European Board of Ophthalmology
excellence in education

NEGATIVE MARKING: PRO´S

CESMA Meeting (April 1st-2nd, 2011)
La Valletta, Malta

Danny G.P. Mathysen

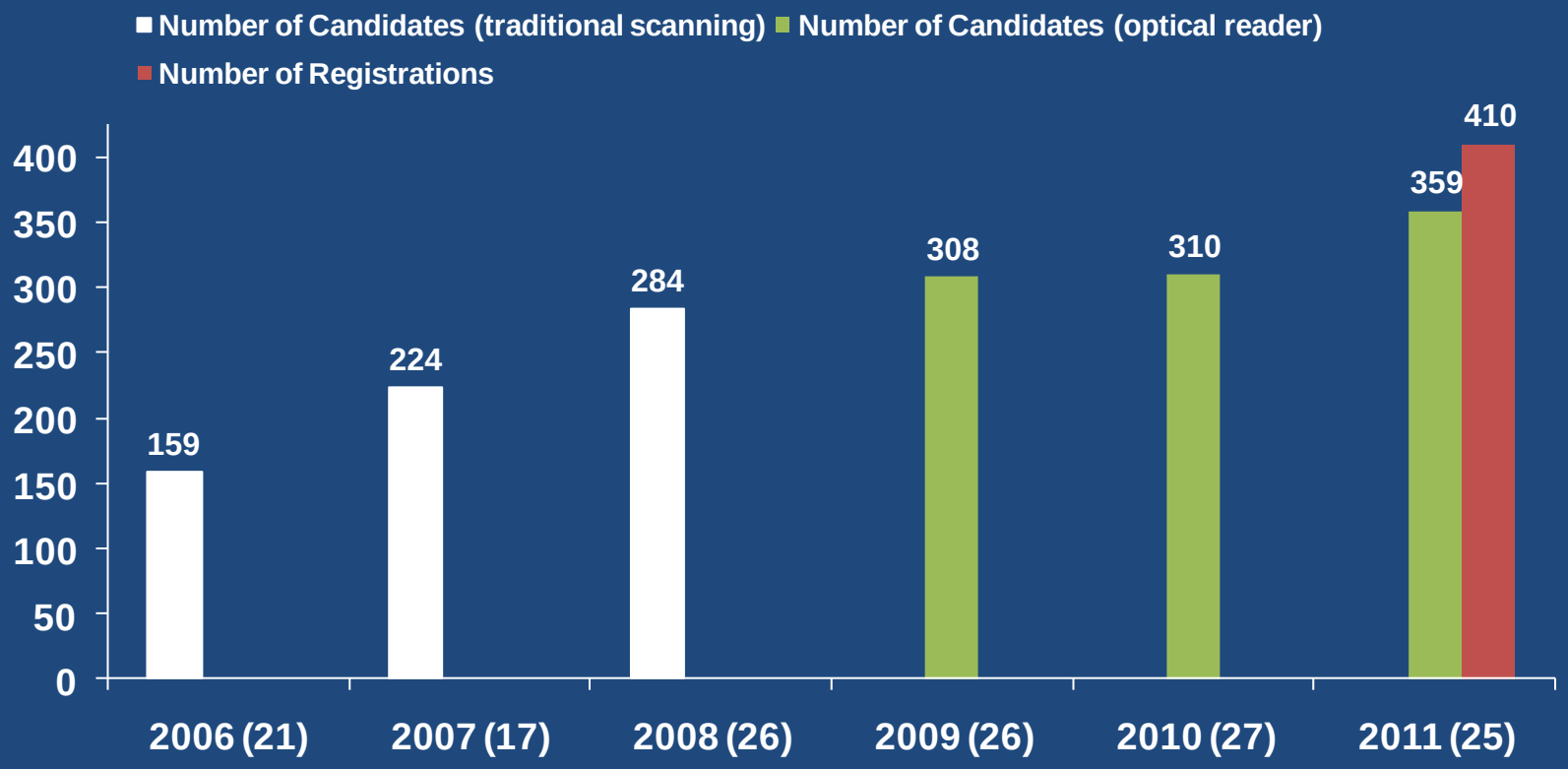
MSc. Biomedical Sciences

