

Exam Analysis

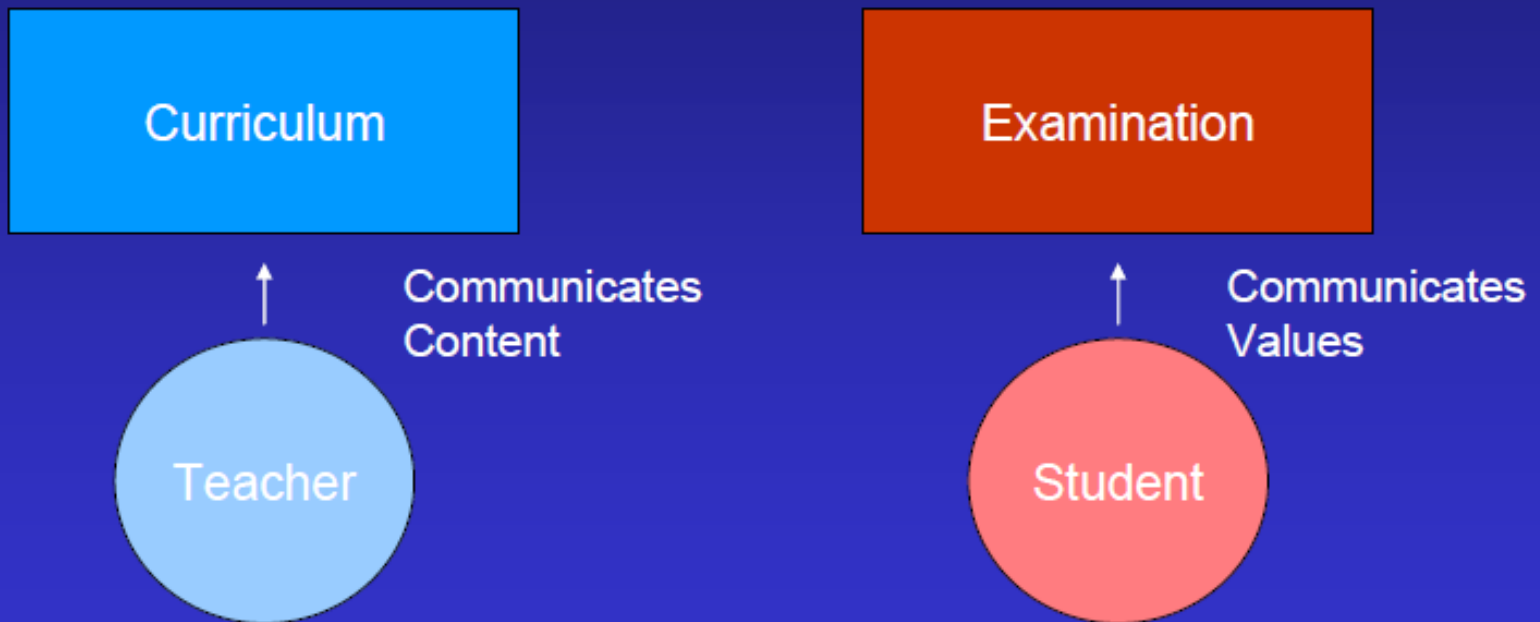
شہریور ۹۵

GHAFFARI, R MD, MMED, PhD

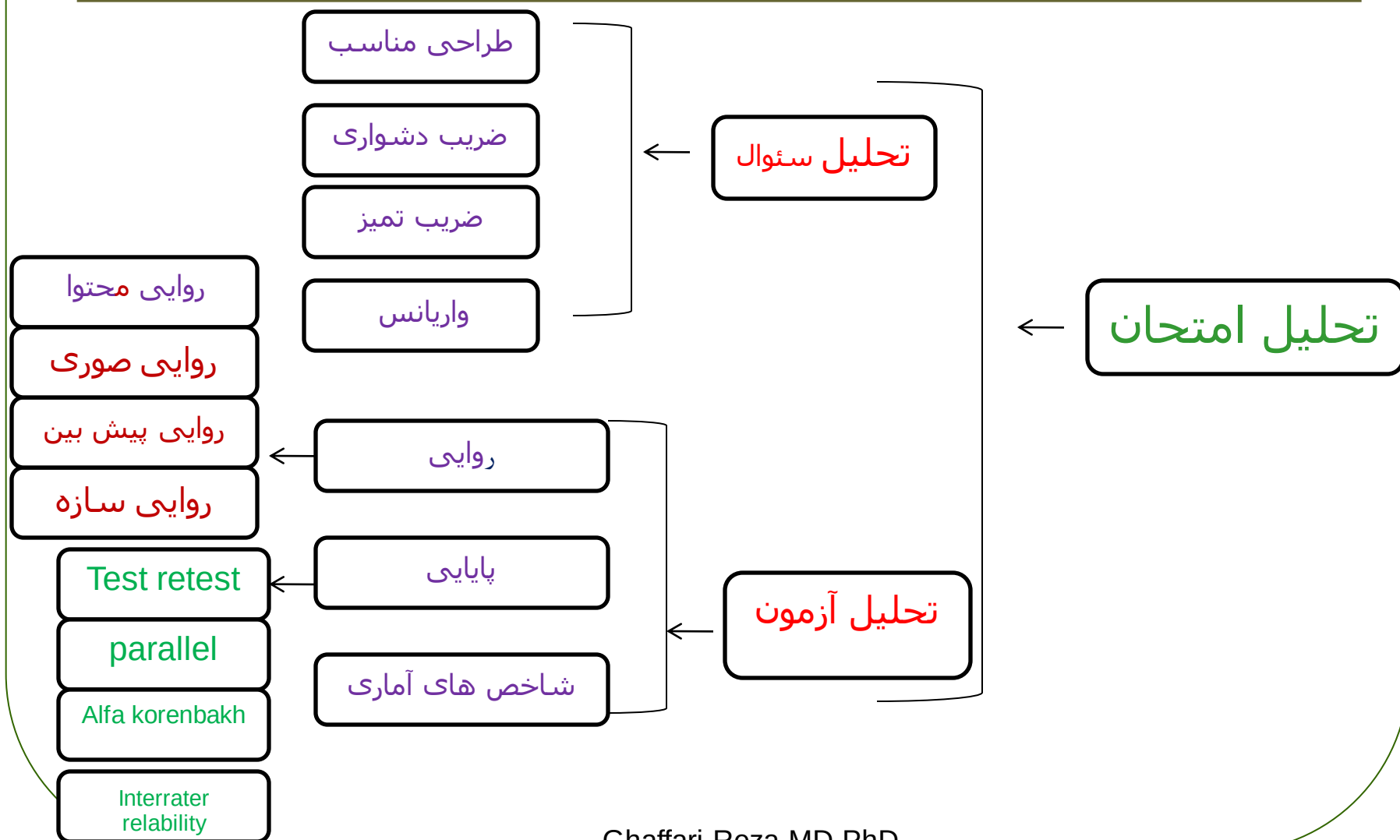
ghafarir@gmail.com

Ghaffari Reza MD PhD

A Teacher's View vs. A Student's View



● تحلیل امتحانات چند گزینه ای



طراحی مناسب (قبل از آزمون)

۱. آناتومی درست (هر چهار جزئ MCQ)
۲. به یکی از دو فرمت **One best answer** یا فرمت **ONLY one correct answer** باشد
۳. بر اساس جدول بودجه بندی آزمون (blueprint) باشد
۴. معیارهای میلن و بخصوص قانون **Cover the options rule** در آن صدق کند
۵. دارای راهنمایی های بیمورد به جواب درست نباشد
۶. واجد پیچیدگی و دشواری گنج کننده نباشد

(۱) آناتومی سؤالات MCQ

۱. تنه (stem)

۲. راهنمای سؤال یا قسمت ارتباط
دهنده (lead in)

۳. گزینه ها (options)

distracters

key



(Stem): A previously healthy 15-year-old boy has cramping periumbilical pain; after several hours, the pain shifts to the right lower quadrant and becomes constant. He vomits several times and is brought to the emergency department. The abdomen is tender on deep palpation of the right lower quadrant. Findings on the chest and abdominal x-ray films are normal. Leukocyte count is 15,000/mm³. Urinalysis shows 3 leukocytes/hpf.

(Lead-in): Which of the following is the most appropriate initial management?

Options:

- A. Supportive treatment at home
- B. Barium enema
- C. CT scan of the abdomen
- D. Intravenous pyelography and cystrography
- E. Surgical exploration of the abdomen^{*(Key)}

گزینه ها

- تعداد معمول گزینه های یک سؤال چند گزینه ای بین ۲ تا ۵ می باشد و هر چه تعداد گزینه ها بیشتر باشد احتمال دادن پاسخ شانس یا اتفاقی پایین می آید .
- طراحی تعداد زیاد گزینه های انحرافی در عمل مشکل بوده و از اینرو سؤالات چهار گزینه ای در بین طراحان از مقبولیت بیشتری برخوردار است

(۲) انواع سؤالات چند گزینه ای

۱. نوع تنها گزینه صحیح (The only correct answer):

در این نوع فقط یک پاسخ صحیح بوده و بقیه گزینه ها کاملاً نادرست می باشند

مثال یک: اگر توان سوم عددی ۲۷ باشد آن عدد کدامست؟

الف: دو

ب: سه

ج: شش

د: هشت

● مثال ۲ : کدامیک از هورمونهای زیر به طور مستقیم روی هسته سلول اثر می گذارد؟

الف : تیروکسین

ب : کورتیزول

ج : هورمون رشد

د: آدرنالین

اشکال مختلف سوالات از نوع تنها یک پاسخ صحیح:

1. C (A / B/ Both/ Neither items)
2. K (complex true/false items)
3. X (simple true/false items)
4. Simulations such as (PMPs)

Sample C-type Item

DIRECTIONS: For each numbered item, select the ONE lettered option that is most closely associated with it and fill in the circle containing the corresponding letter on the answer sheet. Each lettered option may be selected once, more than once, or not at all.

