




# International Conference on Assessment – 2011

CBSE in collaboration with ACER India and  
SSC-Gurgaon Chapter


Gabrielle Matters (ACER)

25 January 2011

Australian Council *for* Educational Research




Linking school-based assessment with  
public examinations:  
The role of moderation

- 
1. Locus of control of assessment
  2. Validating teacher judgments
  3. Social and statistical moderation
  4. Purpose, form, techniques
  5. Scaling school-based assessment to external examinations
  6. Application of the formula for scaling
  7. Limitations and strengths



- Locus of control
  - Who sets the assessment?
  - Who composes the marking schemes?
  - Who marks the student work?



|                          | <b>Sets the assessments</b> | <b>Composes the marking schemes</b> | <b>Marks the student work</b> |
|--------------------------|-----------------------------|-------------------------------------|-------------------------------|
| Public examinations      | E                           | E                                   | E                             |
| School-based assessments | I                           | I                                   | I*                            |
| “Hybrid”                 | E                           | E                                   | I*                            |



- In this discussion, the term external assessment is reserved for subject-specific examinations set by a body external to the school. Such examinations are devised to assess student achievement in a particular subject, whether by multiple-choice, or conventional written, oral or practical questions. All the questions refer to a syllabus.



- Internal (school-based) assessment is devised, constructed and implemented by schools based on an official syllabus. Teachers have to be trained to become consistent judges of the quality of student work and there has to be a quality assurance process in place to guarantee comparability of results.



- How can we validate teacher judgments in high-stakes school-based assessment?
- → Moderation
  - A set of processes designed to achieve comparability across judges (teacher-assessors) and across sites (schools or regions)
  - Social moderation
  - Statistical moderation



- Conditions
  - Common curriculum framework
  - Different assessments/assessors
  - Different sites



- **Social moderation**
  - Human beings
  - Validation of teacher judgments by have those teachers' judgments reviewed by their peers
  - Underpinned by descriptors of standards, evidence of learning and consensus amongst [expert] judges



- Statistical moderation
  - Computers
  - Validation of teacher judgments by scaling those teachers' judgments to a common examination/test
  - Underpinned by a mathematical formula




- Purpose ✓
- Form →
- Technique →



| <b>Form</b>       | <b>Technique</b> | <b>Guild of professionals</b> |
|-------------------|------------------|-------------------------------|
| Social moderation | Teacher meetings | Peer review                   |
|                   | Review panels    | Expert review                 |
|                   | Visitation       | Bureaucratic review           |



| <b>Form</b>            | <b>Technique</b>  |
|------------------------|---|
| Statistical moderation | Using results of an external exam                             |
|                        | Using results of an anchor test (aptitude test, omnibus test) |
|                        | Other subjects (iterative)                                    |



Statistical moderation is the process of adjusting the school-based assessment distribution for a subject so that the average and spread of the school-based assessments match the average and spread of the school's distribution on the external examination scores or scores on a common test.



- Statistical moderation
  - Linear transformation
  - “Scaling”
- Adjusts SBA distribution for a subject in a school to have the same mean and standard deviation as results on the external measure for students in that subject-group at that school

Adjusted score<sub>student 1</sub> =

$$\{[(SBA_{\text{student 1}} - \text{Mean } SBA_{\text{schl}}) / SD SBA_{\text{schl}}] \times SD E_{\text{schl}}\} + \text{Mean } E_{\text{schl}}$$

$$\{[(SBA_{\text{student 1}} - \text{Mean } SBA_{\text{schl}}) / SD SBA_{\text{schl}}] \times SD E_{\text{schl}}\} + \text{Mean } E_{\text{schl}}$$



Adjusted score<sub>student 1</sub> =

$$\{[(SBA_{\text{student 1}} - \text{Mean } SBA_{\text{schl}}) / SD SBA_{\text{schl}}] \times SD E_{\text{schl}}\} + \text{Mean } E_{\text{schl}}$$



- For a given student in a given subject,
  - Adjusted score (“scaled score”) may be higher or lower than the original school score depending on the location and spread of the school scores
  - Mean and standard deviation work together [against cheating]



- Illustrations
  - School-based assessments and adjusted scores in a subject
  - Combined scores for adjusted scores and school-based assessments in that subject


| Student | School | Exam |
|---------|--------|------|
| 1       | 78     | 60   |
| 2       | 66     | 63   |
| 3       | 16     | 50   |
| 4       | 64     | 61   |
| 5       | 99     | 99   |
| 6       | 50     | 50   |
| 7       | 92     | 92   |
| 8       | 51     | 61   |
| 9       | 66     | 81   |
| 10      | 79     | 80   |
| 11      | 65     | 61   |
| 12      | 65     | 53   |
| 13      | 54     | 50   |
| 14      | 74     | 71   |
| 15      | 50     | 55   |
| Mean    | 64.6   | 65.8 |
| SD      | 19.8   | 15.6 |

