EBOD 2009 Examination: Statistical evaluation of the results

SOE C14 (EDU) – How to be successful at the EBO examination? – June 14th, 2009

MSc. Danny G.P. Mathysen

EBOD Assessment Officer

Antwerp University Hospital
Department of Ophthalmology
Wilrijkstraat 10 – B-2650 Edegem
✉️ danny.mathysen@uza.be

No potential conflict of interest to disclose ...
EBO Diploma Examination

- Test designed to assess **knowledge** and **clinical skills** requisites to deliver a **European standard** of ophthalmologic care in hospital and private settings.

- Organised on a **yearly basis** since its introduction in June 1995
Components of education

- **Knowledge**
  - Textbooks, guidelines, articles, (online) courses, clinical ophthalmological practice, ...
  - **EBOD MCQs + EBOD Viva Voce**

- **Skills**
  - Practical/technical skills
    - measure of ophthalmological apprenticeship

- **Professionalism**
  - Personal skills: empathy, attitude, relation with related paramedicals, ...
Structure of EBOD

- **Written paper**
  - 52 MCQs with 5 T/F items each (260 items)
  - 10 topics (see EBO website)
  - 40 percent of total candidate score
  - Languages: **English, French, German**
  - Assessment of knowledge

- **Oral examination (Viva Voce)**
  - 4 different topics (see EBO website)
  - 60 percent of total candidate score
  - Languages: **English** (native language)
  - Assessment of knowledge
1. The age of onset of presbyopic symptoms
   a. Is earlier for a patient with a small amplitude of accommodation **TRUE**
   b. Is earlier for a hyperopic patient who wears contact lenses rather than spectacles **FALSE**
   c. Is earlier for a myopic patient who wears contact lenses rather than spectacles **TRUE**
   d. Is earlier for a myopic than a hyperopic patient who wears spectacles with full distance correction **FALSE**
   e. Is earlier for a short than a tall patient **TRUE**

Correct answers (2 out of 5) are rewarded with 1 point

⇒ This candidate receives 2 points
Written paper (MCQs)

- Scoring of the written paper
  - For each candidate a **total test score** is calculated (theoretical range: 0 – 260)
    - sum of all items answered correctly by the candidate
  - The **average total test score** with according **standard deviation** is calculated
  - Based on the position of the individual total test score according to the average total test score, the candidate will receive a **mark** that will be situated **between 1 and 10**
  - This MCQ mark counts for **40 percent** of the total EBOD score
Written paper (MCQs)

- Advantages for EBO candidates of T/F items
  - Reliable in case of translation (English, French, German)
    - choice of language will not result in being (dis)advantaged
  - Accessibility (e.g. dyslexia)
    - not too complicated for candidates
  - Duration of the examination
    - stress level of candidates can be kept to a minimum
  - Relatively easy to process
    - results can be presented on-site

- Disadvantage for EBO candidates of T/F items
  - Probability of guessing right = 50 %
    - level of weakest candidates is overestimated (oral examination)
Written paper (MCQs)

How to overcome the disadvantages of T/F items?

- Introduction of negative marking
  - 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

NEGATIVE MARKING AT EBOD 2010
Written paper (MCQs)

1. The age of onset of presbyopic symptoms
   a. Is earlier for a patient with a small amplitude of accommodation **TRUE**
   b. Is earlier for a hyperopic patient who wears contact lenses rather than spectacles **FALSE**
   c. Is earlier for a myopic patient who wears contact lenses rather than spectacles **TRUE**
   d. Is earlier for a myopic than a hyperopic patient who wears spectacles with full distance correction **FALSE**
   e. Is earlier for a short than a tall patient **TRUE**

Correct answers (2 out of 5) are rewarded with 1 point
Incorrect answers (3 out of 5) result in −1 point
⇒ This candidate receives −1 point
Written paper (MCQs)

Scoring of the written paper:

- For each candidate a **total test score** is calculated (theoretical range: -260 – 260)
  - sum of all items answered correctly by the candidate minus sum of all items answered incorrectly by the candidate

- The **average total test score** with according **standard deviation** is calculated

- Based on the position of the individual total test score according to the average total test score, the candidate will receive a **mark** that will be situated **between 1 and 10**

- This MCQ mark counts for **40 percent** of the total EBOD score

**NEGATIVE MARKING**

*YOUR CHANCES TO PASS WILL NOT DECREASE BY INTRODUCING NEGATIVE MARKING*
Oral examination (Viva Voce)
Oral examination (Viva Voce)

- Scoring of the oral examination
  - For each candidate a **viva voce score** is given for each topic (theoretical range: 1 – 10)
  - Each **individual viva voce score** counts for **15 percent** of the total EBOD score

\[ EBOD_i = [0.4 \times MCQ_i] + [0.15 \times \left( VVa_i + VVb_i + VVc_i + VVd_i \right)] \]