- A. *Plasmodium vivax* malaria
- B. *Plasmodium falciparum* malaria
- C. Both
- D. Neither

1. A combination of primaquine and chloroquine is treatment of choice for acute attack.
2. Clinical attacks suppressed by ingestion of chloroquine once a week while in an endemic area.
3. Permanently cured by treatment with chloroquine.
4. Infection prevented by ingestion of chloroquine once a week.

Sample K-type Item

Directions Summarized

Directions Summarized

A	B	C	D	E
1, 2, 3 only	1, 3 only	2, 4 only	4 only	All are correct

1. A child suffering from an acute exacerbation of Rheumatic fever usually has

- (1) an elevated sedimentation rate
- (2) a prolonged PR interval
- (3) an elevated antistreptolysin O titer
- (4) subcutaneous nodules

(X type true false)

Which of the following is/are X-linked recessive conditions?

- 1. Hemophilia A (T)
- 2. Cystic fibrosis (F)
- 3. Duchenne's muscular dystrophy (T)
- 4. Tay-Sachs disease (F)

The options can be diagramed as follows.

2

1

4

3

Totally Wrong

Totally Correct

Options

Options

true/false items must satisfy the following rules:

- Stems must be clear and unambiguous.
- Options must be absolutely true or false;

۲- نوع بهترین گزینه (one best answer)

- در این نوع سؤال همه گزینه ها صحیح بوده ولي يکي از آنها از بقیه صحیحتر يا بهتر است و از آزمون شونده انتظار مي رود آنها انتخاب کند.

مثال: دیابت از نظر ژنتیکی جز کدام دسته از بیماریهای زیر است؟

الف: تراتوژنیک ب: پلی ژنیک

ج: کروموزومی د: تک ژنی

در این سؤال گزینه های الف و ب هم صحیح هستند ولي گزینه ج بهترین پاسخ است

Stem

- A 32-year-old man has a 4-day history of progressive weakness in his extremities. He has been healthy except for an upper respiratory tract infection 10 days ago. His temperature is 37.8 C (100 F), blood pressure is 130/80 mm Hg, pulse is 94/min, and respirations are 42/min and shallow. He has symmetric weakness of both sides of the face and the proximal and distal muscles of the extremities. Sensation is intact. No deep tendon reflexes can be elicited; the plantar responses are flexor. (USMLE2004)



Lead-in

Which of the following is the most likely diagnosis?

options

- A. Acute disseminated encephalomyelitis
- B. Guillain-Barré syndrome
- C. Myasthenia gravis
- D. Poliomyelitis
- E. Polymyositis

D C A E B

Least
correct

Most
correct

۳- بودجه بندی سوالات
جدول مشخصات آزمون
(Blueprint)

اقدامات قبل از برگزاری آزمون

۱. تدوین اهداف

۲. تعیین حیطه و محتوی

۳. تدوین بودجه بندی سوالات

۴. انتخاب سوال

۵. طراحی سوال

۶.

Blueprint

- خاستگاه آن از صنعت و معماری می باشد.

تعريف

An assessment blueprint is a method for translating learning objectives into assessment practice.(2)

جدول مشخصات آزمون به روایی محتوایی آزمون دلالت می کند.

یک آزمون خوب پیشرفت تحصیلی آزمونی است که بهترین
شکل منعکس کننده تمامی هدفهای آموزشی و همه محتوای
برنامه درسی باشد. (روایی محتوایی)

مراحل تدوین بودجه بندی سوالات

- تعیین حیطه های ارزیابی و ابزار های آن
- تعیین اهداف یادگیری
- تعیین عناوین و درجه اهمیت و ضرورت و تاثیر آنها در وظایف شغلی برای هر یک از عناوین
- نمونه گیری از زیر عنوان ها
- تعیین وزن و تعداد سوال هر زیر عنوان
- تعیین تاکسونومی مورد انتظار برای هر سوال
- تعیین طراح سوال

چگونگی توزیع سوالها به بخشهای مختلف

آزمون ساز باید تعیین کند که سوالهای آزمون خود را به چه صورتی در میان هدفها و عنوانهای درس توزیع نماید.

روش های توزیع سوالات در بخشهای مختلف محتوی

- اهمیت نسبی هر یک از مطالب
(Must to know)
(Better to know)
(Nice to know)

- حجم یا تعداد صفحات امتحانی کتاب مرجع

- میزان ساعات تدریس یا زمان اختصاص داده شده برای

هر موضوع

- نوع فرا گیر و هدف از ارزیابی

چنانچه بخشهای مختلف محتوا از اهمیت یکسانی برخوردار باشند و در توزیع سئوالها در میان بخشهای مختلف تنها بخواهیم مقدار ساعتهای تدریس را برای هر بخش منظور کنیم ، در آن صورت می توانیم از فرمول زیر استفاده نمائیم :

تعداد کل ساعتهای تدریس / تعداد ساعتهایی که صرف تدریس بخش شده =
نسبت ساعت های تدریس هربخش

نحوه محاسبه تعداد سوالات آزمون

- (۱) در آزمونهای فینال و نهایی تعداد سوالات از قبل مشخص شده است
- (۲) در صورتیکه تعداد سوالات از قبل مشخص نیست برای تعیین تعداد سوالات میتوان از حاصل ضرب تعداد اهداف آموزشی کلی در عدد $۰/۷$ استفاده کرد
- (۳) معمولاً تعداد متوسط اهداف آموزشی کلی برای هر واحد درسی نظری بین ۶۰ تا ۱۰۰ متغیر بوده و از اینرو تعداد سوال چند گزینه ای مورد نظر بیت ۴۰ تا ۷۰ سوال خواهد بود

محاسبه زمان کل آزمون

- متوسط زمان پاسخگویی به سوالات چند گزینه ای با تاکسونومی یک حدود ۳۵-۴۰ ثانیه بوده و برای سوالات با تاکسونومی دو و سه به ترتیب ۴۵-۵۰ و ۵۵-۶۰ ثانیه در نظر گرفته میشود
- بنابراین زمان کل آزمون با توجه به تعداد سوالات با تاکسونومی ۱ و ۲ و ۳ قابل احتساب است
- در آزمونهای دوره های عمومی رشته های علوم پزشکی بهتر است ۳۵٪ سوالات با تاکسونومی یک ، ۴۰٪ با تاکسونومی دو و ۲۵٪ با تاکسونومی سه طرح گردد
- در آزمونهای دوره های تکمیلی و تخصصی این نسبت به ترتیب ۳۰، ۳۵ و ۳۵ درصد میباشد

- اگر بیش از یک منبع برای آزمون وجود دارد بسته به میزان اهمیت و محتوی منابع **تعیین درصد سوالات از هر منبع** آزمون را عادلانه تر میکند

استفاده از بیش از یک طراح سوال

- امنیت آزمون را افزایش میدهد
- امکان انتخاب سوالات با کیفیت و ساختار بهتر بیشتر میشود
- توزیع سوالات در زیر عنوان ها را افزایش میدهد

طراحی جدول مشخصات آزمون

جدول مشخصات آزمون یک جدول دوبعدی است با تعدادی ردیف و ستون ، متناسب با محتوا و هدفهای آموزشی درس.

در ردیف بالای جدول که بعد محتوا نام دارد اجزای مطالب درس نوشته می شوند و در ستون کنار آن که بعد هدف نامیده می شود فهرست هدفهای آموزشی قرار می گیرند.

طراحی مناسب (قبل از آزمون)

۱. آناتومی درست (هر چهار جزئ MCQ)
۲. به یکی از دو فرمت **One best answer** یا فرمت **ONLY one correct answer** باشد
۳. بر اساس جدول بودجه بندی آزمون (blueprint) باشد
۴. معیارهای میلن و بخصوص قانون **Cover the options rule** در آن صدق کند
۵. دارای راهنمایی های بیمورد به جواب درست نباشد
۶. واجد پیچیدگی و دشواری گیج کننده نباشد

جدول بودجه بندی سوال

به نام خدا

جدول پیشنهادی برای تعیین بودجه بندی سوال های امتحان

موضوع / درس مورد امتحان :

تاکسونومی ۳ (۲۵٪ - ۲۰٪):

تاکسونومی ۴ (۴۰٪ - ۵۰٪):

تاکسونومی ۱ (۳۵٪ - ۲۰٪):

تعداد کل سوال :

نوع آزمون :

تعداد تجمعی سوال	عنوان اصلی	زیر عنوان تخصصی	تعداد سوال	تاکسونومی ۱	تاکسونومی ۲	تاکسونومی ۳
	۱					
		(۲-۱)				
		(۲-۲)				
		(۲-۳)				
		(۲-۴)				
		(۲-.....)				
	۲					
		(۳-۱)				
		(۳-۲)				
		(۳-۳)				
		(۳-۴)				
		(۳-.....)				
	۳					
		(۳-۱)				
		(۳-۲)				
		(۳-۳)				
		(۳-۴)				
		(۳-.....)				

