Perceptions of the Use of Blueprinting in a Formative Theory Assessment in Pharmacology Education

*Tejas Patel, Manoj K. Saurabh, Parvati Patel

Abstract: Objectives: This study aimed to assess perceptions of the use of a blueprint in a pharmacology formative theory assessment. Methods: This study took place from October 2015 to February 2016 at a medical college in Gujarat, India. Faculty from the Department of Pharmacology used an internal syllabus to prepare an assessment blueprint. A total of 12 faculty members prepared learning objectives and categorized cognitive domain levels by consensus. Learning objectives were scored according to clinical importance and marks were distributed according to proportional weighting. A three-dimensional test specification table of syllabus content, assessment tools and cognitive domains was prepared. Based on this table, a theory paper was created and administered to 126 pharmacology students. Feedback was then collected from the faculty members and students using a 5-point Likert scale. Results: The majority of faculty members agreed that using a blueprint ensured proper weighting of marks for important topics (90.00%), aligned questions with learning objectives (80.00%), and weighted distributed questions according to clinical importance (100.00%) and minimised inter-examiner variations in selecting questions (90.00%). Few faculty members believed that use of the blueprint created too many easy questions (10.00%) or too many difficult questions (10.00%). Most students felt that the paper had a uniform distribution of questions from the syllabus (90.24%), that important topics were appropriately weighted (77.23%), was well organised (79.67%) and tested in-depth subject knowledge (74.80%). Conclusion: These findings indicate that blueprinting should be an integral part of written assessments in pharmacology education.

Keywords: Medical Education; Educational Measurement; Perceptions; Students; Faculty; Reliability and Validity.
Written examinations are the most commonly employed method to assess knowledge in medical education and are used to test recall abilities as well as higher-order cognitive functions, such as the interpretation of data and problem-solving skills. Valid assessment methods are necessary to determine whether students have learned the required information. Content validity gauges the extent to which an assessment covers a representative sample of the material which should be assessed; for example, if examination questions cover the learning objectives of the syllabus, the examination is considered to have content validity. In contrast, construct validity covers all aspects of subject knowledge such as application, data gathering and interpretation as a collection of interrelated components, which together allow for the assessment to make sense.

Construct under-representation and construct irrelevance variance are two important challenges to construct validity. If an examiner creates a test with overly easy or difficult questions, asks well-known questions or uses unnecessarily complex language, it would mean that there is construct irrelevance variance, which can lead to the inflation or deflation of test scores. Content under-representation in a paper can occur due to the inadequate weighting of marks for clinically relevant topics, unequal distribution of course content across the assessment or examiner bias, such as a tendency to focus on popular topics. Moreover, the teachers who deliver lectures are usually not the ones who create the assessments; this can reduce the content and construct validity of an assessment, which in turn can lead to a mistrust in the assessment system on the part of the students. Content imbalance may also result in students focusing less on key areas of learning during revision. For an assessment paper to be valid, it should match course content, have proportional weighting of content according to clinical importance, consist of questions which are neither overly difficult nor easy and have multiple tools to determine various types of information.

Blueprinting can be defined as the creation of a template to determine the content of a test; it lists the number and type of questions across the course content, with learning objectives and relative weighting given to each topic. A blueprint provides a systematic multi-step approach to an assessment, defining the purpose (e.g. formative/summative and written/practical) and scope (e.g. for undergraduate or postgraduate students) of the test in order to subsequently determine content and method of assessment. Based on the content, learning objectives and their domains are identified and different assessment tools are chosen, such as short answer questions (SAQs), essay questions (EQs) or multiple choice questions. The content of the assessment is then proportionally weighted as per clinical importance, learning domains and methods of assessment. The total weighting of the number of items to be included in the assessment is decided and a three-dimensional (3D) table of test specifications is prepared to align content, learning domains and assessment tools, as well as prepare individual questions.

Blueprinting is increasingly used in the field of medical education worldwide. In the UK, assessments created using a blueprint are considered essential to enable future doctors to meet mandatory standards; the assessments are prepared in such a way that students who have not met important learning outcomes are not able to graduate. Previous research indicates that blueprinting optimises student assessment, making a positive impact and helping them to focus on key areas in an examination, thereby improving performance. This approach can reduce inter-individual variability by providing guidance to examiners; moreover, set question papers are usually more valid and reliable than those created without a blueprint. However, studies conducted on the use of blueprinting in India are scarce.

