



Professional Development Course for External Examiners

Participant Handbook

Name: _____

February 2019

Table of Contents

About the course	3
Introduction to the course	4
Course completion	4
Course facilitators	4
Course aim	4
Learning outcomes	5
The UK Professional Standards Framework	5
Useful information	6
Part 2 Programme	6
Session 1: Introduction	7
Task 1: Review of Part 1 Activity 2 scenarios	8
Follow up of Part 1 Activity 2: Scenarios with comments	10
Summary log	14
Session 2: Variability in academic standards	15
Summary log	17
Session 3: People as a source of variation in standards	18
Summary log	21
Session 4: Tools and tasks (Part A)	22
Summary log	25
Session 5: Tools and tasks (Part B)	26
Session 5: Evaluating tools in assuring standards within programmes and between programmes in different institutions – research findings.....	28
Summary log	33
Session 6: Professional practice in the external examiner role and decision making	34
Task 6: Dilemmas in professional practice.....	34
Summary log	36
Task 6: Dilemmas with comments.....	37
Additional set of dilemmas	44
Session 7: Social moderation and calibration of standards	62
Task 7: Social moderation	62
Case study: Summary of a social-moderation process to improve the reliability of dissertation assessment in a university department.....	63
Summary log	65
Example of calibration: a project in accounting education in Australia.....	65
Session 8: Conclusion	68
Task 8: Completing the summary log and identifying key points	68
What you should get out of it.....	68
Summary log	69
Follow-up activities.....	70
Glossary of terms	72
References	72

About the course

Thank you for participating in the course for external examiners. We hope that you enjoy the course and find it a stimulating opportunity to enhance your contribution to safeguarding UK higher education standards and improving the student experience.

The course has been developed as part of The Degree Standards project, which is led by Advance HE and managed by the Office for Students on behalf of England and the devolved administrations in Northern Ireland and Wales. The Project has explored sector-owned processes focusing on the professional development of external examiners. You can find out more about the project at <https://www.heacademy.ac.uk/degree-standards>.

The course has been developed through two rounds of piloting and evaluation to ensure that it is a valuable professional development experience for those taking part. The pilot stages involved partnership with eight UK higher education providers.

Oxford Brookes University
University of Liverpool
Royal Northern College of Music
Queen's University, Belfast
Cardiff Metropolitan University
University of Edinburgh
Northern Hub (Higher Education in Further Education): Newcastle College, New College Durham,
Sunderland College)
The Open University

As part of ongoing evaluation, we will be seeking feedback from participants so that the course can be improved further. Please do not hesitate to contact us for further information or if you have any queries about the course or the project more generally. Send an email to external.examining@heacademy.ac.uk.

Introduction to the course

The design of the course reflects an active learning approach, with an emphasis on:

- building understanding through valuing participants' experiences and prior knowledge;
- active participation, through a variety of activities (e.g. reading, exercises, discussion);
- ensuring direct relevance of content to participants' examining role, with the use of tasks that are representative of the kinds of activities with which, as examiners, they will engage;
- exchanging good practice, particularly by sharing the expertise of experienced examiners with participants who have less experience.

The course comprises two essential parts.

➤ Part 1: Online activities and reading

Part 1 comprises a series of activities and reading which are essential to making a constructive contribution and gain the maximum benefit from Part 2 because they lay the foundation for the activities and discussions. The reading required and any optional reading you have undertaken will help take an informed perspective on being an external examiner.

You should have access to or a record of your work that you have entered on the VLE. Other participants have reported feeling frustrated when they find themselves working in a group where some colleagues have not completed Part 1 or forgotten to bring their part 1 work with them.

➤ Part 2: Face-to-face day

Part 2 comprises 8 sessions that build directly on part 1 and include individual and group activities, discussions and reflection. It is a full, structured and intensive day.

Course completion

The blended course mode is designed as a coherent whole aimed to enable participants to meet the learning outcomes. Consequently, participants should only register for a course date when they are able to complete both Part 1 and Part 2 of the programme. Recognition of completion of the course on the HEA's list of those who have completed the external examiner professional development course will only be made based on this full attendance.

We recognise that on occasion unforeseen circumstances or transport arrangements necessitate participants arriving late or leaving early. Where this is unavoidable, the HEA as the course provider will provide the participant with independent study material to replace the missed session. Only one session can be covered in this way. Submission of the completed supplementary material will enable the participant to register full attendance. Absence from more than one session will lead to non-completion of the course.

Course facilitators

The course is delivered by two facilitators. These will be either members of the core project team or facilitators who have been trained to run the course.

Course aim

The aim of the course is to enable aspiring, new or experienced examiners to:

- understand the role of the external examiner as articulated in the UK Quality Code for Higher Education and be confident to undertake it;

- develop a deeper understanding of the nature of academic standards and professional judgement, and explore the implications for external examining;
- use evidence-informed approaches to contribute to impartial, transparent judgements on academic standards and the enhancement of student learning.

Learning outcomes

Having completed the course, participants will be able to do the following.

1. Explain and discuss the nature and purpose of the external examiner role, its function for quality assessment in higher education, including the importance within it of their contribution to safeguarding academic standards.
2. Explain the nature of standards in the higher education context.
3. Draw on practical and scholarly knowledge of assessment as appropriate to the role, including
 - a. professional judgement;
 - b. assessment reliability;
 - c. assessment validity;
 - d. purposes of assessment;
 - e. principles of assessment;
 - f. programme coherence in assessment.
4. Recognise the varied provenance and uniqueness of individuals' standards and the challenge this brings to examiners representing the standards of their subject, discipline and/or professional community.
5. Explain the importance and use of key reference points for academic standards in the relevant subject, discipline and/or professional area.
6. Explain the purpose and value of ongoing calibration activities in supporting the use of common 'discipline community' standards.
7. Recognise the importance of their continuing professional development in assessment and external examining.

The UK Professional Standards Framework

The [UK Professional Standards Framework](#) (UKPSF) is a comprehensive set of professional standards and guidelines for everyone involved in teaching and supporting learning in higher education.

Participants for this course are likely to be involved in external examining or may be leading external examining policies across a faculty, programme or team. If you are considering preparing an application for fellowship of the HEA, you should look at how your external examining experience relates to the dimensions of the UKPSF. For each level of fellowship, you must provide evidence of relevant experience and impact. For example, if you are applying for SFHEA you should show evidence of leadership (e.g. influencing and mentoring other colleagues' practices). There are follow-up activities in this Handbook, as well as online guidance <https://www.heacademy.ac.uk/ukpsf> on the HEA's website on aligning your external examining experience to the UKPSF.

If you work for a UK University, your institution may offer opportunities to achieve HEA fellowship through an accredited programme or scheme. You should check your accredited scheme and be aware of their requirements. If your institution does not have an accredited scheme, you can apply directly to the HEA for fellowship; more information is available on our [website \[https://www.heacademy.ac.uk/individuals/fellowship\]](https://www.heacademy.ac.uk/individuals/fellowship).

HEA fellows are expected to engage in appropriate [continuing professional learning](#) [https://www.heacademy.ac.uk/system/files/downloads/Code_Of_Practice.pdf] throughout their career, and attendance at HEA events offers an opportunity to do this.

Useful information

Higher Education Academy (2019) *Fundamentals of External Examining*. Available at:
<https://www.heacademy.ac.uk/knowledge-hub/external-examining>

Quality Assurance Agency (QAA) The UK Quality Code for Higher Education. Available at:
<https://www.qaa.ac.uk/quality-code>

Higher Education Funding Council for England (HEFCE) (2015) A Review of External Examining Arrangements Across the UK: Report to the UK Higher Education Funding Bodies by the Higher Education Academy. Available at:
<https://www.heacademy.ac.uk/project-section/review-external-examining-arrangements>

Part 2 Programme

Courses beginning at 9:30

09:00 – 09:30	Registration
09:30 – 10:10	Session 1: Introduction
10:10 – 11:00	Session 2: Variability in Standards
11:00 – 11:15	Break
11:15 – 11:45	Session 3: People
11:45 – 12:30	Session 4: Tools and tasks (Part A)
12:30 – 13:10	Lunch
13:10 – 13:40	Session 5: Tools and tasks (Part B)
13:40 – 14:25	Session 6: Professional practice in the external examiner role
14:25 – 14:40	Break
14:40 – 15:20	Session 7: Social moderation and the calibration of standards
15:20 – 15:30	Session 8: Reflection and concluding comments
15:30 – 15:45	Final matters

Courses beginning at 10:00

09:30 – 10:00	Registration
10:00 – 10:40	Session 1: Introduction
10:40 – 11:30	Session 2: Variability in Standards
11:30 – 11:45	Break
11:45 – 12:15	Session 3: People
12:15 – 13:00	Session 4: Tools and tasks (Part A)
13:00 – 13:40	Lunch
13:40 – 14:10	Session 5: Tools and tasks (Part B)
14:10 – 14:55	Session 6: Professional practice in the external examiner role
14:55 – 15:10	Break
15:10 – 15:50	Session 7: Social moderation and the calibration of standards
15:50 – 16:00	Session 8: Reflection and concluding comments
16:00 – 16.15	Final matters

Session 1: Introduction

This session aims to provide you with background to the course and debate matters associated with the external examiner's role.

Introduction

- Welcome
- Facilitators
- Degree Standards Project
- Learning outcomes
- Chatham House rule



AdvanceHE

Introduction continued

- Experience of participants
- Handbook
- Glossary
- The course: what it is not
(Note: see *Fundamentals of External Examining*)
- Intensive day



AdvanceHE

Follow up on part 1

1. Moderating assignments – discuss your comments
2. Scenarios on examiner role – discuss
3. Assessment briefing
 - use to assess a module
 - inform all discussions



AdvanceHE

Academic standards

- **Academic standards:** output measure focusing on student achievement
- **Quality standards:** input and process measures focusing on all other aspects of the assessment cycle (and other resources, procedures, etc.)

Reflects the structure of the revised Quality Code (2018)

AdvanceHE



Quality Code (2018)



Expectations for standards - 'external expertise':

'The academic standards of courses meet the requirements of the relevant national qualifications' frameworks'

'The value of qualifications awarded to students at the point of qualification and over time is in line with sector-recognised standards.'

