Analysis of content of theory question papers in preliminary examinations and marks obtained by first-year MBBS students in physiology

Srabani Bhattacharya, Rucha Wagh, Aniruddha Malgaonkar and Sundaram Kartikeyan

Abstract
This complete enumeration, cross-sectional comparative record-based study was conducted at a medical college in Maharashtra state, India to analyse the coverage of various topics in Physiology question papers for the preliminary written theory examination and to analyse the patterns of marks scored by the First-year MBBS students during the six-year period of study (2011-2016). Though all the questions belonged to the “must know” category, questions on special senses, central nervous, gastro-intestinal, reproductive, endocrine, and renal systems predominated as compared to that on cardiovascular and respiratory systems, nerve-muscle, blood, body temperature and nutrition. Though female students had higher scores in most batches, the gender differences in mean scores for Batches 2011 (P=0.04) and Batch 2013 (P=0.007) were statistically significant. Though methods, such as, blueprinting are available it is essential that university guidelines are established for distribution of marks to various topics so that consistency can be ensured in every question paper.

Keywords: Content analysis, gender differences, physiology examination

1. Introduction
Medical education endeavours to impart holistic training at all stages of education in a way that includes the five levels of Bloom’s Taxonomy of Educational Objectives [1]. Assessment becomes an important component of medical education and an integral part of curriculum of a course because it reveals the process of students’ learning [2]. For assessing various domains of learning, different methodologies are being explored [3]. The assessment tools commonly employed in the theory paper are reasoning (long answer type of questions) and answers to specific questions (short answer type of questions). However, a few disadvantages such as less number of questions, unfair representation or distribution of topics, different difficulty level of questions, vague questions are often encountered in question papers of written examinations.

Content validity of an assessment refers to the extent to which a test actually measures the intended content area [4]. Adequate coverage of the course content is necessary for the validity of assessment. Setting a balanced question paper with respect to content validity would be the first priority of any assessment. As per the guidelines of the Maharashtra University of Health Sciences, affiliated medical colleges should conduct one term-ending examination and one preliminary examination for First-year MBBS students before they appear for First MBBS University examinations. The preliminary examination question papers and pattern of scores were selected for this study because the entire Physiology syllabus is covered. Moreover, the students appearing for preliminary examination have already had prior experience of term-ending examination which is conducted on a similar pattern but covers only a part of the total syllabus for the subject of Physiology. The study period (academic years 2011-2016), was chosen because the same set of teachers taught various topics in the subject of Physiology and thus the effect of confounding variables due to “teacher effect” would be negligible. This study was conducted to analyse the coverage of various topics in Physiology question papers for the preliminary written theory examination and to analyse the patterns of marks scored by the First-year MBBS students.
2. Materials and methods
2.1 Type of Study
Complete enumeration, cross-sectional comparative record-based study

2.2 Study Setting
This study was conducted at Rajiv Gandhi Medical College, Kalwa, Thane, located about 30 kilometres from Mumbai, Maharashtra state, India. This medical college, affiliated to the Maharashtra University of Health Science, has an intake capacity of 60 students during each academic year.

2.3 Pattern of examination
As per university guidelines, the written theory preliminary examination consists of two theory papers (paper-1 & paper-2) carrying a total of 80 marks. Each paper has optional questions carrying 12 marks. Both paper-1 & paper-2 comprise short answer and long answer questions.

2.4 Procedure
After obtaining permissions from the Institutional Ethics Committee (IEC) and institutional authorities for conducting the study, the theory question papers of preliminary examinations during the study period (2011-2016) were analyzed for content. Preliminary examination marks pertaining to the same period, were analyzed to determine the patterns of marks scored by the First-year MBBS students.

2.5 Statistical analysis
The data were statistically analyzed using EpiInfo Version 7.0 (public domain software package from the Centers for Disease Control and Prevention, Atlanta, GA, USA). Continuous data were presented as Mean and Standard Deviation (SD), 95% Confidence interval (CI) was stated as: [Mean-(1.96)* Standard Error] - [Mean+(1.96)* Standard Error]. Student's t-test was used to determine the statistical significance of difference between two means. Statistical significance was determined at P<0.05

3. Results & Discussion
357 students appeared for preliminary examinations in the subject of Physiology during the study period (2011-2016). The question paper content of preliminary examinations during the study period and the marks scored by these 357 students were analyzed.

