HKDSE Liberal Studies IES
Moderation Mechanism
Some Misunderstandings and Clarifications

December 2010
Why Moderation

Misunderstandings:
- SBA is school based ⇒ Marking should be totally determined by school teachers WITHOUT any moderation
- SBA is school based ⇒ Comparability across schools is NOT necessary or may even result in adverse effects

Clarifications:
- SBA results are counted as a component in public assessment
- Public exam results accepted as a valid measure by local and overseas organizations ⇒ Comparability across schools is ESSENTIAL
Misunderstandings:
- School raw IES scores are already valid and comparable across schools, as the scores are thoroughly checked by
  - School teachers
  - School administrators (e.g., Panel heads, Principals)
  - Students of the school

Clarifications:
- ALL these stakeholders of a school are NOT fully aware of the standards of performance across all schools
- An analogy: Thermometers in different scales (e.g., in Celsius and in Fahrenheit)
Why Moderation (Cont’d)

Misunderstandings:
- Some LS teachers expressed that mark moderation reveals that the HKEAA does NOT trust their judgment

Clarifications:
- Trust on a school’s IES marks by OTHER SCHOOLS is essential
- IES results are recognized by LOCAL and OVERSEAS organizations as indispensable
- Moderation of SBA scores adopted in CE & AL for a long time and will be adopted in all HKDSE subjects
What would be Moderated

Misunderstandings:
- IES marks are divided into Task marks and Process marks; Process marks are based on student and teacher interactions which could not be appropriately reflected in students’ Task work or exam results ⇒ Why could Process marks be moderated!

Clarifications:
- Process marks would NOT be moderated in 2012
- Only Task marks would be moderated in 2012
- Certain measures:
  - Review sample Process activities and assessment records
  - Provide feedback
  - Follow up of any problems found
How to Moderate

In 2012, moderation is done:
By adjusting
- The **AVERAGE** of IES Task scores on school basis
- The **SPREAD** of IES Task scores on school basis
- WITHOUT changing the internal ranking within a school

With reference to:
- the **SCHOOL EXAM** scores
- Supplemented with **SAMPLE REVIEW** of students’ work as a safeguard measure: Checking students’ performance against moderated outcomes based on exam
How to Moderate (Cont’d)

Misunderstandings:
- No need for sample review: It is NOT uncommon that the performance of students in SBA could be totally different from that of the school exam performance level.

Clarifications:
- Statistical moderation based on exam scores is justifiable since there is a **reasonable degree of correlation** between school exam scores and the performance of students in SBA.
- Otherwise, moderation by Expert Judgment should be employed instead.
Why Sample Review

Two MAIN reasons:

- The school exam result could NOT cover all factors that determine the student IES performance level; thus, from a statistical point of view, the chance of having some schools whose student IES performance is significantly different from the moderated outcome based on exam could NOT be precluded.

- Small candidature of a particular school could cast the reliability of the corresponding moderated outcome (solely based on school exam result) in doubt.
Queries on the Arrangements of Sample Review

Misunderstandings:
- **Sample proportion** is too low (6 samples out of a school with some 200 students)

Clarifications:
- Sampling variation does NOT depend on sample proportion; but it depends on
  - **Sampling mechanism**
  - (Absolute) **Sample size**
  - **Actual variation** of student performance in the school
Sampling Mechanism

Misunderstandings:
- The samples to be collected should be selected/determined by the school.

Clarifications:
- To have a representative sample, random sampling covering a full spectrum of students’ work should be implemented such that every piece of student work has a non-zero chance of being selected.
- **Stratified random sampling** (by HKEAA to reduce teachers’ workload):
  - Students of a school is divided into a number of strata of equal sizes.
  - In each stratum, students’ IES performance level are similar with each other.
  - Some samples are selected from each stratum in order to cover a full spectrum of student works.
Sampling Accuracy

Sample size: To avoid too much disturbance, it is agreed that 6 samples of student Products (in Stage 3 of IES) are collected from each school AFTER final submission of school IES marks in S6.

Estimated sampling accuracy: A simulation study

Let’s consider the following scenario:
- Full Mark = some 50
- Most of the student performance in a school is confined to: [10,50]
- Sample size = 6; School size=50;100;150; 200; 250

In most of the times (95% Confidence Interval), the difference between sample raw mean and actual school raw mean are some 2 marks; i.e., less than the 5% of the Full Mark.
### Sampling Accuracy: Simulation Results

<table>
<thead>
<tr>
<th>School Size</th>
<th>True School Average</th>
<th>Mean of Sample Averages (from 10,000 Trials)</th>
<th>95% CI (from 10,000 Trials)</th>
<th>Difference between True Value and Lower Bound</th>
<th>Difference between True Value and Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>26.6</td>
<td>26.6</td>
<td>[25.0, 28.0]</td>
<td>1.6</td>
<td>1.4</td>
</tr>
<tr>
<td>100</td>
<td>28.6</td>
<td>28.6</td>
<td>[26.9, 30.8]</td>
<td>1.7</td>
<td>2.2</td>
</tr>
<tr>
<td>150</td>
<td>29.2</td>
<td>29.2</td>
<td>[27.5, 31.1]</td>
<td>1.7</td>
<td>1.9</td>
</tr>
<tr>
<td>200</td>
<td>28.8</td>
<td>28.8</td>
<td>[27.2, 30.8]</td>
<td>1.6</td>
<td>2.0</td>
</tr>
<tr>
<td>250</td>
<td>29.1</td>
<td>29.1</td>
<td>[27.4, 31.0]</td>
<td>1.7</td>
<td>1.9</td>
</tr>
</tbody>
</table>

Note: Sampling variations (and marking variations) would be taken into consideration when sample review results are checked against the moderated results based on exam.
Marking Accuracy

Misunderstandings:
- District Coordinators are not the appropriate persons to re-assess student sample works
- Subjectivity in marking students’ work ⇒ Unreliable

Clarifications:
- Arrangement of marking exam scripts would be largely followed
- Experienced LS teachers would be invited to re-assess student sample work
- Other arrangements would be in place; e.g., standardization meetings, trial marking, and check marking, mark adjustments and involvement of other assessors on the same piece of student work (if necessary)
- Experiences in 2010 ASL LS:
  - Correlation between the closest pairs of marks was 0.8
Comparison of sample review results with statistical moderated outcomes

- **Tolerance limits** would be set up by taking sampling and marking variations into account.

- If the difference between statistical moderated outcome and sample review result is within the tolerance limits, NO adjustment would be applied to statistical moderated outcomes.

- Otherwise, adjustments would be applied so as to reflect the genuine school performance level.
Thank You!