EBOD Assessment and Executive Officer

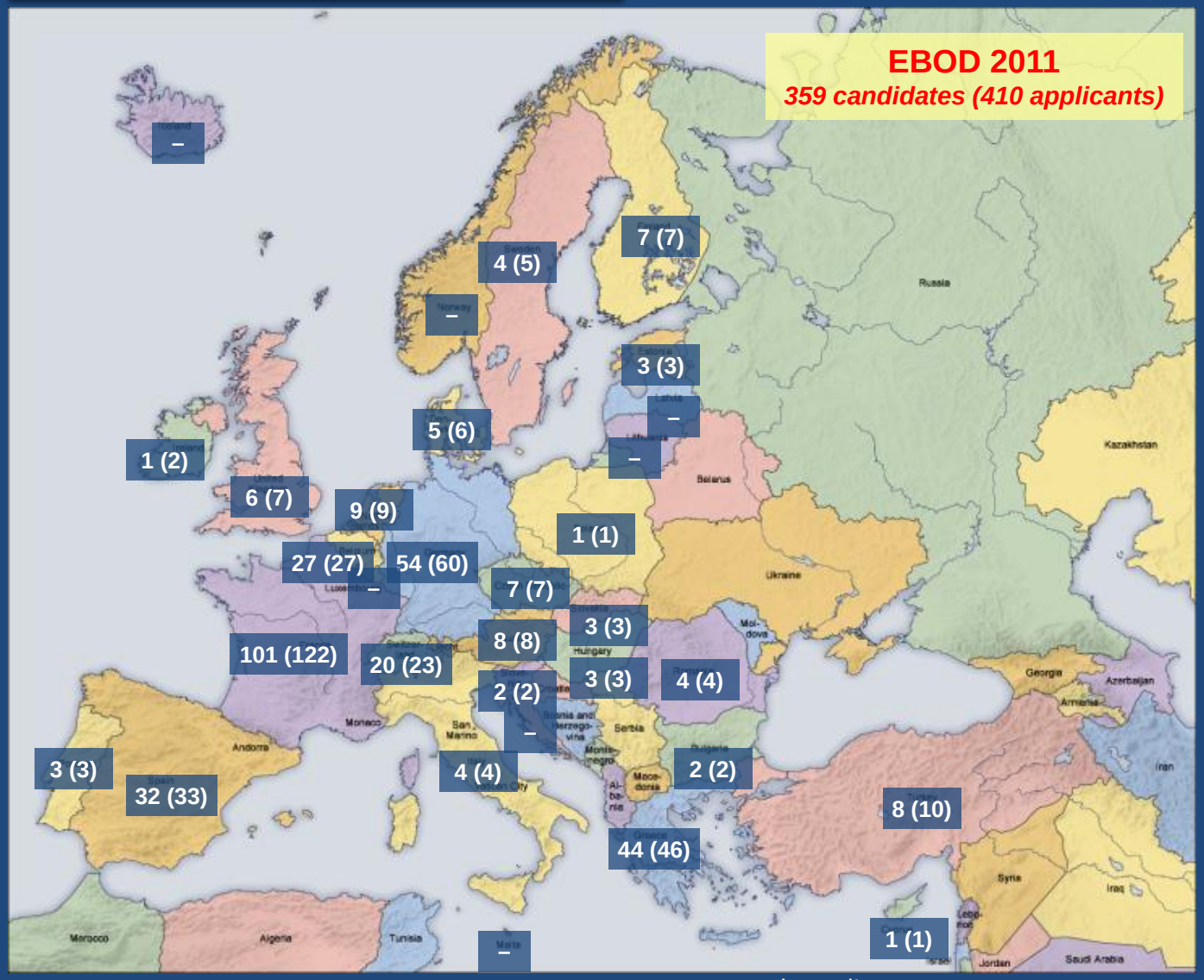
Antwerp University Hospital, Department of Ophthalmology
Wilrijkstraat 10, B-2650 Edegem (Antwerp), Belgium
E-mail: danny.mathysen@uza.be

