

# Moderation Strategies

---

It is not often necessary to employ more than a few strategies, but to consider what are the likely causes of, or contributors to, a lack of comparability in the particular assessment and to implement strategies that will have the greatest impact. For example, the following situations may call for different moderation strategies:

- Assessment tasks are not common to all students
- Assessment occurs on different campuses
- Assessments involves multiple markers
- Assessment is largely subjective
- One person marks many papers

## **1. Make assessment and marking criteria explicit, provide exemplars**

Assessment criteria describe the key characteristics of differing standards of performance and are usually standards defined by marking or classification bands. The development and dissemination of appropriate assessment criteria is a key element of consistent assessment procedures. For criterion-based assessment, grading criteria (broad statements of student achievement associated with each grade band and/or level) must be established prior to students attempting the assessment. Provide students with clear marking criteria or scoring rubrics prior to them attempting any assessment task. Making exemplars available (especially of “good” pieces of work) is also extremely useful, particularly for those students from cultural backgrounds and academic traditions different from those in Australian universities.

## **2. Distribute marking keys/guides**

Marking keys should be collaboratively developed or at least distributed to all markers for comment prior to the assessment being attempted by students; in time for feedback and concerns to be addressed.

## **3. Conduct comparability meetings (‘consensus moderation’)**

Conducting comparability meetings involves having a number of staff coming together, individually marking the same pieces of work and then comparing the results. Through discussion and clarification, a group of staff gradually come to a ‘consensus’ on the marks/grades they would award for a piece of work. This is often necessary even where marking criteria are very explicit, since some markers tend to mark ‘easier’ or ‘harder’. Also, some markers tend to use the full range of marks on a scale (on a scale of 10, say from 0 to

10) while others tend to use a restricted range (say from 4 to 8). By discussing differences of opinion about the same pieces of work, gradually markers come to a common understanding of “the standards”. This is a particularly useful exercise for units with very large student cohorts.

#### **4. Monitoring markers**

From time to time through the assessment process, the single marker should review papers marked earlier (e.g. the day before) to ensure that their marking standard has not changed. When a single marker is new to the task or not part of an assessment team, for example a sessional academic, the unit coordinator or another member of the assessment team should independently re-mark a representative sample of submissions marked by the single marker from across the grade ranges (for example two from the each of the pass/fail boundary, the middle range and the top of the range) in order to confirm that the marker is marking according to the agreed standard. If a single marker is not maintaining a consistent standard, the unit coordinator should ensure that any work already assessed by that marker is re-marked to the agreed standard and that ongoing assessment of submissions assigned to that marker are also marked to the agreed standard.

#### **5. Second-marking**

Second-marking involves a sample of marked assessment items being marked a second time by another person. It may involve remarking or simply a verification of the marks allocated by the first marker. It is used to validate assessment standards across a group of markers. The unit coordinator or other experienced marker may second-mark a sample of the assessed work from each marker in a team across the mark range (for example two from the each of the pass/fail boundary, the middle range and the top of the range). The number of items chosen for second-marking should be representative. The usual is to choose a number equal to the square root of the number of assessed items, or five items, whichever the larger (e.g. if there are 36 students, second-mark six of them, if there are 16 students, second-mark five of them). There should be clear criteria for remarking all the work if there are sufficient inconsistencies or inaccuracies within a sample.

#### **6. Having one person acting as ‘moderator’ for transnational contexts**

Where student cohorts are studying in different locations and/or under different circumstances, and with different groups of markers, the unit coordinator who visits the various locations can perform a very important ‘moderating’ role. S/he can, for example, compare samples of assessed tasks to ascertain whether criteria are being applied consistently across different locations, and in a manner comparable with the Australian situation. Such a person can assist others to develop common understandings of assessment standards.

## **7. Anonymous assessment**

Anonymous assessment attempts to remove marker bias. It only works where the identity of students is not revealed in other ways e.g. through writing styles, references in the text etc. Bias can result, unconsciously or not, from prior dealings with different students and be based on attitude, behaviour, gender, race, disability etc.

Having students' papers anonymous may contribute to students' and others' beliefs of fairness or lack of bias in the marking process, in practice, it is not pedagogically very satisfactory. Some staff in transnational locations in particular are unable to derive the diagnostic benefits of assessment and will find it difficult to assist student learning if unaware of the difficulties which specific students are experiencing

## **8. Double-blind marking.**

Double blind marking means that two markers independently mark the work of each student. This is preferred for extended assessment tasks such as research projects or where the assessment involves highly subjective judgements. Consistency limits and ways of resolving differences must be prescribed. The following is an example:

*If the marks given by the two markers for a particular student are within 10% of each other, the final mark is taken as the average of the two. If the marks differ by more than 10%, a third marker examines the work. The markers may then discuss the marks to arrive at a consensus result or the median of the three marks is taken as the final mark.*

## **9. Panel marking**

Panel marking involves independent but concurrent assessment by two or more markers. It may be used for oral presentations, performances, exhibitions or other transient assessment work. Audio-visual recording of the assessment might also be undertaken. Comparability of assessment is enhanced if the composition of the panel is the same for all students doing a particular assessment task. If the panel is not the same for all students then one member of the assessment panel should chair all panels in order to promote comparability.

## **10. External moderation**

Consistency and quality of assessment standards can be established by having some assessment tasks common to those in another course and/or university. Collaborative or cross-marking with assessors in other courses or universities contributes to maintenance of assessment standards as well as students' perceptions of impartiality. Employment of external markers can also ensure comparability of standards.

## 11. Statistical moderation

Statistical moderation involves:

- checking that different assessors have marked students' work in accordance with pre-determined criteria and marking guidelines, and
- where different markers have been shown to have employed different marking standards, adjusting students' marks accordingly.

While statistical moderation is a recognised process, it should be a last resort, used when moderation procedures have failed or have simply been impractical to implement.