Core practices related to the external examiner role

- The provider ensures that the threshold standards for its qualifications are consistent with the relevant national standards. [Maintainer of academic standards]
- The provider ensures that students who are awarded qualifications have the opportunity to achieve standards beyond the threshold level that are reasonably comparable with those achieved in other UK providers [Guardian of national standards]
- The provider uses assessment and classification processes that are reliable, fair and transparent [Process checker]

AdvanceHE

From UKSCQA & QAA (2018) UK QC for HE, Advice & Guidance: External Expertise QAA, p3

Quality Code (2018)



Expectations for Quality – External expertise

Courses are well-designed, provide a high-quality academic experience for all students and enable a student's achievement to be reliably assessed

'External examiners play an important role in identifying good practice and making recommendations for enhancement of assessment policies and procedures' [Critical friend]

AdvanceHE

From UKSCQA & QAA (2018) UK QC for HE, Advice & Guidance: External Expertise QAA, p4

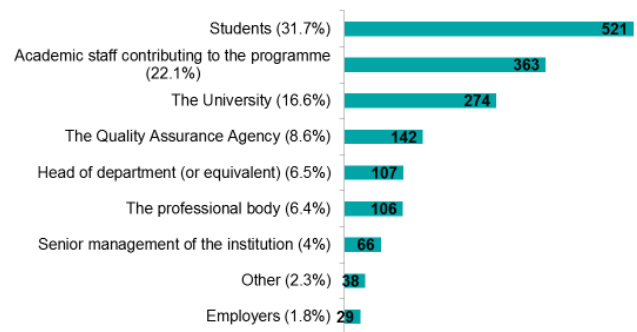
National picture: Role and 'process checking'

- 43% of examiners think the role has been reduced to **'process checking'** (for quality officers; 13%)
- 35% of examiners think there is too much **variability** in the role (for quality officers; 23%)
- Examiners believe they are perceived as (in rank order*):
 - 'Critical friend' (1.89)
 - 'Authority in the subject' (2.48)
 - 'Peer' (2.77)
 - 'Authority in quality assurance' (3.01)
 - 'Other' (4.21)
- Examiners are generally positive about their contribution to programmes (80% agreement, 9% disagreement)

AdvanceHE

* Examiners were asked to rank 1-5 with 1 being the highest
HEFCE Review of external examining report, 2015

Examiners' perception of their primary duties towards:



AdvanceHE

Role challenges

Institutions

- Appoint external examiners for their courses
- Train and set out expectations for these examiners
- Pay fees to these examiners
- Are regarded as autonomous with responsibility for the standards they use.



And yet the external examiner role suggests an element of impartiality and the need for independence from the institution.

Examiners need to balance a good working relationship with relevant staff while retaining sufficient distance to exercise their role.

AdvanceHE

Task 1: Review of Part 1 Activity 2 scenarios

Purpose

To make clear the tension in the role that can arise from attempting to balance 'process checker' and 'critical friend' with 'maintainer' and 'safeguarder' of standards.

What you should get out of it

The session will help you to understand how the different roles are revealed in practical examiner decision making, and to recognise the importance of privileging standards in your decision making.

Relevance to external examining

Understanding the role in action is central to your professional development as an examiner. Examiners should recognise the importance of privileging the maintenance and safeguarding of standards. A focus on checking processes or being a critical friend should not be a substitute for, or at the expense of, the examiner's role in relation to academic standards.

Description of task

In Part 1 you were asked to consider three scenarios illustrating potential tensions in the role of external examiner. We hope the scenarios and our responses to the various options have helped you to consider how the different parts of the role outlined in Chapter B7 of the UK Quality Code are revealed in examiner decision making.

Above all, make sure to recognise the importance of privileging the maintenance and safeguarding of standards and that a focus on checking processes or being a critical friend should not be at the expense of the examiner's role in relation to academic standards.

There is now an opportunity for you to spend 5–7 minutes sharing your option choices with other participants. If you have not brought copies of the scenarios with you, you can find them on the next page.

Follow up of Part 1 Activity 2: Scenarios with comments

The following table provides comments on the different scenarios presented in Activity 2. You may wish to consider your choice in the light of these comments.

Role Scenario A		
<p>You have been invited to a meeting at the beginning of your tenure as external examiner. The course leader wants to brief you about the course, get to know you and establish a working relationship. At the meeting the course leader makes clear that there are no issues around the assessment standards being used as they compare favourably with what he sees as an external examiner at another institution. He identifies two areas where he would appreciate your help. First, he confesses he has concerns about whether two other members of the course team are really committed to using criteria and moderation processes, so he wants you to focus on commenting on assessment processes. Second, he has started to prepare for the revalidation of the course due to take place next academic year and would appreciate suggestions from you about how to improve the course.</p> <p>You are aware that all the institution requires from you is a 'tick box' report.</p>		
Optional responses	Comments on the options	
1	<p>You interpret the tick box report as a lack of interest in your expertise. You see the revalidation as an opportunity to make the job interesting, useful and supportive for colleagues in this institution and so agree that you will help.</p>	<p>A 'tick box' report should not indicate a lack of interest. It is likely that the report pro forma is based on the advice required from examiners set out in Indicator 2, Chapter B7 of the UK Quality Code. This code explicitly requires examiners to advise on threshold standards, comparability of standards, as well as assessment processes. So it would be inappropriate to accept the course leader's assertion that assessment standards should not concern you. Your remit does include consideration of assessment processes and so you should be alert to where those processes are unsatisfactory wherever that might be. If there are two individuals not complying with university processes this is for you to judge.</p> <p>Working with colleagues and offering advice on improvements and the creative process of course design can be rewarding. Chapter B7 recognises that the external examiner could play the role of critical friend to enhance practice and courses, but this is a secondary role to that of advising on standards and thereby contributing to the external examiner system as a whole.</p>
2	<p>You make it very clear that it is up to you how you carry out your role and where you put your focus and resent the direction from the course manager.</p>	<p>This suggests that you have a very clear view of the independence of your role and intend to keep an objective position. You may want to think about how you make this clear, as the effectiveness of the role of examiner often works better when relations between examiner and course leader, team and institution are cordial.</p> <p>Research on examiners showed that a few viewed their independence as the right to impose their views without consideration of the institution's rules, regulations and explications of standards. External examiners should use their knowledge and expertise to advise, but</p>

		recognise the autonomy of institutions in relation to standards.
3	Having looked at the course documentation before the meeting you can see a range of ways to improve it and think that if the course can be improved the assessment standards will look after themselves (i.e. better learning processes and a higher level of challenge will ensure the quality of work improves). So you plan in the first instance to focus on improvement of the programme, which will also support the re-validation, and after that to focus on any modules not reaching high standards.	Even though your personal assessment is similar to the course leader it would be inappropriate to disregard the assessment standards in current use. Your main role is to advise on standards and you should be reporting on these year on year. It is not acceptable to ignore this duty even if you anticipate improvements being made. While it may be true that improvements to a course may facilitate better learning, and better understanding of assessment and quality work for staff and students, there is no guarantee that inputs bring about particular outputs (e.g. if benchmark statements are used in course planning this does not guarantee assessment standards used to mark student work will comply with that statement).
4	You thank the course leader for being honest about possible problems in the course and you bear in mind what he has told you.	Open and honest communication between you and the course leader, team and institution help facilitate an effective external examiner system to enable discussions about concerns and how they are addressed effectively. Within that communication you must make clear that your role is to observe the assessment standards and practices in use in order to make judgements about whether they are appropriate and aligned with the rest of the sector, so offering only to bear in mind the course leader's comments is far as you can go.

Role Scenario B

You are one of two external examiners on an exam board, each covering a different set of compulsory and optional modules within the same course. The other examiner is very experienced and has served as external examiner at several institutions across the sector.

The other examiner makes clear that the quality of the student work she has seen is not high enough and it has not been marked rigorously enough: that is, expectations of quality are too low. However, in fairness to the students she is prepared to allow the marks to go through because the criteria used to judge the work were those given to the students. She asks for your support for this position at the exam board. You are satisfied that on the modules for which you have responsibility for the marking standards being used are appropriate.

Optional responses		Comments on the options
1	You have the students' best interest at heart. The other examiner is not proposing action that will disadvantage the students therefore you are prepared to support the external in agreeing the marks.	Accepting the students as the main stakeholder in the external examining system aligns with the views of many other external examiners and the emphasis on serving the student in Chapter B7 of the Quality Code. However fairness for students is not a straightforward matter. This option is concerned with fairness for students based on transparency and due process. It accepts that marking using local reference points is paramount and overrides national standards. Institutions have responsibility for their assessment standards which are revealed

		through local interpretation and this option suggests that these should not be questioned especially if they have guided the students in preparation of their work.
2	You have the students' best interest at heart and have seen the high-quality work produced by students in the modules you oversee. In fairness to them you cannot support other students receiving higher marks than they should.	Accepting the students as the main stakeholder in the external examining system aligns with the views of many other external examiners and the emphasis on serving the student in Chapter B7 of the Quality Code. However fairness for students is not a straightforward matter. Fairness here is based on equity through the consistent application of appropriate assessment standards. An assumption is being made that the assessment standards being used by the other examiner are largely the same as your own.
3	You regard your responsibility as external examiner as ensuring that the standards of your subject are upheld and therefore cannot support the other examiner.	This position acknowledges that the subject community is the key stakeholder for the external examiner. Subject communities and professional bodies have a strong and vested interest in maintaining the standards of their subject(s). Consequently, as an expert in your subject you have a responsibility to ensure the standards of the subject are upheld. This in turn means that you must hold universities to account for the standards they use. As external examiner, you are able to compare standards used at different institutions which should reflect the subject standards. Underlying this position is a belief that it is not appropriate that students should gain qualifications if they have not achieved the requisite standards.
4	You are aware that criteria alone cannot clarify the quality of work expected and that marking is an imprecise 'science'. You seek to have a more detailed discussion with the examiner before the exam board begins later that day.	Although allowing the marks to be approved may be a pragmatic solution this situation has revealed a major issue that needs to be explored in depth. There are several questions that need to be asked and a range of factors that should be considered including comparison of assessment standards being applied, assessment standards used elsewhere and guidance provided for students.

Role Scenario C

You have been external examiner at an institution for two years. You have developed a good rapport with members of the course team although you think the course manager lacks leadership skills and does not have the capacity to bring about much change. You are facing a dilemma. You can see that some students are being given pass marks for very weak work that does not really provide evidence that they have met the learning outcomes. You know that if you challenge the staff about these judgements you are likely to damage rapport and you still have two years to serve. You are aware how hard members of staff work, and for many students the 'value added' is to be applauded. You know that staff are under pressure to ensure a good retention rate. If the rate drops the course will be under threat of closure.