At the Gujarat Medical Education & Research Society (GMERS) Medical College in Gotri, Vadodara, Gujarat, India, a five-year undergraduate medical education programme results in a Bachelor of Medicine and Bachelor of Surgery (MBBS). Pharmacology education is mainly taught to second-year MBBS students for a total of three semesters of six months’ duration each. The assessment system consists of one summative and two formative examinations, with two written papers in the summative examination and one written paper in one of the formative examinations. Each paper is scored out of 58 marks, although students now receive up to 40 marks due to optional questions. The papers include constructed-response open-ended questions including EQs (4 marks each), short EQs (3 marks each) and SAQs (2 marks each); the maximum number of marks for EQs, short EQs and SAQs are 20, 24 and 14, respectively. In order to pass, a student must get at least 50% on the examination. Traditionally, the assessment content is determined by a paper setter who selects the questions according to the syllabus and question format. This study aimed to prepare a blueprint for a written theory paper and to analyse subsequent feedback on its use in a formative examination by pharmacology faculty members and students.
Methods

This study was conducted between October 2015 and February 2016 at the Department of Pharmacology of GMERS Medical College. In order to familiarise departmental faculty members with the concept of preparing a theory paper via blueprinting, a pilot test consisting of 16 SAQs worth two marks each was administered to 12 members of the faculty. A half-day interactive session was then conducted in three sessions: the first focused on assessment methods and tools, the second on validity and reliability in assessments and the third on the purpose and implementation of blueprinting, including the weighting of topics and assessment methods and how to prepare tables of test specifications. Faculty members were then re-tested using the same SAQs in order to determine the learning outcomes of the interactive session.

The syllabus of the first formative theory examination for second-year MBBS students was used to prepare a blueprint for the written paper, consisting of seven topics: general pharmacology; the autonomic nervous system; the peripheral nervous system; the respiratory system; the gastrointestinal tract; autacoid-related drugs; and drugs affecting blood and blood formation. A literature review of the undergraduate regulations, vision documents, essential drug lists and national health programmes of the Medical Council of India as well as pharmacology textbooks and previous test papers was undertaken to determine learning objectives for each topic.14–20 Each learning objective was categorised into either recall or reasoning domain categories as per Miller’s pyramid of competence.21 Learning objectives and cognitive domain categorisations were then discussed by the same 12 faculty members who had participated in the pilot test. Each faculty member scored learning objectives individually according to clinical importance, with a score of 3 indicating high importance, a score of 2 indicating moderate importance and a score of 1 reflecting little/no clinical importance.7

Mean scores for each learning objective were calculated using a Microsoft Excel spreadsheet, Version 2010 (Microsoft Corp., Redmond, Washington, USA). Differences were resolved via consensus. Based on the learning objective scores, total scores were calculated for each topic as well as the overall syllabus. Proportional weighting was calculated for each topic by dividing the topic score with the total syllabus score.3,7 Table 1 shows the learning objectives, total scores, proportional weighting and mark distribution calculations. Based on the literature review and faculty consensus, a total of 292 learning objectives were identified. The total syllabus score was 663 and the maximum number of marks on the theory test paper was 58, with marks in each topic weighted proportionally.

In the next phase, a 3D table of test specifications for the distribution of marks was prepared by aligning content, assessment tools and cognitive domain categories. Table 2 shows the 3D table of test specifications, representing the mark distribution for each topic according to assessment tool (e.g. EQ, short EQ or SAQ) and cognitive domain category (i.e. recall or reasoning). In terms of assessment tools, EQs were used only to assess reasoning abilities (20 marks; 100.00%). Short EQs were more often used to assess reasoning (15 marks; 62.50%) rather than

Table 1: Learning objectives, mean total scores, proportional weighting and distribution of marks per syllabus topic