EBOD: GENERAL INFORMATION

THE EBOD EXAMINATION



European Board of Ophthalmology
excellence in education



AUSTRIA	8 / 8
BELGIUM	27 / 27
BULGARIA	2 / 2
CROATIA	-
CYPRUS	1 / 1
CZECH REPUBLIC	7 / 7
DENMARK	5 / 6
ESTONIA	3 / 3
FINLAND	7 / 7
FRANCE	101 / 122
GERMANY	54 / 60
GREECE	44 / 46
HUNGARY	3 / 3
ICELAND	-
IRELAND	1 / 2
ITALY	4 / 4
LATVIA	-
LITHUANIA	-
LUXEMBOURG	-
MALTA	- / 1
NETHERLANDS	9 / 9
NORWAY	-
POLAND	1 / 1
PORTUGAL	3 / 3
ROMANIA	4 / 4
SLOVAK REPUBLIC	3 / 3
SLOVENIA	2 / 2
SPAIN	32 / 33
SWEDEN	4 / 5
SWITZERLAND	20 / 23
TURKEY	8 / 10
UNITED KINGDOM	6 / 7
OUTSIDE EU	**- / 11**

www.ebo-online.org

European Board of Ophthalmology
excellence in education

	2006	2007	2008	2009	2010	2011	
AUSTRIA	-	-	2	5	2	8	(8)
BELGIUM	26	42	23	25	20	27	(27)
BULGARIA	-	-	-	4	4	2	(2)
CROATIA	-	-	-	-	2	-	
CYPRUS	-	-	-	-	-	1	(1)
CZECH REPUBLIC	1	1	2	2	2	7	(7)
DENMARK	1	1	4	6	3	5	(6)
ESTONIA	2	3	3	2	1	3	(3)
FINLAND	2	4	7	2	3	7	(7)
FRANCE	57	69	92	96	127	101	(122)
GERMANY	18	33	44	59	48	54	(60)
GREECE	4	7	10	19	25	44	(46)
HUNGARY	-	-	1	2	1	3	(3)
ICELAND	-	-	1	-	-	-	
IRELAND	1	3	6	5	3	1	(2)
ITALY	1	-	4	6	7	4	(4)
LATVIA	1	1	2	1	1	-	
LITHUANIA	-	-	1	1	1	-	
LUXEMBOURG	1	-	-	-	-	-	
MALTA	-	-	-	-	1	-	
NETHERLANDS	2	10	7	7	3	9	(9)
NORWAY	-	-	-	1	-	-	
POLAND	-	-	1	2	2	1	(1)
PORTUGAL	2	-	1	-	2	3	(3)
ROMANIA	1	-	1	-	3	4	(4)
SLOVAK REPUBLIC	1	3	1	1	-	3	(3)
SLOVENIA	1	10	6	5	1	2	(2)
SPAIN	3	5	14	17	17	32	(33)
SWEDEN	4	3	6	5	7	4	(4)
SWITZERLAND	27	25	32	29	17	20	(23)
TURKEY	-	4	11	5	6	8	(10)
UNITED KINGDOM	2	-	2	1	1	6	(7)

THE EBOD EXAMINATION

- The **written part** is a 2.5 hours multiple choice questions (MCQ) examination covering 52 subjects in ophthalmology. For each subject (stem) there are 5 questions (leaves) (52 × 5 = 260 responses)

THE EBOD EXAMINATION

- The **oral part** consists of 4 viva voce of 15 minutes each with 2 examiners, covering 4 major topics in ophthalmology
  - Topic A: Optics, Refraction, Strabismus, and Neuro-ophthalmology
  - Topic B: Cornea, External Diseases, Orbit, and Ocular Adnexa
  - Topic C: Glaucoma, Cataract , and Refractive Surgery
  - Topic D: Posterior Segment, Ocular Inflammation, and Uveitis
- Each **viva voce** is seen to be a discussion among specialists in ophthalmology between the candidate and the two examiners

THE EBOD EXAMINATION

The **EBO written MCQ** examination:

- For each correct answer 1 point is given, therefore a maximum of 260 points can be obtained.
- Since 2010, **negative marking** has been introduced for incorrect answers (–0.5)

THE EBOD EXAMINATION

- Question Number (1 → 52)
- Item Number (A → E)