Optional responses

Comments on the options

1	You strongly sympathise with staff and you think by working with the team you can achieve a lot that will help students learn more effectively. You do not want to jeopardise this opportunity by	This option aligns closely with widely held positions of external examiners. There is a strong preference to play a critical friend role, especially in contrast to a role focused on maintaining standards that potentially puts them in conflict with colleagues. This option also shows the
---	---	--

	criticising staff for occasionally too easily giving the benefit of the doubt to students. For the moment, you decide to focus on helping them improve the course, as the course offers great employment opportunities for successful students.	student as a major stakeholder in the external examiner role.
2	You are clear that the quality of work required to pass is not as high as for students at your home institution. You decide to refuse to agree the marks on several modules and report the problem in your report.	This position acknowledges the responsibility of the external examiner to advise institutions about their threshold standards and to make comparisons with standards at other institutions. The importance of the working relationship between examiner and course team is not considered, but it would be advisable to discuss the position with the course leader or chair of the exam board, which may allow resolution before the exam board.
3	You are aware that courses in your subject area have different emphases, with some courses being academic and others more practice focused. While the course documentation does not make it very clear, in talking to the staff the focus of the course is on developing students' practical skills. You think that this is acceptable within the parameters of the subject benchmark statement so you feel you can overlook the lack of academic prowess that seems to be underlying poor performance.	This position accepts that higher education courses are not uniform, although they should align with national reference points. The benchmark statement is identified, but it would probably be wise to check the Framework for Higher Education Qualifications (FHEQ) in relation to academic expectations. There does seem to be a lack of clarity about the emphasis of the course, and there should be a clear alignment between the course being delivered and the validated course documentation. This might be worth raising with the team.
4	In order to try to maintain the rapport you have with the team you decide to have a quiet word with the course leader to suggest something is done about marking standards.	This position indicates that the course team is seen as the main stakeholder, so the need to maintain good relations and help them improve is a priority rather than using formal channels to register your disquiet. It could also be seen as conflict avoidance, especially as you suspect the course leader will not act. It does allow you to have 'done something about it' even though what might be done is not clear.

Summary log

This course is not formally assessed except for participation. However, we have created a brief process for you to reflect on and gather notes about your learning from the course in a summary log, with an opportunity for you to self-assess after the course has finished.

Therefore on several occasions during Part 2 we give you an opportunity to make these notes in response to a number of prompts that we have included in this Handbook. In this first opportunity we would like you to use the following pro forma to consider some of the key concepts in the external examiner role that have been explored in the introduction.

In this Handbook, you can find further sections of the summary log to complete for each session.

Session 1 summary log: Undertaking the role of the external examiner		
Role	How the role contributes to the external examining system	Evaluation of the role and potential pitfalls
Maintainer of academic standards		
Process checker		
Guardian of national standards		
Critical friend and enhancement		

Session 2: Variability in academic standards

Task 2a: External examining of student work

Purpose

Here you will experience an external examining task which illustrates the difficulty in reaching consistent and comparable judgements about student achievement.

What you should get out of it

You will see that variation is the norm, and start to identify some of the causes of variation.

Relevance to external examining

External examiners are expected to make consistent and comparable judgements but, essentially, this is no different from marking judgements and is difficult to achieve.

Description of task

In Part 1, we provided you with three marked items of student work from a postgraduate module on higher education assessment, accompanied by information on the course outcomes, level, assessment task and marking scheme.

We asked you to write a short statement about each of the three assignments, as shown in this box

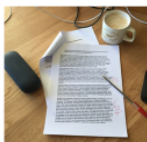
What feedback would you give to the course team on the academic standards of this module based on the sample of assignments? Please comment briefly on the extent to which you think the internal markers have made appropriate judgements about how the three items should be ranked, the marks awarded and the decision to refer one as 'not yet meeting threshold standards'.

In your group, please share your opinions of the work and the marking, and try and achieve a consensus in relation to:

- the rank order of the assignments;
- whether you agree with the fail;
- whether the marking standards are in line with the assessment criteria.

Evidence of individual differences

Many studies, over many years, have found considerable discrepancy in the grades accorded to the same or similar assignments.



Hartog and Rhodes, 1935; Laming et al., 1990; Wolf, 1995; Leach, Neutze, and Zepke, 2001; Elander and Hardman, 2002; Newstead, 2002; Baume, Yorke, and Coffey, 2004; Norton, 2004; Hanlon et al., 2004; Read et al., 2005; Price, 2005; Shay, 2004 and 2005; Brooks, 2012; O'Hagan and Wigglesworth, 2014; Bloxham et al., 2015 (this list is not exhaustive)

AdvanceHE

Variation in examiners' ranking of student work

	Five assignments in each subject – A to E				
	A	B	C	D	E
Psychology	3 rd – 5 th	1 st – joint 2 nd /3 rd	1 st – 5 th	1 st – 5 th	1 st – joint 4 th /5 th
Nursing	1 st – joint 3 rd /4 th	1 st – 5 th	joint 1 st /2 nd – 5 th	Joint 1 st /2 nd – 4 th	1 st – joint 3 rd /4 th
Chemistry	1 st – 5 th	joint 1 st /2 nd – joint 4 th /5 th	joint 1 st /2 nd – 5 th	1 st – 3 rd	1 st – 5 th
History	joint 1 st /2 nd – 3 rd	joint 1 st /2 nd – 4 th	joint 2 nd /3 rd – 5 th	joint 2 nd /3 rd – 5 th	1 st – joint 1 st /2 nd

AdvanceHE

Implications of variation in marks

- Difficult to claim that shared standards exist in subject communities
- Fairness to students
- Student dissatisfaction (e.g. NSS)
- Variation means we cannot assume that external examiners, even though experts in their fields, will necessarily have the same view of standards
- This is important for the external examiners' role in commenting on both threshold standards and comparability of standards across programmes and time



AdvanceHE

Task 2b: Complex scenario: what are the issues relating to variability in academic standards?

Purpose

Variability in academic standards can arise from several sources. This task aims to illustrate the complexity of standards within the context of external examining.

Description of task

Read the scenario given below then, working with your group, identify possible issues that may lead to variability in academic standards. It is not necessary to make proposals about resolving the issues.

Can you see any patterns in the issues identified? (10 mins).

Scenario: Designing a webpage

This module serves several programmes in different faculties and departments. Therefore, students bring different knowledge and skills to their assignment work including differing proficiency with IT. The module handbook includes module learning outcomes, an assessment rubric, institution-wide level descriptors, graduate attributes and transferrable skills. While these are largely consistent with one another it is necessary to look very carefully at the multiple sources to establish key qualities expected in student work. In line with the institution's drive to develop graduate capabilities, the assignment requires the student to design a webpage to engage an interested and informed audience about a controversial topic in the subject. The students were given support in web design (half day workshop) including technical skills and ideas on how to use creativity in design. The assessment criteria are focused largely on the subject content rather than webpage design and this reflects the fact that web design is not part of the module or course learning outcomes.

The task has clearly engaged the students, but they seem to have put more effort into the webpage design than into selecting and organising content, and you know that they have requested that web design is more heavily weighted in the judgement of the work. Each piece has been marked by one marker from the team and there seems to be inconsistency between markers on what is valued in the student work, with some strongly influenced by the quality of the web design. The feedback shows quite different levels of knowledge about the technicalities of web design, views on creativity and expectations of the knowledge content. You find yourself at odds with the evaluations of quite a few pieces of work, particularly in relation to the creativity demonstrated, and consequently you disagree with some of the marks awarded.

You have been provided, as usual, with a brief module report about marks including the average mark, range and a distribution curve. Comparing this with the one from the previous year when the assignment task was different, you find that there is little difference between the profile of marks. The report does not give any information about cross-marker moderation of grades.

Summary log

Session 2 summary log: Variability in standards

What issues does the problem of variability in academic standards raise for you as an external examiner?

Identify three sources of variation that you could focus on in ameliorating variation in standards.

- 1.
- 2.
- 3.

You can also use understanding from this session to complete your summary logs for sessions 3–5.

Session 3: People as a source of variation in standards

Session 3 People as a source of variation in standards

AdvanceHE

People as a source of variation in standards

Important because the external examining system relies on individual examiners holding similar standards

- Examiners should be confident that their standards are acceptable within their subject community.
- Most assume we hold shared standards but this is not supported by research.



AdvanceHE

Nature of standards

- Our understanding of standards is shaped by many influences
- Standards do not exist independently of those who use them – they develop their meaning in use through our social involvement in academic and professional communities



Unlike standards for cycle helmets, the language of academic standards is usually open to different interpretation

Standards are socially constructed

AdvanceHE

How we know our standards – tacit knowledge

Tacit knowledge: something you know that cannot be put into words e.g. “I know a 2.i. when I see it”.

It is fundamental to expertise and judgement-making across the professions, including judgements on student work.

- We learn the tacit knowledge of marking standards ‘on the job’;
- We gradually become less conscious of everything we are taking into account when making judgements;
- These create the conditions for differences in standards which we may not be aware of.

AdvanceHE

What shapes your standards?

Many differences in our experience:

- Your institution now and in the past
- Your values and beliefs
- Your specialist knowledge and knowledge of your profession
- Academic socialisation processes
- Knowledge of, and relationships with, students

There is little evidence that differences are tempered by membership of your subject or professional community



AdvanceHE

Implication: Need for calibration

We cannot rely on our informal learning about standards to ensure that we share standards with the rest of our subject community.



“Consistent assessment decisions among assessors are the product of interactions over time, the internalisation of exemplars, and of inclusive networks. Written instructions, mark schemes and criteria, even when used with scrupulous care, cannot substitute for these” (HEQC, 1997)

AdvanceHE

Task 3: Individual provenance of standards

Some of us believe we are rational beings and that we base all our professional judgements entirely on the framework of criteria and level descriptors for an assignment. Unfortunately, research suggests that our attempts to do this are limited. We have all been subject to individual experiences and influences that have contributed to our tacit knowledge and shaped what we think is important. This short activity should help tease out what has influenced and become important to you in making marking judgements and to help you consider the individual nature of these influences.

Purpose

To help you recognise the range of influences on their standards and consider the likelihood of variation in standards across different examiners.

What you should get out of it

You should identify some of the sources of your standards and the impact that different sources may have on you, such as the institutions you have worked in, the standards of any mentors and the quality of student work that you have been exposed to.

Relevance to external examining

External examiners need to beware of assuming that they can consistently reflect community subject standards just by virtue of being a member of that community.

Description of task

The following list, grouped into categories, is based on research that explored influences on external examiners' academic standards. It prompts you to reflect on possible influences that have shaped and continue to shape your academic standards.

Spend five minutes scanning the list. Can you pick out items that have influenced you? Do you know when they have influenced you? As you think about this you may be able to identify more distant, but equally influential items.

