Consensus Moderation

Table of Contents
What is consensus moderation? ................................................................. 1
Consensus moderation model ........................................................................ 2
Level 1: Unit level assessment planning ................................................... 2
Level 2: Marking student work ................................................................. 2
Level 3: Grading student work ................................................................. 3
Level 4: Standards across courses (internal and external) ......................... 3
Level 5: Standards over time ................................................................. 4
Stages of consensus moderation .............................................................. 4
Pre-assessment moderation ....................................................................... 4
Peri-assessment moderation ..................................................................... 6
Post-assessment moderation ..................................................................... 7
Unit continuity over time ........................................................................... 8
Further reading and resources ................................................................. 9

Acknowledgment: This resource combines three key discussions of consensus moderation:
1. *A quick guide to consensus moderation of assessment*, prepared by the ‘Internal Moderation of Assessment’ project team, ACU Faculty of Health Sciences, September, 2012. (‘FHS quick guide’)
2. ‘Consensus moderation’, from Griffith University’s *Assessment Matters* website. (‘Griffith’)

What is consensus moderation?
Moderation is about ‘comparability’ or ‘equivalence’ between what happens in one class, assessment or program as compared with all others by the same name. (Centre for Learning and Development, Edith Cowan University, 2012, p. 2)
[Consensus moderation] Conducting comparability meetings involves having a number of staff coming together, individually marking the same pieces of work and 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
In order to ensure that there is equivalence between different markers, there are a number of quality control processes that can be implemented, many of which include peer review.

**Consensus moderation model**

The Griffith consensus moderation model describes these quality control processes at 5 levels of assessment practice.

**Level 1: Unit level assessment planning using constructive alignment**

Unit level assessment planning is at the heart of your overall learning and teaching strategy. It describes the … assessment tasks that will be used in a [unit] to support student learning to make valid judgments about a student’s performance against the specified learning outcomes. Your [unit] level assessment plan should be underpinned by principles for good assessment practice.

(‘Level 1: Course Level Assessment Planning’, Griffith)

The ACU Assessment Policy identifies these principles of good assessment practice.

This will ensure that the assessment tasks together with the assessment schedule (number, timing, weighting and sequence of assessment) are designed to support and validly certify student learning.

The purpose of consensus moderation at unit planning level is to ensure appropriateness, alignment with intended learning outcomes, and clarity in the specification of all the assessment methods used: their type, purpose, timing, weighting, and the task specifications.

Consensus moderation of assessment planning [ensures] that the assessment tasks and schedule promotes the achievement of the unit's specified learning objectives.

(‘Level 1: Course Level Assessment Planning’, Griffith)

**Level 2: Marking student work**

Consensus moderation of marking occurs in two areas.

1. Moderation of marking schemes
2. Moderation of the way markers mark i.e. the marking strategy

Moderation of the marking schemes, guides or rubrics occurs to ensure that the criteria and standards for performance are appropriate and clearly understood by students and markers.

Moderation of the marking strategy occurs so that these standards are consistently applied when judging the level of a student's performance in the assessment task. That is, all markers will judge a student's work the same way, for the same reasons, giving rise to the same mark. At this point we say that the markers are 'calibrated' with each other.
Consensus moderation of marking occurs to ensure that valid, reliable and appropriate standards are being used and consistently applied when judging the level of learning achievement demonstrated by students who have completed an assessment task.

(‘Level 2: Marking student work’, Griffith)

**Level 3. Grading student work**

Grading refers to the grouping of student academic work into bands of achievement and should be an accurate reflection of a student's *overall* level of achievement in a unit.

Grades are commonly determined by adding up the raw data of marks for all assessment tasks in a unit, and converting this to a band of achievement (High Distinction, Distinction, Credit, and Pass). Hence, the grading process is usually intrinsically related to the marking process.

We want to ensure that the grade accurately reflects the quality of all of the student's work. Occasionally the total marks convert to a grade that is unexpected: that is, it does not seem to reflect the quality of work the student has produced over the entire unit. If this occurs, it is imperative to go back and review the actual collection of student work to make adjustments (not simply manipulate the marks).

