

# SOP- Standard Setting Using Angoff Methodology (Healthcare Programmes)

## Guiding Principles

1. All written questions must be standard set before being used in a summative examination.
2. There is no requirement to standard set questions that have been previously used, if they are unchanged and have demonstrated good metrics at previous diets.
3. The Angoff procedure is one of the most popular procedures for setting the cut (passing) score in assessments.
4. The Angoff procedure and its variants are examples of criterion referenced standard setting which is accepted best practice in standard setting in medical education assessment.
5. Modified Angoff method is used for all IEHMS written examinations, it is the method which has been used the longest and has the most data to support it.
6. Ideally, borderline regression method should be used for OSCE examinations, however, this is not possible with small cohort sizes (<50 students), and therefore Angoff methodology should be used.
7. The Angoff method is defensible in the eventuality of academic or legal challenge.
8. This SOP does not consider the examination blueprint and / or question selection.
9. Standard Setting meetings should usually be chaired by the exam coordinator.
10. A key element in the assessment process and in documentation of assessment quality is the establishment of performance level descriptors.

## Sequence of a Standard Setting Meeting for SBA, Extended Matching Question (EMQ), SAQ & OSCE questions

### Before the meeting

Identify judges for the panel. A panel should comprise at least 6 judges with sufficient knowledge to make a judgment on the test items being reviewed. The Standard Setters should reflect a range of disciplines and seniorities and be competent to make academic judgements on test material e.g., system/block/module leads, academics, trainees. Where possible, at least one Standard Setter should be drawn from a different part of the curriculum to provide internal externality.

Standard Setters given access to exam papers to prepare for the meeting; on MSCAA Exam-Write platform if online exam. Standard Setters should submit their scores online or come to the standard setting meeting with their judgement on each question.

## At the meeting

Exam Coordinator to chair meeting.

- Outlines the name and structure of the assessment, what stage the students are in the curriculum, what other assessments they may be sitting at the same time, and the consequences of success and failure.
- Presents data from the last sitting of this examination (overall standard set pass mark and number of students passing / histogram of marks for the whole exam).
- Leads short discussion on attributes of an “excellent” student, a “terrible” student and then finally a borderline student – i.e., one who is minimally competent and equally likely to pass or fail. This discussion is important so that judges understand that they are not setting a passing score, but instead a judgement on how a borderline student would perform on a particular item.

For single best answer question/extended matching question format:

- Each Standard Setter presents their judgement on the probability of a borderline candidate scoring the item correctly i.e., 1.00 all borderline candidates would answer question correctly and 0.00 no borderline candidates would answer question correctly. An easy item might have a probability of 0.9, whereas a very difficult item around 0.3.
- If there is a large discrepancy between the marks, there should be a discussion between Standard Setters.
- Standard Setters are then asked if they wish to modify their probability.
- Once final judgements are received, the average probability is taken.
- Repeat for each item.
- Calculate the sum of the agreed probabilities and then divide by the number of questions and multiply by 100 to get the percentage passing score (passing score =  $100 \times \text{sum of probabilities} / \text{number of questions}$ ).
- Judges should then be asked if they think any changes are required and can revisit questions.

For short answer type questions / OSCE format:

- Each Standard Setter presents their judgement on the score that a borderline candidate would get for the whole question. Normally, this score should be above half the total.
- If there is a large discrepancy between the scores, there should be a discussion between Standard Setters.
- Standard Setters are then asked if they wish to modify their score.
- Once final judgements are received, the average score is taken.
- Repeat for each item.

- The percentage passing score is calculated as the sum of the average scores for each short answer question, divided by the total number of available marks multiplied by 100.
- Standard Setters should then be asked if they think any changes are required and can revisit questions.

Occasionally, despite careful review of questions in advance of the meeting, flawed questions will be identified during the standard setting meeting. Unless a flawed item can be corrected very quickly, it should be set aside and replaced.