<table>
<thead>
<tr>
<th>Syllabus topic</th>
<th>Learning objectives</th>
<th>Mean total score</th>
<th>Proportional weighting</th>
<th>Distribution of marks</th>
</tr>
</thead>
<tbody>
<tr>
<td>General pharmacology</td>
<td>87</td>
<td>201</td>
<td>0.30</td>
<td>17.40 (17)</td>
</tr>
<tr>
<td>Autonomic nervous system</td>
<td>70</td>
<td>163</td>
<td>0.25</td>
<td>14.50 (15)</td>
</tr>
<tr>
<td>Peripheral nervous system</td>
<td>24</td>
<td>45</td>
<td>0.07</td>
<td>4.06 (4)</td>
</tr>
<tr>
<td>Gastrointestinal tract</td>
<td>21</td>
<td>44</td>
<td>0.07</td>
<td>4.06 (4)</td>
</tr>
<tr>
<td>Respiratory tract</td>
<td>17</td>
<td>43</td>
<td>0.06</td>
<td>3.48 (4)</td>
</tr>
<tr>
<td>Autacoid-related drugs</td>
<td>31</td>
<td>71</td>
<td>0.11</td>
<td>6.38 (6)</td>
</tr>
<tr>
<td>Drugs affecting blood and blood formation</td>
<td>42</td>
<td>96</td>
<td>0.14</td>
<td>8.12 (8)</td>
</tr>
<tr>
<td>Total</td>
<td>292</td>
<td>663</td>
<td>1.00</td>
<td>58.00 (58)</td>
</tr>
</tbody>
</table>

*Learning objectives were scored by faculty members according to clinical importance, with a score of 3 indicating high importance, a score of 2 indicating moderate importance and a score of 1 reflecting little/no clinical importance.

†Proportional weighting was calculated by dividing the mean total score per topic by the total syllabus score.

‡Distribution of marks was determined by multiplying the proportional weighting by the total number of marks overall. The numbers within the brackets show the actual marks allotted per topic in the test paper.
recall (9 marks; 37.50%). In comparison, SAQs were more frequently used to test recall (8 marks; 57.14%) rather than reasoning (6 marks; 42.86%). Overall, the distribution of recall (17 marks; 29.31%) to reasoning (41 marks; 70.69%) marks was approximately 30:70.

Based on the blueprint, a paper setter prepared individual test questions following good practices. The questions were framed using directive verbs from the revised Bloom’s taxonomy of learning domains to assess the appropriate cognitive domain according to the blueprint and to give students clear directions for their responses. The verb “remember” was used to frame recall questions, while the verbs “understand”, “apply” and “analyse” were used for reasoning questions. The verbs “evaluate” and “create” were not used. The paper setter selected questions of high, moderate and no/little clinical importance in a ratio of 60:30:10. The paper was then administered to a total of 126 second year pharmacology students.

The 12 faculty members were requested to provide feedback regarding their perceptions of the quality and use of the theory assessment paper designed using blueprinting. They were provided with two previous assessment papers for comparison. Feedback was collected using closed- and open-ended questionnaires. For the former, participants were given 14 statements and asked to respond on a 5-point Likert scale using the following responses: strongly disagree, disagree, neutral, agree and strongly agree. In the open-ended questionnaire, they were asked to describe benefits and difficulties in the preparation and implementation of the blueprint as well as the feasibility of this approach for future implementation in formative assessments. The perceptions of the students regarding the paper were determined one week after the assessment. A questionnaire consisting of nine statements was distributed and the students were asked to respond on the previously described 5-point Likert scale.

For the pilot test, the SAQ scores of the faculty members were presented as means and standard deviations and compared using a paired t-test. Results from the feedback questionnaires were presented as percentages. Strongly disagree and disagree responses were merged, as were the agree and strongly agree responses. Statistical analysis was performed using GraphPad Prism, Version 6.0 (GraphPad Software Inc., La Jolla, California, USA). A P value of <0.050 was considered statistically significant.

This study received ethical approval from the Institutional Human Ethics Committee of GMERS Medical College, Gotri (IHEC #101/2015/Pharmacology-16). Informed consent was obtained from the faculty and undergraduate medical students before data collection.

