

## 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' Note the Authors of this work were Nulty, D. and Shapland, N.)
3. Nulty, D., & Ashford-Rowe, K.H. (2010). Developing consensus moderation practices to support comprehensive quality assurance of assessment standards, internal project, Griffith University.

## What is consensus moderation?

Perhaps it is best to draw a distinction between “Moderation” and “Consensus Moderation” to ensure we can be clear about what consensus moderation is not, and also why it is important.

Moderation (not “Consensus moderation”) is about achieving ‘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). While this definition allows for a lot of scope, and can underpin good practice, generally “moderation” includes, and sometimes focuses on, the post-hoc adjustment of marks or grades. This is not good practice. An example might be ensuring that the distribution and mean of marks awarded to one group of students studying a subject on one campus is comparable to the distribution and mean of marks awarded to another group of students studying the same subject on a different campus. In this example, moderation can be mathematically applied as an adjustment. Moderation of that kind does not require any external benchmark or reference point, nor does it require the markers on the two campuses to agree. In fact, the markers themselves do not even have to be consulted. Similarly it does not have to involve any direct scrutiny of the students work and the associated marking. Moderation of this kind creates the illusion that the marking was undertaken in a way that had consensus about the standards as an underpinning. It does not mean that consensus was ever actually there.

“Consensus moderation” is different. In the example above, *consensus moderation* would require the markers on the two campuses to reach an agreed common position (a consensus) about what marks to give, and why, ideally before they undertake any marking.

Achieving consensus can be approached in a variety of ways. For example, it commonly involves the marking staff participating in “comparability meetings” with other staff. In these meetings participants individually and independently mark the same pieces of work and compare the results. Through discussion and clarification, the staff members gradually reach ‘consensus’ on the marks/grades they would award for a piece of work and the reasons for them. (‘Moderation strategies’, p. 1, ALTC Assessment Moderation Toolkit 2010).

Importantly however, there are other approaches to achieving consensus. Because of this there are choices that can be made about which approach to adopt. Some are influenced by the context, others apply across multiple contexts. The main point is that there is scope for flexibility, and a single “one-size-fits-all” approach is not required.

Recognising that there are different ways to achieve consensus among marking staff, it is important to keep in view the common goal. In short, the common goal is to ensure that the marks awarded to a piece of student work are a valid reflection of the standard of that work, and are comparable regardless of who does the marking, or when, or where. Thus, there are three essential components: the validity of the marks, the standards applied to marking, and the comparability of marks between markers, locations, and times.

Broadly, these three may be thought of in terms of fairness for students, and in

terms of equivalence between different markers. To achieve this, there are a number of quality control processes that can be implemented. Most of these involve direct peer review, others involve indirect peer review (e.g. via the use of credible benchmarks – such as the consistent use of a common textbook, which is a peer reviewed resource regarded as ‘a standard’.)

### **Consensus moderation model**

The Griffith University consensus moderation model (developed by Nulty, D. D. and Shapland, N. but built on foundational work by Sadler, R.) describes these quality control processes at 5 levels of assessment practice. The “levels” are actually phases in the cycle of planning, implementing and reviewing units (courses) for teaching and the quality assurance of the associated assessment practices. They are therefore presented in temporal order.

It is important to note that the order for these phases is cyclical. The significant advantage of this cycle is that each loop brings an improvement. The degree of quality assurance achieved increases with each iteration. In practical terms it is important to note that the first iteration is the most challenging because it cannot benefit from participants having been through the process before. On second and subsequent iterations, the participants can be reasonably expected to carry forward at least some of their learnings from previous rounds. When speaking of consensus moderation, people who have reached consensus have adopted a common position. The terms ‘calibration’ and ‘calibrated’ are used to describe this. Markers who have previously achieved consensus may be described as being in ‘calibration’ or as having ‘calibrated’ their respective positions with each other, and in alignment with accepted authoritative benchmarks. Such people will achieve consensus in subsequent rounds more rapidly, and with more enduring result. Thus, the workload burden for calibration to achieve consensus is higher for those just starting out, and/or in teams with changing membership, and lower for more experienced markers and/or in stable teams. This means that it is not usually necessary for all markers to engage in the same process to sustain consensus.

