


MULTIPLE CHOICE

Elien Sabbe
Educational Quality Assurance Office

- 
- 1. Best option?**
 - 2. Form**
 - 3. Content**
 - 4. Guessing?**
 - 5. Quality Control**

1. BEST OPTION?

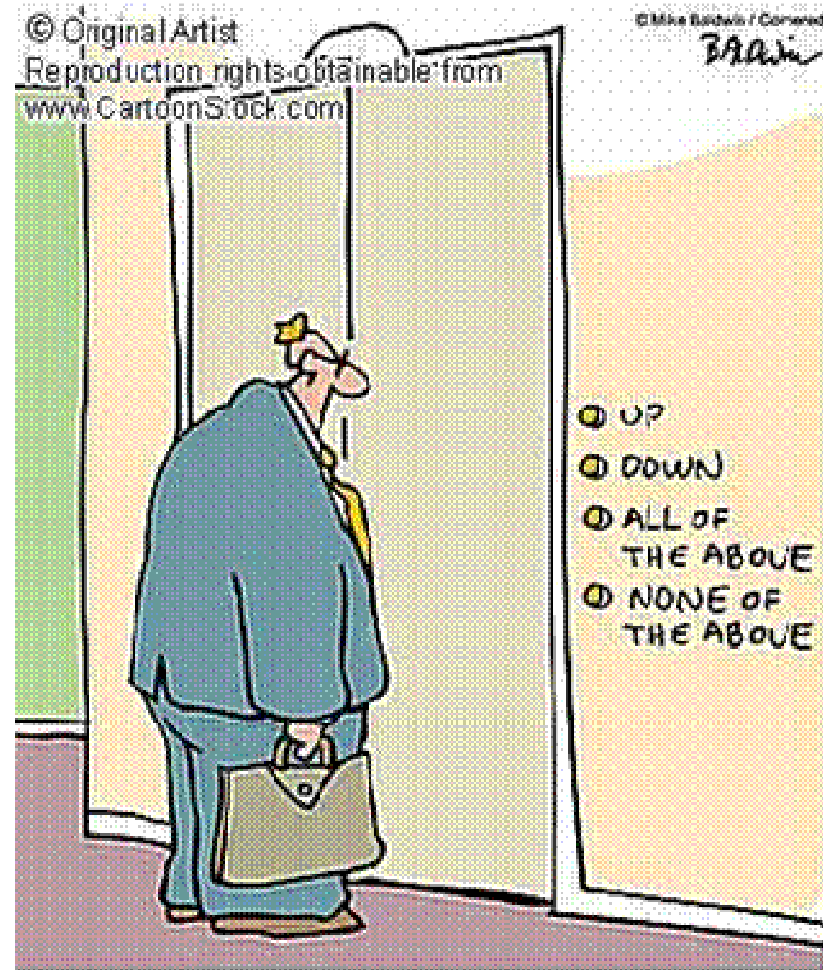
MC-EXAMS REQUIRE LESS TIME AND WORK FROM LECTURERS THAN OTHER EVALUATION METHODS

MC-EXAMS ONLY TEST FACTUAL KNOWLEDGE

MC-EXAMS ARE THE MOST OBJECTIVE EVALUATION METHOD

2. FORM

- Deception and confusion
 - guidelines on form (Haladyna et al, 2002)
- Clues for the correct answer
 - test wiseness



THE ADMINISTRATIVE STAFF OF A REAL ESTATE AGENCY HAS BEEN INSTRUCTED NOT TO ENTER THE WAITING AREA DURING OPENING HOURS. NEVERTHELESS, ONE OF THEM DOES ENTER THE AREA, CARRYING A PILE OF ORDERS THAT BLOCKS HIS VIEW. HE RUNS OVER A CUSTOMER, WHO BREAKS HER ARM BECAUSE OF THIS. TO WHOM CAN THE CUSTOMER TURN FOR THE DAMAGE SUFFERED?

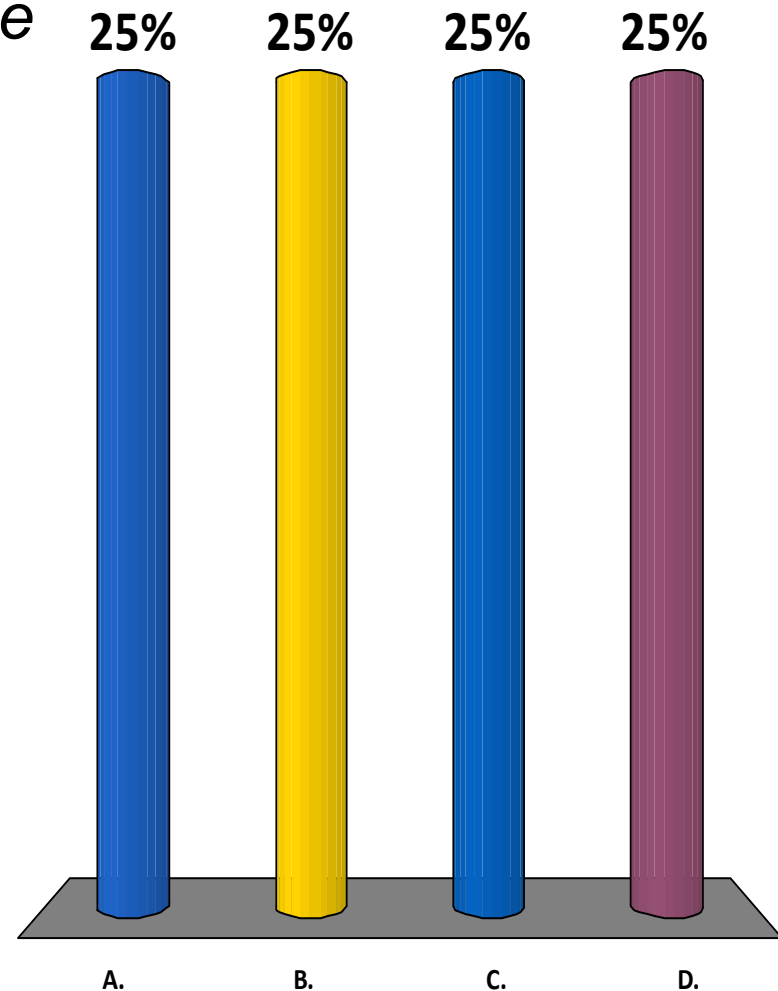
- A. *The customer can only hold the employee liable, because he entered the waiting area against the express instruction of his employer.*
- B. *The customer can only hold the employer liable, because there was no premeditation on the employee's behalf.*
- C. *The customer can hold both the employer and the employee liable, on grounds of tort liability and vicarious liability, respectively.*