مثال بودجه بندی

به نام خدا

جدول پیشنهادی برای تعیین بودجه بندی سوال های امتحان

موضوع / درس مورد امتحان : اورولوژی

تاکسونومی ۳ : ۲۸

تاکسونومی ۲ : ۷۴

تاکسونومی ۱ : ۲۸

تعداد کل سوال : ۱۵۰

نوع آزمون : ارتقاء دستیاری

تعداد تجمعی سوال	عنوان اصلی	زیر عنوان تخصصی	تعداد سوال	تاکسونومی ۱	تاکسونومی ۲	تاکسونومی ۳
	1) Anatomy		۲		۲	
۱		1-1) Surgical anatomy of the retroperitoneum, adrenals, kidneys and ureters	۱		۱	
۲		1-2) Anatomy of the lower urinary	۱		۱	
	2) Clinical decision making		۲		۱	۱
۳		2-1) Evaluation of the urologic patient: history, physical examination,...	۲		۱	۱
	3) Basic of urologic surgery		۶	۱	۳	۲
۶		3-1) Core principles of perioperative care	۲		۱	۱
۷		3-2) Fundamentals of instrumentation and urinary tract drainage	۱		۱	
۸		3-3) Principles of endoscopy	۱		۱	
۱۰		3-4) Fundamentals of laparoscopic and robotic urology surgery	۲	۱	۱	
	4) Infections and inflammation		۸	۲	۵	۱
۱۳		4-1) Infection of the urinary tract	۳	۱	۱	۱
۱۵		4-2) Prostatitis and related conditions, orchitis and epididymitis	۲	۱	۱	
۱۶		4-3) Sexually transmitted infections	۱		۱	
۱۷		4-4) Urologic aspects of AIDS and HIV infection	۱		۱	
۱۸		4-5) Tuberculosis and other opportunistic infection of the genitourinary system	۱		۱	

جدول پیشنهادی برای تعیین بودجه بندی سوال های امتحانی

موضوع درس / درس مورد امتحان : ارتقاء گواهینامه گروه روانپزشکی دانشگاه علوم پزشکی تبریز

نوع آزمون: ارتقا دستیاری تعداد کل سوال: ۱۵۰ تاکسونومی ۱: ۴۱ تاکسونومی ۲: ۷۴ تاکسونومی ۳: ۳۵

تعداد تجمعی سوال	عنوان اصلی	زیر عنوان تخصصی	تعداد سوال	تاکسونومی ۱	تاکسونومی ۲	تاکسونومی ۳
	1) Neural Sciences		5	5		
1		1-1) Neural Development And Neurologogenesis	1	1		
2		1-2) Novel Neuro Transmitters	1	1		
3		1-3) Psychoneuro endocrinology	1	1		
4		1-4) Populudion Genetics and Genetic Epidemiology In Psychiatry	1	1		
5		1-5) Basic Science of Sleep	1	1		
	2) Neuropsychiatry and Behavioral Neurology		5		5	
6		2-1)Neuro Psychiatric Aspect of Cereloro vascular Disorders	1			
7		2-2) Neuro Psychiatric Aspect s of Epilepsy	1			
8		2-3) Neuro Psychiatric Concequences of Traumalties Brain Injury	1			
9		2-4) Neuro Psychiatric Aspects of Movement Disorders				
10		2-5) Neuro Psychiatric Aspects of Headache		1		
	3) Contribution of the Psychological Sciences					
11		3-1) Piaget And Cognitive Development				
12		3-2) Biology of Memory				

جدول پیشنهادی برای تعیین بودجه بندی سوال های امتحانی

موضوع درس / درس مورد امتحان

نوع آزمون: طب اورژانس، روزن ۲۰۱۰ تعداد کل سوال: تاکسونومی ۱ تاکسونومی ۲ تاکسونومی ۳

طراح سوال	عنوان اصلی	زیر عنوان تخصصی	تعداد سوال	تاکسونومی ۱	تاکسونومی ۲	تاکسونومی ۳
	1) Fundamental and clinical concepts	(۱-۱) Critical Management Principles: Chapter 6 and 7	3			
		(۲-۱) Cardinal Presentations: Chapter 12-35	5			
			8	2	4	2
	2) Trauma		24	6	12	6
		(۱-۲) General Concepts: Chapter 36-40	4			
		(۲-۲) System Injuries: Chapter 41-48	7			
		(۳-۲) Orthopedic Lesions: Chapter 49-58	10			
		(۴-۲) Soft tissue injury: Chapter 61 Mammalian bite	2			
		(۵-۲) Violence and Abuse: Chapter 65-69	1			

مشخصات آزمون (Blueprint) برای ارزیابی مهارت‌های بالینی

عنوان درس یا دوره: رشته و مقطع تحصیلی
فراگیران:

عناوین و محتوای یادگیری	توانمندیها یا مهارت‌های اصلی در مراقبت بالینی	مهارت بالینی	آزمون های مناسب
کلی	تعداد سوال یا درصد نمره	مسائل یا موارد بالینی	درصد نمره یا سوال
دستگاه گوارش		اقدامات عملی پیچیده و در شرایط دشوار	
		اقدامات عملی اختصاصی	
		اقدامات عملی عادی	
		ارزنیابی	
		پیگیری	
دستگاه عصبی		آموزش	
		اقدامات عملی	
		درمان	
		تشخیص	
		اقدامات عملی عادی	
		اقدامات عملی عادی	
		اقدامات عملی عادی	
		اقدامات عملی عادی	
		اقدامات عملی عادی	
		گرفتن شرح حال	

بودجه بندي ازمون OSCE

- تهیه بلو پرینت اولین گام برای اطمینان از روایی محتوای آزمون است.
- در بودجه بندي آزمون به دو جنبه باید دقت کرد:
- انواع موضوعات و محتواهایی که باید مورد سنجش قرار گیرد
- انواع مهارت هایی که باید مورد سنجش قرار گیرد

بودجه بندی برای OSCE

محتوای ایستگاه								مهارت های مورد ارزیابی
موضوع اساسی ۸	موضوع اساسی ۷	موضوع اساسی ۶	موضوع اساسی ۵	موضوع اساسی ۴	موضوع اساسی ۳	موضوع اساسی ۲	موضوع اساسی ۱	
								شرح حال
								معاینه
								مهارت های ارتباطی
								تفسیر اطلاعات و استدلال بالینی
								مهارت های عملی
								نگرش

خلاصه

یکی از ویژگیهای آزمون استاندارد روایی آن می باشد. بودجه بندی blueprint سوالات ابزار ارزیابی روایی محتوایی و عدالت آزمون می باشد.

علمی ترین و مطلوب ترین روش توزیع سوالات ،بودجه بندی بر اساس اهداف آموزشی و اهمیت نسبی این اهداف میباشد.

وفاداری به بودجه بندی در طی روند آزمون از طرف هیات ممتحنه یکی از شاخص های آزمون عادلانه است. پایش آن از طریق موسسه آموزشی ضرورت دارد.

طراحی مناسب (قبل از آزمون)

۱. آناتومی درست (هر چهار جزئ MCQ)
۲. به یکی از دو فرمت **One best answer** یا فرمت **ONLY one correct answer** باشد
۳. بر اساس جدول بودجه بندی آزمون (blueprint) باشد
۴. معیارهای میلن و بخصوص قانون **Cover the options rule** در آن صدق کند
۵. دارای راهنمایی های بیمورد به جواب درست نباشد
۶. واجد پیچیدگی و دشواری گیج کننده نباشد

۴- طراحی سؤالات چندگزینه ای

چک لیست میلمن برای طراحی سؤالات چند گزینه ای
مناسب

چک لیست میلمن جهت بررسی رعایت اصول ساختاری در سئوالات

چند گزینه ای:

سوال	بلی	خیر
۱- آیا بخش اعظم اطلاعات در ساقه سوال گنجانده شده است؟		
Cover the options rule		
۲ - آیا سوال یک هدف اختصاصی یادگیری را مورد ارزیابی قرار می-دهد؟		
۳ - آیا لغات استفاده شده در ساقه یا گزینه‌ها، شفاف و مستقیم بیان شده‌اند؟		
۴ - آیا از کاربرد گزینه منفی برای ساقه منفی خودداری شده است؟		
۵ - آیا از کاربرد گزینه‌های نظیر همه موارد هیچکدام و گزینه‌های ترکیبی خودداری شده است؟		
۶ - آیا از کاربرد گزینه‌های متضاد یکدیگر خودداری شده است؟		

چک لیست میلمن جهت بررسی رعایت اصول ساختاری در سئوالات چند گزینه ای:

سوال	بلی	خیر
۷- آیا لغات مثبت در ساقه سوال استفاده شده است یا در صورت منفی بودن ساقه سوال لغایت منفی مشخص شده اند؟		
۸- آیا هر سوال مستقل از سئوالات دیگر می باشد؟		
۹- آیا گزینه ها از نظر طول، ساختار لغوی و سبک نگارش هم سنگ هستند؟		
۱۰- آیا تا حد امکان از کاربرد ساختار لغوی و سبک نگارش هم سنگ هستند؟		
۱۱- آیا کلمات به کاررفته در ساقه و یا گزینه ها از نظر املائی صحیح هستند؟		
۱۲- آیا گزینه ها بطور عمودی لیست شده اند.		

اصول Millman در خصوص طراحی MCQ

۱- پایه یا ساقه سوال باید مسائل اصلی، هدف آموزشی مورد پرسش و کمیتها را در برگیرد.

۲- هر Item باید تا حد امکان کوتاه باشد (ضمن حفظ وضوح جملات).

۳- از ذکر سئوالات منفی در پایه حتی المقدور خودداری شود. در صورت انجام این امر زیر جمله منفی خط کشیده شود یا با حروف درشت نوشته شود.

اصول Millman در خصوص طراحی MCQ

۴- مطالب اصلی سوال را به طور کامل در تنه سوال بنویسید.

پایه یا ساقه سؤال باید به نحوی تنظیم شود که بدون کمک گرفتن از گزینه‌ها، بیان کننده مسئله اصلی باشد. در غیر اینصورت آزمون شونده، پیش از خواندن گزینه‌ها متوجه نخواهد شد که سوال درباره چه مطلبی طرح شده است.