Moderation of grading occurs to ensure that all markers agree on the reasons why a particular grade is awarded and its appropriateness. At this point we can say that the markers' grading judgments are 'calibrated' with each other.

The purpose of consensus moderation at this level is to ensure that the standard of the final unit grades are appropriate and consistent. Grades awarded to students need to be a valid reflection of their overall level of learning achievement in a unit, as illustrated by the full collection of their assessment work for that unit.

(‘Level 3: Grading student work’, Griffith)

**Level 4. Standards across Units (internal and external)**

As part of the assurance of assessment quality and academic standards we use consensus moderation to ensure that the standards required of students taking similar courses within any Institution are comparable. This means that the grading judgments are consistent. The performance required for a High Distinction in Chemistry in Program A, would look the same as that required for a High Distinction in Chemistry in Program B (which may be in the same Institution, or another institution, even internationally).

Furthermore, part of the consensus moderation process is to ensure that the standards used to reach the consistent judgments are themselves appropriate. Two cognate courses may have consistent standards, yet if the standards used in both are inappropriate; this is not an assurance of quality.

The standards between your course and cognate (similar courses), at ACU and other institutions need to be comparable.
The purpose of consensus moderation at this level is to ensure comparable levels of achievement standards for students between your course and cognate courses, and that these achievement standards are both appropriate and consistent.

(‘Level 4: Standards across courses’, Griffith)

Level 5. Standards over time

As a part of the assurance of quality, consensus moderation is also used to ensure there is no ‘slippage’ of assessment standards and judgments over time. When courses are delivered over years, it is not uncommon for small changes made over time to compromise assessment standards and quality. Staff turnover can be a further factor to consider when assuring standards over time.

Assessment judgments need to remain consistent, to ensure that the level of performance required for a High Distinction in a course in any year is comparable to previous years’ and future years.

Furthermore, as with assuring standards across courses, part of the consensus moderation process is to ensure that the standards used to reach the consistent judgments are themselves appropriate.

Purpose

The purpose of consensus moderation at this level is to ensure that assessment standards, including achievement standards applied to students’ work are appropriate and consistent over time.

(‘Level 5: Standards over time’, Griffith)

Stages of consensus moderation

The ‘Internal moderation of assessment’ team from ACU’s Faculty of Health Sciences developed an alternative overview of the quality controls for consensus moderation by looking at consensus moderation as a staged process. They termed these stages pre-, peri- and post-assessment.

They list some good practice considerations for each stage of the process.

Forward planning is paramount: moderation practices should be planned well in advance, and all members of the marking team should be aware of their requirements as a marker, and their required attendance at various key meetings for marking staff. Here are three short lists of some ways to implement consensus moderation and ensure equivalence. Best practice considerations including course content, teaching and assessment choices are also addressed.

(FHS Quick guide, p. 2)

Pre-assessment moderation

Pre-assessment moderation: Moderation practices which ensure the unit content, assessment tasks and associated learning outcomes are aligned, and that the assessment tasks are valid, reliable, fair and equitable, and equivalent across sites.

Pre-assessment moderation is also about maintaining academic standards – that is, ensuring that the assessment tasks reflect the year of study being undertaken by the students (first, second, final year).
**Good practice examples:**

- Moderation of the unit outline
- Constructive alignment
- Meetings with marking teams to ensure understanding of marking standards
- Shared rubric and marking guide development

**Simple steps to moderate pre-assessment and ensure best practice**

1. If the unit is supported by a textbook, use well-recognised textbook in your field to benchmark both assessment items and unit content.

2. Use constructive alignment (Biggs & Tang, 2007): make learning objectives and outcomes align with teaching and assessment tasks, at the appropriate level. Be able to demonstrate the links and keep a record of, for example, which learning objective matches which lectures and assessment items.

3. Plan in your timeline where marking calibration meetings with other teaching staff will occur, along with a proposed date and time. You may need to book a room or videoconference room (inter-campus calibration) well in advance to ensure these occur.