### Results

Among the 12 faculty members, the mean test scores before and after the interactive session were 7.72 ± 4.19 and 19.29 ± 5.29, respectively, out of a maximum score of 32 (paired t-test t-value: 10.23; degree of freedom: 11; P <0.001). In total, 10 out of 12 faculty members gave their opinions of the blueprint (response rate: 83.33%). The two remaining faculty members were transferred before the formative assessment was conducted. All of the faculty members agreed that the blueprint ensured a uniform distribution of questions across the syllabus topics, helped to maintain a balance between questions in the recall and reasoning domains and assured the distribution of questions according to clinical importance. In addition, the majority agreed

<table>
<thead>
<tr>
<th>Syllabus topic</th>
<th>EQs</th>
<th>Short EQs</th>
<th>SAQs</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reasoning</td>
<td>Recall</td>
<td>Reasoning</td>
<td>Recall</td>
</tr>
<tr>
<td>General pharmacology</td>
<td>8 (2)</td>
<td>3 (1)</td>
<td>3 (1)</td>
<td>2 (1)</td>
</tr>
<tr>
<td>Autonomic nervous system</td>
<td>4 (1)</td>
<td>3 (1)</td>
<td>3 (1)</td>
<td>2 (1)</td>
</tr>
<tr>
<td>Peripheral nervous system</td>
<td>-</td>
<td>-</td>
<td>3 (1)</td>
<td>-</td>
</tr>
<tr>
<td>Gastro-intestinal tract</td>
<td>-</td>
<td>-</td>
<td>3 (1)</td>
<td>2 (1)</td>
</tr>
<tr>
<td>Respiratory system</td>
<td>4 (1)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Autacoid-related drugs</td>
<td>4 (1)</td>
<td>-</td>
<td>-</td>
<td>2 (1)</td>
</tr>
<tr>
<td>Drugs affecting blood and blood formation</td>
<td>-</td>
<td>3 (1)</td>
<td>3 (1)</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>20 (5)</td>
<td>9 (3)</td>
<td>15 (5)</td>
<td>8 (4)</td>
</tr>
</tbody>
</table>

EQs = essay questions; SAQs = short answer questions.

### Table 2: Table of test specifications showing the distribution of marks for each topic according to assessment tool and cognitive domain category
that the blueprint resulted in the adequate weighting of important topics (90.00%), aligned questions with learning objectives (80.00%), ensured well-organised theory test papers (70.00%), tested in-depth subject knowledge (60.00%) and minimised inter-examiner variations in selecting questions (90.00%). Few faculty members believed that the blueprint created too many easy questions (10.00%) or too many difficult questions (10.00%) [Figure 1].

The majority of faculty members (90.00%) believed that a blueprint should be incorporated into the design of future examinations and all of them agreed that the blueprint should be an integral part of theory paper framing. Moreover, the majority of faculty members thought that blueprints should be prepared for the entire syllabus (90.00%) and that there was a need to change teaching schedules as per the blueprint (90.00%) [Figure 2]. Table 3 summarises faculty responses to open-ended questions about their perceptions of the blueprint and assessment paper.

A total of 123 out of 126 students provided feedback regarding the formative assessment paper (response rate: 97.62%). The majority of students believed that the distribution of questions was uniform and covered each topic (90.24%) and allowed for proper weighting of clinically important topics (77.23%). In addition, the paper was generally perceived to be well organised (79.67%). Most of the students believed that all of the questions were from the defined syllabus (90.24%) and that the paper tested in-depth knowledge of the subject (74.80%). Overall, few students believed that there were too many easy questions (12.20%) or too many difficult questions (9.75%); indeed, only a minority of the students felt that the paper was exhausting/lengthy (26.83%) or stressful (17.08%) [Figure 3].

**Discussion**

An assessment drives, directs and influences learning and is a tool for educational improvement.²,²³ Students successfully learn when a relationship exists between teaching, assessment and results. In an assessment, every question format (e.g. EQs, short EQs and SAQs) has advantages and disadvantages; using a variety of formats helps to counter the possible bias associated with one individual format.¹ A blueprint helps to

---

**Table 3: Summary of faculty responses to open-ended questions regarding the use of a blueprint in the formation of a pharmacology theory assessment paper (N = 10)**