**40 % MCQ-score**

**60 % Viva Voce score**
Statistical analysis of EBOD 2009

- **SpeedWell**
  - SpeedWell is specialised in organising medical examinations
  - Optical reader system
    - continuous and yearly increase of applications / interest in EBOD
  - Provided software tools
    - Design of the MCQ answer sheet
    - Design of the Viva Voce mark sheets
    - Statistical analysis output (MultiQuest®) based on similar statistical techniques

- **Yearly increase of candidates**
  - 2005: 74
  - 2006: 159
  - 2007: 224
  - 2008: 284
  - 2009: 308
Demographics of EBOD 2009

Many EU countries apply

<table>
<thead>
<tr>
<th>Country</th>
<th>2008</th>
<th>2009</th>
<th>Δ</th>
<th>Country</th>
<th>2008</th>
<th>2009</th>
<th>Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>2</td>
<td>5</td>
<td>↑</td>
<td>Latvia</td>
<td>2</td>
<td>1</td>
<td>↓</td>
</tr>
<tr>
<td>Belgium</td>
<td>23</td>
<td>25</td>
<td>↑</td>
<td>Lithuania</td>
<td>1</td>
<td>1</td>
<td>→</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>4</td>
<td>4</td>
<td>↑</td>
<td>Norway</td>
<td>1</td>
<td>1</td>
<td>↑</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>2</td>
<td>2</td>
<td>→</td>
<td>Poland</td>
<td>1</td>
<td>2</td>
<td>↑</td>
</tr>
<tr>
<td>Denmark</td>
<td>4</td>
<td>6</td>
<td>↑</td>
<td>Slovakia</td>
<td>1</td>
<td>1</td>
<td>→</td>
</tr>
<tr>
<td>Estonia</td>
<td>3</td>
<td>2</td>
<td>↓</td>
<td>Slovenia</td>
<td>6</td>
<td>5</td>
<td>↓</td>
</tr>
<tr>
<td>Finland</td>
<td>7</td>
<td>2</td>
<td>↓</td>
<td>Spain</td>
<td>14</td>
<td>17</td>
<td>↑</td>
</tr>
<tr>
<td>France</td>
<td>92</td>
<td>96</td>
<td>↑</td>
<td>Sweden</td>
<td>6</td>
<td>5</td>
<td>↓</td>
</tr>
<tr>
<td>Germany</td>
<td>44</td>
<td>59</td>
<td>↑</td>
<td>Switzerland</td>
<td>32</td>
<td>29</td>
<td>↓</td>
</tr>
<tr>
<td>Greece</td>
<td>10</td>
<td>19</td>
<td>↑</td>
<td>The Netherlands</td>
<td>7</td>
<td>7</td>
<td>→</td>
</tr>
<tr>
<td>Hungary</td>
<td>1</td>
<td>2</td>
<td>↑</td>
<td>Turkey</td>
<td>11</td>
<td>5</td>
<td>↓</td>
</tr>
<tr>
<td>Ireland</td>
<td>5</td>
<td>5</td>
<td>→</td>
<td>United Kingdom</td>
<td>2</td>
<td>1</td>
<td>↓</td>
</tr>
<tr>
<td>Italy</td>
<td>4</td>
<td>6</td>
<td>↑</td>
<td><strong>Total</strong></td>
<td><strong>284</strong></td>
<td><strong>308</strong></td>
<td>↑</td>
</tr>
</tbody>
</table>

Many EU countries apply.
EBOD 2009: Analysis of MCQs

- MCQ total scores
  - Range of total scores: 154 – 230
  - Mean ± SD total score: 204.11 ± 13.04

EBOD 2009 MCQ Scores with 95% Confidence Intervals

No significant difference!

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>Mean ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residents</td>
<td>220</td>
<td>205.40 ± 12.18</td>
</tr>
<tr>
<td>Specialists</td>
<td>88</td>
<td>200.91 ± 14.41</td>
</tr>
</tbody>
</table>
Residents have **higher total MCQ scores** with **lower standard deviations** when compared to specialists.

In general there are **no statistically significant differences** between **countries**.

<table>
<thead>
<tr>
<th></th>
<th>Belgium</th>
<th>Switzerland</th>
<th>Germany</th>
<th>France</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residents</td>
<td>n = 21</td>
<td>n = 29</td>
<td>n = 39</td>
<td>n = 84</td>
</tr>
<tr>
<td></td>
<td>207.71 ± 10.96</td>
<td>207.97 ± 12.22</td>
<td>209.67 ± 10.46</td>
<td>201.52 ± 11.22</td>
</tr>
<tr>
<td>Specialists</td>
<td>n = 4</td>
<td>n = 20</td>
<td>n = 12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>181.25 ± 20.22</td>
<td>206.10 ± 15.57</td>
<td>200.58 ± 15.20</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>n = 25</td>
<td>n = 29</td>
<td>n = 59</td>
<td>n = 96</td>
</tr>
<tr>
<td></td>
<td>203.48 ± 16.14</td>
<td>207.97 ± 12.22</td>
<td>208.46 ± 12.54</td>
<td>201.41 ± 11.80</td>
</tr>
</tbody>
</table>
# EBOD 2009: Analysis of MCQs

EBOD is not a language test!