3.1 Content Analysis
In the present study, the marks (mean ± standard deviation) allocated to theory questions on various topics during the six-year period of study (2011-2016) were central nervous system (23.2 ± 4.38), gastro-intestinal system (11.2 ± 4.38), reproductive system (10.4 ± 2.19), endocrine system (10.4 ± 3.58), renal system (9.6 ± 2.19), and special senses (9.6 ± 3.58). All the questions in the question papers during the six-year period of study (2011-2016) belonged to the “must know” category. Questions on topics such as cardiovascular system, respiratory system, nerve-muscle were assigned between 4 and 8 marks. Questions on blood and body temperature were consistently allotted 4 marks. The mean marks allocated to questions on nutrition was only 1.6 ± 2.19.

Similar observations have been made in previously published studies, which emphasize the need for guidelines for distribution of marks to various content areas [2-5, 6]. A study conducted at Gujarat University has also reported that the year-wise distribution for various topics was not uniform [7]. The pattern of allocation of marks also influences the content of student learning [8].

Currently, there are no precise university guidelines regarding the marks to be allotted to various topics in the theory question papers. Teachers select questions from various systems according to their own judgement. Conflicting opinions may be obtained from subject experts regarding the allocation of marks to various topics in theory question papers [9].

Adequate coverage of course content is necessary for the validity of the assessment. Blueprinting, a method of linking assessment with learning objectives, stipulates the objectives that are to be tested in an examination and the relative weights to be assigned to each objective. Blueprinting may be considered an obligatory requirement for conducting an examination [10]. Since new topics are added to the syllabus, blueprints should be revised every academic year [11]. Content validity can also be established by framing a specifications grid which should identify the content areas, specify learning outcomes, determine the number of items for each content area and learning objective and ensure that the number of items in each cell is in proportion to the time spent in teaching and learning [12].

3.2 Marks Obtained by Students
The marks (mean ± standard deviation, out of a total of 80) obtained by all students (both genders) during preliminary examinations over the six-year period (2011 to 2016) were as follows – 2011 Batch (n=60): 56.27 ± 11.76 (95% CI: 53.29-59.24); 2012 Batch (n=60): 44.41 ± 11.62 (95% CI: 41.47-47.35); 2013 Batch (n=60): 52.03 ± 10.27 (95% CI: 49.43-54.62); 2014 Batch (n=58): 51.05 ± 9.58 (95% CI: 48.59-53.52); 2015 Batch (n=60): 54.03 ± 10.61 (95% CI: 52.34-56.71) and 2016 Batch (n=59): 54.58 ± 7.58 (95% CI: 52.66-56.49)

3.3 Gender differences
The maximum score out of 80 was higher for female students in all batches except Batch 2012 and Batch 2016. The third quartile was higher for female students in all batches except Batch 2016. The median score for Batch 2014 was identical for both sexes, while it was higher for female students in all batches except Batch 2016. The first quartile was higher for female students in all batches except Batch 2014. The minimum score out of 80 was higher for female students in all batches except Batch 2014 and Batch 2016. (Fig 1) However, only the gender differences in mean scores for Batches 2011 (P=0.04) and Batch 2013 (P=0.007) were statistically significant. (Table 1) Except for Batches 2014 and 2016, the variance in scores (square of standard deviation) was higher for male students, as compared to that for their female counterparts. (Table 1) Researchers [13, 14] have reported that male students show more variability in scores as compared to their female counterparts.
**Table 1:** Gender differences in mean scores

<table>
<thead>
<tr>
<th>Batch</th>
<th>Females</th>
<th>Males</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>n</td>
</tr>
<tr>
<td>2011</td>
<td>59.27</td>
<td>11.24</td>
<td>31</td>
</tr>
<tr>
<td>2012</td>
<td>45.00</td>
<td>11.19</td>
<td>32</td>
</tr>
<tr>
<td>2013</td>
<td>55.01</td>
<td>8.70</td>
<td>35</td>
</tr>
<tr>
<td>2014</td>
<td>50.74</td>
<td>10.68</td>
<td>33</td>
</tr>
<tr>
<td>2015</td>
<td>55.48</td>
<td>9.01</td>
<td>40</td>
</tr>
<tr>
<td>2016</td>
<td>53.89</td>
<td>8.03</td>
<td>31</td>
</tr>
</tbody>
</table>

SD = Standard deviation; n = number of students; *Statistically significant

3.4. Limitations
Due to absence of university guidelines regarding marks to be allotted to each topic in the question papers, the allocation of marks to various topics in theory question papers have not been correlated with marks scored by students or the number of teaching hours allotted to that particular topic.

4. Conclusion
It is imperative that while setting a question paper, some topics should be selected on the basis of importance (the “must know” category of topics). Consistency can be ensured in every question paper if there are established university guidelines for marks distribution to the topics. Evaluation should be used as a tool for promoting learning and should not be used for merely grading and certifying students. Therefore, methods such as blueprinting and table of specifications should be employed during the question paper setting process to improve the assessment system.

5. References