

مروری کوتاه بر تحلیل آزمون و
طراحی سوال چهارگزینه ای

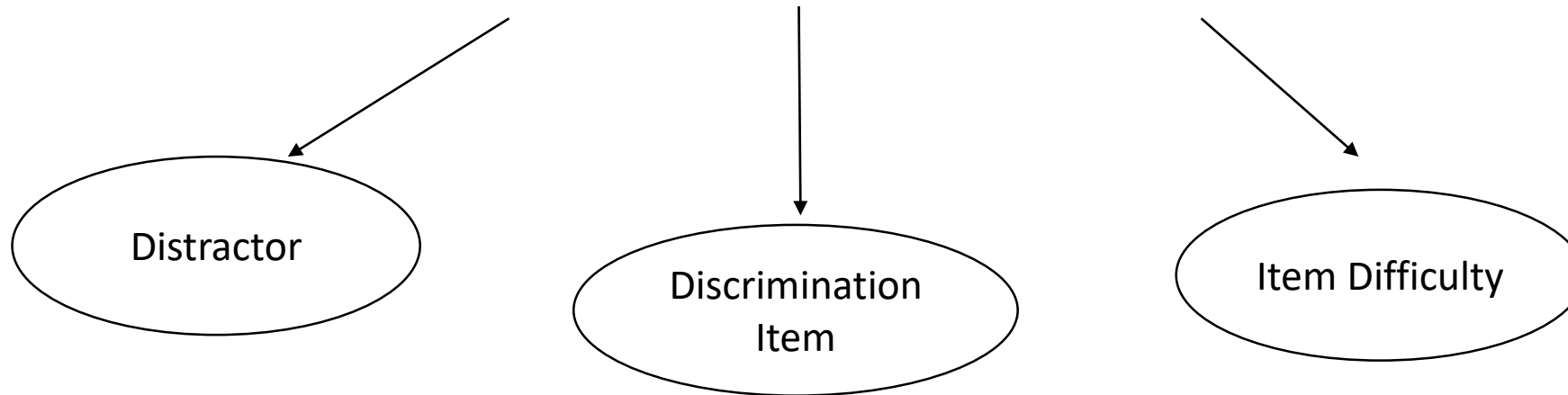
تحلیل سوال Item Analysis

- مجموعه ای از روش هایی است که از طریق آنها اثربخشی و کارایی هر یک از سوالات امتحان و کل امتحان اندازه گرفته می شود. شناسایی عوامل موثر در افزایش اثربخشی و کارآمدی هر یک از سوالات امتحان، ما را قادر می سازد تا آزمونها یا امتحانات استانداردتر و مناسب تری طراحی کنیم.
- تحلیل سوالات این امکان را فراهم می سازد تا ویژگی های هر یک از سوالات امتحان بررسی شود از این حیث اطمینان پیدا کرد که سوالات از استاندارد کافی و مناسب برای سنجش ویژگی یا توانایی مورد انتظار برخوردارند.
- تحلیل سوال رویکردهای متفاوتی دارد که از جمله آنها می توان به دو گروه نظریه های کلاسیک و نظریه های جدید (نظریه سوال-پاسخ و نظریه تعمیم پذیری) اشاره کرد.

انواع تجزیه و تحلیل

• الف) تجزیه و تحلیل کیفی

• ب) تجزیه و تحلیل کمی (CTT) و (IRT): در تجزیه و تحلیل کمی - سوال بر اساس سه پارامتر مورد نقد و بررسی قرار می گیرد:



• دشواری یا سهولت سوال **Item Facility** عبارت از نسبت افرادی است که به یک سوال پاسخ صحیح داده اند.

:Item Difficulty

عبارت از نسبت افرادی است که به یک سوال پاسخ صحیح داده اند.

Difficulty Index/Facility Index: .

• $P =$ تعداد پاسخ های درست

• تعداد شرکت کنندگان

• دامنه درجه دشواری:

• $0\% \text{-----} 30\% \text{-----} 70\% \text{-----} 100\%$

• $0 - 0.3 =$ خیلی سخت

• $0.3 - 0.7 =$ مناسب (نسبتاً دشوار)

• $0.7 - 1 =$ خیلی آسان

• P Low (< 30%) Difficult test

• P moderate (.30 - .70) moderately diff

• P high (> .70) easy test

• کسب نمره بالا می تواند ناشی از " هوش (IQ) بالای فراگیر - تبخیر فراگیر در موضوع - ساماندهی بسیار مناسب موضوعات درسی یا منبع درسی - روش تدریس مناسب (محتوا به صورت عینی ارائه شده باشد) آزمون آسان باشد " باشد.