THE ADMINISTRATIVE STAFF OF A REAL ESTATE AGENCY HAS BEEN INSTRUCTED NOT TO ENTER THE WAITING AREA DURING OPENING HOURS. NEVERTHELESS, ONE OF THEM DOES ENTER THE AREA, CARRYING A PILE OF ORDERS THAT BLOCKS HIS VIEW. HE RUNS OVER A CUSTOMER, WHO BREAKS HER ARM BECAUSE OF THIS. TO WHOM CAN THE CUSTOMER TURN FOR THE DAMAGE SUFFERED?

- A. *The customer can only hold the employee liable, because he entered the waiting area against the express instruction of his employer.*
- B. *The customer can only hold the employer liable, because there was no premeditation on the employee's behalf.*
- C. ***The customer can hold both the employer and the employee liable, on grounds of tort liability and vicarious liability, respectively. ****

LOCAL ANESTHETICS ARE MOST EFFECTIVE IN

- A. *anionic form, working from the outside of the nerve cell membrane.*
- B. *cationic form, working from the inside of the nerve cell membrane.*
- C. *cationic form, working on the outside of the nerve cell membrane.*
- D. *neutral, uncharged form, working from the outside of the nerve cell membrane.*

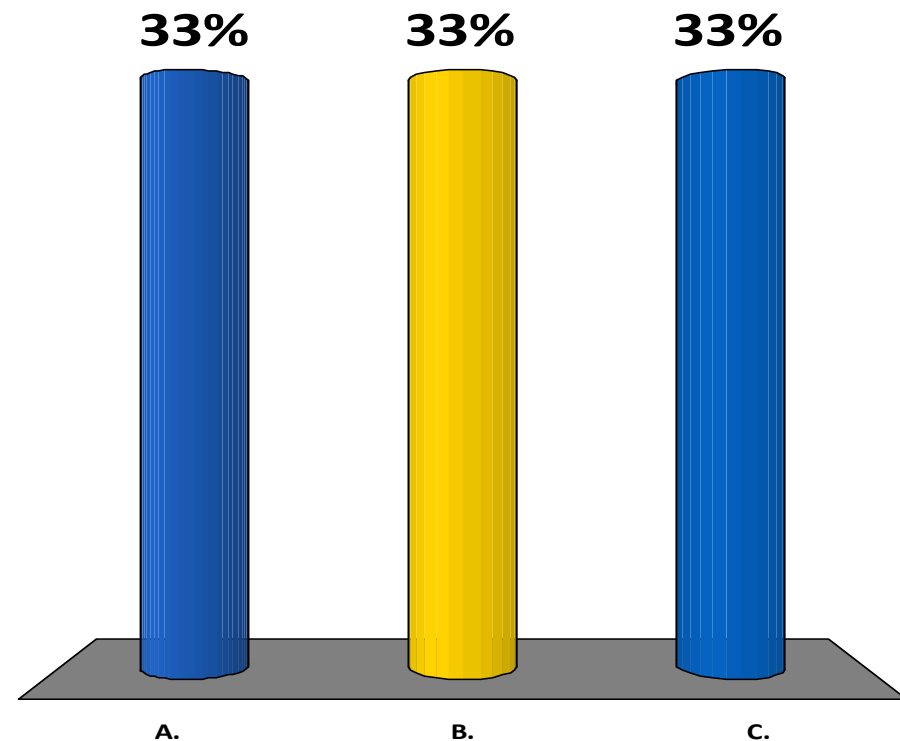


LOCAL ANESTHETICS ARE MOST EFFECTIVE IN

- A. *anionic form, working from the **outside** of the nerve cell membrane.*
- B. ***cationic form**, working from the inside of the nerve cell membrane.*
- C. ***cationic form**, working on the **outside** of the nerve cell membrane. **
- D. *neutral, uncharged form, working from the **outside** of the nerve cell membrane.*

JOINING TWO CONSECUTIVE ACTIVITIES IN THE COMPANY COLUMN IN ONE ORGANIZATION IS:

- A. always a form of parallelization.
- B. *usually a form of integration.*
- C. *definitely a form of specialization.*