Influential people and groups

- Mentor(s)
- Research supervisor(s)
- External examiner(s)
- Expert teacher(s)
- Memorable teacher(s)
- Friends
- Family
- Professional bodies
- Discussion networks
- Learned societies
- Other:

Experience

- Local norms (e.g. standards, processes, regulations at your home institution or others known to you)
- Marking student work
- Your own education (e.g. what and how you were taught, your experience of being a student, of being marked)
- Being an external examiner
- Professional experience
- Continuing professional development
- Helping others learn to mark

Other:

Personal values and beliefs

- Higher education (e.g. What it is for? Who should it serve? Who should have access to it? What sort of knowledge and skills must graduates have?)
- Your subject (e.g. theoretical or practical, difficulty, knowledge-based or enquiry-based)
- Your role as an academic (e.g. is your main identity as a teacher, researcher, student guide, gatekeeper, facilitator?)
- Purpose and nature of assessment (e.g. measurement, hurdle, learning process, collaborative)
- Particular identity (e.g. as a member of a profession or as 'keeper' of national standards)
- Other:

Summary log

Session 3 summary log: Relying on people	
Issues	Your comments on these issues, their implications for variation in standards and how they can be accommodated
Standards rely in large part on tacit knowledge and professional judgement	
Socially constructed nature of standards	
External examining and comparability of standards	

Session 4: Tools and tasks (Part A)

Session 4 Tools and tasks Part A: Formal guidance

**AdvanceHE

Tools – two types

1. **Documented reference points** – sector-wide, profession-based, institution-wide or internal to the department, programme or module



2. **Processes** for quality assuring standards



**AdvanceHE

Reference points for examining decisions - external

Key Reference Points

- The Framework for Higher Education Qualifications
- Qualification characteristics
- Subject benchmark statements
- Degree classification descriptions (under consultation)

Quality Code for UK Higher Education (2018)

Themes in the QC on *External expertise* and *Assessment* are particularly important for external examiners

- There are also other external reference points, such as relevant professional standards

**AdvanceHE

Reference points for examining decisions - internal

These differ by institution and programme, but include information and guidance such as:

- Institutional qualification descriptors
- Institutional grade descriptors
- Assessment regulations, policies and procedures
- Programme specifications
- Programme and module learning outcomes
- Assessment guidance to students
- Assessment criteria and marking schemes

**AdvanceHE

Reference points and external examiners

- Research with practising external examiners has discovered that they claim to make limited use of formal reference points in examining (QAA, 2005; Colley and Silver, 2005; QAA-HEA, 2013)
- There is a tendency to consider that reference points such as benchmarks and qualification level characteristics will have to be considered at the course approval stage (QAA-HEA 2013)
- In general, external examiners claim to be more influenced in their judgements by personal experience of standards and quality assurance procedures than by formal documents (Colley and Silver, 2005; Ross, 2009)


**AdvanceHE

Reference point limitations

- We need to recognise that both national and professional reference points can be interpreted in different ways
- **Example from QAA 'Characteristics statement for masters degrees'**

'All master's degree graduates have in-depth and advanced knowledge and understanding of their subject and/or profession, informed by current practice, scholarship and research. This will include a critical awareness of current issues and developments in the subject and/or profession; critical skills; knowledge of professional responsibility, integrity and ethics; and the ability to reflect on their own progress as a learner.'

**AdvanceHE

The limitations of assessment criteria	Open-ended tasks						
<table border="0"> <tr> <td data-bbox="220 212 391 235">Positive intentions</td> <td data-bbox="497 212 689 235">But</td> </tr> <tr> <td data-bbox="220 275 391 342">Designed to improve consistency of marking</td> <td data-bbox="419 241 778 387"> <ul style="list-style-type: none"> Academic judgement cannot be easily represented by a short set of criteria Can't capture tutors' tacit knowledge Can't represent complexity of assessment tasks Need interpretation of meaning, different interpretation leads to different judgement </td> </tr> <tr> <td data-bbox="220 427 391 472">Staff and students positive about them</td> <td data-bbox="419 398 778 544"> <ul style="list-style-type: none"> Some assessors disagree with, ignore or chose not to use criteria, taking an holistic approach to judgement Use personal criteria (e.g. effort) rather than published criteria Use criteria they may not be aware of, e.g. font size. </td> </tr> </table>	Positive intentions	But	Designed to improve consistency of marking	<ul style="list-style-type: none"> Academic judgement cannot be easily represented by a short set of criteria Can't capture tutors' tacit knowledge Can't represent complexity of assessment tasks Need interpretation of meaning, different interpretation leads to different judgement 	Staff and students positive about them	<ul style="list-style-type: none"> Some assessors disagree with, ignore or chose not to use criteria, taking an holistic approach to judgement Use personal criteria (e.g. effort) rather than published criteria Use criteria they may not be aware of, e.g. font size. 	<ul style="list-style-type: none"> Assessment tasks which generate responses from students are: <ul style="list-style-type: none"> Varying Complex Unpredictable Sometimes personalised – e.g. an essay Higher order and complex performance necessarily involves work with these characteristics It is unfeasible to precisely document marking schemes Marking relies on judgement 
Positive intentions	But						
Designed to improve consistency of marking	<ul style="list-style-type: none"> Academic judgement cannot be easily represented by a short set of criteria Can't capture tutors' tacit knowledge Can't represent complexity of assessment tasks Need interpretation of meaning, different interpretation leads to different judgement 						
Staff and students positive about them	<ul style="list-style-type: none"> Some assessors disagree with, ignore or chose not to use criteria, taking an holistic approach to judgement Use personal criteria (e.g. effort) rather than published criteria Use criteria they may not be aware of, e.g. font size. 						
*AdvanceHE	*AdvanceHE						

Task 4: Reviewing the assessment of a module

Purpose

To model what examiners might do to evaluate the quality of assessment tasks, particularly in relation to effective assessment practice (scholarship), alignment to learning outcomes (validity), coherence of assessment strategy, and alignment with national reference points.

What you should get out of it

- An understanding of what to look for in reviewing assessment documentation at speed.
- An ability to recognise the most important factors to review in relation to examiners' primary responsibilities (maintainer of standards, process checker, safeguarder of national standards and critical friend).
- A recognition that it is difficult to determine the standards of a module just from the documentation – standards are a combination of setting appropriate tasks, appropriate assessment criteria and student performance.

Relevance to external examining

External examiners need to be able to make sound, well informed judgements about the quality of assessment task design in order to report on safeguarding national standards and appropriately robust assessment processes, and to provide commentary on good practice and opportunities for enhancement.


Description of task

You were asked to bring the assessment information for one module with you – typically in the form of a module handbook. The task is to work with a partner to review this assessment information, which represents a key task of external examiners. You have 20 minutes for this task. Consider, in particular:

- the learning outcomes;
- the assessment tasks, both formative and summative (assignments and/or examinations);
- any assessment guidance (e.g. information about how to prepare for or to carry out the task and assessment criteria or the marking scheme).

Use the table below to note the key aspects, enacting your different roles as an external examiner.

Aspect	Your notes
<p>How well do the assessment task(s) provide a coherent 'package' that validly assesses the learning outcomes for the module? <i>(Process checker role)</i></p>	
<p>Does the information indicate that students successfully completing the module will have achieved the appropriate academic standards in relation to:</p> <ul style="list-style-type: none"> the level of the module (e.g. undergraduate year 1, master's level)? the expectations of the subject or professional field? <p><i>(Guardian of national standards role)</i></p>	
<p>Drawing on your scholarly knowledge of effective assessment, what comments could you make to the module team on examples of good practice and opportunities for enhancement? <i>(Critical friend role)</i></p>	


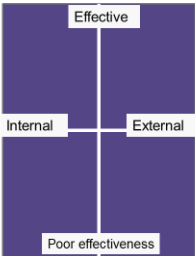
<h3 data-bbox="240 1014 730 1084">Assessment briefs and academic standards</h3> <p data-bbox="240 1093 678 1187">It is an important part of the role to review course documentation including assessment tasks, but such documents have limitations for revealing academic standards.</p>  <p data-bbox="443 1211 730 1375">Standards are only really revealed in the combination of valid assessment tasks, appropriate criteria and students' performance.</p> <p data-bbox="220 1406 338 1429">*AdvanceHE</p>	<h3 data-bbox="858 1025 1155 1061">Summary Session 4</h3> <ul data-bbox="852 1099 1342 1357" style="list-style-type: none"> • External examiners should be familiar with and use relevant external reference points in their role; • Whilst work should be moderated in line with the stated local standards, it is important to judge whether these local reference points reflect national reference points. • Written standards are open to different interpretation – an important reason for examiners to calibrate their interpretation of them with others in their subject community. • External examiners need to pay attention to the combination of assessment tasks, criteria and students' performance. <p data-bbox="836 1406 957 1429">*AdvanceHE</p>
---	---

Summary log

Session 4 summary log: Tools and tasks: Examining assignment information		
Types of tools and task information	Your comments on the usefulness and limitations of the items listed in relation to the external examiner role and variability in standards	
	Usefulness	Limitations
External reference points		
Assessment task instructions		
Identification of learning outcomes being assessed		
Rubrics (including criteria and level descriptors)		
Provision of exemplars		

Session 5: Tools and tasks (Part B)

This session follows on from this morning's session where you considered tasks plus the tools designed to support clarity and consistency in academic standards. This session is concerned with evaluating other tools used to assure academic standards.

<p>Session 5 Tools and tasks Part B</p> <p>AdvanceHE</p>	<p>Focus of the morning</p> <ul style="list-style-type: none"> • Role of an external examiner is multifaceted • Need for more balance in the role to ensure a robust approach to maintaining and safeguarding academic standards • Academic standards are not straightforward • Started to examine three key sources of variability – people, tasks and tools <ul style="list-style-type: none"> – People: multiple influences on their tacit knowledge of standards – Tasks and tools: task briefs and tools for defining academic standards expected. What do they tell you about expected academic standards? <p>AdvanceHE</p>
<p>Focus of the afternoon</p> <ul style="list-style-type: none"> • More on tools • Dealing with dilemmas and tricky problems in external examining • Social moderation and calibration <ul style="list-style-type: none"> – What are they? – How do they work?  <p>AdvanceHE</p>	<p>Task 5: Debrief</p> <ul style="list-style-type: none"> • Have any patterns emerged on your matrix? • Any other findings or comments from doing the matrix? • Which methods took longest to place – involved most debate?  <p>AdvanceHE</p>

Task 5: Evaluating tools for assuring standards within programmes and between programmes in different institutions

Within the higher education sector various approaches have been adopted to improve consistency of marking. The purpose of this task is to critique some common and more innovative mechanisms for assuring academic standards within and across modules, programmes and institutions.

Purpose

To evaluate common tools and processes designed to reduce variation in standards using a framework focused on effectiveness and locus of comparability (national or local).

What you should get out of it

By engaging in careful consideration of a range of methods you will come to a realisation that effectiveness of methods is variable.