• مقدار مناسب درجه دشواری به هدف آزمون بستگی دارد.

• در آزمون های پیشرفت تحصیلی سوالاتی با مقدار P بین ۳۰% تا ۷۰% سوالات بسیار مناسبی هستند.

• گنجاندن چند سوال فوق العاده آسان و دشوار نیز در آزمون های پیشرفت تحصیلی ضروری است.

Item Discrimination

قدرت تمیز بین افراد با توانایی بالا (نمره بالا) و توانایی پایین (نمره پایین)

• یعنی چقدر آزمون توانسته است بین دانشجویان ضعیف و قوی افتراق بگذارد.

• روش های برآورد قدرت تشخیص : گروه (پاسخ دهندگان به سوالات) را به دو گروه قوی (Top group) و گروه ضعیف (Bottom group) تقسیم می کنیم.

• Discrimination index= -1 → +1

• D = 0 سوال فاقد قدرت تشخیص است.

• قدرت تمیز منفی یعنی دانشجویان ضعیف هم به سوال پاسخ صحیح داده اند.

قدرت تمیز دارد ولی کم است = 0.3 - 0

مناسب است (هر چقدر بالاتر باشد بهتر است) = 1 - 0.3

• اگر مقدار قدرت تشخیص پایین بود می تواند به دلایل زیر باشد:

• 1-نحوه طراحی سوال

• 2-عدم همبستگی بعضی از سوالات با سوالات دیگر

Discrimination index= RPbis

- سوال قابل قبول - سوالی با قدرت تمیز ۰,۳ - ۱ است.
- قدرت تمیز بالای ۰,۳ مطلوب است
- اگر قدرت تمیز منفی شود حتما نیاز به تجدید نظر و بررسی دارد.
- شناسایی گزینه های انحرافی Distractor Analysis
- چه تعداد افراد هر گزینه یا پاسخ را انتخاب می کنند ؟ اگر تنها یک پاسخ مناسب وجود داشته باشد ، پس تمام پاسخ های دیگر گزینه انحرافی (response options - distractor) می باشند.
- ایده ال آن است هر فردی که می داند پاسخ درست را بدست آورد و افرادی که نمی دانند پاسخ های نادرست (گزینه های انحرافی) را به طور مساوی انتخاب کنند.

مثال زیر را در نظر بگیرید:

از یک آزمون ۱۰۰ سوالی که ۴۰ دانشجو در آن شرکت کرده اند پاسخ به سوال به صورت زیر بوده است

• 1 – D (2%) - C(8%) – B (90%) – A (0%)

• گزینه B گزینه کلید بوده است و ۹۰ درصد فراگیران به سوال یک این آزمون پاسخ درست داده اند.

گزینه انحرافی A نتوانسته است کسی را به خود جلب کند.

سایر گزینه ها نیز به تعداد اندکی فراگیران را به خود جلب کرده اند.

• بنابراین از ۱۰۰ درصد کل ۹۰ درصد را کم کرده و ۱۰ درصد باقیمانده را تقسیم بر ۳ (سه گزینه) می نمایم.

بنابراین در شرایط ایده آل هریک از گزینه های انحرافی باید حدود ۳٫۳ درصد افراد را به خود جلب می کرد.

اگر بالای ۰٫۳ باشد گزینه انحرافی مطلوبی خواهد بود.

مثال دیگر:

• 2- A (55%) - B (36%) – C (9%) - D (0%)

• گزینه A کلید سوال است .

۵۵ را از ۱۰۰ درصد کم کرده و ۴۵ درصد باقیمانده را بر ۳ گزینه انحرافی تقسیم می نمایم.

در این صورت باید هر یک از گزینه های انحرافی ۱۵ درصد دانشجویها را به خود جلب کرده باشد.
پس گزینه B گزینه فوق العاده خوبی است.

طراح سوال باید گزینه D را اصلاح کند و گزینه ای استفاده کند که مقبولیت بیشتری داشته باشد و توجه دانشجویان را بیشتر جلب کند.

- با بیرون کشیدن سوالات با قدرت تشخیص منفی از مجموعه سوالات خود وضعیت آزمون خود را بهبود ببخشید. این گونه سوالات را حذف کنید و مجدداً آن را بررسی کنید. با این کار پایایی را افزایش دهید.

- **KR20** یعنی پایایی کل آزمون و مقدارش بین ۰ الی ۱ است، حتماً باید بالای ۰٫۸ باشد.