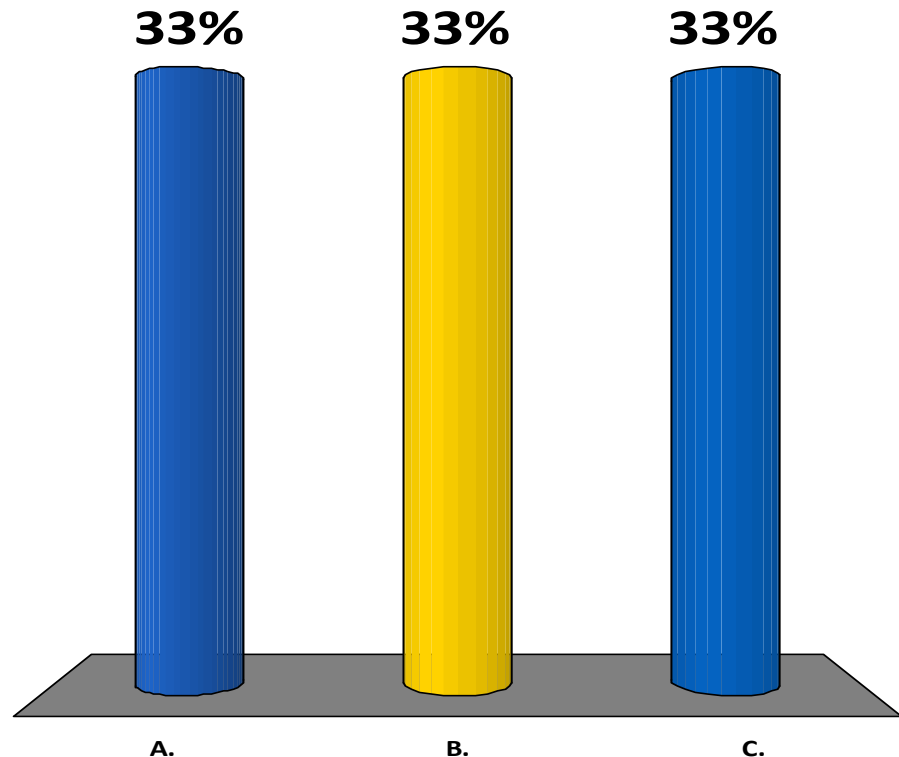


JOINING TWO CONSECUTIVE ACTIVITIES IN THE COMPANY COLUMN IN ONE ORGANIZATION IS:

- A. **always** a form of parallelization.
- B. usually** a form of *integration*. *
- C. **definitely** a form of *specialization*.

POLICE BRING A 60-YEAR-OLD MAN TO THE ER. HE WAS FOUND UNCONSCIOUS ON THE PAVEMENT. HIS BREATH SMELLS LIKE ALCOHOL. AFTER THEY HAVE CHECKED WHETHER HIS AIRWAYS ARE CLEARED, THE NEXT STEP IS AN INTRAVENOUS ADMINISTRATION OF :

- A. an examination of the cerebrospinal fluid.*
- B. glucose with vitamin B1.*
- C. a CT scan of the head.*

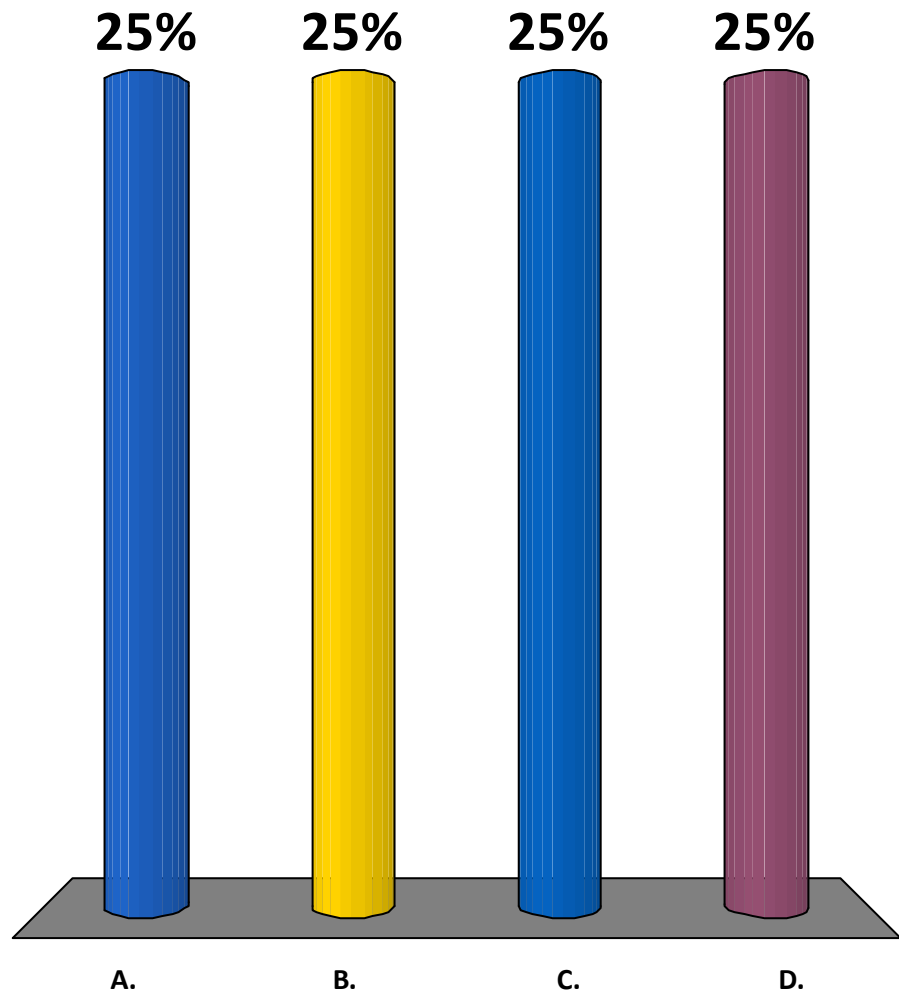


POLICE BRING IN A 60-YEAR-OLD MAN INTO THE ER. HE WAS FOUND UNCONSCIOUS ON THE PAVEMENT. HIS BREATH SMELLS LIKE ALCOHOL. AFTER THEY HAVE CHECKED WHETHER HIS AIRWAYS ARE CLEARED, THE NEXT STEP IS AN INTRAVENOUS ADMINISTRATION OF:

- A. an examination of the cerebrospinal fluid.*
- B. glucose with vitamin B1. **
- C. a CT scan of the head.*

ZARFARKLE, EN KO DAY?