Relevance to external examiners

External examiners need to be able to evaluate effectiveness of tools and process being used.

Description of task

Each small group is provided with a set of cards and a large matrix with two axes. Working together, consider the cards one at a time. Establish a shared understanding of the method described on the card and agree where to place it on the matrix. Consider as many cards as possible in the time and be prepared to comment on the distribution of the cards in the matrix and the reasons for the positioning of the cards (10 minutes).

Use the glossary in the handbook for any unfamiliar terms.

You will then have the opportunity to compare your evaluations with those of other groups and engage in a short plenary.

Session 5: Evaluating tools in assuring standards within programmes and between programmes in different institutions – research findings

Table of methods

Method	Comments on effectiveness	internal (I) or external (E)
Peer scrutiny of module assessment, instructions, criteria etc. (before start of module)	This is dependent on the type of scrutiny. It can be positive if it engenders conversation about the expected quality of work but a simple 'tick box' pro forma may encourage superficial scrutiny. Scrutineer may not recognise their own assumptions about the meaning of assessment brief and/or expected quality of work.	I
Pre-teaching briefing to module team on expectations for the assessment	To be beneficial, the module team need to discuss exemplars of student work (e.g. from a previous cohort) in relation to the criteria, rather than just receive briefing notes on assessment before start of module. This increases the chance that they will develop a similar grasp of the assessment requirements and provide students in different groups with consistent advice about the assessment.	I
Pre-teaching module team exercise to mark and discuss exemplar assignments (e.g. from the previous year)	Examples should open up discussion about academic standards and allow module team to share a common view of key aspects of quality expected before the start of the module. This increases the likelihood of students receiving consistent advice.	I
Whole course team development and enactment of programme assessment strategy	This can ensure that module leaders see how the assessment in their module contributes to meeting programme learning outcomes, particularly where discussion involves expectations about academic standards at each level of the course, progression, balance between assessment of and for learning and the relationship between assessment in other modules.	I/E Inputs at programme planning stage include external reference points and external advisors
Second marking of all work, resolving differences by discussion	Can provide a second viewpoint on work and insight into colleague's judgements by another marker. However, little evidence of effectiveness because second markers are	I

<p>or by averaging</p>	<p>suggestible if they see first marker's grades. Blind second marking can encourage 'defensive' marking to avoid being too far out of line with colleagues.</p> <p>Averaging incompatible with criterion-referenced assessment can advantage or disadvantage the student unfairly and research suggests resulting grade often less accurate than first mark.</p> <p>Discussion likely to lead to greater sharing and alignment of academic standards and the 'right' mark for the work. Discussion outcomes can be influenced by a range of factors such as hierarchical relationship between markers. Resource intensive.</p>	
<p>Blind double marking (BDM) of all work, resolving differences by discussion or by averaging.</p>	<p>BDM means at least two markers apply their independent academic standards.</p> <p>BDM can encourage 'defensive' marking, avoiding high or low marks for fear of making wildly different judgements.</p> <p>Averaging is incompatible with criterion-referenced assessment and could advantage or disadvantage the student unfairly. Research suggests resulting grade often less accurate than first mark.</p> <p>Discussion likely to lead to greater sharing and alignment of academic standards and the 'right' mark for the work.</p> <p>Discussion outcomes can be influenced by a range of factors such as hierarchical relationship between markers.</p> <p>Resource intensive.</p>	<p>I</p>
<p>Moderation discussion after first marking, involving all markers on a module</p>	<p>Discussion of marked student work will help to share and compare academic standards.</p> <p>Completion of all first marking means markers likely to be reluctant to re-mark work.</p>	<p>I</p>
<p>Sample second marking by module leader</p>	<p>Although not all students work is considered, sampling can help to identify inconsistencies between markers.</p> <p>This common method means that the academic standards applied belong to one person. As designer of the assessment, the module leader has the best understanding of the assessment and expected academic standards.</p> <p>Sample marking has the same dangers as second marking: that is, suggestibility of the</p>	<p>I</p>

	<p>first marker judgment. Also samples may not be representative of the whole where open-ended tasks (e.g. essays) are involved. However, if the second marker sees samples from more than one, first marker comparisons can be made.</p> <p>It is in the interest of module leaders to try to share understanding of standards with markers before marking, so that second marking does not show major discrepancies.</p>	
Use of a detailed marking scheme	<p>For complex tasks devising an 'easy to apply' marking scheme which provides for marker consistency is almost impossible. First markers necessarily develop their own understanding of the scheme as many words (e.g. good, excellent) need interpretation.</p> <p>On the positive side, a scheme will give an idea to first markers about key aspects of quality.</p>	I
Provision of model answer	<p>Complex tasks are unlikely to have only one answer. A model answer might restrict high marks to work that conforms to the model answer rather than work that exhibits high academic standards in a different way.</p> <p>Provision of a range of good answers may allow markers to make better judgements.</p>	I
All markers mark and discuss a common sample of work before full marking process.	<p>Provides an opportunity to compare, discuss and agree the basis for the judgements of the remaining work. This is likely to lead to greater consistency and the need for only light touch post-marking moderation.</p> <p>It cannot eliminate the need for post-marking moderation entirely.</p> <p>Some scripts get more attention than others and it marginally increases the marking load of each marker.</p> <p>If sample marking only results in comparison of marks awarded rather than reasons for the different judgements, it is less likely to lead to greater consistency in subsequent judgements. However, it would provide a benchmark and have similar benefits to an exemplar.</p>	I
Team marking session with markers able to discuss decisions particularly about 'unusual' work – <i>marking bee</i>	<p>Discussion and comparison of judgements as they make them should relatively quickly build a shared understanding of academic standards among markers in that context.</p> <p>Often used in conjunction with joint sample marking at the start of the bee.</p> <p>Finding time to mark simultaneously can be difficult. Many markers are used to, and prefer to mark on their own and at home and may be reluctant to participate.</p>	I

Moderation by comparing averages and distribution of marks given by each marker in the team	While this may be a useful initial snapshot of markers' judgements and student achievement, it is insufficient on its own to be a reliable moderation process. Batches of marking will not necessarily include the same range and quality of work rendering averages and distributions non-comparable. Moderation processes for criterion-referenced assessment should be concerned with the extent of achievement of learning outcomes rather than statistical patterns.	I
Exam board consideration of means and standard deviations of module marks	Provides a spurious comparison between modules, given that many factors can affect distributions including abilities and effort of students, nature of the subject matter and teaching methods. May be useful if designed to open up discussion about reasons for variation, including potential differences in academic standards (e.g. where students appear to score consistently higher in some electives compared with others).	I
Institutions require module mark profiles to conform to a reasonable 'curve', requiring justification for variation	Distribution curves are based on 'random' activities, so inappropriate for 'purposive' activity, such as education. Expectations that mark profiles will conform to a 'curve' does not fit well with criterion-referenced assessment. It may lead staff to 'fit' marks to 'requirements'.	
External examining	Depends on institutional expectations and external examiner's perception of, and effectiveness in, the role. Requires examiners to hold knowledge of agreed disciplinary standards.	E
Markers having experience as external examiners or as assessors at other institutions	Can provide a valuable contribution to internal discussions if the experience has helped align their academic standards with a range of disciplinary colleagues across the sector.	E
Markers being members of a learned society or professional body	Such membership should bring markers into contact with others outside their institution. However, discussions may not often be focused on academic standards.	E
Markers being familiar with national reference points	Although it is unlikely to be appropriate to use national reference points directly as part of the assessment process, if assessors are familiar with them they will provide background that can inform discussion.	E

	<p>Like all attempts to make academic standards explicit they cannot ensure consistency. The use of these reference points in the design of a course is often deemed sufficient by assessors and external examiners to ensure that the academic standards required will form the basis of judgement of student work. The opportunity for 'mistranslation' of standards from the design of course process to the assessment of individual assignments is high.</p>	
--	---	--

Summary log

Session 5 summary log: Using tools and processes

Use Activity 5 (including the 'Table of methods' provided) and presentation in Session 5 to fill in the following.

Tool and process	More effective in assuring academic standards	Less effective in assuring academic standards
Internal processes		
External processes		

Session 6: Professional practice in the external examiner role and decision making

The practice of being an external examiner is unlikely to be as straightforward as may have been portrayed so far in this course. It is a complex, multifaceted role that is made messier by academic, organisational and social pressures. We have tried to unpick facets of the role to explore them, but now we would like to consider how these different aspects come together in the reality of the role, how trade-offs might need to be made, and also to be clear about what is non-negotiable.

Task 6: Dilemmas in professional practice

Purpose

To explore how an understanding of the role and academic standards might play out in reality by considering potentially difficult situations that as an external examiner you may face.

What you should get out of it

- An opportunity to consider and discuss with other external examiners how they would deal with dilemmas in carrying out the role.
- Access to expert commentary applying theoretical knowledge to identified responses to dilemmas.

Relevance to external examiners

External examiners need to be able make the best professional judgements possible 'on the ground'. A consideration of the issues and knowledge of other external examiners' positions helps to develop consistent professional judgement.

Description of task

Your group has been provided with a set of cards. These consist of:

- dilemma cards, where one side includes a description of a dilemma and the other four possible options for action;
- voting cards (A–D).

1. Divide the dilemma cards between you.
2. The first person reads out the dilemma on their card, and may need to do this a couple of times to ensure everyone understands.
3. The same person then reads out the options on the reverse (A–D).
4. Group members silently consider the dilemma and choose one of the options. (Do not share your views at this point.)
5. All group members simultaneously lay down their chosen voting cards. If there are differences, discuss them as a group. If there are no differences go to Step 7.
6. Take a second vote to see if you are closer to a consensus.
7. Repeat the process with the other dilemmas.

Total time (30 mins)

8. Share the main areas that prompted discussion in plenary (10 mins).

There are some additional dilemmas provided below that you can explore in your own time, relating to other situations met by external examiners.

Professional practice

- **Be clear about the most important role of an external examiner** – threshold standards, rigorous process and comparability of standards
- It is **not always possible to please everyone**. Opting to be a critical friend often feels more rewarding and easier, but is not more important than overseeing academic standards
- To be confident about overseeing standards you need to be confident that the **academic standards you use are representative of the subject community** to which you belong

Professional practice

- As **the only check on outputs** the job of external examiner holds a lot of responsibility for stakeholders and UK standards
- It is important for examiners to be **well informed about the evidence around effective assessment** rather than just relying on their own experience of assessment
- Given the **complexity of clarifying academic standards** it is important for external examiners to be able to evaluate the relevance of factors in that complexity

Spend a few minutes now completing your summary log for session 6

Summary log

Session 6 summary log: Professional external examiner practice
External examining is not an easy job. Please identify potential tensions in external examiner decision making
Drawing on the main 'pillars' of professional practice, in what ways can these tensions can be dealt with?