The 5 levels, or phases are:

#### **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 and 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)

The important point to note here is that validity of assessment and consistency of marking standards must be ‘baked in’ from the start. If the assessment tasks themselves are not valid, and/or if there is ambiguity over the purpose of the assessments, and the ways in which they should be marked, those weaknesses will be carried forward throughout the whole learning experience and the marking/grading – *regardless* of all other measures taken to assure quality.

### **Level 1(b) - Teaching**

Strictly speaking, the next ‘Phase’ of quality assurance is to ensure that the whole teaching team have a consensus about the fundamental purpose of the unit (its rationale), the unit’s aim, and how this purpose and aim will be pursued through the learning, teaching and assessment strategy. This was omitted from the Griffith model because at Griffith there were pre-existing institutional mechanisms that already targeted this aspect of quality assurance. Where such mechanisms are not already required, it is important to include the establishment of consensus among the teaching team before teaching commences. If this isn’t established before teaching commences, there is the risk that there will be inconsistencies in the learning outcomes achieved – which will be reflected in the students’ performances on their assessments, and which will result in inconsistency of marking that could be attributable to the inconsistency in teaching and learning and/or inconsistency in marking. It’s important to ensure everyone ‘sings the same song’, and therefore that there is constructive alignment in the way units are designed and delivered. As a point of note, delivery of “teaching” here is also taken to include consistency in the way the teaching staff members interpret and explain the assessments to students.

### **Level 2: Marking student work**

Consensus moderation of marking occurs at two times in the sequence.

1. Consensus moderation of marking schemes (which should occur before the unit is taught for the first time, and then repeatedly reviewed prior to each subsequent delivery).
2. Consensus moderation of the marking strategy i.e. the way markers will mark (which should occur before any marker does any actual marking, and can then, in addition, also occur during and after marking).

Consensus 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, the teaching team, and markers.

Consensus 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 tasks. That is, all markers will judge a student's work the same way, for the same reasons, giving rise to the

same mark. As noted above, at this point we say that the markers are 'calibrated' with each other. Ideally consensus around the marking strategy is first established before marking begins and then "topped up" by ongoing checks throughout the actual marking. This helps to consolidate and maintain consensus throughout, and supports the longevity of that consensus into following marking rounds.

Note: in practice, it is generally not practical nor possible to achieve marking that is "the same" from person to person, place to place and time to time. There is even an argument that says "sameness" is not desirable either – particularly in the creative arts for example. The opposite argument applies in disciplinary contexts where a particular approach is a requirement – particularly in health related or engineering contexts for example. However, if we accept that "to be identical" is generally not practically possible, the achievable purpose of consensus moderation is to seek an acceptable degree of comparability. In some contexts "acceptable degree of comparability" means "very high comparability" and in other contexts it means "broad discipline agreement". Higher degrees of comparability are easier to achieve when marking tasks for which definitive declarative answers are available – for example, there can only be one correct calculation for a drug dose. It is much more difficult when marking tasks for which a high degree of expert judgement is required – for example when evaluating a patient treatment plan, or when marking creative works.

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)

The outcome is marks that are not only comparable, but comparable for reasons that are comparable. This can only be achieved when the markers were in consensus prior to marking, and remain so throughout.

### **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 is 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)

Note: when considering if the sum of the marks given to all works produced by a particular student result in a valid grade, problems with the overall design of the assessment in the unit can be uncovered. For example, in a unit it is possible that the proportion of marks allocated to relatively low-level knowledge may inadvertently be too high – and as a result, the proportion of marks reflecting the desired higher level learning outcomes is too low. In such conditions it is possible for students to accumulate sufficient marks to pass without demonstrating a passing standard on the higher-order learning outcomes. Indeed, in such conditions, all students' marks are elevated because of the relative ease of accumulating marks associated with low-level learning outcomes. Such conditions can easily arise because it is easier to set and to mark assignments that assess low level outcomes, and because there is pressure on all academics to design assessment regimes that they can mark within a very limited amount of time. When this happens, the marks, and the grades do not reflect a valid assessment of the students' learning outcomes. Furthermore, the problem is not the marking – it is the design of the assessment across the unit. This problem is common when, for example, simple weekly quizzes are used, and the proportion of marks allocated to them is too high. It also happens when such approaches are combined with other assessments that incorporate, in part, the assessment of low level outcomes – so that when marks are added across other assessment tasks, the overall proportion of marks given to low outcomes is too high.