- A. *Henkledorf*
- B. *Ricktoffen*
- C. *Ifetain*
- D. *Krator*

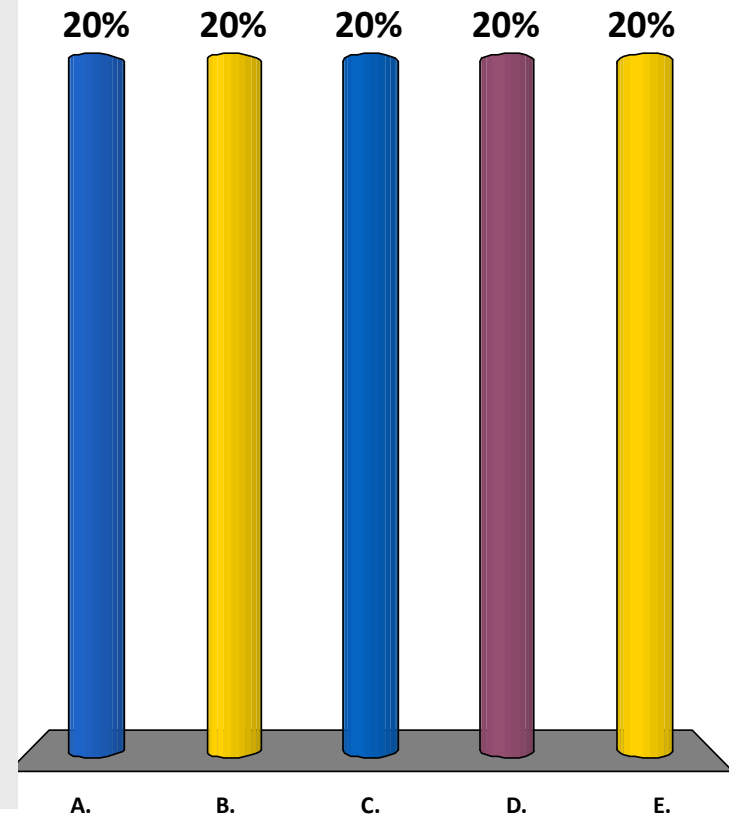


ZARFARKLE, EN KO DAY?

- A. Henkledorf*
- B. Ricktoffen*
- C. Ifetain****
- D. Krator*

CRIME IS

- A. equally represented across the social classes.
- B. *overrepresented among the poor.*
- C. *overrepresented among the middle class and the wealthy.*
- D. *mainly a sign of psychosexual maladaptation.*
- E. *slowly becoming tolerated by the government.*

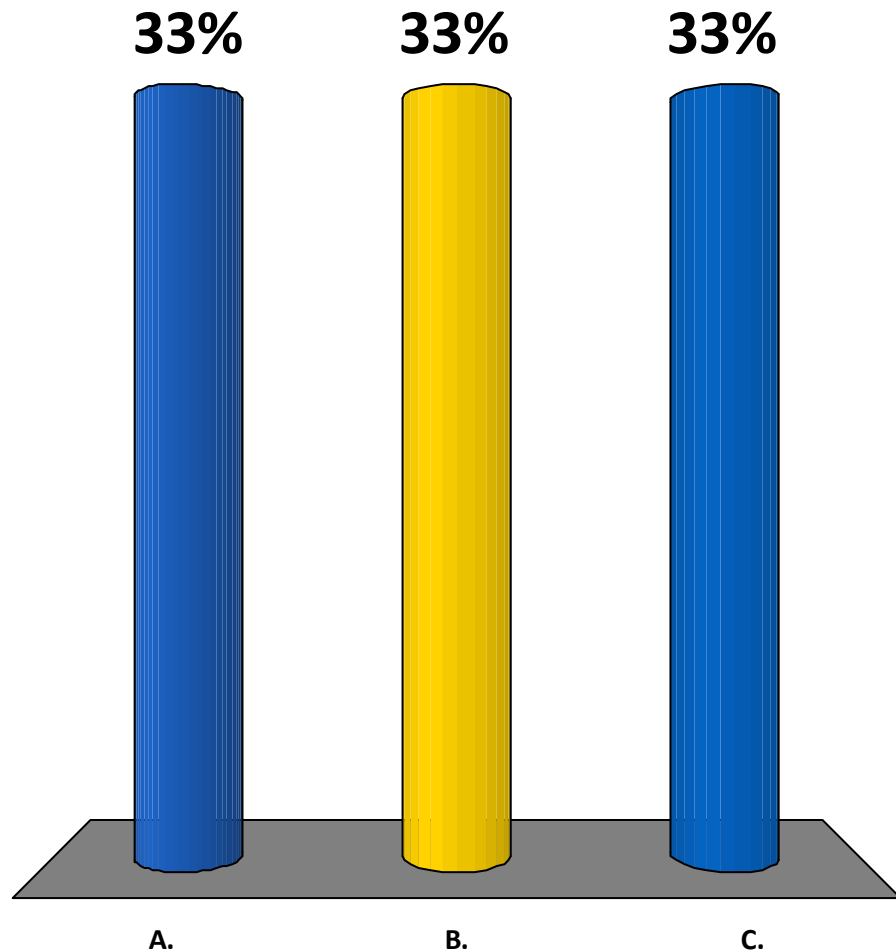


CRIME IS

- A. equally represented across the social classes.
- B. *overrepresented among the poor.*
- C. *overrepresented among the middle class and the wealthy.*
- D. *mainly a sign of psychosexual maladaptation.*
- E. *slowly becoming tolerated by the government.*

WHAT IS MEANT BY THE CONCEPT OF “THE POPULATION ATTRIBUTABLE FRACTION”?