Task 6: Dilemmas with comments

We have reproduced the dilemmas from Task 6 here with comments on each option. You may find it helpful to compare your viewpoints with the comments.

Dilemma – Collaborative partnership

The undergraduate course for which you are examiner is run on several sites, including at one of the institution's collaborative partners, an FE college. You see a selection of student work from these sites and think that the work of the students ($n=50$) studying at the FE college has been marked more generously than other students' work. The module leader does not disagree and argues that because the students are starting from a disadvantaged position they need some time to catch up. It is important not to demotivate them by low marks. The module leader is of the view that as long as they are being judged by the same standards as the other students by the time they reach Level 6 (levels 9-10 in Scotland) there is no need to worry.

The pass mark for the module is 40%.

As is usual with undergraduate programmes it is possible for students to be awarded a certificate or diploma in HE if they leave before completion. This has never happened and is not anticipated for the future. The exam board is due to take place in three hours' time.

Optional responses

Comments on the options

A	All the FE work must be re-marked to ensure fairness.	This will result in equitable treatment but it is a tough option to impose. It is unlikely that re-marking will be possible before the exam board, which means the students will not get their results when they expect (which could be very demotivating) and the exam board process will be disrupted, which will probably not make you very popular with the course team and administrators.
B	The work awarded marks below 50% must be re-marked to ensure threshold standards are met to safeguard the integrity of potential awards of a Certificate in Higher Education.	This will safeguard threshold standards but lead to a disjuncture in standards applied across the grade bands. Marks over 50% will remain inflated, which is likely to mislead students about the quality of their work and not serve them well.

C	You are prepared to allow this for the student work studying at the FE college, but in the interest of fairness to other students you make clear that at Level 6 (levels 9-10 in Scotland) consistent standards must apply.	This response seems to depend on beliefs about how disadvantaged students are best served and motivated. Higher marks may be motivating, but students need to understand the quality of work expected and be supported to achieve that quality. By not marking rigorously these students are being misled about the quality of work required to do well. It would be difficult to tell the students they had been given higher marks than their work deserved.
D	This once you are prepared to allow the marks to stand but make clear it cannot happen again.	This is the easiest option and will allow the exam board to go ahead without disruption, but allowing lower standards creates many problems, such as students being misled about standards, and precedents about acceptable standards set. Also importantly you have a responsibility to advise the institution about whether it is meeting threshold academic standards.

Dilemma – Multiple learning outcomes

There is only one assignment for the whole module. The task is quite complex but well designed, and far more innovative than many others. It challenges the students who clearly are engaged by the task and learning a lot. It incorporates a variety of learning outcomes and the criteria being used to frame the assessment are very varied in order to cover all the module's learning outcomes. You can see from the feedback sheets, which include ticks on a marking rubric grid, that some students failed to achieve a pass on some criteria but did well on others. All students have gained a pass.

Optional responses		Comments on the options
A	Request that work not reaching at least a threshold pass on all criteria should be referred and offered a resit.	If this is the only place where particular learning outcome(s) are assessed and they have not been achieved then the learning outcomes have not been met so referral is appropriate. However, the criteria may cover elements of learning outcomes (e.g. referencing within academic argument, evaluation, synthesis etc.) that students will be required to demonstrate elsewhere in the programme, and so could be overlooked here. The problem with this approach is that there is no guarantee that this will happen, especially if compensation between diverse criteria is common practice.
B	Averaging the achievement across several criteria is common practice in higher education courses. You do not regard this as an issue.	It is true that this is common practice across HE and the course team may consider that you are being 'difficult' to make an issue of this. However when such practices are common across a programmes and there is insufficient checking that programme learning outcomes have been achieved standards may not be being maintained or programme outcomes achieved. External examiners should be confident that the standards are safeguarded.
C	Accept the marks, but seek reassurance that the learning outcomes (e.g. ability to reference correctly) have or will be achieved in other modules.	This is a comfortable option but not one that leads to clear reassurance that standards will be maintained.
D	Accept the marks for this year but request the design of the assignment and criteria are changed next year to ensure this does not happen again.	This solution remains focused on the module, and given that the current assessment task seems to be working well, a request for a full redesign may have unintended consequences. It may be better to suggest a whole-programme view focused on the programme learning outcomes.

Dilemma – Bias

A piece of coursework for a master's module has a strict word limit and is about the contribution of writers in the field to its development, current thinking and debates. Criteria for assessment are quality of research (including referencing), demonstration of contribution of selected writers and identification of current debates. Most students relied on the dominant well-known names in the field but some have looked further afield and identified contributors many of whom have not received much credit for their work and contribution. All but one of the students who have done this are female and most additional contributors to the field that have been cited are female. You have looked at the quality of the sample of work against the marking rubric and conclude that there may not be parity of reward between work of males and females. The female-oriented work seems to have been downgraded and the feedback points to the omission of well-cited male writers as the reason. In the rest of the coursework sample there are examples of sound but unremarkable work that covers well-known names and provides explanation of their contributions. These have been awarded high marks where the student is male and very slightly lower when female. The marks for the exam (marked anonymously) do not show much difference between coursework and exam marks for individual students.

Optional responses

Comments on the options

A	Raise your concerns with the programme leader and recommend an independent re-mark.	In choosing this option you need to be clear about the extent of the evidence that you have. What counts as evidence and how sure are you of bias? The exam marks may not be reliable in this case as females traditionally do less well than males in exam conditions (but better in coursework conditions). Talking to the course leader may provide useful background information and identify policies or process that could be used to deal with this internally. Ultimately it may be appropriate for you to talk to the marker to raise your concerns.
B	Accept that it is possible to justify the marking within the criteria so take no action.	Task instructions and guidance as well as criteria are almost always open to interpretation. Such interpretation needs to be reasonable and aligned with any briefings given to the students. Within that interpretation it is important that bias is avoided and the criteria have been used consistently. As external examiner you need to be assured that this is the case.

C	Talk to the module leader with a view to trying to establish if the bias is intentional or unconscious and change their practice in future.	This will need to be done very delicately. Perhaps by talking about the task and the markers expectations of the approach students would take to the work. The module leader may not be aware that bias exists and be able to justify the marking. If bias does exist it may be unconscious and the marker may be dismayed when it is pointed out.
D	You are sure that this cannot be bias as a male student has been treated in the same way as the females. Take no action.	It is possible that this shows that the marker has not downgraded female students but the question remains about whether the bias exists through downgrading work that highlights female contributions to the field. It would be important to consider how this work should be judged using the criteria. (see 2 above).

Dilemma – Context and quality

The assignment task requires investigation of a topic chosen from a list and production of a poster setting out the key points from the investigation, the linkages between them and one key message. The criteria being used to judge this work are research quality, communication, knowledge and understanding, referencing. The learning outcomes set out ability to analyse real problems, select and use theory to construct an argument, critically evaluate recent research in the subject.

The sample of work provided covers a full range of marks and there is evidence that the criteria have been used, but in your view the work overall deserves higher marks because of the sophistication of the poster designs that demonstrate an understanding and synthesis of theory and application to practice. However you are quite surprised by the quality of the work as similar work, (e.g. a PowerPoint presentation on a set topic) produced by this cohort on other modules, both this and last semester, was not high quality.

You only have the module description and an extract from the module handbook setting out the assignment instructions and criteria. You have no broader information about the module such as its structure, specific content, learning activities and support for completing the assessed work. You assume the students had considerable support in completing this assignment.

Optional responses		Comments on the options
A	You accept the local judgements that are based on the criteria and local knowledge about the context.	As external examiner you will never know the precise interpretation of the criteria or the details of the support the students have been given. However if you do not explore how judgements are made and what is taken into account it is questionable whether you can take a view on the academic standards in use.
B	You recommend raising the marks. The students have produced excellent work and provided good evidence of achieving the learning outcomes.	This option means you are relying entirely on your own view of academic standards. If this view is grounded in the common view of academic standards within your subject community it would be reasonable to make the request. However it will always be useful to discuss the local judgements and reasons for them in order to support calibration of standards used in the programme.
C	You discuss with the programme leader the discrepancy between modules relating to the quality of work and marks awarded.	By exploring differences it may reveal factors about this module or others that explain the differences. In turn this may raise issues about standards and/or appropriate support that needs to be addressed, or identify good practice that should be shared with other modules. After the discussion you may choose to recommend changes in marks.
D	You accept the marks this time, but insist on being provided with much more detailed information about the module in the future so that the level of support provided for students is clear.	Receiving more detailed information about the module may give some indication about the extent to which the assessment has been scaffolded and students given help and support in completing it, but it is unlikely to reveal the whole story. It would probably be worth raising the issue now to ensure the same issue does not arise again next time the module runs. From the point of view of the students, if the work they have produced is good it may be unfair and they may be disadvantaged if the marks are not changed.

Dilemma – Level and expectation

In an undergraduate programme you oversee two modules which use student presentations as part of the summative task. One is in the first year and one in the final year. You notice that the criteria and level descriptors used to mark the presentations are the same. The feedback provided suggests that students' presentation skills are expected to be a little more developed in the final year but the judgements are still largely based on quite 'mechanistic' factors, such as eye contact and voice projection. As a consequence, you think the standard required for the final year module is too low and does not align with the subject benchmark statement. Consequently the marks awarded for presentation skills seem to be too high.

However you are also concerned that the students may not have not understood that an improved quality of work has been expected of them, because the rubric is the same. And you are unaware of any other opportunities that the students have to give presentations between these two instances so you cannot expect the student presentation skills to have greatly improved.

Optional responses

Comments on the options

A	This is too minor an issue to raise. The marks for presentation are unlikely to make a material difference to final honours classifications.	There is an issue that the standards do not seem to be aligned with the national standard for this subject. So while it may not be critical to the students' final awards it is something that should be raised with the course team for them to review.
B	You request that the marks are adjusted (downwards) so that the standards reflect expectations set out in the benchmark statement.	Before making this request it might be wise to ask what sort of guidance students were given to help them understand the quality of presentation expected. It would also be helpful to know what other formative opportunities to make presentations they have been given. If the students were given appropriate guidance and have had practice in giving presentations this option is appropriate. If they only had the rubric to rely on this decision may be considered harsh not only by the students, especially if it affected their final award, but also by the course team.
C	You accept the position for this year as it would be unfair on students to 'change' expectations.	This decision is not one that sits well with the responsibility of the external examiner to uphold national threshold standards. See comments (2) above, it maybe that your assumptions are incorrect and that students were fully aware of expectations.
D	You will raise it with the module leader with a view to establishing the tacit expectations about presentation standards that have influenced the marking to clarify if this is	The fact that the rubric is the same does not mean that the academic standards applied are the same. However, the marks would suggest they are not high enough. It would be useful to find out if the marker(s) is aware of the standards set out in the benchmark

	likely to be an issue in the future.	statement and if so how they have interpreted them. It may be useful to suggest that this issue is considered by the whole course team.
--	--------------------------------------	---

Additional set of dilemmas

These additional dilemmas provide further examples of the issues and decisions that external examiners face in their role. Each one provides you with a description of the dilemma followed by several options. You may like to consider which option you would select and then read the commentary.