- **SE** یا خطای اندازه گیری = خطای کل آزمون (نه تک تک سوالات) = این ابزار چقدر با خطا همراه بوده است؟ چقدر این ابزار قابل قبول است؟

- هر چقدر پایایی یک آزمون بالاتر باشد خطای اندازه گیری آن کمتر خواهد بود و بالعکس
- **Mean P** = میانگین ضریب دشواری کل آزمون است

Mean P = میانگین ضریب دشواری کل آزمون است

• برای مثال اگر میانگین ضریب دشواری کل آزمونی به صورت زیر باشد:

MP= 0.58 یعنی از نظر دشواری مناسب است

• **0 - 0.3 = خیلی سخت**

• **0.3 - 0.7 = مناسب**

• **0.7 - 1 = آسان**

• **MRPbi = تمیز کل آزمون**

• **0.3 - 0 = کم ولی قابل قبول**

• **1 - 0.3 = مناسب**

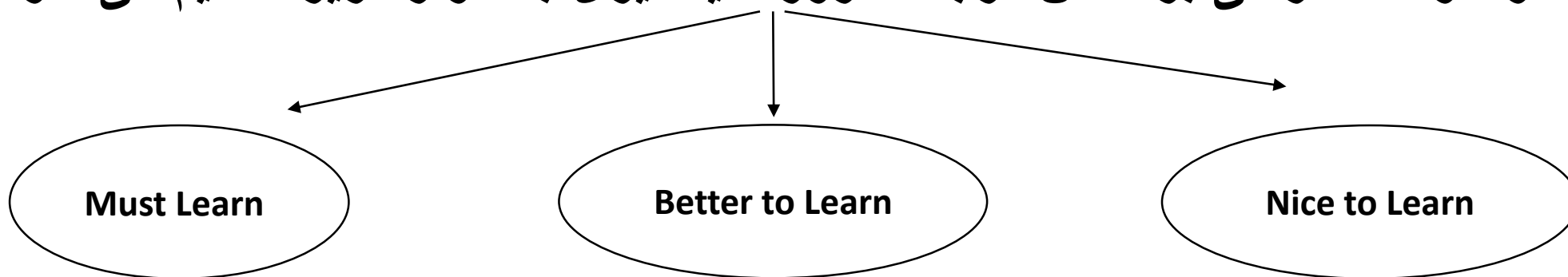
• **منفی = بسیار نامناسب**

• **Alpha = همگنی سوالات**

روایی آزمون چیست؟

- آزمونی که بتواند حیطه های مورد نظر گروه را به درستی ارزیابی نماید، آزمونی رواست.
- چگونه می توان روایی آزمون را بررسی نمود؟
- روایی آزمون را می توان به صورت مستقل از چگونگی پاسخ شرکت کنندگان در آزمون ها مورد ارزیابی قرار داد
- این کار را قبل از برگزاری آزمون انجام داد و نسبت به اصلاح سؤالات اقدام نمود
- آیا روایی از انواع مختلفی برخوردار است؟
- روایی محتوا
- روایی ساختار
- روایی پیشگویی کننده
- روایی ظاهری
- روایی محتوا و روایی ساختار در بررسی روایی یک سؤال از ضرورت بیشتری برخوردار است

• موضوعات درسی براساس درجه ضرورت یادگیری به موارد زیر تقسیم می شوند.



• در یک آزمون موضوعات بیشتری از دسته **Must** (۸۰-۷۰ درصد)،
تعداد کمتری **Better** (۲۰ درصد) و تعداد کمتری **Nice** (۵ درصد) را شامل می شود.

• سوالات مطرح شده از نظر چگونگی روند تفکر در ارائه پاسخ می تواند از حفظ کردن تا انواع عمیق تر مثل فهمیدن، آنالیز، ترکیب و کاربرد باشد.

انواع تاکسونومی سوالات چیست؟

- تاکسونومی I: سؤالی که تنها بر مبنای حافظه باشد: شایعترین تظاهر بیماری X کدام است
- تاکسونومی II: وقتی ذهن خود با انجام یک فرآیند به پاسخ می رسد:
تظاهرات بیماری (علائم و نشانه ها) داده شود و تشخیص خواسته شوند،
البته اگر **Problem Case** فوق قبلا عینا در کلاس درس مطرح شده باشد و دانشجو بتواند با حافظه به آن پاسخ دهد سطح آن I میشود
- تاکسونومی III: وقتی ذهن با انجام دو مرحله فرآیند به پاسخ می رسد
- توزیع سوالات در تاکسونومی های مختلف (I, II, III) بر اساس مخاطب (آزمون شونده) متفاوت خواهد بود
- در سطوح بالاتر آزمون شوندگان درصد تاکسونومی بالاتر باید بیشتر باشد
- نباید سوالات مبتنی بر حافظه تاکسونومی I بیش از ۵۰ درصد سوالات امتحان را شامل شود