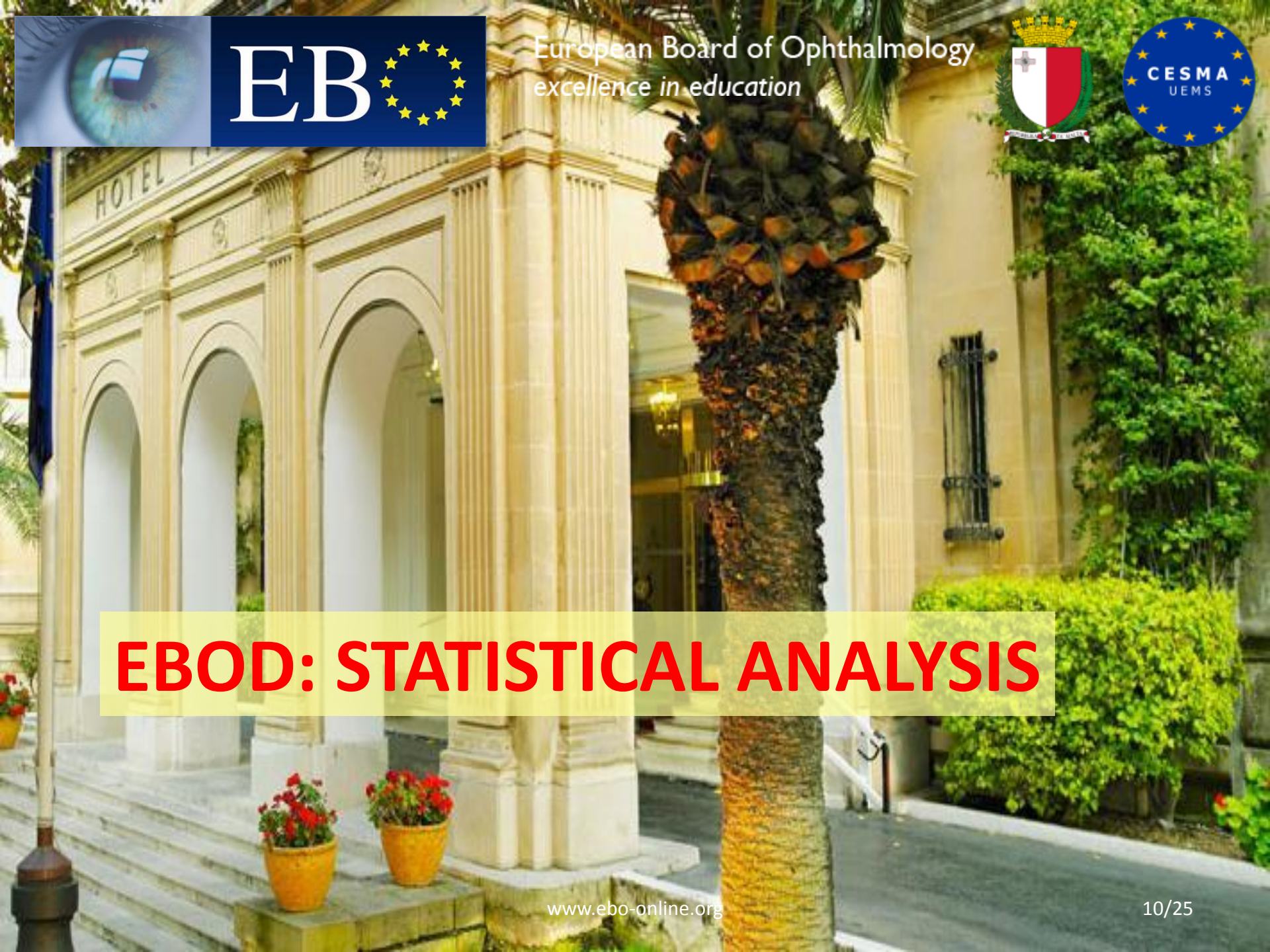
- T (True)
- F (False)
- D (Don't know)

Marks obtained?

