

Using consensus moderation to ensure appropriate academic standards over time

Dr Duncan D Nulty
Griffith Institute for Higher Education

Dr Jill Freiberg
School of Education and Professional Studies

© Duncan D Nulty, 2011. All use of any material within this document must first be discussed with, and approved by, the copyright owner.

Outline

- Background & imperative
- Pre-marking calibration (consensus moderation)
 - Aim
 - Timing
 - What you'll need
 - Steps involved
 - How that leads to consistency over time
- Why should you do this?
- Making it easier

Background

- A course in Education ...
- With 250 students ...
- Across two campuses ...
- With two senior, very experienced full-time academics and three tutors teaching it ...
- Both academics with assessment expertise ...
- Both had an open, positive and respectful working relationship ...
- Both overtly sought marking consistency ...
- ... but still encountered difficulties achieving this!
- → Something had to be done.

Pre-marking calibration

- Pre-marking calibration is a ‘prior-to-marking’ process that addresses the problem of achieving marker consistency across sessional, part-time and full-time academic staff marking across several campus cohorts.
- It can also be used by ‘sole-markers’ by engaging with other academics in other courses and (in principle) other institutions.

Aim

- The aim of pre-marking calibration is to ensure intra-marker and inter-marker consistency and appropriateness of standards.
- Markers learn to:
 - Make the same interpretations of criteria and standards statements (by having a shared conception of “quality” in the work);
 - Recognise the same evidence in student work; and,
 - Attribute the same value to this evidence when making decisions about marks and standards of achievement (by giving the same number of marks).

Pre-marking calibration - timing

- Ideally the pre-marking calibration takes place:
 - prior to *teaching*
 - (so that the teaching is consistent and appropriate)
 - and then again prior to *marking*
 - (so that the marking is consistent and appropriate)

What you will need:

1. An assessment item *task description*,
 - (either the current one you are about to mark, or a previous one that's similar against which you will be benchmarking your standards)
2. The associated criteria and standards (i.e a *rubric or marking guide*)
3. Copies of *a few of the students' scripts or products* (from past and the current marking rounds).

Additional requirements ...

1. *Annotated exemplars* of students' past assignments/answers on similar types of assessment items.
 - You select those that exemplify different standards.So this means:
 - You collect and keep copies of students work that you have marked.
 - And, you identify which of these exemplify particular standards.
 - You ensure that your annotations clearly convey this.

Additional requirements ...

2. A data-base of evidence-based, *feedback comments* relating to each criterion and each of five standards of achievement.
 - The criteria are: those you are judging
 - The five standards are: High Distinction, Distinction, Credit, Pass, Fail (*or* A, B, C, D, F).
 - You build the database using feedback comments provided to students on items of assessment that:
 - You marked in previous years
 - Are similar to the current assessment item (the one you are calibrating standards for).

Process

1. Using **annotated exemplars** of students' assignments/answers on similar types of assessment items, get together with your colleagues to ***discuss*** specific examples of the types of evidence that will count as demonstrations of key criteria, and the distinguishing features of work at different standards.
2. Seek agreement.

The discussion

- *Key criteria* usually include: knowledge of facts, theories, models, concepts, procedures studied in the course, critical reasoning, problem solving, communication skills etc.
- *What will count as evidence* required of students must be clearly identified and discussed.
- How students work will be *marked* must be clarified.
 - e.g. how will text book definitions inserted without elaboration or appropriate exemplification be marked? Will that count as an accurate definition

Process

3. Next, **clean** copy exemplars are considered and markers are asked to identify and discuss evidence of one criterion and the associated features that will distinguish work at A, B, C, D, F standards *on that criterion*.
 - For this activity, markers also have access to a *comments database for the criterion under consideration*. Comments in the database are written so that the required evidence and the distinguishing features at different standards are clearly identified.
4. Repeat this for all the other criteria.

Process

- 5 Each marker is given a copy of one student's work to mark/grade from the *current* marking round. (All markers get a copy of the same one).
- 6 Markers allocate a mark/grade.
- 7 The group then discuss and justify the mark/grade they have given with close reference to specific criteria and standards and the evidence in the student's work.
- 8 Through discussion, discrepancies between markers marks are resolved.
- 9 Repeat.

How does this give consistency over time?

- Consistency of standards is more likely to be maintained from year to year because markers:
 - Use, review and reflect on a variety of assessment instruments that assess the same knowledge, skills and processes, using the same criteria and standards, the same rubrics and the same comments database.
 - Are asked to consider marked assessment items and to mark assignments from previous years using the relevant task sheet and criteria and standards rubric, and therefore:
 - Repeatedly practice, and reflect on, their marking to progressively resolve discrepancies and increase the levels of consistency.

Consistency over time?

- This process includes practice and reflection.
- Following the marking of each assignment, the levels of consistency are established and discrepancies discussed and 'resolved'.
- This process is repeated, points of inconsistency are identified for each marker and this provides a guide for them to self-check their marking and for reference at end of marking moderation.

What's in it for you?

- This practice was developed because of difficulties encountered in which it was clear that the marking standards employed by markers varied by an unacceptably large amount: e.g. the average mark out of 40 by one marker's cohort was 38, for the other marker's cohort it was 26.
- → *All the work (250 students) had to be re-marked!!*
- The calibration process can be regarded as a “risk-reduction” strategy – yes, it takes time, *but it saves more.*

Making it easier

- These practices are made easier when
 - you use Safe Assign (Helps you collect clean unmarked digital samples of students' work)
 - You provide your feedback to students electronically e.g. through “ReMarks” or “MS Word’s track changes”. (Helps you to collect comments that relate to specific features of students’ work, and specific criteria on the marking rubric).
 - You have a pre-specified marking guide or rubric that stipulates the marking criteria and standards (Provides you with a consistent reference point for you to map your feedback to).

Making it easier

- Once you and your colleagues have reached a high level of consistency, you do not need to undertake this exercise every time you mark.
- If you do, it may not take you so long as the first few times.



☎ Phone: (+61-7) 3735 6813

✉ e.mail: d.nulty@griffith.edu.au

🌐 <http://www.griffith.edu.au/gihe/staff/duncan-nulty>



☎ Phone: (+61-7) 3735 5843

✉ e.mail: j.freiberg@griffith.edu.au

🌐 <http://www.griffith.edu.au/education/school-education-professional-studies/staff/academic-staff/jill-freiberg>

You can contact us.