- A. The number of infected cases in the whole population that can be attributed to the risk factor.
- B. *The risk factor in the emergence of the disease.*
- C. *The chance of preventing the disease.*

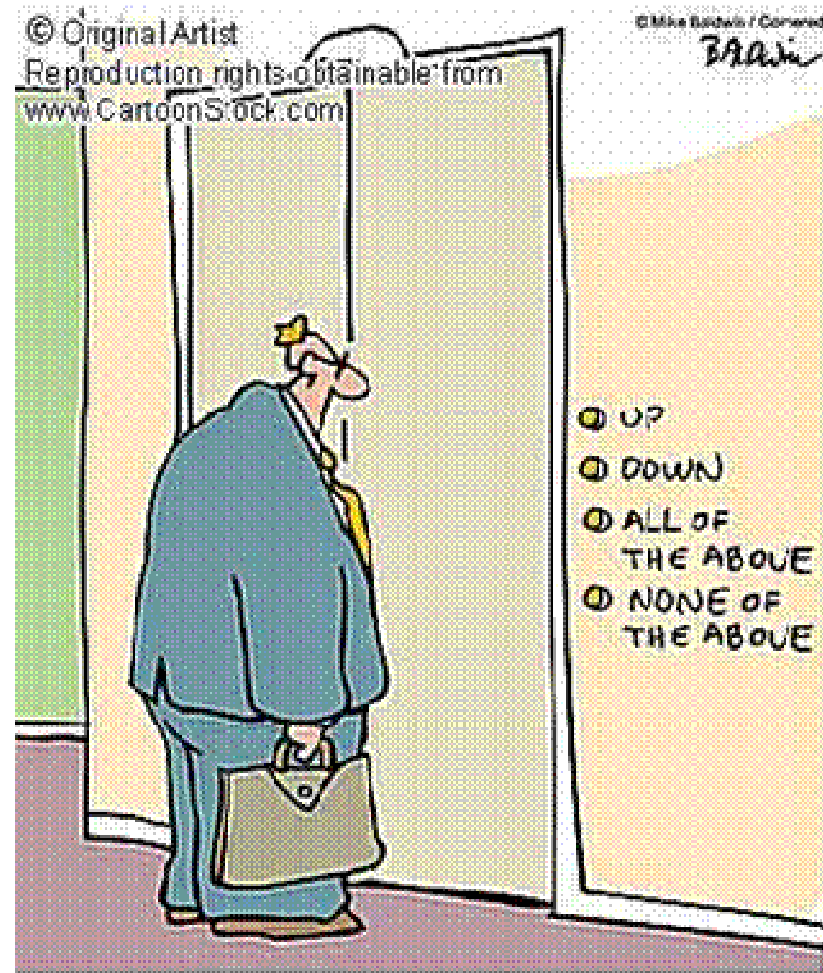


WHAT IS MEANT BY THE CONCEPT OF “THE POPULATION ATTRIBUTABLE FRACTION”?

- A. The number of infected cases in the whole population that can be attributed to the risk factor.***
- B. The risk factor in the emergence of the disease.*
- C. The chance of preventing the disease.*

FORM

- Deception and confusion
 - guidelines on form (Haladyna et al, 2002)
- Clues for the correct answer
 - test wiseness



3. WHAT TO EVALUATE?

- Material → objectives
- Representative reflection
- More important topics deserve more questions



"THEN, AS YOU CAN SEE, WE GIVE THEM SOME MULTIPLE CHOICE TESTS."

TEST MATRIX

Content/ material	Level	% items	Number of items	Which items
Market demand	Reproduction	20%	4	1,5,8,14
Market supply	Reproduction	20%	4	2,3,4,7
Market demand/ supply	Connections/ insight	30%	6	6,9,10,11,12,13
Pricing	Application	30%	6	15,16,17,18,19,20

MORE THAN FACTUAL KNOWLEDGE? YES!

- Characteristics, examples of a concept
- Recognizing a concept, principle, theory or formula
- Connections, differences, similarities

- Predictions, cause/consequence
- Interpreting data
- Applying a formula/principle
- Diagnosis, decision, model, problem-solving

In which example does “acquisition under universal title” apply in private law?

- a. The municipality of Ede gains ownership of a piece of land expropriated from farmer De Vries.
- b. Together with his brother, Mr Berends inherits his father’s house.*
- c. Mrs Arnts gains ownership of the home she bought after the conveyance deed has been registered in the cadastre.

Various sound laws turned the Proto-Germanic form *aiz “metal” into Old English ar/a.r/. However, Old English ar/a.r/ can also be traced back to Proto-Germanic *air “messenger”. This second development also occurred according to various sound laws.

What is the name of the phonological process illustrated by the relationship of these two examples?

- a. ersatzdehnung
- b. merger*
- c. monophthongization
- d. rhotacism

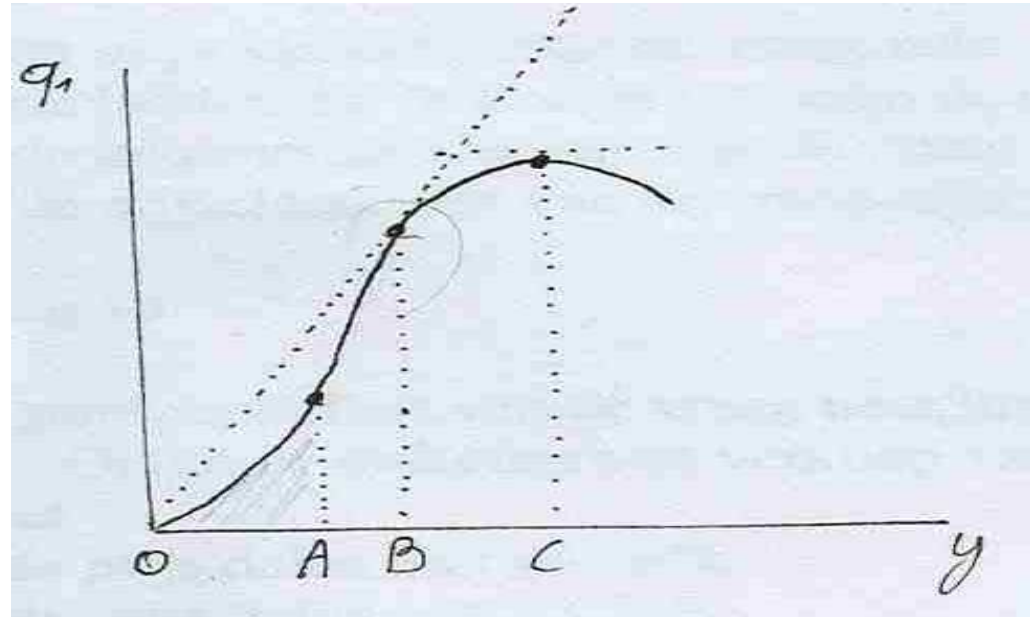
Which of the following statements is indicative of Mill's utilitarianism, but not Bentham's?

