

## How the National Board Establishes Its Pass-Fail Standards for Part III CSE

In the Part III-Clinical Skills Examination (CSE), the candidate examines a different standardized patient at each of four 30-minute stations, in the performance of 20 clinical skills.

Although this examination primarily measures psychomotor skills, it contains an assessment of communication and affective (i.e., clinical habits and attitudes) skills, as well as interpretation of clinical findings.

In August 2012, the Board conducted a comprehensive standard setting study to establish the cut score for the 2013 administration of Part III CSE. Twenty panelists were recruited from among the remote and in-house examiners who actively score Part III-CSE. An additional two panelists from the Board of Directors also were invited to participate which brought the total panel to 22.

The Standard Setting study was conducted in two phases. In the first phase, panelists were asked to independently rank and weight each of the 20 skills. These results were used to establish how much each should contribute to the total score. For this task, panelists were asked to base their judgments on “both the importance of the skill for safe, effective clinical practice and the manner in which it is tested in CSE.” The newly established skill weights are summarized in Table 1.

*Table 1. Skill Weights for Part III-CSE*

	Points	Weight
1 Case History/Patient Communication	76	7.6%
2 Patient Education	49	4.9%
3 Near Cover Test	27	2.7%
4 Binocular Extraocular Muscle Motility Evaluation	30	3.0%
5 Static Confrontation Visual Fields	29	2.9%
6 Pupil Testing	47	4.7%
7 Blood Pressure Measurement	34	3.4%
8 Ophthalmic Lens Evaluation	35	3.5%
9 Biomicroscopy	80	8.0%
10 Goldmann Applanation Tonometry	62	6.2%
11 3-Mirror Gonioscopy	51	5.1%
12 Collagen Implant Insertion and Removal	33	3.3%
13 Soft and GP Contact Lens Insertion, Evaluation, and Removal	57	5.7%
14 Retinoscopy	54	5.4%
15 Distance Subjective Refraction	73	7.3%
16 Heterophoria and Vergence Testing at Distance	27	2.7%
17 Accommodation Testing	24	2.4%
18 Binocular Indirect Ophthalmoscopy	75	7.5%
19 Dilated Biomicroscopy and Non-Contact Fundus Lens Evaluation	76	7.6%
20 Injections	47	4.7%
General items	16	1.4%

In the second phase of the Standard Setting project, all of the items in the assessment were reviewed and differentially weighted based on their criticality. The item criticality descriptors are summarized in the

Table 2. Each of the 22 panelists provided judgments for 350+ items which were aggregated to determine the item’s final weight.

Table 2. Rating Scale for Item Criticality- Clinical Skills Examination

<b>Essential</b>	Essential for satisfactory care; poor performance would result in deficient patient care
<b>Very Important</b>	Very likely to affect the quality of patient care; only extraordinary measures could compensate for poor performance
<b>Important</b>	Likely to, but would not necessarily, affect the quality of patient care
<b>Desirable</b>	Desirable activity, but would not necessarily affect the quality of patient care if done incorrectly, or not at all

The scoring and standard setting methodology are integrated. For each of the 20 skills, panelists were asked to identify which of the “Important” items a theoretical, minimally qualified candidate (MQC) would likely miss (i.e., not get correct). The characteristics of the MQC were discussed and it was emphasized that this candidate was a barely “passer” but not a “failer.”

The 22 panelists provided standard setting judgments for each of the 20 skills in CSE which were aggregated and analyzed. These results were comparable with a more holistic approach which was adopted by the NBEO Board of Directors. In the holistic approach, points from 50% of the “Important” items and 100% of the points from the “Desirable” items were deducted from the total skill score. This skill threshold is summarized in the following formula.

$$Skill\ Threshold = MAX_{points} - \frac{1}{2} Important_{points} - Desirable_{points}$$

A skill threshold score was determined to be excessively harsh if it initially was calculated to be 90% or more. In this case, the threshold was truncated downward to less than 90%.

Summing the threshold scores across the 20 clinical skills yields the Part III pass-fail cutoff score. Historically, the Clinical Skills rounded cutoff score has ranged from 80% to 84%. The results of both the Panelist and formula (“Importance”) derived weights were consistent with each other and were within the historical range. The NBEO Board of Directors adopted the above formula for the 2013 CSE administration because it is validated by panelist ratings and offers flexibility in establishing cut scores for future administration of Part III-CSE.

Candidates should be aware that because of the item criticality scoring weights, it is possible that “missing” a relatively small percentage of the items may result in missing a relatively large percentage of the skill points. The items on CSE are scored dichotomously. That is, candidates receive the full point value for items performed correctly, and no points (zero) for items either performed incorrectly, or omitted. In addition, candidates are reminded that the differential weighting of skills will also affect the final score and that the number of items needed to successfully perform a skill does not influence its ultimate weight in scoring.