+1
In case ONLY the correct answer was completed

0
In case ONLY the D option was completed

–0.5
In case ONLY the incorrect answer was completed
In case T AND F were completed
In case NOTHING was completed (blank item)
In case D was COMBINED with T and/or F


EB
European Board of Ophthalmology
excellence in education
CESMA
UEMS
EBOD: STATISTICAL ANALYSIS

European Board of Ophthalmology
excellence in education

BENEFITS OF NEGATIVE MARKING AT THE EUROPEAN BOARD OF OPHTHALMOLOGY DIPLOMA (EBOD) EXAMINATION, BOTH FOR ORGANISER & CANDIDATES

Danny G.P. Mathysen, Marie-José Tassignon , Catherine Creuzot-Garcher, Marko Hawlina, Peter J. Ringens, and Wagih Aclimandos

***Correspondence**: D.G.P. Mathysen, Antwerp University Hospital, Department of Ophthalmology, Wilrijkstraat 10, B-2650 Edegem (Antwerp), Belgium, E-mail: danny.mathysen@uza.be*

INTRODUCTION AND RESEARCH QUESTIONS

Introduction:

The European Board of Ophthalmology Diploma (EBOD) examination consists of a **written examination** (presented in this poster) followed by an oral examination. The written part of EBOD contains 52 **MCQs** with 5 **multiple true-false** items each (260 answers to be given by the candidates) with a pre-defined distribution of 10 **topics** within ophthalmology (more details on EBO website).

Research questions:

1. Does the introduction of **negative marking** at EBOD (*to avoid wild guesses with a probability as high as 50 % to be correct*) have a **positive effect** on the **statistical performance** parameters of all EBOD test items **in total** and test **items individually** and?
2. Does **negative marking** have a **discriminative** effect towards **female candidates** (main argument against negative marking according to literature)?

INTRODUCTION OF NEGATIVE MARKING

- Hypothesis on the influence of negative marking
  - Average scores will drop (punishment of incorrect answers)
  - Spread of candidate scores will enlarge (→ room for discrimination)
  - Rit-value of individual items will increase
  - Reliability of EBOD will increase
- Argument against negative marking expressed in literature
  - Negative marking is discriminating towards female candidates

INTRODUCTION OF NEGATIVE MARKING

The **EBO written MCQ** examination:

- The pass mark (6) of the MCQ paper is set at the average minus 1 standard deviation (SD):

	2005	2006	2007	2008	2009	2010
MCQ Mean ± SD	193 ± 15	189 ± 14	191 ± 15	184 ± 15	204 ± 13	**146 ± 25**

- The result of the MCQ paper counts for 40 % of the final score

INTRODUCTION OF NEGATIVE MARKING

The **EBOD examination success rate**:

- As the level of candidates tends to be good the overall EBO examination success rate is usually high:

	2005	2006	2007	2008	2009	2010
Success Rate	87.6 %	88.1 %	89.2 %	90.8 %	88.6 %	**92.0 %**

STUDY POPULATION, DATA ANALYSIS AND RESULTS

Study population:

2009: 308 candidates (M: 185, F: 123) took part at EBOD **without negative marking**
2010: 310 candidates (M: 168, F:142) took part at EBOD **with negative marking**

Data analysis and Results (Statistical performance parameters):

Statistical Performance Parameter	Parameter: Rule of thumb	2009	2010
General statistical performance of EBOD (i.e. all items) • **Cronbach-α value (internal consistency)** *to be considered as the degree to which all test items are measuring the same (i.e. knowledge of candidates)*	**Cronbach-α ≥ 0.80**	0.78	**0.87**
Statistical performance of individual EBOD items (average) • **P-value (percentage of correct answers)** *to be considered as an estimation of the level of difficulty (or facility) of test items* • **Rit-value (correlation of item score with EBOD score)** *to be considered as the degree to which a test item has an added value to the total examination*	**0.10 < P-value < 0.90** **Rit-value ≥ 0.15**	**0.79** 0.14	**0.66** **0.18**

EBOD: INTERNAL CONSISTENCY

- Cronbach's coefficient alpha (r) = 0.87 (2009: 0.78)
  - Estimator of the lower bound of the **internal consistency** (degree to which all MCQs leaves (n) are measuring the same i.e. knowledge of candidates) of EBOD 2010 (95% CI: 0.86 – 0.89)

$$r = \frac{n}{n-1}\left[1 - \frac{\sum_{i=1}^{n} \sigma_{Yi}^{2}}{\sigma_{x}^{2}}\right]$$

internal consistency of EBOD MCQ-test is good

EBOD: DEGREE OF DIFFICULTY

- EBOD 2009
  - Degree of Difficulty (P-value) of 0.79 (over-estimated due to guessing)
  - Estimation of a large proportion of candidates guessing (> 33 %)
- EBOD 2010
  - Introduction of the “Don’t know” option
    → reduction of wild guesses
    → used on average for 15 % of items (or 39 items) per candidate
  - Degree of Difficulty (P-value) of 0.66