| Student | School | Exam | Adjusted |
|---------|--------|------|----------|
| 1       | 78     | 60   | 76.3     |
| 2       | 66     | 63   | 66.9     |
| 3       | 16     | 50   | 27.7     |
| 4       | 64     | 61   | 65.3     |
| 5       | 99     | 99   | 92.8     |
| 6       | 50     | 50   | 54.3     |
| 7       | 92     | 92   | 87.3     |
| 8       | 51     | 61   | 55.1     |
| 9       | 66     | 81   | 66.9     |
| 10      | 79     | 80   | 77.1     |
| 11      | 65     | 61   | 66.1     |
| 12      | 65     | 53   | 66.1     |
| 13      | 54     | 50   | 57.5     |
| 14      | 74     | 71   | 73.2     |
| 15      | 50     | 55   | 54.3     |
| Sum     | 969    | 987  | 987.0    |
| Mean    | 64.6   | 65.8 | 65.8     |
| SD      | 19.8   | 15.6 | 15.6     |

| 20:80 | 30:70 | 40:60 | 50:50 | 60:40 |
|-------|-------|-------|-------|-------|
| 63.3  | 64.9  | 66.5  | 68.2  | 69.8  |
| 63.8  | 64.2  | 64.6  | 64.9  | 65.3  |
| 45.5  | 43.3  | 41.1  | 38.8  | 36.6  |
| 61.9  | 62.3  | 62.7  | 63.2  | 63.6  |
| 97.8  | 97.1  | 96.5  | 95.9  | 95.3  |
| 50.9  | 51.3  | 51.7  | 52.2  | 52.6  |
| 91.1  | 90.6  | 90.1  | 89.7  | 89.2  |
| 59.8  | 59.2  | 58.6  | 58.1  | 57.5  |
| 78.2  | 76.8  | 75.4  | 73.9  | 72.5  |
| 79.4  | 79.1  | 78.8  | 78.6  | 78.3  |
| 62.0  | 62.5  | 63.0  | 63.6  | 64.1  |
| 55.6  | 56.9  | 58.2  | 59.6  | 60.9  |
| 51.5  | 52.2  | 53.0  | 53.7  | 54.5  |
| 71.4  | 71.7  | 71.9  | 72.1  | 72.3  |
| 54.9  | 54.8  | 54.7  | 54.7  | 54.6  |




| <b>School</b> | <b>Exam</b> | <b>Adjusted</b> | <b>20:80</b> | <b>50:50</b> |
|---------------|-------------|-----------------|--------------|--------------|
| 70            | 43          | 49.7            | 44.3         | 46.3         |
| 70            | 58          | 49.7            | 56.3         | 53.8         |
| 1331          | 919         | 919             | 919          | 919          |
| 66.55         | 45.95       | 45.95           | 45.95        | 45.95        |
| 20.4          | 22.0        | 22.0            | 20.8         | 20.2         |

- 
- The order of the teacher's rankings from the school is maintained after the school assessment has been moderated.
  - The top student from the school assessment is the top student after the school assessments have been moderated
  - The second top will still be the second top and so on ... to the lowest-achieving student
  - The relative differences between students is maintained



- When the sum of the marks from school-based assessment is close to the sum of the marks from the external exam, the change in the marks from school-based assessment is minimal after moderation.
- And if, at the same time, the standard deviations are equal, there will be virtually no change in the school-based assessments.



Possible explanations for “spread of internal scores less than spread of external scores”

1. Pushing up low achievers because ...
2. Squashing of marks ...



- When the teacher uses a marking scale that is much easier than the external exam, all students' marks will be lower after moderation.
- When teacher's marking is too hard relative to the external exam, all students' marks will be higher after moderation.



- Advantages of statistical moderation
  - Cheapest method for achieving comparability
  - Does not involve movement of materials
  - Does not involve movement of personnel
  - Efficient because it relies on algorithms and computer processing of data



- Limitations of statistical moderation

Since the two sets of marks refer to different assessment performances, the performances may not reference the same underlying characteristics.

Furthermore, even if they did reference the same characteristics (e.g. where the school assessment mirrors the external measure in their content and form), the standard of performances could be quite different.



- Limitations continued

Therefore, the scaling merely realises an expectation that one set of results mirror the other set of results.

The process of scaling adopts the assumption of equivalence but cannot verify it.

Also using an algorithm does not promote interventions for anomalies



- Some observed negative consequences
  - Time management demands
  - Over-assessment
  - Tensions can arise between formative and summative purposes
  - Teachers can misinterpret the function of statistical moderation
  - Schools can be suspicious of results produced by other schools
  - The general public can lose faith in the credibility of teachers as assessors



- Some observed positive consequences
  - Enhanced teacher confidence in reliability of their reported results
  - Enhanced public confidence in reliability of school-based assessment



## Social moderation

- About internal assessment
  - privileges contextualised judgments
  - susceptible to grade inflation
  - assumes consistency in teacher judgments
- Side-effects of social moderation
  - teacher dialogue
  - up-skilling of the profession
  - shared understanding of standing of standards



- With the increased emphasis on teacher judgment in high-stakes assessment has come the role of moderation in ensuring comparability of standards.
- Social moderation and statistical moderation are two ways of linking results of one assessment to those of another. And there are variations within each of these techniques in practice.
- I have outlined the theoretical underpinnings of statistical moderation and social moderation, and illustrated the application of statistical moderation to scaling school-based assessments to public examination results.
- Thank you for your attention.
- I am happy to take questions.

[matters@acer.edu.au](mailto:matters@acer.edu.au)