Dilemma – Standards in innovative assessment

An assignment task for students who have returned from a placement year involves them in working together to plan, organise and run an engaging exhibition/fair for placement employers and students currently seeking a placement. The instructions give considerable freedom to the students in the theme, timing, and format of the exhibition and how they organise themselves to deliver it. It is stipulated that in the final part of the assessment there must be a peer assessment of each student's contribution to the process. The first task they must complete is to make a proposal, and get staff agreement for the criteria and level descriptors for judging the students' work including those for the final task, a peer assessment.

You have been provided with a copy of the criteria and level descriptors and they are similar to other practice-based final year modules. You also have documentation such as detailed plans, minutes of meetings and an evaluation report of the event, plus samples of commentary on individual student's contributions to the process used in peer evaluation. These are the evidence base used to grade work. The marks overall do not seem to be particularly high, except for a couple of students, and you have a spreadsheet giving the details of both tutor marks and peer assessment marks. There seems to be consistency between the two.

While you admire the creativity and innovation in this approach to assessment you want to think very carefully about it.

Optional responses	Comments on the options
<p>A You cannot see what this assessment has got to do with academic learning. You will propose that next time there are many more constraints on the students, including a requirement to write a short piece making links between their activities and relevant theory, and that the criteria are set by staff and issued to the students. Only then do you think you can be confident about whether threshold standards have been met.</p>	<p>To stipulate such constraints is likely to alter the nature of the task and learning that can be derived from it. It is important to look at the learning outcomes of the module to see if the assessment 'fits'. You also need to look at the agreed criteria and consider whether they are unsuitable or have advantaged the students. It may be more important that the criteria focus on problem solving using such academic skills as analysis and evaluation rather than requiring the formal inclusion of theoretical content. Rather than being concerned about your comfort you need to be assured that there is value in this assignment and that the academic skills being used have been judged appropriately.</p>
<p>B It seems that the design of the assessment could easily be open to abuse, with students proposing criteria and using peer assessment to inflate their marks. You will raise your concerns with the exam board.</p>	<p>These assumptions do not appear to be borne out by the evidence. It is clear that the criteria must be agreed with staff, not that staff just accept the student proposal. You, as external examiner, can also make a judgement about whether the criteria and level descriptors seem appropriate and whether they have been used appropriately in the marking. The detailed mark information does not provide evidence of students inflating their marks. It is not clear how a couple of students could 'game' the system. To raise this formally at the exam board would not be appropriate before at least discussing your concerns with the module leader who may be able to allay them further.</p>
<p>C As you are satisfied that the criteria are aligned with the learning outcomes and are comparable to other modules at that level, your only concern is the possibility of subjective judgements as the work cannot be anonymised. You want to speak to the markers about how they guarded against this possibility. You also have a recommendation about improving the peer assessment by incorporating an element of feedback, but otherwise you are happy with the assessment.</p>	<p>Being satisfied that there is constructive alignment and comparability is a good place to start for such an authentic assessment. You are also right to think about fairness and bias. There is a number of ways in which the tutors could try to reduce the possibility of bias, such as a system of double marking (with a range of pairs if there are several tutors) and awareness and discussion about the issue. It is good that you can suggest an enhancement in keeping with the task design.</p>
<p>D The couple of students who have particularly high marks took the</p>	<p>You may want to talk to the module leader about how students come to play those</p>

<p>lead organising roles for the whole exhibition. You are concerned that such roles may unfairly enable them to get much higher marks. You want to know what safeguards exist to prevent unfair advantage.</p>	<p>roles (self-appointed, elected by cohort, appointed by module leader, etc.) to ensure equity of opportunity. However if the peer assessment gives these student consistently high marks and there is evidence in the documentation of a professional approach it would suggest that these 'leaders' did a good job and deserved high marks. It seems unlikely that they were able to 'bully' the rest of the cohort, especially if peer review was anonymous. Regular meetings with tutors to report on their work could also provide some monitoring</p>
---	--

Dilemma – Converting grades for exchange students

A number of students completed modules overseas at a partnership institution through the Erasmus Programme. The marks awarded by the partner institution have been confirmed through their procedures, but you have responsibility for the quality and standards of the whole programme. The marks awarded locally have been converted using a University conversion table before being entered in the student record system for the exam board. You notice that the converted marks awarded to students at the partnership institution are, on average, quite a bit higher than those awarded at home. The modules taken by students who study overseas are at the same level and the module descriptions were examined when the partnership was established. Staff leading the exchange programme in the home institution have been to the partner institution to work alongside the module teams when the partnership was first agreed.

Optional responses

Comments on the options

A	<p>You are aware that usually the students who choose to study overseas are more motivated to do well, so it is not surprising they all got good grades. So you do not see any issues here.</p>	<p>True, often these students are keen to do well, but it does not mean they are consistently more able or motivated than the rest of the cohort. Being taught and assessed in a different language, together with the cultural challenges and new learning methods students will have encountered may mean it is less likely that all will perform very well. It is probably worth asking some questions to find out more about this.</p>
B	<p>You think it should be acknowledged that the students are working in a second language and therefore reasonable to take account of this in awarding marks. So if there has been compensation you think it is appropriate.</p>	<p>This raises a big problem of consistency and fairness. If different academic standards relating to the subject of study are being applied to different students this is likely to be deemed unfair. A comparative situation is international students working in a second language in the UK. Would you or should you accept lower academic</p>

		standards in their work?
C	This situation concerns you enough to get more information. Given your main concern as external examiner is academic standards you want to check whether students' work has been marked too generously at the partnership institution. You ask for a sample of work from the partner institution to be made available to you.	You are being asked to confirm marks for work that you have not reviewed. You are in a difficult position, having to agree marks without having seen a full sample of work. You cannot know if the marking is generous unless you see a sample. However that sample will be in another language, but translation is possible. Bear in mind that it would also be difficult to question the academic standards of another institution. So this may not be the place to start as a partnership agreement exists to cover the exchange programme. Another route may be to discuss the processes that exist to guard against this and seek reassurance that they are working well.
D	This situation concerns you enough to get more information. Given there has been some attempt to align standards between the institutions in the partnership you think the problem is likely to arise from the conversion table, i.e. it provides generous conversions. At your own institution marks are not transferred but students gain credit and the basis for their final classification is altered i.e. they effectively get an average of their home marks for all modules taken abroad. You think this has a sound basis to it and you are not at all sure about the conversion table idea. Your recommendation is to review the conversion table itself and perhaps the system for bringing in marks from elsewhere.	Given that all the overseas marks are generous, it seems likely that the problem lies with the conversion table rather than the marking at the placement institution. It would be prudent to ask questions about marks resulting from the use of the conversion table in past to see if this is a consistent problem. Suggest that if the table does consistently provide generous conversions then the institution may want to review the table, or even the way overseas marks are incorporated by the institution.

Dilemma – Credit entry standards

Many of the modules for the programme are taken by students overseas at collaborative partner institutions. Students who achieve good grades are offered a place to complete the programme in the UK and the number of students that achieve this helps with recruitment at the collaborative partner institution. There are locally based external examiners that look at the marking and quality of student work but you have responsibility for the marks for the whole module, which appear on the mark sheet. You notice that the marks awarded to students at the partnership institutions are on average higher than those awarded at home. The assessment task, marking criteria and level descriptors are the same for all the institutions. Staff leading the modules in the overseas institutions have been

to the home institution to work alongside the module team here.

Optional responses		Comments on the options
A	You realise that the students are keen to do well so you do not think it is surprising or noteworthy they have got good grades.	True, students are keen to do well but it doesn't mean they all will do better and lead to a higher average. It seems likely that there are other factors affecting the marks
B	You suspect that the locally based external examiners may not be attuned to the academic standards required by this programme or feel pressure from the partner institution to not question generous marking. You decide to ask the course leader to investigate recent trends in cohort mark profiles as well as let you know the selection and induction procedures for external examiners. To address the issue in the short term you ask if there are samples of work you can see.	You cannot make that assumption that local examiners are not using appropriate academic standards so the need for more information is important. If, like the staff they are very familiar with the modules and assessment and have been well briefed it does not follow that they would be generous. However you are in a difficult position having to agree marks without having seen a full sample of work. Seeing a sample of work may help to reassure you but you might also suggest that some work from the home institution is sent to the externals to act as benchmarks for them.
C	You are wondering whether the problem lies with hard marking in the home institution. You have looked at a sample of the work and felt that the demands being made in relation to exact referencing (e.g. losing marks for a missing full stop) maybe too harsh. You will raise the anomalies you have noticed in the pattern of the marks with the module leader to get the insider perspective.	It would sensible to discuss the issue as a diagnosis and remedy may already have been worked out. In discussion you can evaluate the various positions to inform your view.
D	The module teaching methods and assessment being used at the partnership institution reflect UK education methods and are likely to be very different from those students have experienced before joining the programme. They will work in their second language so it seems reasonable not to be quite so demanding of them especially in early stages of the programme. So generous marking is appropriate.	This raises a big problem of consistency and fairness. If different standards are expected from different students this is likely to be deemed unfair. However some might argue that in the name of fairness there should be some dispensation to account for language especially where it is not part of the assessment criteria.

Dilemma – Bias

A module within an online course has two parts to the assessment. One part requires students to write a reflective statement about their learning process within the module. Some of the statements are fairly descriptive about the process and they identify ways in which it helped them learn. Other students have taken a more critical stance and sought to evaluate the contribution of the module to their learning. Feedback on the reflective work does not make it very clear what constitutes quality work. However, where critiques are positive the work has been awarded high marks and it seems that where students have been critical of the learning approach used by the module the marks in both assessments are quite low.