4. Ensure unit materials are clear and will allow an equivalent delivery between groups, both by the same or different teaching staff.

5. Make sure assessment pieces are appropriately weighted with increasing summative (not formative) assessment as the unit progresses. The majority of marks need to be given on student samples from the latter part of the semester, but it is important that earlier formative assessment gives student feedback and calibrates their expectations of expected academic standards. The sequence of assessment should also be logical with expected knowledge and skill acquisition.

6. A variety of different types of assessment items need to be used. This allows more complex skills to be demonstrated by students (e.g. critical analysis, self-reflection). For instance, some examples include individual or group assignments, individual or group presentations, practical skills exams, written or multi-choice exams and self-reflection or peer review tasks.

7. Produce and distribute materials to students that clearly outline expectations of expected learning outcomes and explicit assessment criteria. Identify how students will receive feedback for each assessment piece, and how this feedback will be given so it will support students to improve their work quality.

8. Compare unit content and assessment with equivalent units at other universities and use these to benchmark your unit.

9. Assessment items need to be designed to be fair and equitable. For example, are there options for delivery or marking that overcome disadvantages of students from different groups such as students with a disability, cultural groups, technology access or skills?

10. In your planning, use any ‘lessons learnt’ from previous offerings of the unit or appraisals of graduate outcomes (e.g. by employers or on practicum placements), when determining assessment tasks and how they will be marked.
11. Finally, use internal peer, co-ordinator and faculty of the final assessment plan as well as any rubrics and criteria for individual pieces. External review (by a peer at another university in the same teaching area or a professional body) can be used as well. This might be especially helpful when there are few teaching staff in your field at your university. “Internal” peers include “academic colleagues, teaching team members, discipline head, head of school, learning & teaching committees” (Griffith University, 2012, ‘Peer review’). External peers include “convenors of related units and courses from other institutions; industry liaison groups” (Nulty, 2011, p. 25) and “professional experts” (Griffith University, 2012, ‘Peer review’). Accreditation processes also can form a type of peer review (Nulty, 2011, p. 25)

**Peri-assessment moderation**

**Peri-assessment moderation:** Moderation practices which take place during the marking process to facilitate comparability of standards between different markers.

Peri-assessment moderation provides opportunities for members of a marking team to meet and review their mark and grade allocations early in the marking process to ensure they are allocating marks and grades in a consistent manner.

**Good practice examples:**
- Marking team meetings part-way through the marking process to review grade allocations to date
- Blind marking part-way through the marking process to ensure comparability (both on-campus and between campuses)
- Coordination of marking days on campus for members of the marking teams to encourage discussion about ‘difficult’ mark allocations
- Mentoring of new or sessional marking staff via regular meetings during the marking process to ensure they are familiar with application of the marking criteria, and the relevant academic standards

**Simple steps to moderate peri-assessment and ensure best practice**

1. Facilitate calibration of marking meetings by bringing the marking team together on campus in a venue conducive to marking and associated discussion (Note: Dates should be identified and rooms booked well in advance)

2. A small proportion of student assessments should be double-marked early in the marking process (both on-campus and between campuses) to ensure comparability of grades being allocated by a range of markers

3. If there are new members of the marking team, additional support should be offered – these staff should have the opportunity to be mentored by their lecturer in charge, and should be provided with guidance and feedback on their interpretation and understanding of the rubrics and marking guidelines, and the subsequent allocation of marks and grades, to ensure they are marking to the appropriate academic standards

4. When discrepancies between markers are identified during the marking process, it should be referred to the National Team Leader. The NTL has a role here to rectify the discrepancy, either (preferably) through discussion with the members of the
marking team involved to ensure a common understanding of the rubrics and marking guides or by reallocating marking when a discrepancy cannot be resolved in this manner.

**Post-assessment moderation**

**Post-assessment moderation:** Moderation practices which occur after marking has been completed for the purposes of continuous improvement and the maintenance of academic standards.