#### **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 a simpler overview of the quality controls for consensus moderation by looking at consensus moderation as a process involving three broad stages. They termed these stages pre-, peri- and post-assessment. Importantly, these still reflect all six of the levels reported above from the Griffith model developed by Nulty and Shapland.

### Pre-assessment consensus moderation

These are practices which ensure the unit rationale, aim, learning outcomes, content, and assessment tasks are aligned, and that the assessment tasks are valid, reliable, fair and equitable, and equivalent across sites.

These practices are also about maintaining academic standards – that is, ensuring that the assessment tasks reflect the standard of learning associated with the year level of study being undertaken by the students (first, second, final year).

#### Good practice examples

- Consensus moderation of the unit outline
- Constructive alignment of all unit components
- Marking teams engage in activities that ensure a consensus about marking standards.
- Teaching and marking staff share in the development of a marking rubric / marking guide to ensure a shared understanding of it.

Simple steps to support pre-assessment consensus moderation and 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/explain this alignment and keep a record of, for example, which learning objective matches which lectures and assessment items.
3. Plan, schedule and conduct marking calibration meetings for those staff members who need this to achieve calibration or who need to update their calibration. Note: This is good practice for all staff but may not be required for all staff. Note 2: You may need to book a room or videoconference room (inter-campus calibration) well in advance to ensure these meetings can occur and that the required participants can be present.
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 weighting to 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. A useful model is to consider a progression across declarative (factual) knowledge, to conceptual knowledge (understanding), and then to skills of application.

6. It is generally good practice to use a variety of different types of assessment. This allows for individual differences among students and supports assessment of a broader array of learning outcomes in a more triangulated manner. Some examples are 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, students from different cultural groups, students with different access to technology, and/or students with different (but equally valuable) 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. When possible, use external review (by a peer(s) at another university in the same teaching area or a professional body). This might be especially helpful when there are few teaching staff in your field at your university. Note: "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 consensus moderation**

These are practices which take place during the marking process (i.e. throughout a teaching period such as a semester) to facilitate comparability of standards between different markers.

These practices provide 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:**

- Discussions among marking team members throughout the marking process to review marks awarded to date. (NB: these can either be formally organised meetings, or discussions organised between more experienced markers other individual markers, or groups of markers using a structured sampling frame to ensure those staff members who need this most receive the greatest support.)
- Blind double marking and review part-way through the marking process to ensure comparability (both on-campus and between campuses). Again, a stratified sampling approach can be taken to give more support to less experienced markers.
- Coordination of marking days on campus for members of the marking teams – encourages and supports discussion/ review of 'difficult' or 'unusual' marking challenges and mark allocations.
- Mentoring of new and less experienced marking staff via any structured support mechanism (e.g. meetings, briefings, support resources, individual or group discussions) to ensure they are familiar with the 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 discussions. This can include by bringing the marking team together (virtually or physically), or by targeted sampling approach which gives those members of the marking team who need this support the most more support.
2. Double mark and review a small proportion of student assessments early in the marking process (both within a single 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, a higher level of 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 consensus moderation**

These are practices that occur after the marking for a particular teaching period (e.g. a semester) has been completed. The purposes are for continuous improvement and the maintenance of academic standards.

These practices include assurance that grades which have been allocated are fair and representative of the overall standard of work produced by each and every individual student.