گزینه‌ها نیز باید مرتبط به موضوع اصلی سوال و حتی المقدور مستقل از یکدیگر باشد. بکوشید تا همه گزینه‌های یک سوال متجانس و مربوط به موضوع واحدی باشند. در چنین حالتی سوال False-True در می‌آید که از دیگر گزینه‌ها مستقل است. نمونه معمول این گونه سوالها، سوالی است که تنه آن این جمله را شامل می‌شود: کدام یک از مطالب زیر درست است؟

اصول Millman در خصوص طراحی MCQ

۴- مطالب اصلی سوال را به طور کامل در تنه سوال بنویسید. گزینه ها نیز باید مرتبط به موضوع اصلی سوال باشد:

■ در مورد رشد و تمایز دوران جنینی کدام گزینه صحیح می باشد؟

الف) در طول سه ماهه سوم وزن جنین سه برابر و قد او دو برابر می شود.

ب) در هفته ۱۰ جنس دستگاه تناسلی خارجی بخوبی قابل تشخیص است.

ج) از هفته ۹ فقط تعداد سلولهای بدن جنین افزایش می یابد.

د) در انتهای هفته ۸ وزن جنین بطور متوسط ۹۰ گرم می باشد

■ در معاینه نوزاد نارس با اختلال رشد داخل رحمی ، ارگانومگالی، ایکتر، پتشی و کلسیفیکاسیون پری و نتریکولار در سی تی اسکن مغزی، گزینه صحیح را انتخاب کنید:

الف- انجام PCR از نمونه ادراری در روزهای اولیه بعد از تولد

ب- اندازه گیری و ثبت تیتراهای فزاینده IgG در طی سال اول زندگی بر عفونت مادرزادی دلالت میکند

ج- اندازه گیری IgM در تشخیص بیاری کودک تستی با حساسیت و ویژگی بالا ست

د- مثبت بودن IgG در سرم این افراد در تشخیص بیماری بسیار کمک کننده است

اصول Millman در خصوص طراحی MCQ

۵- بهترین پاسخ باید خواسته شود یا از عبارت بیشترین یا بهترین یا نزدیکترین در صورت سوال استفاده شود. (در صورتیکه بیش از یک پاسخ نسبتاً صحیح داشته باشد)

۶- در پایه سؤالاتی که جای خالی گذاشته می شود تا با یکی از گزینه ها پر شود، قسمت خالی حتی المقدور نباید ابتدای جمله گذاشته شود.

اصول Millman در خصوص طراحی MCQ

۷- دشواری‌های زبانی گزینه‌ها باید پایین باشد ، از ایجاد ابهام در بیان ساقه سوال و یا گزینه ها پرهیزید :

اگر فراگیران در جواب دادن به این نوع سوالها موفق نباشند معلوم نمی شود که علت شکست آنان ندانستن جواب سوال یا ناتوانی در درک متن سوال است.

میزان دشواری یک سوال باید مربوط به محتوا و هدف آموزشی مورد سنجش باشد نه به فهم کلمات و عباراتی که در آن به کار می روند.

اصول Millman در خصوص طراحی MCQ

۷- دشواری‌های زبانی گزینه‌ها باید پایین باشد. از بکار بردن دو قید یا فعل منفی در ساقه یک سوال و یا گزینه‌های منفی برای سوال با ساقه منفی اجتناب شود. (منفی مضاعف یا منفی در منفی نباشد)

■ MRA (MR آنژیوگرافی) در تمام موارد زیر کنترا اندیکاسیون دارد، بجز:

الف) گرافتهای عروقی

ب) شانت های اینتراکرانیال

ج) استیمولاتورهای طناب نخاعی

د) کاشت حلزون Cochlear Implant

اصول Millman در خصوص طراحی MCQ

۷- دشواری‌های زبانی گزینه‌ها باید پایین باشد. از بکار بردن دو قید یا فعل منفی در ساقه یک سوال و یا گزینه های منفی برای سوال با ساقه منفی اجتناب شود. (منفی مضاعف یا منفی در منفی نباشد)

■ در مورد فلج مغزی (Cerebral Palsy) نمی توان گفت که :

- الف) شایعترین علت آن آسیفیکسی زایمانی (Hypoxic Ischemic Encephalopathy) نیست.
- ب) در اغلب موارد همراه با عقب ماندگی ذهنی نیست.
- ج) اشکال حرکتی مرکزی و غیرپیشرونده بوده که معمولاً تا قبل از یک سالگی پدید نمی آید.
- د) تشنج در این بیماران بصورت همراه دیده نمی شود.

اصول Millman در خصوص طراحی MCQ

۸- با هر گزینه فقط یک نکته یا هدف آموزشی را باید مورد سؤال قرار داد.

هر سوال را به گونه ای انتخاب کنید که یک موضوع مهم آموزشی را اندازه بگیرد. در طرح هر سوال ، نویسنده سوال باید از خود بپرسد که آیا دانستن یا ندانستن پاسخ سوال در تعیین صلاحیت آزمون شونده، از لحاظ مطلب مورد آزمایش تاثیر دارد یا نه؟ اگر جواب این سوال مثبت بود آن گاه باید به گنجاندن موضوع در سوال اقدام کند .

در بیمارانی که آرام کردن (SEDATION) آنها هنگام تهویه مکانیکی مشکل است و به تناوب دچار آریتهاسیون و OVER SEDATION می شوند تجویز کدامیک از داروهای زیر را توصیه می کنید؟ (جراحی ارتقا

۸۹-تبریز)

الف) Dexmedetomidine

ب) Midazolam

ج) Propofol

د) Haloperidol

این سوال به نظر میرسد برای رزیدنت بیهوشی از اهداف مهم باشد. (نه برای رزیدنت جراحی عمومی)

قانون ۸- با هر گزینه فقط یک نکته یا هدف آموزشی را باید مورد سؤال قرار داد.

هر سوال را به گونه ای انتخاب کنید که فقط یک موضوع مهم آموزشی را اندازه بگیرد . بنابراین ، هر سوال باید یک هدف آموزشی را بسنجد یا شامل یک بخش مهم از محتوای درس باشد. اگر بیش از یک مطلب یا یک هدف در سوال مطرح شود، نه تنها سوال را پیچیده می کند بلکه اگر آزمون شونده در پاسخگویی به سوال موفق نباشد معلوم نیست که کدام یک از مطالب را ندانسته و علت عدم موفقیت او چه بوده است.

خانمی ۲۵ ساله در هفته ۶ حاملگی بعلت حملات تشنجی تحت درمان با فنی توئین قرار گرفته است در مورد مرحله جنینی و اثرات تراتوژنیک دارو کدامیک صحیح می باشد؟ (ارتقا زنان ۸۹- تبریز)

الف) بعلت مرحله Pre implantation هیچ خطری جنین را تهدید نمی کند

ب) بعلت مرحله organogenesis با بیشترین خطر ناهنجاری ساختمانی همراه می باشد

ج) در مرحله embryonic بوده و خطری جنین را تهدید نمی کند

د) بعلت مرحله Fetal امکان ایجاد آسیب در اعضای جنین وجود ندارد

اصول Millman در خصوص طراحی MCQ

۹- حتی المقدور از تکرار کلمات در گزینه‌ها خودداری شود مگر توالی منطقی وجود داشته باشد.

■ کودک ۲ ساله ای به علت اختلال رشد تحت بررسی قرار گرفته است. در معاینه فشار خون ۹۰/۱۴۰ میلی متر جیوه بوده و پتاسیم 2.8 meq/Lit میباشد. کدام یک از تشخیص‌های زیر بیشتر مطرح می باشد؟

الف- سندرم Liddle

ب- سندرم Barter

ج- سندرم Gitelman

د- سندرم Dent

اصول Millman در خصوص طراحی MCQ

۱۰- گزینه های انحرافی باید منطقی و جالب توجه باشد.

(برای اینکه پایه سؤال بتواند درک و فهم واقعی از اهداف آموزشی را اندازه گیری نماید، گزینه ها باید ظاهری منطقی داشته باشند و در یک بررسی سطحی درست جلوه کنند).

۱۱- تمام گزینه ها از نظر دستور زبان و اصول گرامر باید مطابق با پایه سؤال باشد یعنی اگر پایه سؤال جمع است گزینه ها نیز همه جمع باشند و ...

این ناهماهنگی در رابطه با گزینه های غلط بیشتر به چشم می خورد .

این ناهماهنگی در رابطه با گزینه های غلط بیشتر به چشم می خورد . . .

■ کودکی با تشخیص کاوازاکی تحت درمان ایمونوگلوبولین وریدی با دوز 2g/kg قرار گرفته است. در واکسیناسیون وی . . . :

- الف- واکسن ویروسی زنده را ۳ ماه باید تعویق انداخت.
- ب- واکسن ویروسی زنده را ۸ ماه باید تعویق انداخت.
- ج- واکسن ویروسی زنده را ۱۱ ماه باید تعویق انداخت.
- د- نیازی به تغییر در برنامه واکسیناسیون ندارد.

اصول Millman در خصوص طراحی MCQ

۱۲- همه گزینه ها (اصلی و انحرافی) از نظر طول جمله، دشواری فنی و کاربردی یکسان باشند.

■ نوزاد ۱۰ روزه به دنبال افتادن بند ناف مایع شفاف زرد رنگی با حجم نسبتاً زیادی از نافش خارج می شود. دور ناف اریتماتو نیست. کدام یک از اقدامات زیر برای رفع علایم وی مناسبتر است ؟

الف- سوزاندن محل پایه بند ناف با نیترات نقره

ب- شروع دو آنتی بیوتیک وسیع الطیف وریدی

ج- سیستوگرافی

د- انجام باریوم انما

اصول Millman در خصوص طراحی MCQ

۱۳- پایه و گزینه‌ها باید از نظر زبان نگارش، قواعد دستوری، محتوای موضوعی و شکل ظاهری یکنواخت و همگن باشد.

۱۴- از توالی منظم پاسخ صحیح در مجموعه سؤالات امتحانی خودداری شود.

