری یا تشخیص بین گروه قوی و گروه ضعیف آزمون شوندگان مشخص می کند . »

تفسیر ضریب تمیز :

هر قدر ضریب تمیز سوال بزرگتر باشد قوه تمیز آن بیشتر و هر قدر این ضریب کوچکتر باشد قوه تمیز آن کمتر است . ضریب تمیز صفر حاکی از آن است که سوال بین گروه قوی و ضعیف تمایز ایجاد ننموده است و اگر در محاسبه با ضریب تمیز منفی روبرو شویم به معنای آن است که در این سوال گروه قوی ، ضعیف عمل کرده است و گروه پایین در کلاس آن را بهتر پاسخ داده اند. بدیهی است این سوال اشکال اساسی دارد و باید به کلی کنار گذاشته شود و یا در آن تجدید نظر اساسی به عمل آید.

روش محاسبه ضریب تمیز در آزمون های وابسته به ملاک :

a: تعداد آزمون شوندگان که سوال را درست پاسخ داده اند اما نمره کمتر از حد تسلط گرفته اند.

b: تعداد آزمون شوندگان که سوال را درست جواب داده اند و در آزمون نیز به حد تسلط رسیده اند.

c: تعداد آزمون شوندگان که سوال را غلط جواب داده اند و در کل به حد تسلط نرسیده اند.

d: تعداد آزمون شوندگان که سوال را غلط جواب داده اند و در آزمون به حد تسلط رسیده اند.

تفسیر ضریب تمیز بدست آمده از این روش همانند شیوه محاسبه قبلی است . در آزمون های وابسته به ملاک ضریب تمیز صفر (زمانی که کلیه فراگیران به سوال پاسخ صحیح داده اند) نشان دهنده کاستی در سوال نیست چرا که ممکن است بر اساس اهداف آموزش کلیه فراگیران باید به این سوال پاسخ صحیح بدهند.

بنابراین در آزمون های وابسته به ملاک تفسیر ضریب دشواری و تمیز باید توسط مدرس و با توجه به اهداف آموزشی صورا گیرد.

تحلیل سوال های آزمون های وابسته به ملاک :

نتیجه آرمانی آزمون های وابسته به ملاک این است که در آنها اکثر آزمون شوندگان اکثر سوال ها را درست جواب دهند. در این حالت ضریب دشواری بزرگ و ضریب تمیز بسیار کوچک خواهد بود.

۱- تحلیل گزینه های انحرافی :

برای دستیابی به اطلاعات لازم در مورد سوال های آزمون ، علاوه بر تعیین ضریب دشواری و تمیز برای هر سوال می بایست نحوه پراکندگی پاسخ ها مربوط به گزینه های انحرافی نیز بررسی گردد.

گزینه ای به عنوان گزینه انحرافی مناسب است که تعداد افرادی را که از گروه ضعیف به خود جلب می کند بیش از گروه قوی باشد .

تجدید نظر و اصلاح آزمون :

زمانی که ضریب دشواری و ضریب تمیز سوال های یک آزمون تعیین شد باید به اصلاح سوالاتی که نیاز به تغییر دارند اقدام نمود . سوال هایی که ضریب دشواری پایین یا بالا دارند و یا ضریب تمیز آنها خیلی کم است باید مورد تجدید نظر قرار گیرند. این موارد شامل گزینه های انحرافی معیوب نیز می شود که باید اصلاح یا عوض شوند .

منحنی ویژگی سوال :

یکی از روش هایی دقیق برای تحلیل سوال آزمون مخصوصاً در آزمون های وابسته به ملاک است. در این روش ویژگی های سوال را به صورت نمودار نشان می دهند منحنی ویژگی سوال که برای هر یک از سوال های آزمون رسم می شود باز نمایی از رابطه بین احتمال پاسخ درست دادن به۸ سوال و موفقیت آزمون شونده در صفت مورد اندازه گیری توسط آزمون است . برای تهیه این منحنی نسبت یا درصد آزمون شوندگانی که آن سوال را درست جواب داده اند در برابر نمره کل آزمون انها رسم می شود . به سخن دیگر محور افقی نمره کل آزمون و بر روی محور عمودی نسبت آزمون شوندگان که به سوال پاسخ درست داده اند مشخص می شود .

از روی منحنی ویژگی سوال می توان ضرایب دشواری و تمیز سوال را تعیین نمود. ضریب دشواری عبارت است از نمره معیاری که در آن ۵۰ درصد آزمون شوندگان سوال را درست پاسخ داده اند و ضریب تمیز عبارتند از شیب منحنی ویژگی سوال.

در این نمودار افزایش تدریجی شیب منحنی نشانگر آن است که با افزایش نمرات آزمون درصد کسانی که به سوال پاسخ صحیح داده اند افزایش می یابد .

این روش تحلیل سوال این خصوصیت را دارد که به دقت و تفکیک معین می کند که سوال در چه محدوده ای ، از قابلیت تفکیک مناسبی برخوردار است. زمانی که استاد ، برگزار کننده آزمونی با معیار مشخصی است می بایست قابلیت تفکیک و تمیز آزمون در محدود معیار مشخص شده ، بیشترین میزان باشد تا بتواند فراگیران بالاتر از معیار و کمتر را، از هم تمیز دهد.

منحنی ویژگی یک سوال خوب آن نوع منحنی است که دارای شیب مثبت و ضریب دشواری متوسط باشد. اگر قرار است از یک آزمون برای انتخاب داوطلبانی که در بالا یا پائین یک نمره معین(معیار) قرار دارند استفاده کنیم. سطح دشواری باید برابر با همان نمره باشد و منحنی ویژگی سوال نیز در نمره معین نشان دهنده قدرت تمیز بالا باشد . اگر بخواهیم از میان آزمون شوندگان دارای نمرات بالا دست به انتخاب بزنیم شیب منحنی باید در قسمت انتهای آن بیشتر باشد .

پایایی آزمون :

آزمونی که در پایان دوره تحصیلی میزان یادگیری فراگیران را می سنجد باید پایا باشد. به این معنی که اگر این آزمون را در فاصله کوتاهی چندین بار به گروه واحدی از افراد بدهیم نتایج حاصل نزدیک به هم باشد. اگر آزمون در هر بار اجرا نتایج متفاوتی بدهد آن آزمون پایا نخواهد بود و در واقع چیزی را به درستی اندازه نخواهد گرفت و در این حالت آزمون اطلاعات مفیدی را به ما نخواهد داد. روش های متعدد محاسبه پایایی وجود دارد اما با توجه به امکانات بهترین روش برای اندازه گیری پایایی یک آزمون روش های ذیل است :

۱- **<u>روش گودرریچاردسون :</u>** در این روش آزمون یک بار اجرا می شود اما در این روش متوسط ضریب پایایی آزمون بر اساس ضریب همبستگی در نیمه های آن از طریق همه روش های ممکن دو نیمه کردن بدست می آید. عدد بدست آمده پایایی و همسانی درونی محتوای آزمون را نشان می دهد .

۲- **روش دو نیمه کردن آزمون (اسپیلت هاف) :** آزمون یک بار با گروه واحدی از آزمون شوندگان اجرا می شود و سپس آن را به دو نیمه تقسیم می کنیم . بهترین روش دو نیمه کردن آزمون بصورت جدا نمودن سوالات فرد و زوج است . برای محاسبه ضریب پایایی در این روش باید همبستگی بین نیمه های آزمون رابدست آوریم.

۳- **روش ضریب آلفا :** این روش زمانی از روش گودرریچارسون متفاوت خواهد بود که آزمون از تعدادی خرده آزمون تشکیل شده باشد .