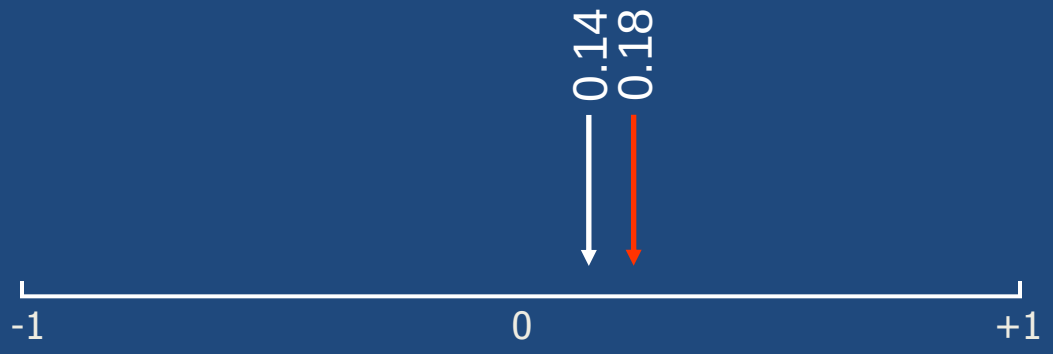
EBOD: POINT BISERIAL CORRELATION

- Point biserial correlation coefficient (Rit)
  - Estimator of the correlation between
    the **individual item scores Y_i** (either -0.5, 0 or 1) and
    the **total MCQ scores X_i** (ranging from 61.5 to 209) of the candidates

$$Rit_i = \frac{\sum_{i=1}^{n} \left(\frac{X_i - \overline{X}}{\sigma_X} \right) \left(\frac{Y_i - \overline{Y}}{\sigma_Y} \right)}{n - 1}$$

EBOD: POINT BISERIAL CORRELATION

- Point biserial correlation coefficient (Rit)

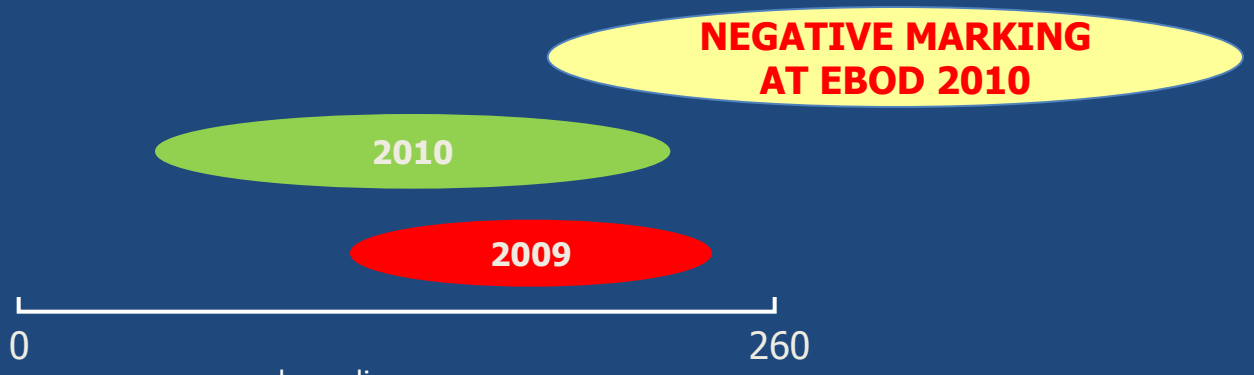


correlation between item and total MCQ score

EBOD 2010: SPREAD OF SCORES

- Negative marking to overcome disadvantages of multiple True/False items?
  - Increase of **discriminative power** of examination
  - Reduction of **guess factor**
    - wild guesses will be punished (weakest candidates)
    - guesses by reasoning (partial knowledge) will be rewarded

	2009	2010
Min	154	61.5
Max	230	209
Mean	204.11	145.99
Stdev	13.04	24.76



DATA ANALYSIS AND RESULTS, DISCUSSION

Data analysis and Results (Male versus Female):