<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th>German</th>
<th>French</th>
</tr>
</thead>
</table>
| **Residents** | n = 58  
205.98 ± 12.54 | n = 61  
209.46 ± 11.51 | n = 101  
202.60 ± 11.62 |
| **Specialists** | n = 53  
200.08 ± 12.71 | n = 21  
205.67 ± 15.27 | n = 14  
196.93 ± 17.06 |
| **Total**   | n = 111  
203.16 ± 12.96 | n = 82  
208.46 ± 12.54 | n = 115  
201.91 ± 12.55 |

Residents have **higher total MCQ scores** with **lower standard deviations** when compared to specialists.

In general there are **no statistically significant differences** between **languages**.
EBOD 2009: Analysis of MCQs

- Cronbach’s coefficient alpha ($r$) = 0.78
  - Estimator of the lower bound of the **internal consistency** (degree to which all MCQs leaves are measuring the same, i.e. knowledge of candidates) of EBOD 2009 (95% CI: 0.75 – 0.81)

\[
r = \frac{260}{260 - 1} \left[ 1 - \frac{\sum_{i=1}^{260} \sigma_i^2}{\left( \sum_{i=1}^{260} R_{it_i} \sigma_i \right)^2} \right] = 0.78
\]
Point biserial correlation coefficient (Rit) = 0.14

Estimator of the correlation between the individual item scores $X_i$ (either 0 or 1) and the total MCQ scores $Y_i$ (ranging from 154 to 230) of the candidates

$$R_{it} = \frac{1}{n-1} \sum_{i=1}^{n} \left( \frac{X_i - \bar{X}}{s_X} \right) \left( \frac{Y_i - \bar{Y}}{s_Y} \right)$$

correlation between item and total MCQ score
EBOD 2009: Analysis of MCQs

- Assessment of the degree of difficulty
  - Average P-value ≈ 0.79
    - Indication of items answered incorrectly by guessing ≈ 0.21
    - Estimation of items answered correctly by guessing ≈ 0.21
    - Estimation of percentage of candidates guessing ≈ 0.42
    OR Estimation of percentage of candidates knowing ≈ 0.58
Residents have **higher Viva Voce scores** with **lower standard deviations** when compared to specialists.

**No statistically significant differences** between the **topics**.

### EBOD 2009 Viva Voce Scores

<table>
<thead>
<tr>
<th>Topic</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Optics, Refractions, Strabismus and Neuro-ophthalmology</td>
<td>7.62 ± 1.32</td>
</tr>
<tr>
<td>B. Cornea, External diseases and Ocular adnexa</td>
<td>7.59 ± 1.29</td>
</tr>
<tr>
<td>C. Glaucoma, Cataract and Refractive surgery</td>
<td>7.45 ± 1.24</td>
</tr>
<tr>
<td>D. Posterior segment, Ocular inflammation and Uveitis</td>
<td>7.83 ± 1.28</td>
</tr>
</tbody>
</table>

**EBOD scores are high!**
Residents have higher Viva Voce scores with lower standard deviations when compared to specialists.

EBOD scores are comparable for MCQ and Viva Voce!
Residents have **higher total scores** with **lower standard deviations** when compared to specialists.

No significant differences are observed between the countries.
EBOD 2009: Success rate

Success rate of EBOD is much higher as compared to other medical specialties (60-70 %)

<table>
<thead>
<tr>
<th>Year</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>Success Rate</td>
<td>87.6%</td>
<td>88.1%</td>
<td>89.2%</td>
<td>90.8%</td>
<td>89.6%</td>
</tr>
</tbody>
</table>

EBOD success rate is quite stable over the years and quite high as the level of candidates usually tends to be good.

18 Residents (out of 220: 8.2%) and 14 specialists (out of 88: 15.9 %) failed at EBOD 2009. As there were 308 candidates the general failure rate was 10.4 %.
In conclusion…

- There are absolutely **no reasons** for you **not to participate** at EBOD since …
  - … EBOD examination is a chance to show your **excellence** in ophthalmology on a **European** level
  - … EBOD has a **high success rate** (compared to European examinations of other medical specialties)
  - … It has been shown by EBO that the level of candidates participating is usually very good, which results in **very good individual marks**
In conclusion...

- There are absolutely **no reasons** for you **not to participate** at EBOD since …
  - … It has been demonstrated that **language** nor **country of origin** (dis)advantage candidates
  - … It has been shown that **residents** perform well at prior EBOD editions
  - … Introduction of **negative marking** will only be **beneficial** for good candidates! Furthermore introduction of negative marking **will not decrease** your **chances to be successful** at EBOD
In conclusion...

... Therefore EBO hopes to welcome you all at EBOD 2010!