<table>
<thead>
<tr>
<th>Open-ended responses</th>
<th>Faculty responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>What are the advantages of using a blueprint?</td>
<td>• It makes the assessment more objective&lt;br&gt;• There is less chance of subjectivity, bias and prejudice&lt;br&gt;• It makes the theory test examination fairer&lt;br&gt;• Students are tested on different cognitive domains&lt;br&gt;• It is easy for the examiner to prepare question papers&lt;br&gt;• It creates a well-organised paper and syllabus</td>
</tr>
<tr>
<td>What are the disadvantages of using a blueprint?</td>
<td>• It is time-consuming to score learning objectives by clinical importance&lt;br&gt;• It takes time to adjust marks according to recall and reasoning domains&lt;br&gt;• It is sometimes difficult to reach a consensus between faculty members&lt;br&gt;• It will require periodic revisions in future&lt;br&gt;• The choice of questions is restricted&lt;br&gt;• There is a chance that questions in the paper will be more predictable</td>
</tr>
<tr>
<td>What were the differences between the paper created with a blueprint and previous years’ papers?</td>
<td>• The questions in the blueprint-created paper were more specific&lt;br&gt;• The distribution of marks and questions followed the recommendations of the blueprint&lt;br&gt;• There was greater clarity in student assessment&lt;br&gt;• The blueprint-created paper allowed for more questions in a higher-function cognitive domain&lt;br&gt;• The paper devised with a blueprint was well structured</td>
</tr>
</tbody>
</table>

---

*One faculty member did not respond to these statements.*
ensure the balance between different format questions by specifying the content, learning domains across the syllabus and assessment tools of an examination in a rational manner. In the current study, a blueprint was used to create a written assessment paper for pharmacology students; learning objectives were identified and syllabus topics were marked using a proportional weighting system based on clinical importance to ensure content validity. Moreover, the paper assessed not only knowledge recall, but also understanding, application and critical analysis abilities using the interrelated cognitive domain of reasoning, thus ensuring construct validity.

In the current study, the majority of faculty members and students agreed that the use of a blueprint in the formation of a pharmacology formative assessment paper resulted in a uniform distribution of questions and adequate weighting of clinically important topics and testing of in-depth subject knowledge. Faculty members also agreed that the test paper struck an appropriate balance between questions in the recall and reasoning domains and that the questions were aligned with the learning objectives of the syllabus. These findings suggest that the test paper created through the blueprint was able to avoid construct under-representation. Previous research has also indicated that the use of a blueprint in assessment ensures coverage of all aspects of the curriculum, learning objectives and educational domains. Very few faculty members and students in the current study believed that the assessment contained questions which were too easy or too difficult; this suggests that the use of a blueprint resulted in a rational paper and avoided construct irrelevance variance. Moreover, the blueprint seemed to make the assessment more ‘fair’ in the eyes of both faculty members and students. The experience of an authentic assessment seems to be a motivating factor; among medical students, a fair assessment has been reported to stimulate a deep approach to learning when tailored to curricular objectives. This may subsequently affect clinical practice and patient care once the students have graduated.

Earlier reports in the literature have provided evidence that blueprints make it easier for an examiner to select questions, set papers according to accepted norms and standards, test higher-order cognitive domains and create well-organised theory papers. These findings were also reflected in the current study during faculty responses to open-ended questions concerning the benefits of blueprinting. However, certain challenges to the use of blueprinting were also identified; many faculty members stated that the preparation of a blueprint was time-consuming and that it was occasionally difficult to reach a consensus. Nevertheless, despite these barriers, the majority of faculty members found blueprinting to be a promising approach for preparing assessment papers and indicated that they would use this approach in future examinations and change their teaching patterns accordingly. Overall, the use of a blueprint in a formative assessment paper had a positive impact on faculty perceptions; it is therefore realistic to recommend the inclusion of this approach as a valid tool to frame theory papers for summative assessment. Nevertheless, the data from this study would be more valuable if they were combined with evidence of increased academic achievement due to the blueprint; unfortunately, the study was limited in this regard as the students’ performance could not be compared to students in previous years due to the differences in syllabus.

**Conclusion**

According to the perceptions of both pharmacology faculty members and students, the use of blueprinting in the formation of a written formative assessment paper was found to result in high content and construct validity. As this approach helps to align the content, cognitive domains and assessment tools of a paper in a rational way, it should be implemented as an integral part of the framing of theory assessments in future.

**CONFLICT OF INTEREST**

The authors declare no conflicts of interest.

**FUNDING**

No funding was received for this study.
ACKNOWLEDGEMENTS

The authors wish to thank all of the faculty members and students who participated in this study. In addition, they are grateful to Dr. Neena Doshi of the GMERS Medical College, Gotri, for critically reviewing the manuscript before publication.

References