Post-assessment consensus moderation can also review how successful the unit has been in achieving the stated learning outcomes and the overall design of the assessment. 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 totaled, 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

- Adams, R., & Roessler, T. (2012). TEQSA: Course accreditation (AQF qualifications) and CRICOS registration (Presentation). Retrieved from <http://www.teqsa.gov.au/sites/default/files/Course accreditation and CRICOS registration session 27 03 12.pdf>
- Australian Catholic University. (1998). Mission Statement. Retrieved from <http://www.acu.edu.au/45153>
- Australian Catholic University. (2009). Australian Catholic University students: Graduate attributes. Retrieved from <http://students.acu.edu.au/204356>
- Australian Catholic University. (2012). Australian Catholic University assessment review draft, 23 March.
- Australian Qualifications Framework Council. (2013). Australian Qualifications Framework (2<sup>nd</sup> ed.). South Australia: Australian Qualifications Framework Council.
- Biggs, J., & Tang, C. (2007). Teaching for quality learning at university: What the student does (3<sup>rd</sup> ed.). Buckingham, England: Open University Press.
- Boud, D., & Associates. (2010). Assessment 2020: Seven propositions for assessment reform in higher education, 1-4. Retrieved from [http://www.iml.uts.edu.au/assessment-futures/Assessment-2020\\_propositions\\_final.pdf](http://www.iml.uts.edu.au/assessment-futures/Assessment-2020_propositions_final.pdf)
- Brennan, D. J. (2008). University student anonymity in the summative assessment of written work. *Higher Education Research & Development*, 27(1), 43–54. doi:10.1080/07294360701658724
- Carless, D. (2007). Learning-oriented assessment: conceptual bases and practical implications. *Innovations in Education and Teaching International*, 44(1), 57–66. doi: 10.1080/14703290601081332
- Centre for Learning and Development, Edith Cowan University. (2012). Moderation of assessment in higher education. A literature review. Retrieved from [http://intranet.ecu.edu.au/\\_data/assets/pdf\\_file/0008/329255/Moderation-ECU-Literature-Review-120702.pdf](http://intranet.ecu.edu.au/_data/assets/pdf_file/0008/329255/Moderation-ECU-Literature-Review-120702.pdf)
- Crisp, G. (2007). *The e-assessment handbook*. London: Continuum.
- Crisp, G. (2010). Interactive e-assessment: Practical approaches to constructing more sophisticated online tasks. *Journal of Learning Design*, 3(3), 1-10. Retrieved from <https://www.jld.edu.au/article/view/57>
- Crisp, G. (2012). Assessment in virtual learning spaces: Concepts for the modern environment. In M. Keppell, K. Souter & M. Riddle (Eds.), *Physical and virtual learning spaces in higher education: Concepts for the modern learning environment* (pp. 199-218). IGI-Global. doi: 10.4018/978-1-60960-114-0.
- Crisp, G. (2012). Integrative assessment: reframing assessment practice for current and future learning. *Assessment & Evaluation in Higher Education*, 37(1), 33–43. doi:10.1080/02602938.2010.494234
- Crisp, V., & Novakovic, N. (2009). Are all assessments equal? The comparability of demands of college-based assessments in a vocationally related qualification.

Research in Post-Compulsory Education, 14(1), 1-18.  
doi:10.1080/13596740902717366

Ewan, C. (2009). Learning and teaching in Australian Universities. A thematic analysis of Cycle 1 AUQA audits. AUQA Occasional Publications, Number 18. Retrieved from [http://www.auqa.edu.au/files/publications/learning\\_and\\_teaching.pdf](http://www.auqa.edu.au/files/publications/learning_and_teaching.pdf) Faculty of Health Sciences, Australian Catholic University. (n.d.). Interprofessional learning implementation committee (IPLIC) preassessment moderation review checklist. Australian Catholic University.

Gamliel, E., & Davidovitz, L. (2005). Online versus traditional teaching evaluation: mode can matter. *Assessment & Evaluation in Higher Education*, 30(6), 581–592. doi: 10.1080/02602930500260647

Griffith University. (2012). Consensus moderation. Assessment matters. [Note, the principal authors of this document were Nulty, D. D. and Shapland, N.] Retrieved from <http://app.griffith.edu.au/assessment-matters/docs/consensus-moderation>.