(جواب صحیح سؤالات متوالی به ترتیب الف، ب، ج، د نباشد
یا اکثریت با جواب ج نباشد)

اصول Millman در خصوص طراحی MCQ

۱۵- بازای هر موضوع حداقل ۴ گزینه داشته باشید: یک گزینه صحیح و سه گزینه انحرافی.

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

■ تست های سه گانه سرمی آلفا فتوپروتئین ، β -HCG و استریول غیر کونژگه در سه ماهه دوم حاملگی در کدامیک از خانمهای حامله باید انجام شود؟ (ارتقا ۸۹ کودکان - تبریز)

الف) خانم های حامله بالاتر از ۳۵ سال

ب) خانم های حامله زیر ۳۵ سال

ج) خانم هایی که قبلا یک فرزند مبتلا به سندرم Down داشته اند

د) کلیه خانم های باردار

اصول Millman در خصوص طراحی MCQ

۱۵- بازای هر موضوع حداقل ۴ گزینه داشته باشید: یک گزینه صحیح و سه گزینه انحرافی.

تا حد امکان از کاربرد گزینه "هیچ یک از موارد بالا" پرهیز کنید. صاحب نظران سنجش و اندازه گیری استفاده از گزینه "هیچکدام" یا "همه موارد فوق" را پیشنهاد نمی کنند.

■ در بیماران تحت درمان با داروهای ایمنوساپرسیو، عفونت HIV، نقص ایمنی مادرزادی و سندرم نفروتیک تزریق کدامیک از واکسن های پنوموکوک زیر توصیه می شوند؟

(ارتقا ۸۹ کودکان-تبریز)

الف) واکسن پنوموکوک ۲۳ ظرفیتی

ب) واکسن پنوموکوک ۲۳ ظرفیتی + واکسن پنوموکوک ۷ ظرفیتی

ج) هیچکدام

د) واکسن پنوموکوک ۷ ظرفیتی

اصول Millman در خصوص طراحی MCQ

۱۵- بازای هر موضوع حداقل ۴ گزینه داشته باشید: یک گزینه صحیح و سه گزینه انحرافی.

دو گزینه متضاد را که یکی از آنها درست است به کار نبرید. در این حالت احتمالاً آزمون شونده انتخاب خود را به یکی از دو گزینه محدود خواهد کرد زیرا تصور می کند که ممکن نیست هر دو پاسخ غلط باشند. بهتر است یا به طور کلی از کاربرد متضادها صرف نظر شود یا اینکه دو زوج متضاد برای سوالهای چهار گزینه ای مورد استفاده گیرد.

■ تغییرات PH مایع مغزی نخاعی نسبت به خون در کدامیک از موارد زیر مقاومتر است؟

- گزینه های بهتر:
- الف) آلكالوز متابوليك
- ب) اسيدوز متابوليك
- ج) آلكالوز تنفسي
- د) اسيدوز تنفسي

- الف) آلكالوز متابوليك
- ب) هيپوناترمي
- ج) اسيدوز متابوليك
- د) هيپوكلسمي

اصول Millman در خصوص طراحی MCQ

۱۶- از به کار بردن عباراتی که به نحوی تکرار یا تشابه بین پایه و سؤال باشد، باید خودداری کرد.

■ عصب رادیال به کدامیک از عضلات زیر عصب می دهد؟

الف) گاستروکنمیوس

ب) سولئوس

ج) براکیورادیالیس

د) تیبیالیس قدامی

اصول Millman در خصوص طراحی MCQ

۱۷- از بکاربردن عین عبارت کتاب خودداری شود.

۱۸- از بکار بردن پایه سئوالاتی که پاسخ به سئوال دیگری در همان آزمون است، خودداری شود.

اصول Millman در خصوص طراحی MCQ

۱۹- گزینه‌ها نباید مترادف یا شامل یکدیگر بوده یا همپوشانی داشته باشند.

۲۰- از بکاربردن شاخصهای قطعی، معلوم و خاص مثل همیشه و یا هرگز خودداری شود.

۲۱- در پرسش راجع به فهم و درک یک اصطلاح یا مفهوم، ابتدا اصطلاح را ارائه نمود و سپس با یک سری مشخصه و تعاریف گزینه‌ها را انتخاب نمود.

علاوه بر این ۲۱ قانون

۲۲- هیچکدام از موارد فوق برای پرسشهای ریاضی یا محاسبه محض الزاماً مفید نیست.

۲۳- استفاده از محتوای زیاد در گزینه ها خودداری شود.

۲۴- از نوشتن پرسش به صورت ترتیبی مثل TEXT خودداری شود.

طراحی مناسب (قبل از آزمون)

۱. آناتومی درست (هر چهار جزئ MCQ)
۲. به یکی از دو فرمت **One best answer** یا فرمت **ONLY one correct answer** باشد
۳. بر اساس جدول بودجه بندی آزمون (blueprint) باشد
۴. معیارهای میلن و بخصوص قانون **Cover the options rule** در آن صدق کند
۵. دارای راهنمایی های بیمورد به جواب درست نباشد
۶. واجد پیچیدگی و دشواری گنج کننده نباشد

تاکسونومی سؤالات چند گزینه ای

سطوح یادگیری در حیطه شناختی



Cognitive domain evaluation

- Recognition recall
- Interpretation
- Problem solving

Taxonomy 1

- Knowledge
- Remember
- Recall
- Recognition

تاکسونومی یک (محفوظات)

شایع ترین تومور مغزی کدام است ؟

الف - مینژیوما

ب - استروسیتوما

ج - همانژیوما

د - اولیگودندروگلیوما

۳۵-۴۵ درصد از مجموع سوالات را تشکیل میدهند

زمان جواب دادن ۳۰-۳۵ ثانیه

تاکسونومی یک (محفوظات)

« سطوح ارزشیابی دانشجو » Recognition Recall



مربوط به پله اول حیطه شناختی (Knowledge)

Taxonomy 2

- درک مطلب
- تفسیر مطلب
- آنالیز مطالب
- تشخیص

تاکسونومی دو (تشخیص و تفسیر)

نوجوان ۱۵ ساله ای پس از برگشتن از مسابقه فوتبال شکایت از درد شدید در بیضه سمت چپ و شکم همراه با تهوع و استفراغ به مدت ۳ ساعت دارد. در معاینه شکم نرم و سمع طبیعی داشت و محتوی کیسه بیضه سفت، دردناک و کبود بود. بعلت درد، معاینه حلقه اینگوینال میسر نشد. گرافی ساده شکم و آزمایش کامل ادرار طبیعی بود تشخیص چیست؟

الف - اورکیت

ب - فتق اینگوینال مختنق

ج - تورشن تستیکول

د - هماتوم ناشی از ضربه

۲۰-۳۰ درصد از سوالات و زمان پاسخ دهی ۴۵-۵۰ ثانیه

تاکسونومی دو (تشخیص و تفسیر)

تفسیر اطلاعات بمنظور کاربرد ، توانایی تفسیر و بکارگیری دانسته‌های قبلی
در موقعیت جدید.

مثال:

کاربرد - عمل می‌کند - آماده می‌سازد.

محاسبه می‌کند - ارتباط می‌دهد.

نمایش می‌دهد - معاینه می‌کند.

آزمایش می‌کند - تفسیر می‌کند.

نشان می‌دهد - تولید می‌کند.

مربوط به پله سوم حیطه شناختی Interpretation for Application

Taxonomy 3

- ارزشیابی
- قضاوت (بالینی)
- حل مساله
- قدرت و توان پیش بینی
- قدرت تصمیم گیری در تنظیم برنامه درمانی
- قدرت ارایه راه حل‌های عملی

توکسونومی ۳ (تصمیم گیری)

چه تصمیمی برای این جوان میگیرید ؟
الف - تجویز مسکن و پیگیری تا ۲۴ ساعت

ب - دتورشن فوری بیضه

ج - اورکیکتومی الکتیو

د - هر نیورافی فوری

۳۰-۳۵ درصد سوالات را تشکیل میدهند.

زمان پاسخ دهی ۵۵-۶۰ ثانیه

توکسونومی ۳ (تصمیم گیری)

توانایی قضاوت براساس معیارهای معین Problem Solving

مثال:

ارزیابی - برآورد
تشخیص - انتخاب
مقایسه - انتقاد
تخمین - رتبه بندی
ارزیابی - قضاوت
اندازه گیری - تجدید نظر
جداسازی - رده بندی

مربوط به پله آخر حیطه شناختی Evolution

مثال

- یافته های گرافی PA قفسه سینه چیست؟ تفسیر کنید.
- این گرافی از قفسه سینه مربوط به آقای ۸۰ ساله معتاد به تر یاک و سیگار که در حمام لیز خورده و شکایت از درد در سینه و تنگی نفس بمدت ۴ ساعت دارد ، چه تصمیمی برای وی میگیرید ؟

یادآوری مهم

- بعضی از سوالات ظاهری شبیه تاکسونومی ۳ دارد در حالیکه فقط محفوظات را می سنجند:
- جوان ۲۲ ساله ای شکایت از چند توده بدون درد در ناحیه سوپرا کلاویکولار سمت چپ به مدت ۴ ماه دارد در معاینه ۴ توده سفت و به هم چسبیده بدون درد در لمس بطور متوسط در اندازه های ۳ در ۴ سانتی متر یافت شد . در صورتیکه جواب بیوپسی از ضایعات لنفوم هوجکین باشد کدامیک از انواع ذیل پیش اگهی بهتری دارد ؟

یاد آوری مهم

- بعضی سوالات ظاهری شبیه تاکسونومی ۱ دارند ولی نیاز به تجزیه و تحلیل دارند !!!
- در معاینات بدو استخدام فردی که در معدن سنگ آهن فعالیت می کند چه تستهایی را پیشنهاد می کنید؟

طراحی مناسب (قبل از آزمون)

۱. آناتومی درست (هر چهار جزئ MCQ)
۲. به یکی از دو فرمت **One best answer** یا فرمت **ONLY one correct answer** باشد
۳. بر اساس جدول بودجه بندی آزمون (blueprint) باشد
۴. معیارهای میلن و بخصوص قانون **Cover the options rule** در آن صدق کند
۵. دارای راهنمایی های بیمورد به جواب درست نباشد
۶. واجد پیچیدگی و دشواری گیج کننده نباشد

(۵) تجزیہ و تحلیل کمی سؤالات آزمون:

تجزیه و تحلیل کمی سؤالات آزمون:

- محاسبه ضریب دشواری سؤال (P)
- محاسبه ضریب تمیز سؤال (d)
- واریانس سؤال
- ضریب همبستگی رشته ای دو نقطه ای
- *IRT (item response theory)*

تعريف ضريب دشواري سؤال (*Difficulty Index*)

درصد کل آزمون شوندگاني است که به يك
سؤال جواب درست مي دهند و با حرف P
نمايش داده مي شود.