2009 – Male versus Female candidates (χ^2 test)		2010 – Male versus Female candidates (χ^2 test)	
Difference between pass – fail?	p = 0.909 (NS)	Difference between pass – fail?	p = 0.286 (NS)
Difference between scores (1–10)?	p = 0.430 (NS)	Difference between scores (1–10)?	p = 0.264 (NS)
		Difference between "Don't know"	**p = 0.02 (S)**

Discussion:

- The **rationale** behind **negative marking** relies upon the fact that with true-false test items the **probability** of a **correct answer** by **guessing** is as high as **50 %** due to which the **level** of the non-able or **borderline candidates** is generally assumed to be **over-estimated**. Hence the space available to **discriminate** able from borderline candidates is (too) **limited**.
- The main **argument against** negative marking as described in literature is the **assumption** that negative marking would be **discriminative towards females**.

STATISTICAL EVALUATION EBOD 2010

- Analysis of failed candidates …
  - 24 Failed candidates at EBOD 2010:
    * Residents: 12 (236 in total) or 5.1 % (2009: 8.2 %)
    * Specialists: 12 (74 in total) or 16.2 % (2009: 15.9 %)

2009	Fail	Pass
Residents	18	202
Specialists	14	74

2010	Fail	Pass
Residents	12	224
Specialists	12	62

* Odds ratio 2009: 0.47 (18 × 74) / (14 × 202)
* Odds ratio 2010: 0.28 (12 × 62) / (12 × 224)

DISCUSSION AND CONCLUSION

Discussion:

- Cronbach-α value: - has **improved** after introduction of negative marking
- P-value: - was **no longer over-estimated** with negative marking
- Rit-value: - has **improved** after introduction of negative marking
- Males vs. Females: - Female candidates are **less keen to guess** (significance) (female candidates choose more for "Don't know")
  - Female candidates have the **same chances to pass** EBOD as male candidates (no significant difference in scoring)

Conclusion:

The introduction of **negative marking** for EBOD has proven to be **very successful** with **benefits** for both:

- the **organiser** of the examination (statistical performance parameters), and
- the **candidates** (better discrimination with borderline candidates).

THANK YOU VERY MUCH FOR YOUR ATTENTION

Faleminderit shumë (Albanian) Shterakravetsun (Armenian) Eskerrik asko (Basque)
Mnogo blagodarya (Bulgarian) Dzãkujã (Cassubian) Moltes gràcies (Catalan) Merastawhy (Cornish)
À ringraziavvi (Corsican) Hvala lijepa (Croatian) Děkuji (Czech) Mange tak (Danish) Dank u wel (Dutch)
Thank you (English) Ic sæcge eow Þancas (English, old) Dankon al vi (Esperanto) Aitäh (Estonian)
Paljon kiitoksia (Finnish) Merci beaucoup (French) Tanke wol (Frisian) Graciis (Friulian) Grazas (Galician)
Mèrczi (Gallo) Merci (Gascon) Vielen dank (German) Merci villmahl (German: Zurich Switzerland)
Ευχαριστώ (Greek) Toda raba (Hebrew) Nagyon köszönöm (Hungarian) Takk fyrir (Icelandic)
Gratias (Interlingua) Qujanaq (Inuttut) Go raibh mile maith agaibh (Irish Gaelic) Gratias tibi ago (Latin)
Liels paldies (Latvian) Mouchou gratzia (Lingua Franca) Labai achiu (Lithuanian) Merci (Luxembourgish)
Grazzi hafna (Maltese) Gura mie mooar ayd (Manx) Merçì (Monegasque) Gràzzie (Napulitano)
Dziękuję (Polish) Obrigado (Portuguese) Mercé plan (Provencal) Nais tuke (Romani: gypsy)
Oven saste (Romani) Mulţumesc (Romanian) Grazscha (Romansch) Спасибо (Russian)
Giitus eanat (Saami Lappish) Moran taing (Scottish Gaelic) Grazzii (Sicilian) Dakujem vám (Slovak)
Hvala lepa (Slovenian) Dz'akujo so (Sorbian) Muchas gracias (Spanish) Dankeschee (Swabian)
Tackar så mycket (Swedish) Çok tesekkür ederim (Turkish) Moltes gracies (Valencian)
Merci (Walloon) Diolch yn fawr iawn (Welsh) A dank aych (Yiddish)