- a. Individuals are responsible for working towards greater happiness in our society.*
- b. Individual self-interest is the basis of any kind of moral action.
- c. Another person's happiness is merely incidental.

What is the effect of increasing the sample size or the number of test subjects in an experiment if H_0 is untrue and all facts concerning the population remain unchanged with an increasing n ?

- a. The confidence interval becomes narrower and the p-value drops.*
- b. The confidence interval becomes narrower and the p-value rises.
- c. The confidence interval becomes broader and the p-value drops.
- d. The confidence interval becomes broader and the p-value rises.

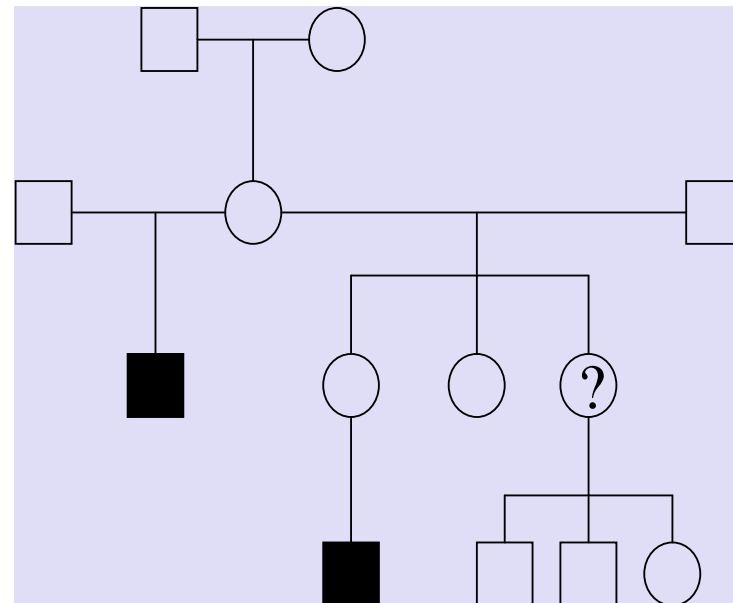
Which of the following statements **CANNOT** be deduced from this Engel curve?



- Only when the income level is between O and A, is the income elasticity larger than 1.
- For all incomes between B and C, good 1 is a necessary good.
- For incomes between A and B, the budget share of good 1 increases as the income rises.
- When the income is B, then the income elasticity is 1.*

This family suffers from an X-linked recessive disorder.
What are the odds that individual III-4 is a carrier?

- a. $1/2$
- b. $1/4$
- c. $1/5^*$
- d. $1/8$
- e. $1/10$



A 34-year-old woman has had severe, watery, non-bloody diarrhea for four days now. Two months ago, she suffered from infectious mononucleosis (mononucleosis infectiosa). She has a history of (intravenous) drug abuse and is HIV-positive. A physical examination shows that she is dehydrated and suffers from muscle weakness.

What is the best way to identify the cause of the diarrhea?

- a. Colonoscopy with a biopsy to detect Giardia Lamblia.
- b. Culture of the oral cavity for Candida Albicans.
- c. Duodenal biopsy to detect Entamoeba histolytica.
- d. Faeces sample to detect Cryptosporidium.*
- e. Stomach aspiration to detect Mycobacterium avium-intracellulare.

INSPIRATION?

- Students' mistakes
- Variations with respect to the answer
- Write down ideas after class
- Variations with respect to the root
- Templates
- Errors in reasoning

TEMPLATES

- What is another term for ...?
- What best defines...?
- What is a characteristic of ...?
- What is an example of...?
- Which of the following is the best illustration of ...?
- What is the most effective way to ...?
- What is essential in the procedure...?
- What is the cause/reason/purpose of ...?
- What happens if ...?
- What is a consequence of...?
- Calculate...

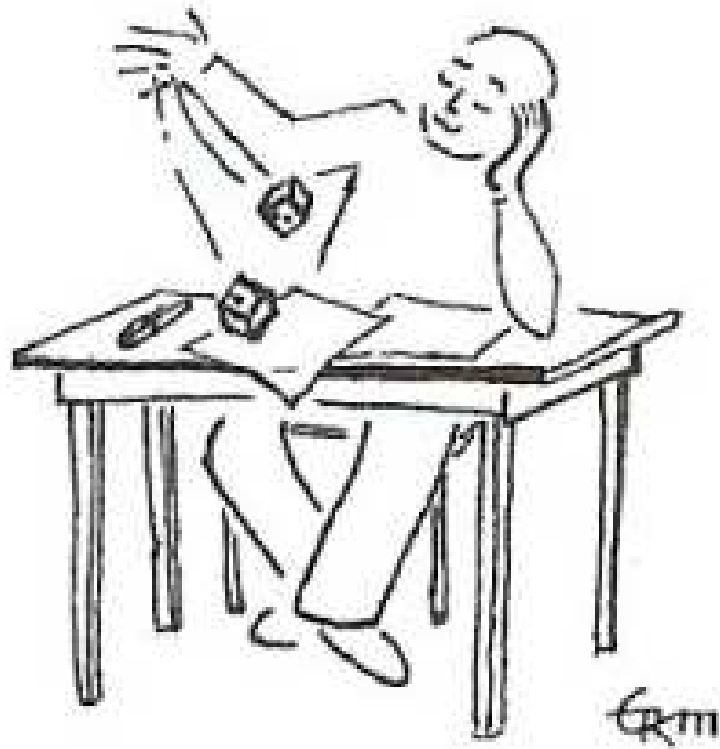
Which visual fields are lost when the optic chiasm is damaged?