محاسبه ضریب دشواری سؤال

الف : اگر در آزمونی کلیه برگه های امتحانی دخالت داده شوند برای محاسبه ضریب دشواری سؤال کافی است تعداد کل افرادی را که به آن سؤال جواب درست داده اند را به تعداد کل آزمون شوندگان تقسیم کنیم (دقیقترین روش)

$$P = \frac{R}{T}$$

گامهاي محاسبه ضريب دشواري

الف: دانشجويان را بر اساس نمرات آزمون
بترتيب Rank و دسته بندي مي نماييم.

ب: دانشجويان را بعد از رتبه بندي به دو گروه
قوي (بالا) و ضعيف (پايين) بشرح اسلايد
بعدتقسيم مي کنيم.

۱) اگر تعداد کل دانشجویانی که در آزمون شرکت کرده اند کمتر یا مساوی ۲۰ نفر باشد کل آنها را به دو گروه بالا و پایین ده نفری تقسیم نمایید.

۲) اگر تعداد کل دانشجویان بین ۲۰ تا ۴۰ نفر باشد ۱۰ برگه بالا و ۱۰ برگه پایین را انتخاب کنید.

۳) اگر تعداد کل آزمون شوندگان بیش از ۴۰ نفر باشد بهترین رقم برای گروه بالا و پایین ۲۷% کل برگه ها خواهد بود.

تذکر:

شرایط فوق الذکر در صورتی است که توزیع
نمرات **هنگار** باشد اما در آزمونهای غیر هنگار
و کلاس هر يك از ارقام گروههای بالا یا پایین
بهتر است ۳۳ درصد انتخاب شود. (نیتکو ۱۹۹۳)

مثال

- در یک آزمون چهار گزینه ای با ۴۰ سوال که از ۶۰ نفر از دانشجویان دوره اکسترنی بعمل آمد نصف دانشجویان به تمام سئوالات پاسخ صحیح دادند به نظر شما ضریب دشواری آزمون کدامیک از موارد زیر است؟

الف: ۵/۰

ب: ۶۶/۰

ج: ۲/۰

د: هیچکدام

• در یک آزمون چهار گزینه ای با ۴۰ سوال که از ۶۰ نفر از دانشجویان دوره اکسترنی بعمل آمد نصف دانشجویان به سئوال دوم پاسخ صحیح دادند به نظر شما ضریب دشواری آن سئوال کدامیک از موارد زیر است؟

الف: ۵/۰

ب: ۶۶/۰

ج: ۲/۰

د: هیچکدام

نکته مهم

محاسبه میانگین ضریب دشواری برای کل
آزمون بدون توجه به ضریب دشواری تک
تک سوالات به تنهایی ارزشی ندارد

ب - اگر تعداد کل آزمون شوندگان زیاد است آنها را به دو گروه بالا و پایین مطابق دستورالعمل فوق الذکر تقسیم و ضریب دشواری سئوال را بشرح زیر محاسبه می نماییم.

$$P = \frac{H + L}{n + n'}$$

$$n = n' \Rightarrow P = \frac{H + L}{2n}$$

H=تعداد کسانی که از گروه بالا به سئوال پاسخ صحیح داده اند

L=تعداد کسانی که از گروه پایین به سئوال پاسخ صحیح داده اند

n=تعداد کل افراد گروه بالا

n'=تعداد کل افراد گروه پایین

سئوال

- حداکثر و حداقل مقدار ضریب دشواری یک سئوال چقدر می تواند باشد؟
- در سئوالات آسان و مشکل انتظار دارید P چه مقدار عددی را بخود اختصاص دهد؟

بر اساس توضیحات بالا می توان نتایج زیر را استنباط کرد:

۱. ضرب دشواری هر سؤال عددي است بين **صفر** و **يك**
۲. هر اندازه ضرب دشواری يك سؤال بزرگتر باشد (به يك نزديكتر باشد) سؤال **آسانتر** است.
۳. هر اندازه ضرب دشواری يك سؤال كوچكتر باشد (به صفر نزديك باشد) ، سؤال **دشووارتر** است.

مثال: ضریب دشواری سئوال زیر را محاسبه نمایید

● نام دیگر ویتامین B1 کدامست؟

الف) ریبوفلاوین ب) تیامین * ج) کوبالامین د) پیرویدوکسین

بدون پاسخ	د	ج	ب	الف	
۱	۰	۳	۶	۰	گروه بالا
	۰	۳	۲	۵	گروه پایین

مثال

- اگر در آزمونی ۶۰٪ دانشجویان گروه قوی و ۳۰٪ از دانشجویان گروه ضعیف به سؤال مورد نظر جواب درست داده باشند ضریب دشواری این سؤال را محاسبه نمایید

محاسبه واریانس سؤال

واریانس سؤال $V_Q = P \times (1 - P)$

مثال $P = 0.2 \Rightarrow V_Q = (0.2)(0.8) = 0.16$

از نظر تئوریک بهترین ضریب دشواری برای یک سؤال ... می باشد که در آن صورت واریانس سؤال برابر خواهد بود.

حداکثر واریانس سؤال فوق:

$$V_Q = (0.5)(0.5) = 0.25$$

تفسير ضريب دشواري

سؤال

در آزمونهاي وابسته به هنجار (Normative) هر چه پراکندگي نمرات بيشتر باشد آزمون بهتري ميباشد

به عبارت ديگر هر چه واريانس نمرات حاصل از يك آزمون وابسته به هنجار بزرگتر باشد آن آزمون مطلوبتر مي باشد.

پس از نظر تئوریک:

بهترین شرایط در يك آزمون این است که همه سؤالات يك آزمون با یکدیگر همبستگی کامل داشته باشد و ضریب دشواری همه آنها 0.5 باشد یعنی نیم بالایی آزمون شوندگان در آزمون نمره کامل و نیم پایینی دیگر نمره صفر بگیرند.

سوال

به نظر شما مطلوبترین محدوده عددی ضریب دشواری
سوال چقدر است؟

سطح موفقیت با حدس زدن؟

- میزان احتمال انتخاب پاسخ درست در **یک** سوال چند گزینه ای

سطح موفقیت با حدس زدن برای سئوالهای **چهار** گزینه ای است

سطح موفقیت با حدس زدن برای سئوالهای **پنج** گزینه ای است

در آزمونهاي چند گزینه اي سطح بهينه دشواري سؤال اندكي کمتر از
وسط فاصله بين يك و سطح موفقیت با حدس زدن است.

سطح موفقیت با حدس زدن براي سؤالهاي ۴ گزینه اي $0.25/0$ است.

سطح موفقیت با حدس زدن براي سؤالهاي ۵ گزینه اي $0.2/0$ است.

پس سطح دشواري مناسب سؤال در آزمونهاي ۴ گزینه اي برابر خواهد
بود با :

$$P_i = \frac{1 + 0.25}{2} = 0.625$$

در آزمونهاي ۴ گزینه اي

$$P_i = \frac{1 + 0.2}{2} = 0.6$$

در آزمونهاي ۵ گزینه اي

خلاصه:

ضریب دشواری بین $0/3$ تا $0/7$ حداکثر اطلاع را درباره تفاوت بین آزمودنیها بدست می دهند.

تبصره ■

در شرایط ویژه ضریب دشواری مطلوب فرق خواهد کرد.

- مثلاً در آزمونی که به صورت آزمون ورودی دکتری برای انتخاب بهترین داوطلبان است و حدوداً ۵٪ بالای افراد انتخاب خواهند شد باید سئوالات دشوار بکار رود

P=0.2

- در مقابل ممکن است در آزمونی سئوالات آسان برای تشویق دانشجویان مورد استفاده قرار گیرد (**P=0.9**).

Relating Item Difficulty to Testing Purpose

difficulty index (p)

30-70% =acceptable

50-65% = ideal

more than 70% =very easy

less than 30% =very difficult.