## Standard Setting for Anatomy Objectively Structured Practical Exam (OSPE)

### Guiding Principles

1. All OSPE questions must be standard set before being used in a summative examination.
2. Because of the intrinsic variability in individual test items (e.g., cadaveric specimens; placing of arrows), it is good practise to standard set each diet.
3. An OSPE may require more than one circuit (or “site”) for logistical reasons, e.g., to minimise testing time and the chances of students who have completed the test interacting with those who have not. In such cases, the exam coordinator must ensure that multiple circuits have consistent construction and presentation of material. NB: Most exam diets (except re-sits) have 3 parallel running sites.
4. The Angoff procedure and its variants are examples of criterion referenced standard setting which is accepted best practice in standard setting in medical education assessment.
5. The Angoff method is defensible in the eventuality of academic or legal challenge.
6. This SOP does not consider the examination blueprint and/or question selection.
7. Standard Setting should usually be led by the Exam Coordinator.

### Sequence of a Standard Setting Process

#### Before the meeting

1. Proof-read the questions in advance. Ensure that the “house style” of items is adhered to (e.g., consistency of tense, correct grammar, UK spelling, and number of words in the question). See further reading suggestion below.
2. Occasionally, despite careful review of questions in advance of the meeting, flawed questions will be identified. A panel of anatomy teaching staff should review these before the standard setting process so changes can be made if necessary.
3. Identify judges for the panel. A panel should where possible comprise at least 4-6 judges with sufficient knowledge to make a judgment on the test items being reviewed. Judges should reflect a range of disciplines and seniorities and be competent to make academic

judgements on test material. Where possible, at least 1 judge should be drawn from a different part of the curriculum to provide internal externality.

4. Obtain learning guides in case the standard setting panel needs it.
5. Ensure that copies of questions are available for each judge.
6. The standard setting process is to be carried out after the exam has been set up; either just before the exam itself or just after completion of the exam.
7. Judges may carry out the standard setting process individually or as a group. Most exam diets (except re-sits) have 3 parallel running sites. For uniformity, the same site is to be used by all judges for standard setting.

#### The standard setting process

1. Exam coordinator to brief each judge before they start the standard setting process:
  - a. Outline name and structure of the assessment, stage of students in the curriculum, other assessments in the exam diet, and consequences of success and failure.
  - b. Lead a short discussion on attributes of an “excellent” student, a “terrible” student and finally a borderline student who is minimally competent/equally likely to pass or fail. This discussion is important so that judges understand that they are not setting a passing score, but instead a judgement on how a borderline student would perform on a particular item.
2. Each judge independently reads each question and (if relevant) inspects associated specimen/model/image and makes a judgment on the probability of a borderline candidate scoring the item correct. An easy item might have a probability of 0.9, whereas a very difficult item around 0.3. A probability of 1.0 would indicate that all borderline students would answer question correctly.
3. Scores are systematically combined (totalled and averaged) to determine a passing score.
4. As part of the standard setting process, standard setting groups can be asked to set both the passmark and the point at which students display ‘excellent’ knowledge or skills, at either the A5 or A1 level. At the A5 level, a probability of 1.0 would indicate that all excellent A5 level students would answer question correctly.

## Appendix 1: Anchor statements for standard setting IEMDS using Angoff methodology

Unsatisfactory	Demonstrates poor or no understanding of required medical / dental knowledge. Unsafe. Frequent errors, omission or misunderstandings.
Borderline	Demonstrates adequate medical / dental knowledge to be a <b>just safe</b> practitioner. Some significant errors, omissions and misunderstandings. Patchy performance. Often formulaic in approach with limited attempt to adapt to situation.
Satisfactory	Demonstrates adequate medical / dental knowledge to be a safe practitioner. Acceptable performance despite some omissions, errors or misunderstanding. Can be formulaic in approach at times.
Highly satisfactory	Demonstrates thorough medical / dental knowledge. Few minor and non-essential omissions / errors.
Excellent	Excellent demonstration of medical / dental knowledge. Wide breadth of knowledge.