- a. left visual field of the left and right eye.*
- b. right visual field of the left and right eye.*
- c. left visual field of the left eye and right visual field of the right eye.**
- d. left visual field of the right eye and right visual field of the left eye.*

A lecturer needs to evaluate 400 students. S/he can spend quite some time on the development of the test, but aims to limit the time spent correcting it. The evaluation itself will take 30 minutes. The learning objectives mainly involve factual knowledge.

Which type of evaluation is most suitable here?

- a. oral exam with open questions*
- b. written exam with open questions*
- c. written exam with multiple-choice questions**



4. GUESSING?

December 6th, 2013 the Board of Governors decided to replace negative marking with standard setting.

NEGATIVE MARKING

- Scoring formulas
- Students \neq rational test takers
- Willingness-to-gamble effect
 - Overcorrection
 - Undercorrection
- Transparency
- Tempo

Correct: +1

Incorrect: $-1/n-1$

Open: 0

STANDARD SETTING

- In students' best interest to answer all questions
 - ▶ no tactical considerations in answering questions
 - ~~▶ willigness to gamble~~
- Cut-off score $> 50\%$ to compensate for points gained due to guessing

HOW TO DETERMINE THE STANDARD?

- Basis: number of options
- Formula: $N(n+1)/(2n)$
- Reasoning: students have as much chance of passing a test with a higher cutoff score as a test with negative marking

CONCRETELY...

2 options	3 options	4 options	5 options	6 options
75%	66.67%	62.50%	60%	58.33%

HOW ARE THE GRADES RECALCULATED?

Principles

- Cutoff score = 10
- All questions right = 20
- All questions wrong = negative marks are turned into 0

$$z = 10 + 10(y-C)/(N-C)$$

EVIDENCE BASE OF THE DECISION

- Review of international literature
- Quasi-experimental study at the UGent on guessing behavior (1500 students): NM-SS
 - Students with low guessing tendency have sign. higher marks with SS vs. NM (especially women)
 - For students with high guessing tendency: no difference in marks
- Probability analysis by UGent statisticians:
 - Formulas based on this condition: $NM=SS$

EVALUATION OF THE IMPLEMENTATION

Impact on results

- Success rate = but +5% 15/20
- +4,1% ♀ graduates (% ♂=)

5. QUALITY CONTROL?

- **Validity:** explicit goals and test matrix
- **Reliability**
 - rules of thumb
 - revise after a period of time
 - let a colleague revise your test (cf. 4-eyes principle)
 - item analysis
- **Transparency:** example questions

	/350	%	Question1	Question 2	Question 3	Question 4	Question 5
Average score	221.30	63.23	6.78	8.32	8.29	3.80	5.24
Average score L27	167.56	47.87	4.74	7.44	6.79	2.95	4.10
Average score L50	188.42	53.83	5.55	7.81	7.40	3.49	4.25
Average score U50	254.18	72.62	8.01	8.84	9.18	4.11	6.23
Average score U27	269.10	76.89	7.95	8.85	9.36	4.10	7.05

% v. alle deelnemers met antwoord	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
% v. alle deelnemers zonder antwoord	--	--	--	--	--	--	--	--	--	--	--	--	--
% v. alle deelnemers met als antwoord A	67.81	15.07	9.93	24.66	52.40	4.45	8.90	13.70	9.93	26.37	40.41	19.18	
% v. alle deelnemers met als antwoord B	7.19	83.22	3.08	56.85	31.85	47.60	1.03	0.68	5.82	13.70	12.67	70.21	
% v. alle deelnemers met als antwoord C	4.11	1.37	83.22	5.82	7.53	9.93	0.34	82.19	2.40	55.48	17.81	9.93	
% v. alle deelnemers met als antwoord D	20.89	0.34	4.11	12.67	8.22	38.01	89.73	3.42	81.85	4.45	29.11	1.03	
% v. alle deelnemers met als antwoord E	--	--	--	--	--	--	--	--	--	--	--	--	
% v. L27 met antwoord	98.73	98.73	98.73	98.73	98.73	98.73	98.73	98.73	98.73	98.73	98.73	98.73	98.73
% v. L27 met als antwoord A	46.84	21.52	13.92	45.57	40.51	8.86	10.13	21.52	16.46	29.11	34.18	29.11	
% v. L27 met als antwoord B	11.39	73.42	8.86	25.32	35.44	45.57	3.80	--	13.92	24.05	8.86	45.57	
% v. L27 met als antwoord C	5.06	3.80	67.09	15.19	12.66	15.19	1.27	72.15	6.33	35.44	21.52	20.25	
% v. L27 met als antwoord D	35.44	--	8.86	12.66	10.13	29.11	83.54	5.06	62.03	10.13	34.18	3.80	
% v. L27 met als antwoord E	--	--	--	--	--	--	--	--	--	--	--	--	
% v. L50 met antwoord	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
% v. L50 met als antwoord A	55.48	19.86	13.01	39.73	42.47	5.48	10.96	19.18	13.01	30.14	36.30	23.29	
% v. L50 met als antwoord B	8.90	78.08	5.48	35.62	37.67	47.95	2.05	1.37	10.27	17.12	8.90	59.59	
% v. L50 met als antwoord C	6.85	2.05	74.66	9.59	10.96	11.64	0.68	75.34	4.11	45.21	21.23	15.75	
% v. L50 met als antwoord D	28.77	--	7.53	15.07	8.90	34.93	86.30	4.11	72.60	7.53	33.56	2.05	
% v. L50 met als antwoord E	--	--	--	--	--	--	--	--	--	--	--	--	
% v. U50 met antwoord	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
% v. U50 met als antwoord A	80.14	10.27	6.85	9.59	62.33	3.42	6.85	8.22	6.85	22.60	44.52	15.07	
% v. U50 met als antwoord B	5.48	88.36	0.68	78.08	26.03	47.26	--	--	1.37	10.27	16.44	80.82	
% v. U50 met als antwoord C	1.37	0.68	91.78	2.05	4.11	8.22	--	89.04	0.68	65.75	14.38	4.11	
% v. U50 met als antwoord D	13.01	0.68	0.68	10.27	7.53	41.10	93.15	2.74	91.10	1.37	24.66	--	
% v. U50 met als antwoord E	--	--	--	--	--	--	--	--	--	--	--	--	
% v. U27 met antwoord	98.73	98.73	98.73	98.73	98.73	98.73	98.73	98.73	98.73	98.73	98.73	98.73	98.73
% v. U27 met als antwoord A	78.48	10.13	6.33	6.33	69.62	6.33	8.86	3.80	6.33	15.19	45.57	13.92	
% v. U27 met als antwoord B	5.06	87.34	--	83.54	20.25	41.77	--	--	--	12.66	20.25	83.54	
% v. U27 met als antwoord C	1.27	1.27	92.41	--	1.27	10.13	--	94.94	--	70.89	12.66	1.27	
% v. U27 met als antwoord D	13.92	--	--	8.86	7.59	40.51	89.87	--	92.41	--	20.25	--	
% v. U27 met als antwoord E	--	--	--	--	--	--	--	--	--	--	--	--	