Items ->	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	T
Key	b	c	a	d	d	a	a	b	b	b	a	d	c	c	a	a	a	b	d	b	c	21
s1	b	c	a	a	d	b	a	b	b	b	a	d	c	c	a	c	a	b	d	b	c	18
s2	b	c	b	a	d	c	a	b	b	c	a	d	b	b	a	b	a	b	d	b	c	14
s3	c	c	b	a	d	d	c	b	a	c	a	c	a	a	c	a	c	b	d	b	c	9
s4	a	c	b	a	d	d	a	d	a	c	d	d	d	d	b	a	a	b	d	b	c	10
s5	a	c	b	a	c	c	a	d	b	b	c	a	d	c	c	a	b	d	a	a	a	6
s6	c	c	b	a	c	b	b	d	b	b	b	c	d	c	b	c	b	d	d	a	c	6
s7	b	c	b	a	c	b	d	d	c	b	a	c	c	c	b	b	b	c	a	a	c	7
s8	b	c	a	a	d	c	a	b	b	b	a	d	c	c	a	a	a	b	d	b	c	20
Diff Level																						
Disc Index																						

ضریب تمیز سئوال (discriminative Index)

- عبارت است از قدرت سئوال در تمایز گذاری یا تشخیص بین گروه قوی و گروه ضعیف آزمون شوندگان.
- ضریب تمیز سئوال مشخص می کند، که تا چه اندازه سئوال توانسته گروه قوی را از گروه ضعیف جدا سازد
- ضریب تمیز سئوال را با حرف d نمایش می دهند.

$$d = \frac{H - L}{n}$$

H = تعداد کسانی که در گروه بالا به سؤال پاسخ صحیح داده اند
 L = تعداد کسانی که در گروه پایین به سؤال پاسخ صحیح داده اند
 n = تعداد افراد يك گروه (بالا یا پایین)

پس ضریب تمیز هر سؤال عددی است بین و

تفسیر ضریب تمیز سوال

- هر قدر ضریب تمیز يك سوال بیشتر باشد قوه تمیز سوال **بیشتر** بوده و سوال **مناسبتري** است.
- مثال: اگر ضریب تمیز سوالی 0.9 باشد نشان می‌دهد که سوال بخوبی توانسته بین افراد رده بالا و پایین تمایز ایجاد کند.
- ضریب تمیز **صفر** حاکی از این است که سوال نتوانسته هیچ تمایزی بین گروه بالا و پایین ایجاد کند.
- ضریب تمیز **منفی** نشان می‌دهد که در آن سوال گروه قوی بدتر و ضعیفتر از گروه ضعیف عمل کرده‌اند که نشانگر معایب اساسی در سوال بوده و چنین سوالی باید حذف گردد.

پرسش: اگر سوال داراي اشكال فني نبوده ولي ضريب
تميز منفي داشته باشد نشانگر چیست؟

- ۱) اشكال عمده در سوال
- ۲) افراد گروه بالا آن سوال را غلط آموخته بودند.
- ۳) احتمال تقلب در آزمون
- ۴) احتمال نادرست بودن كليد آزمون
- ۵) اشكال عمده در تدريس

Discrimination Index (D)

The bigger the discrimination # the better

- ❖ **Excellent: > 0.35**
- ❖ **Good: $0.25 - 0.35$**
- ❖ **Acceptable: $0.15 - 0.25$**
- ❖ **Deleted: < 0.1**

مثال

- اگر در آزمونی ۶۰٪ دانشجویان گروه قوی و تنها ۳۰٪ از دانشجویان گروه ضعیف به سؤال مورد نظر جواب درست داده باشند ضریب تمیز این سؤال را محاسبه فرمایید

Quick Reference?

- Use the following table as a guideline to determine whether an item (or its corresponding instruction) should be considered for revision.

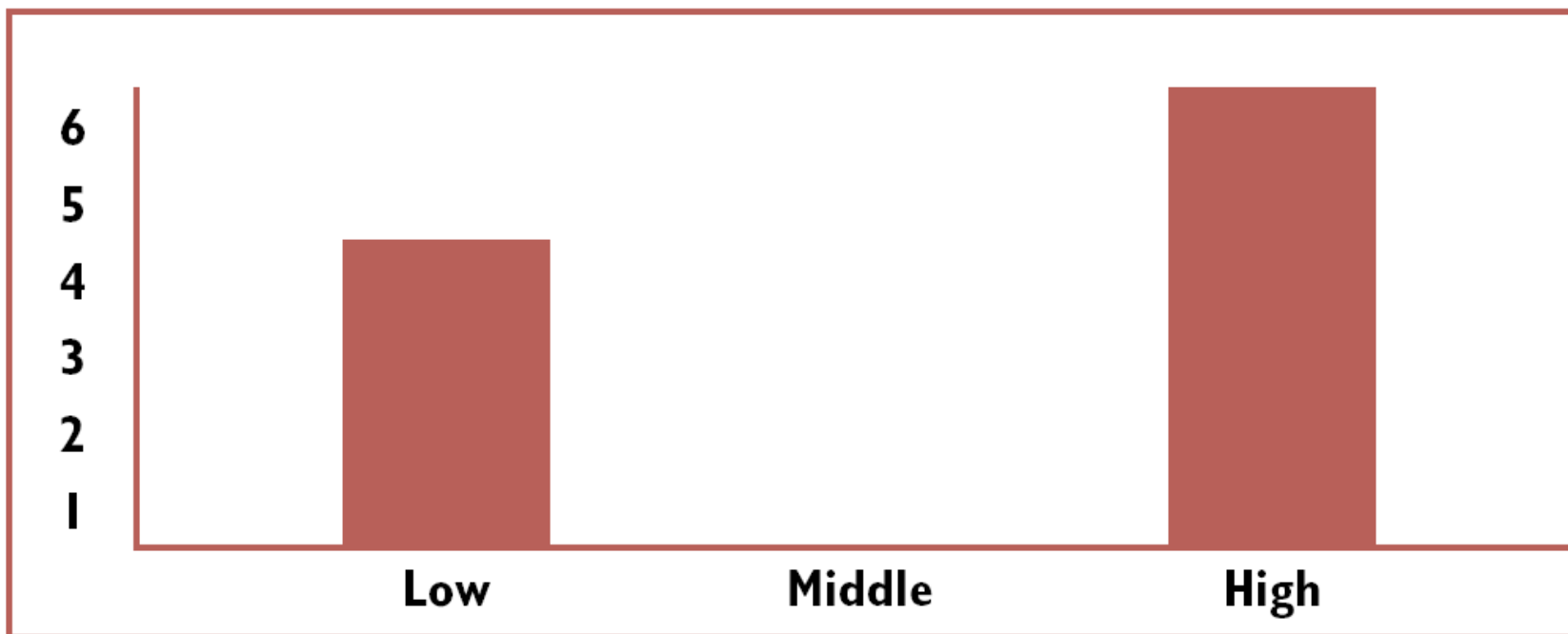
Item Discrimination (D)	Item Difficulty		
	High	Medium	Low
$D \leq 0\%$	<i>review</i>	<i>review</i>	<i>review</i>
$0\% < D < 30\%$	<i>ok</i>	<i>review</i>	<i>ok</i>
$D \geq 30\%$	<i>ok</i>	<i>ok</i>	<i>ok</i>

Items ->	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	T
Key	b	c	a	d	d	a	a	b	b	b	a	d	c	c	a	a	a	b	d	b	c	21
s1	b	c	a	a	d	b	a	b	b	b	a	d	c	c	a	c	a	b	d	b	c	18
s2	b	c	b	a	d	c	a	b	b	c	a	d	b	b	a	b	a	b	d	b	c	14
s3	c	c	b	a	d	d	c	b	a	c	a	c	a	a	c	a	c	b	d	b	c	9
s4	a	c	b	a	d	d	a	d	a	c	d	d	d	d	b	a	a	b	d	b	c	10
s5	a	c	b	a	c	c	a	d	b	b	c	a	d	c	c	a	b	d	a	a	a	6
s6	c	c	b	a	c	b	b	d	b	b	b	c	d	c	b	c	b	d	d	a	c	6
s7	b	c	b	a	c	b	d	d	c	b	a	c	c	c	b	b	b	c	a	a	c	7
s8	b	c	a	a	d	c	a	b	b	b	a	d	c	c	a	a	a	b	d	b	c	20
Diff Level																						
Disc Index																						

يكي ديگر از راههاي محاسبه ضريب تميز
سوالهاي يك آزمون محاسبه ضريب همبستگي
بين نمره آزمون شونندگان در آن سوال با
نمرات آنان در كل آزمون است.

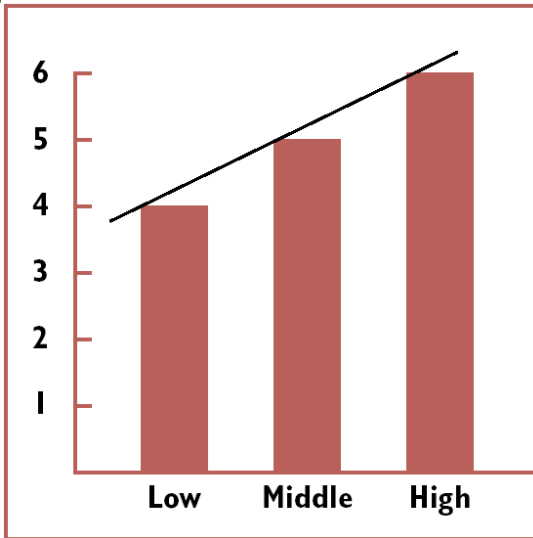
Correct answer responses for item 1

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
1	0	1	0	1	1	1	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	1	1	1	1	1	1	15