Cited components include:

Definition and explanation of terms. Retrieved from

<http://app.griffith.edu.au/assessment-matters/docs/consensus-moderation/definition>

Peer review. Retrieved from <http://app.griffith.edu.au/assessment-matters/docs/consensus-moderation/peer-review>

Level 1: Course level assessment planning. Retrieved from

<http://app.griffith.edu.au/assessment-matters/docs/consensus-moderation/model/level1>

Level 2: marking student work Retrieved from

<http://app.griffith.edu.au/assessment-matters/docs/consensus-moderation/model/level2>

Level 3: grading student work Retrieved from

<http://app.griffith.edu.au/assessment-matters/docs/consensus-moderation/model/level3>

Level 5: standards over time Retrieved from

<http://app.griffith.edu.au/assessment-matters/docs/consensus-moderation/model/level5>

Higher Education Standards Framework (Threshold Standards) (Cth) (Austl.). Retrieved from <http://www.comlaw.gov.au/Details/F2012L00003/Download>

Kennedy, G., Dalgarno, B., Bennett, S., Gray, K., Waycott, J., Judd, T., ... Chang, R. (2009). Educating the net generation - A handbook of findings for practice and policy. Retrieved from <http://www.netgen.unimelb.edu.au/downloads/handbook/NetGenHandbookAll.pdf>

Kennedy, G., Judd, T., Churchward, A., Gray, K., & Krause, K. (2008). First year students' experiences with technology: Are they really digital natives? *Australasian Journal of Educational Technology*. 24(1), 108–122.

Keppell, M., Suddaby, G., & Hard, N., (Eds.) (2011). Technology-enhanced Learning and Teaching: Good Practice Report. Retrieved from <http://www.olt.gov.au/resource-good-practice-report-technology-enhanced-learning-and-teaching-2011>

- Klenowski, V., & Adie, L. (2009). Moderation as judgement practice: reconciling system level accountability and local level practice. *Curriculum Perspectives*, 29(1), 10–28.
- Krause, K. (2007). E-learning and the e-Generation: The changing face of higher education in the 21<sup>st</sup> Century. In J. Lockard & M. Pegrum (Eds.), *Brave new classrooms: Educational democracy and the internet* (pp. 125–140). New York: Peter Lang Publishing.
- Lawson, K., & Yorke, J. (2009). The development of moderation across the institution: a comparison of two approaches. Paper presented at the Australian Technology Network Assessment Conference 2009: Assessment in different dimensions, Melbourne, 19–20 November 2009.
- Mahmud, S., & Sanderson, G. (2011). Moderation for fair assessment in transnational learning and teaching. Final Report 2011. Retrieved from <http://www.olt.gov.au/resource-fair-assessment-unisa-2011>
- McAllister, S., Lincoln, M., Ferguson, A., Davidson, B., Hill, A., Davenport, R., ... & Tedesco, H. (2011). Benchmarking COMPASS® for curriculum renewal. Establishing infrastructure and collaborative processes for cross-institutional benchmarking of student clinical performance in speech pathology. Final Report. Retrieved from <http://eprints.jcu.edu.au/19925/>
- McMahon, T. (1999). Using negotiation in summative assessment to encourage critical thinking. *Teaching in Higher Education*, 4(4), 549–554.
- Nulty, D. (2011). Practices to support comprehensive quality assurance of assessment standards at Griffith [Presentation].
- Nulty, D. D. (2010). Curriculum review tools for QAA - Quality assurance of assessment: Part 1 - The framework. Retrieved from [http://www.griffith.edu.au/\\_data/assets/pdf\\_file/0005/344444/Curriculum-Review-Tools-for-QAA-Part-1.pdf](http://www.griffith.edu.au/_data/assets/pdf_file/0005/344444/Curriculum-Review-Tools-for-QAA-Part-1.pdf)
- Nulty, D. D. (2010). Curriculum review tools for QAA - Quality assurance of assessment: Part 2 - A four step approach to QAA of assessment planning. Retrieved from [http://www.griffith.edu.au/\\_data/assets/pdf\\_file/0007/344446/Curriculum-Review-Tools-for-QAA-Part-2.pdf](http://www.griffith.edu.au/_data/assets/pdf_file/0007/344446/Curriculum-Review-Tools-for-QAA-Part-2.pdf)
- Nulty, D. D., Short, L. M., & Johnson, N. W. (2010). Improving assessment in dental education through a paradigm of comprehensive care: a case report. *Journal of Dental Education*, 74(12), 1367–1379.
- Office of Assessment, Teaching and Learning, Curtin University, (2010). Developing appropriate assessment tasks. In *Teaching and Learning at Curtin* (pp. 22-44). Perth: Curtin University.  
Retrieved from [http://otl.curtin.edu.au/local/downloads/learning\\_teaching/tl\\_handbook/tlbookch ap5\\_2012.pdf](http://otl.curtin.edu.au/local/downloads/learning_teaching/tl_handbook/tlbookch ap5_2012.pdf)
- Robertson, D. (2011). A review of current assessment and related policies in Australian universities; literature in the area of assessment in higher education. Australian Catholic University.
- Sadler, D. R. (2009). Grade integrity and the representation of academic