روایی ساختاری چیست؟

- شامل مواردی است که باید در ساختار یک سؤال چند گزینه ای رعایت شود
- در دو قسمت ساقه سؤال (stem) و گزینه ها (options) ارائه می شود
(به ترتیب گزینه های چک لیست)

ویژگیهای ساقه سؤال چیست؟

- قسمت اصلی اطلاعات باید در ساقه سؤال باشد نه در گزینه ها. پس بطور کلی سؤالاتی که به صورت:
- در مورد x کدام یک از جملات زیر درست / نادرست است؟ **قابل قبول نیست**
- در ساقه هر سؤال بهتر است یک هدف اختصاصی یادگیری (SOB) مورد پرسش قرار گیرد
- نباید هم در مورد تشخیص و هم در مورد درمان سؤال شود ، زیرا در آنالیز پس آزمونی تحلیل دو SOB ایجاد مشکل می نماید
- **(معلوم نمی شود مشکل یادگیری در کدام SOB وجود داشته است)**
- ادبیات بکار رفته برای ساقه سؤال و گزینه ها باید شفاف و بدون ابهام باشد
- کاربرد کلماتی مثل " **تصور کنید، درست تر و امروزه** " مطلوب نیست
- کلمات مطلق مانند: " **همیشه، قطعا** " ممکن است سبب اشتباه دانشجو شود
- در تنه سؤال از لغات مثبت استفاده شود
- اگر از لغات منفی استفاده شود باید مشخص و متمایز باشد یا **زیر آنها خط کشیده شود و یا درشت تر و با خطی متفاوت** باشد
- از بکار بردن گزینه منفی برای ساقه منفی خودداری شود- **درک سوال و پاسخ دادن را مشکل می کند**

گزینه ها:

- مانند ساقه سوال درگزینه ها هم باید از ادبیات شفاف استفاده شده و کلمات مطلق تا حد امکان استفاده نشود
- درگزینه ها نباید از « هیچکدام» یا گزینه های ترکیبی مانند « مورد ۱ و ۲ » یا « هر دو» یا « همه موارد» استفاده شود
- از به کاربردن گزینه های با مفهوم متضاد خودداری شود
- به کار بردن گزینه های هایپرتنشن یا هایپوتنشن در دو گزینه یک سوال صحیح نیست
- هر سوال مستقل از سوالات دیگر باشد
- سوال " در مورد بیمار معرفی شده در سوال قبل اقدام درمانی بعدی چیست" صحیح نمی باشد
- گزینه ها از نظر طول، ساختار لغوی (کاربرد لغات انگلیسی و فارسی) از ساختار مشابه برخوردار باشد
- استفاده ازگزینه های متفاوت ممکن است دانشجو را دچار اشتباه نماید.
- عبارت های تکراری در گزینه ها که فقط در چند کلمه متفاوتند موجب گمراه شدن فراگیر می شوند.

- اشکالات املایی در ساقه یا گزینه ها به راحتی پیش از برگزاری آزمون برطرف می شوند گاهی این اشکالات مفهوم سؤال را تحت تاثیر قرار داده و فراگیر را گمراه می کند
- برای ساقه منفی نباید از گزینه منفی استفاده شود.
- پاسخ گزینه صحیح باید به روشنی مشخص باشد
- نباید از موارد بحث برانگیز که بین منابع مختلف تفاوت نظر وجود دارد و یا پاسخهای نزدیک و مبهم که موجب به اشتباه انداختن دانشجو می شود استفاده کرد
- گزینه های فرعی نیز نباید برای به اشتباه انداختن دانشجو بویژه دانشجویان مطلع استفاده شوند
- قرار دادن موارد بحث برانگیز در گزینه ای انحرافی که ممکن است در کتاب مرجع خاصی (که مورد قبول مدرس نبوده) آمده باشد برای آشفته کردن دانشجویی که آن کتاب را مطالعه نموده مطلوب نیست
- چینش گزینه ها باید عمودی باشد تا دانشجو سریعتر به آنها دسترسی داشته باشد
- از به کاربردن کلمات و لغات منفی خودداری شود اگر با رعایت شرایط از آنها استفاده می شود بهتر است تعداد این سؤالات محدود باشد
- پاسخ صحیح باید در سؤالات مختلف به صورت تصادفی بین گزینه ها توزیع شود