Post-assessment moderation includes assurance that grades which have been allocated are fair and representative of the standards of work produced by students.

Post-assessment moderation can also review how successful the unit has been in achieving the stated learning outcomes. It is an evaluation and a self-improvement process for the unit, enabling changes to be made for subsequent iterations.

**Good practice examples:**

- Review of borderline assessment tasks and high distinction and fail grades to ensure grades reflect academic standards, and have been allocated appropriately
- Double-marking of a sample of assessment tasks (both on-campus, and between campuses)
- Review of rubrics and marking guides in consultation with entire marking team for ease of use and of interpretation
- Review of grade distributions to inform subsequent teaching practice
- Review of unit evaluations by students
- Feedback from the marking team

**Simple steps to moderate post-assessment and ensure best practice**

**At the marking stage**

1. Plan consensus moderation activities (when and where) if they were not embedded initially. Ideally this would have occurred in unit planning.

2. Use a marking guide or rubric with explicit criteria. Develop these guides prior to or early in the marking stage and circulate them for critique and peer review. When finalised, disseminate these to students.

3. Conduct a peer review meeting with all markers for different assessment items, especially for those of higher weighting. Use a de-identified sample of students’ responses marked by all markers. In the meeting determine the application of criteria and rubrics to students’ responses so marking is equivalent between markers. Meetings can occur before, during or after marking to address any concerns of equivalence.

4. Following marking, double-mark or peer review (either internally or externally) assessment pieces where marks border on different achievement levels, and those which will lead to a fail or high distinction grade. Levels for which this review is
necessary are set by the lecturer in charge, and this can be undertaken at the calibration meeting, as well.

5. With a new marker, a lecturer in charge could remark a subset of responses at different grade levels and check standards are consistent. If more training is needed, address this in order to ensure marks are consistent for all markers.

6. Ideally use de-identified student work to allow fair consideration of responses.

7. Double or panel marking needs to be used for assessment items with high weighting (e.g. theses) and be blinded if possible.

8. Remember any adjustments to marks awarded following consensus moderation activities need to be finalised before returning students’ work.

9. **If you are a ‘sole operator’, you still need to use consensus moderation.** If you are at a different campus, peer review can still occur, for instance with de-identified scanning and emailing and double marking with internal or external peers. Marking can be benchmarked using a textbook with example tasks and assessment items and criteria or rubrics and weighting can be checked with internal or external peers during unit planning or at the marking stage. At grading, you will also need to use internal or external peers to ensure fair distribution.

**At the grading stage**

1. Assign grades based on actual student work. Do not average or adjust marks (for example, using the bell curve method) or use behaviour that is not actually a demonstration of skills or knowledge such as attendance or participation.

2. Peer review of grading of all assessment items for a selection of students is recommended. Grades need to reflect overall achievement, and with appropriate pre-assessment moderation and a strong assessment plan this is likely to occur. If once totalled, all marks do not assign a grade that seems reasonable, it may be necessary to peer review a student’s items and check all the marks assigned. It is usual to choose a selection of graded items at different levels as well as focus on students at borders of grades and those who are at a fail or high distinction grade level.

3. If grading seems consistently inappropriate for the sample of students chosen for review in a unit, the overall weighting of all assessment items may need to be reviewed and modified.

(FHS Quick guide, pp. 4-11)

**Unit continuity over time**

Keep records of annotated unit outlines, assessment plans, rubrics, selection criteria and a sample of marked student work across grade levels. Use these to inform your unit design and marking between semester offerings each year and for later assessment of multiple unit offerings using peer review. The goal is consistent standards over time.

(FHS Quick guide, p. 12)
Further reading and resources


Australian Catholic University. (2012). Australian Catholic University assessment review draft, 23 March.


Faculty of Health Sciences, Australian Catholic University. (n.d.). Interprofessional learning implementation committee (IPLIC) preassessment moderation review checklist. Australian Catholic University.


Cited components include:


University of Southern Queensland. Moderation of assessment.