- achievement. *Studies in Higher Education*, 34(7), 807–826. doi: 10.1080/03075070802706553
- Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education*, 34(2), 159–179. doi:10.1080/02602930801956059
- Sadler, D. R. (2010). Assuring academic achievement standards at Griffith University (2<sup>nd</sup> ed.). Retrieved from [http://www.griffith.edu.au/\\_data/assets/pdf\\_file/0010/375337/AssureAcadStndrds-Consultation-Paper.pdf](http://www.griffith.edu.au/_data/assets/pdf_file/0010/375337/AssureAcadStndrds-Consultation-Paper.pdf)
- Sadler, D. R. (2010). Fidelity as a precondition for integrity in grading academic achievement. *Assessment & Evaluation in Higher Education*, 35(6), 727–743. doi:10.1080/02602930902977756
- Sadler, R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, 30(2), 175–194.
- Sadler, R. (2011). Introduction to the process of moderating assessments. Retrieved from [http://www.griffith.edu.au/\\_data/assets/pdf\\_file/0003/225795/Introduction\\_process\\_of\\_moderating\\_assessments.pdf](http://www.griffith.edu.au/_data/assets/pdf_file/0003/225795/Introduction_process_of_moderating_assessments.pdf)
- Sadler, R., Davies, L., Moni, R., Stewart, D., & Brown, C. (2011). Good practice in assessment case study – School of Public Health: Assuring academic achievement standards and investigating grade inflation. Retrieved from [http://www.griffith.edu.au/\\_data/assets/pdf\\_file/0005/256811/Public-Health-Case-Study.pdf](http://www.griffith.edu.au/_data/assets/pdf_file/0005/256811/Public-Health-Case-Study.pdf)
- Sendziuk, P. (2010). Sink or swim? Improving student learning through feedback and self- assessment. *International Journal of Teaching and Learning in Higher Education*, 22(3), 320–330.
- Student and Academic Services, University of South Australia. (2012). Moderation and variation. In *Assessment Policies and Procedures Manual* (pp. 18–20): University of South Australia.
- Tertiary Education Quality and Standards Agency Act 2011 (Cth) (Austl.). Retrieved from <http://www.comlaw.gov.au/Details/C2011C00582>
- Tertiary Education Quality And Standards Agency Act 2011. Explanatory statement. (Cth) (Austl.) Tertiary Education Quality Standards Agency. (2012). Tertiary education quality standards agency: Thematic reviews. Retrieved from <http://www.teqsa.gov.au/thematic-reviews>
- Torrance, H., & Pryor, J. (2001). Developing formative assessment in the classroom: Using action research to explore and modify theory. *British Educational Research Journal*, 27(5), 615–631.
- Tremblay, K., Lalancette, D., & Roseveare, D., (2012). Assessment of Higher Education Learning Outcomes Feasibility Study Report (Vol. 1: Design and implementation). OECD. Retrieved from <http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume1.pdf>
- University of Southern Queensland. Moderation of assessment.
- University of Tasmania. (2012). Pre-assessment moderation. Retrieved from <http://www.utas.edu.au/education/learning->

and-teaching/policies-and- procedures/moderation/pre-  
assessment