THANK YOU!



REFERENCES

- ABAD, F.J., OLEA, J. & PONSODA, V. (2001). Analysis of the optimum number alternatives from the item response theory. *Psicothema*, 13(1), 152-158.
- ACE-FERRER, A.J., BACKHOFF-ESCUADERO, E. & PENA-GONZALEZ, L.I. (2001). Effects of the number of choices on the psychometric properties of multiple-choice tests. *Revista Mexicana de Psicología*, 18(3), 337-345.
- ATTALI, Y. & BAR-HILLEL, M. (2003). Guess where: the position of correct answers in multiple-choice test items as a psychometric variable. *Journal of Educational Measurement*, 40(2), 109-128.
- BEN SHAKKAR, G. & SINAI, Y. (1991). Gender differences in multiple-choice tests: the role of differential guessing tendencies. *Journal of Educational Measurement*; 28(1), 23-35.
- BUDESCU, D. & HAR-HILLIL, M. (1993). To guess or not to guess: a decision-theoretic view of formula scoring. *Journal of Educational Measurement*, 30(4), 277-291.
- BURTON, R.F. (2002). Misinformation, partial knowledge and guessing in true/false tests. *Medical Education*, 36, 805-811.
- CHAMBERS, D.W. (2007). Correcting for guessing on multiple-choice exams. *Journal of Dental Education*, 71 (2), 193-196.

- DE GRUIJTER, D. (2008). *Toetsing en toetsanalyse*. ICLON, sectie onderwijsontwikkeling
- DIAMOND, J.E. & EVANS, W. (1973). The correction for guessing. *Review of Educational Research*, 43, 181-191.
- DOWNING, S.M. (2004). On guessing corrections. *Medical Education*, 38, 112-113.
- ESPINOSA, M.P. & GARDEAZABAL, J. (2010). Optimal correction for guessing in multiple-choice tests. *Journal of Mathematical Psychology*, 54 (5), 415-425.
- GAFNI, N. & MELAMED, E. (1994). Differential tendencies to guess as a function of gender and lingual-cultural reference group. *Studies in educational evaluation*, 20, 309-319.
- GIERL, M.J. & HOLLIS, L. (2015). Using automated processes to generate test items and their associated solutions and rationales to support formative feedback. *Interaction Design and Architecture(s) Journal*, 25, 9-20.
- HALADYNA, T.M., DOWNING, S.M. & RODRIGUEZ, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, 15(3), 309-334.
- HARASYM, P.H., LEONG, E.J., VIOLATO, C., BRANT, R. & LORSCHIEDER, F.L. (1998). Cuing effect of ‘all of the above’ on the reliability and validity of multiple-choice test items. *Evaluation and the Health Professions*, 21(1), 120-133.
- FRARY, R.B. (1991). The none-of-the-above option: an empirical study. *Applied Measurement in Education*, 4, 115-124.

- LESAGE, E., VALCKE M. SABBE E. (2013). Scoring methods for multiple choice assessment in higher education: is it still a matter of number right scoring or negative marking? *Studies in Educational Evaluation*, 39(3), 188-193.
- LESAGE, E., VALCKE, M. & SABBE, E. (2015) The Effect of Guessing Tendency on Student Performance: A Comparative Study between Negative Marking, Number Right Scoring, Standard Setting and Retrospective Correction for Guessing. (in review)
- OTTOY, J.-P., THAS, O. & DE NEVE, N. (2013). Probabiliteitsanalyse van gokken bij meerkeuzevragen. UGent-nota voor de Onderwijsraad.
- SABBE, E. & LESAGE, E. (2012). *Meerkeuzetoetsen: praktische handleiding voor leerkrachten en docenten*. Antwerpen: Garant
- SCOULLER, K. (1998). The influence of assessment method on students' learning approaches : multiple choice question examination versus assignment essay. *Higher Education*, 35, 453-472.
- STANGER-HALL, K. F. (2012). Multiple-choice exams: an obstacle for higher-level thinking in introductory science classes. *Life Sciences Education*, 11, 294-306.



VAN BERKEL, H. & BAX, A. (2002). *Toetsen in het hoger onderwijs*. Houten/Diegem: Bohn Stafleu Van Loghum.