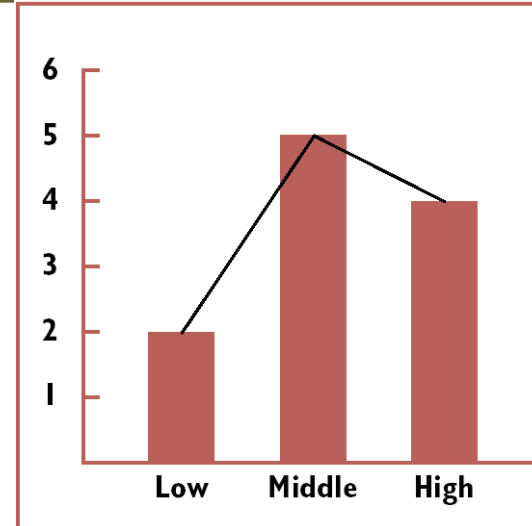


Correct answer responses for L, M and H groups on items 1, 4, 19, and 21

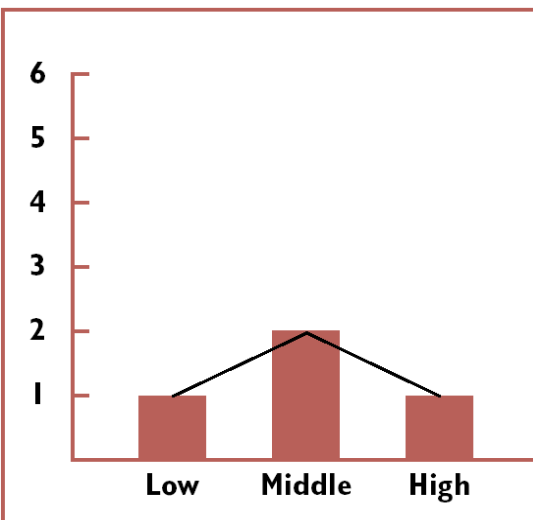
Item 1



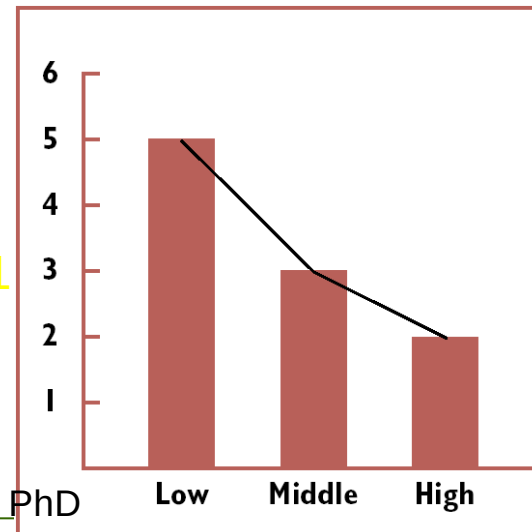
Item 4



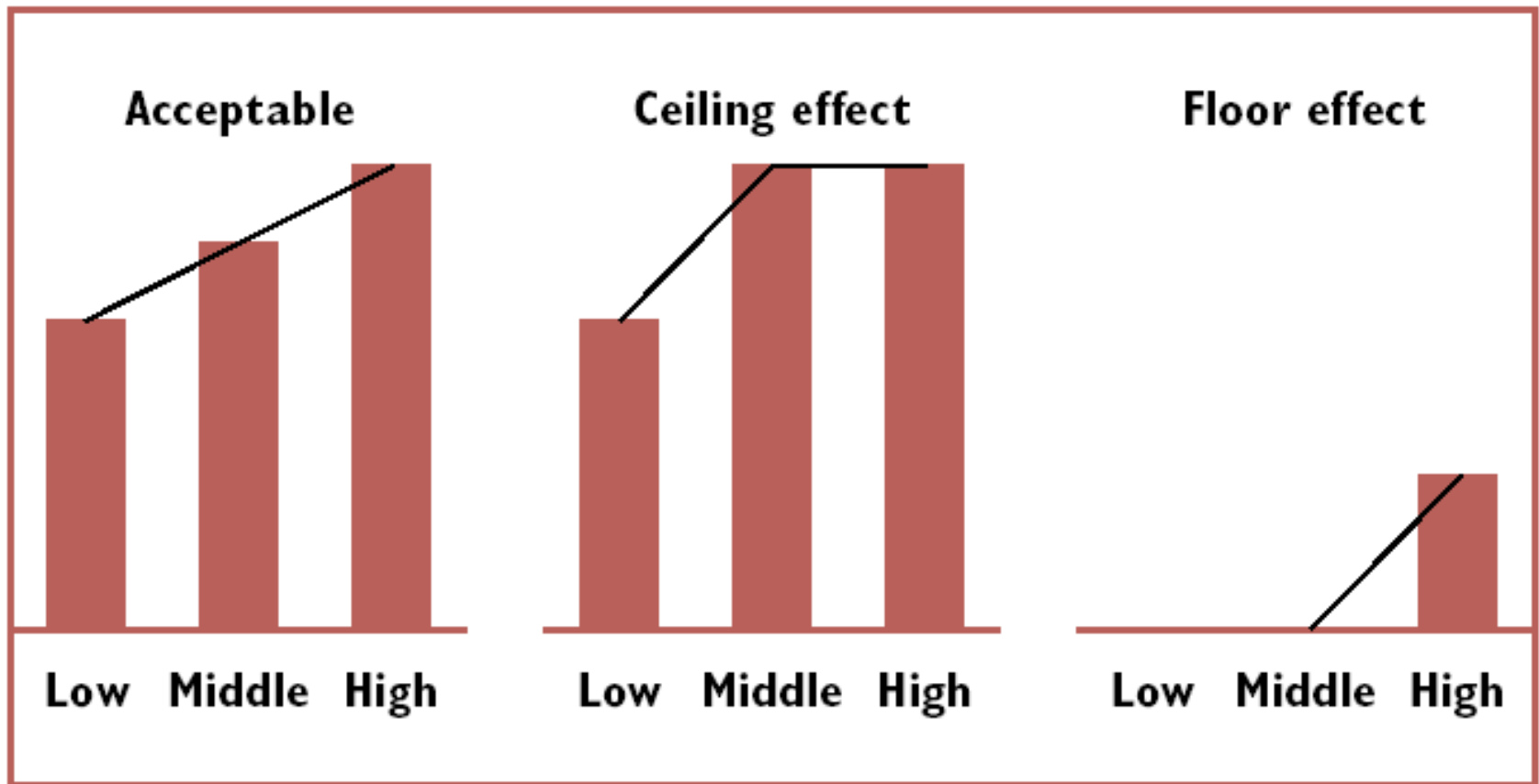
Item 19



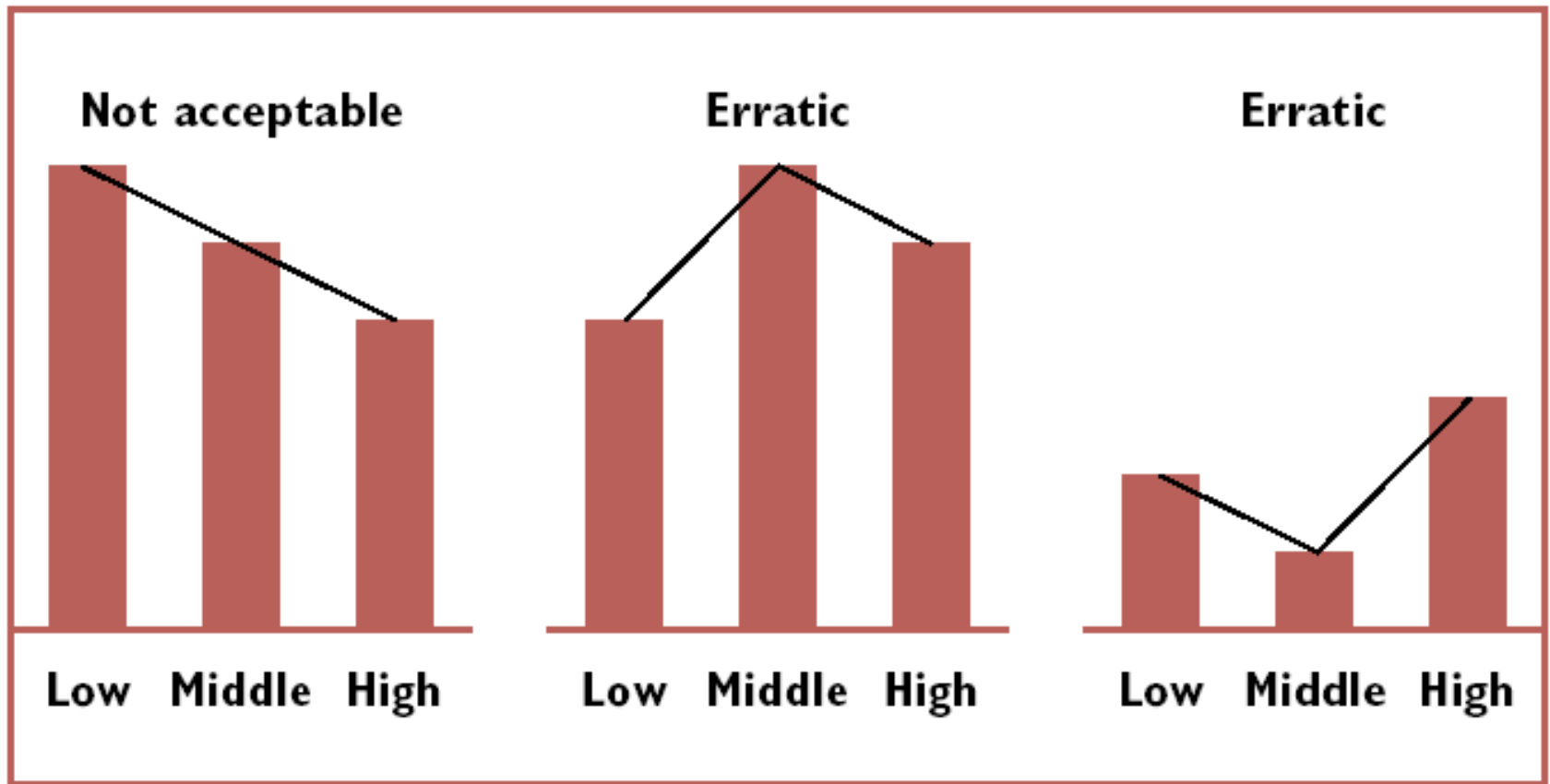
Item 21



Acceptable and may be acceptable correct answer response patterns

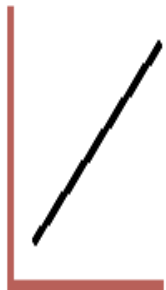


Unacceptable correct answer response patterns



Patterns for correct responses

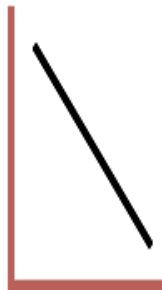
Correct responses



OK



?



?



?



?



OK?



OK?