Optional responses

Comments on the options

A	Online courses lack the personal touch of a face-to-face course where the student can easily identify the tutor and develop a trusting relationship. This is not the fault of the tutors, so where students have not taken this into account in their reflections it is likely to be of less high quality. You do not see any need to follow this up.	Online learning uses different learning techniques from traditional face-to-face approaches, but this does not make it less valid or effective. In fact some students prefer online approaches. Rapport and trust with tutors is built in different ways. You should not assume that new approaches to learning are bound to be less satisfactory and students' dissatisfaction is normal and acceptable. For this assessment, it is the quality of reflection rather than the students view of the course that should determine the grade.
B	Reflection is often poorly understood. It is not just a description or expression of feelings but, done well, requires critique of issues and evaluation of their impact. There is, of course, no right answer and quality will depend on the quality of reasoning and evaluation. You review the assignment guidance including learning outcomes and criteria, which is sound. You think this is a valuable assignment and would like to see it retained so you plan to talk about what you look for in reflective pieces with the module leader.	This could be a tricky conversation. If the module leader is sensitive to criticism from students it may well be the case with you too. You could use examples in the student work to talk about how student reflections on their work could be used to improve the module 'even further'.
C	Each student responds to learning approaches differently, therefore this type of reflection is bound to engender some positive and some negative evaluations. However whether the evaluations are positive or negative the students should	This might be a good approach instead of making accusations about bias. The discussion would inevitably be about the academic standards being used to judge the work which in turn would uncover inconsistency.

B	Given the university regulations you cannot see how this module leader can cover all the outcomes in one exam. You sympathise with the impossibility of some regulations and accept the status quo.	Many universities have restrictive rules around assessment that constrain staff from designing the most relevant assignments. As an individual external examiner it is difficult to bring about changes, although you could raise it in your report.
C	You are aware that focusing all the attention on one module does not take the students' holistic experience of the programme into account. You seek reassurance that the formative work, particularly on presentation, will be built on and assessed in subsequent compulsory modules. If there is an assessment strategy that shows a well-planned route for student learning and assessment across their programme you will feel less concerned about this one module.	Recognising the need for a coordinated and planned approach across a programme is good practice. It helps the students build on their learning and feedback, prepares them for summative assessment and helps to motivate students to engage in formative assessment. However it only addresses the issues of the 'home' programme. This module serves multiple programmes, which are likely to rely on it to have summatively assessed all the learning outcomes. If the outcome cannot be covered it would be a good idea to have a discussion with the module leader and programme leader about changing the learning outcomes (e.g. shifting the assessment of skills to the next module) to tidy up the anomaly.
D	Trying to hold this module together and retain a committed team of staff is a challenge. You are sympathetic to this and fear requesting too much change might make it impossible. You will point out the issues but then make it clear that you will give the module team the time (maybe over several runs of the module) they need to work out a solution themselves.	It is likely that the involvement of an enthusiastic team contributes greatly to the success of the formative assessment. It would not be appropriate as an external examiner to increase pressure on staff. The team could be encouraged to think creatively about the problem. You could offer some suggestions drawing on your experience elsewhere.

Dilemma – Double marking

In a sample of master's dissertations you have access to the first and second markers' original marks and comments, as well as the final agreed mark and feedback sheet. You begin to realise that where there is a significant difference in marks between the first and second marker the final mark is an average of the two. All dissertation marking is carried out by a small number of markers (all research active staff) in the very busy department

Optional responses

Comments on the options

A	You find this unacceptable. Giving the student a 'compromise'	The use of widely varying academic standards among markers is well established and
----------	---	--

	mark may advantage some students while disadvantaging others. The mark should be determined according to the quality of the work. You raise it with the exam board.	dissertations, being complex pieces of work, are likely to be subject to varied judgements. These differences must be resolved through consideration of the qualities of the work and not a mere arithmetic calculation. 'Splitting the difference' in marks means the student is awarded a mark that is not a result of judgement. Dialogue between the markers is required to reveal what is valued and why, in order to reach a closer judgement.
B	It seems inevitable that when a small number of staff must mark all the dissertations they will not have time to have detailed discussions. You know that marking is an inexact science and the average mark is as likely to be as accurate as either of the two markers so you do not query the marks and approve the final marks.	Staff are facing increasing levels of demand on their time, but making the best judgements on student work, especially on a major piece of work such as a dissertation, is important. It is correct that marking is an inexact science, but that does not mean that efforts to try to ameliorate difference should be abandoned. 'Splitting the difference' in marks means the student is awarded a mark that is not a result of judgement. Staff should certainly find the time to consider the dissertations where there is a large difference between first and second marker. The more they discuss differences the fewer large variations should occur as they will gradually develop a more common understanding of what is valued.
C	You know that in a modular programme the students' marks are subject to statistical manipulation in order to reach the classification. This master's programme offers only pass or distinction in the final award. It seems to you that the inaccuracy in marks awarded for the dissertation created by this approach is not significant enough to affect the final outcomes so you decide not to make an issue of it.	Individual decisions form the basis of final awards and therefore should be treated with respect. Multiple small decisions could make a big difference. So, effort should be made to come to a final judgement on the quality of a dissertation. Also, a student is likely to have invested a large amount of effort and time in their dissertation and rightly expects it to be given due consideration and fair judgement.
D	Using your experience from elsewhere you suggest that a section be included on the forms that requires an explanation of how the final mark was agreed, in the hope that your inference is wrong or that it will stop the practice of averaging.	This may act as a failsafe in the future. It would be wise to explain why it is needed and the benefits it would bring. You should also check that the students affected in the current cohort are not disadvantaged by the averaging.

Dilemma – Moderation method

You have some concern about inconsistency in the marks awarded on the sample of work you have seen. The sample has clear evidence of having been through a moderation process. You ask about the moderation process and the module leader explains that each marker ($n=8$, six of whom are experienced on this module) is asked to provide information about the average and range of marks for their batch of marking ($n=25$). The module leader compares the averages and ranges and moderates the batches that are out of line either by raising or lowering all marks in a particular batch. The module leader suggests that, because markers have a random selection of work to mark rather than the work of their own seminar groups, this system of moderation is acceptable especially given the size of the module.

Optional responses

Comments on the options

A	Accept that it is likely that each batch will include a typical set of student work so accept the moderation is appropriate.	It is a dubious assumption that all the batches will be similar. It is easily possible that one batch will include a large proportion of very good work or very poor work, even if randomly assigned. Using this method of moderation is at odds with the criterion-reference assessment used for the original judgements. Also changing the marks that have been arrived at using criteria is likely to make them out of synch with the feedback provided. This method should not be accepted.
B	You are aware that marking will always be somewhat inconsistent but wonder whether this has been exacerbated by the two new markers on the team. Ask about, and discuss, the average and range of marks for their batches of work. Marks could be altered as a result.	Basing your discussion around the profile of marks suggests you think this method of moderation is appropriate and aligns with criterion-referenced assessment used by first markers. Similarly, by focusing on the two new markers you are assuming that there are unlikely to be other causes of inconsistency to address e.g. there is no pre-marking preparation for the team to consider key aspects of the assessment or opportunity to compare their standards. Just because markers have worked on the module before does not mean they will take the same approach to marking, especially if their previous marking has been moderated on the basis of mark profiles.
C	Acknowledge that moderation can be time consuming and comparing averages might be a starting point. But express concern about the incompatibility of this as a moderation approach with criterion-referenced assessment. Suggest that moderation should also involve comparing the quality	Accepting that the average mark of each batch provides useful evidence for comparison is dangerous given it is based on a dubious assumption that all the batches will be similar. You are right that there is an inconsistency between this method of moderation and criterion-referenced assessment. There is a need to discuss the purpose of moderation and understandings of criterion-referenced assessment as well as

	of the work. Recommend moderation of the marks using a different method.	suggesting more appropriate approaches to moderation. It may be helpful to discuss the position of the new markers who need to assimilate any shared understanding of academic standards that exist between the experienced team members.
D	Request that the non-moderated marks be reinstated because these are more likely to reflect the quality of the work judged using the criteria.	There is logic in this. There is a team of markers many of whom have marked work from this module before so may have developed a common view of academic standards. However there are two new team members so the preparation for marking is an important factor. Preparation for marking through sharing of standards through the common marking of an exemplar or a small sample of work, or if markers had marked at a marking bee (all together), would have helped to reinforce experienced markers and developed new markers' views, increasing the likelihood of consistent marking. Without such processes there are bound to be a few anomalies in the first marking.

Dilemma – Power and marking

You have looked at a sample of work from a module that has three markers. You agree with most of the marks. Two thirds of the scripts have been marked scrupulously using all the criteria and have feedback addressing the criteria. The other marker has given quite detailed feedback but rarely covering all the criteria, and the justification of the mark is not always clear. You have established that the team of markers consists of two doctoral students and a professor. It is the professor's marking that is out of line with the other two markers. The method used for moderating marks is unclear.

Optional responses

Comments on the options

A	You recognise that there are different approaches to marking, including holistic professional judgement. The students are getting some useful, albeit different styles of feedback regardless of the marker so you don't raise any concerns.	If you don't think the marking reflects the use of appropriate or consistent academic standards then as external examiner you should be concerned. Experienced examiners often draw on internalised standards that can be out of line with other markers, and they may pay less attention to the published criteria that the students have been given. It might be helpful to ask about the moderation process as a way of pointing out inconsistencies in the judgements and feedback.
B	You accept that the professor's knowledge of the subject and experience as an assessor means the marks are likely to be	Knowledge of the subject or experience as an assessor does not always lead to use of appropriate and consistent academic standards in judging student work. Experienced markers often look for very particular qualities in work that is not reflected in the

	sound and hope that the doctoral students will learn a lot from working in this team.	criteria. A marking team should work together to try to come close to a common understanding of how to judge student work. This does not appear to have happened. It may be difficult for doctoral students to question a professor, and as module leader the professor oversees the assessment. As external examiner you can use your role to try and improve practice.
C	You need to discuss the differences in approach and what feedback the students receive with the module leader (the professor) with a view to ensuring consistency of academic judgements.	It may be difficult for doctoral students to question a professor and, as module leader, the professor oversees the assessment. As external examiner you can use your role to try and improve practice.
D	Rather than directly address the differences you have seen you decide to act on this by asking more generally about the approach taken to assessment on the module, e.g. how the team of markers is briefed, how moderation is carried out, and what attempts are made to ensure consistency of approach.	This may make the conversation easier, but it is important to recommend changes to ensure the same thing does not happen again next year. Asking about moderation would be a useful route into discussing differences in approach and academic standards. It could also open up conversation about guiding and working with newer staff.

Dilemma – Different reasons for marks

A master's level synoptic module has been blind double marked. You have been provided with all the first and second marker marks and the final agreed marks. There are several instances where first and second marker have given the same mark. You have access to all the scripts so you look at some of these perfect matches as you believe that it will give an insight into the academic standards that are commonly used in this module. What you frequently find is that although the overall mark is the same the judgements made against each criterion are very different and the final mark just happens to be the same. This also leads to the two markers giving different and sometimes conflicting feedback.

Optional responses

Comments on the options

A	You know that professional judgement is usually made holistically with a retro fitting to the marking criteria. Consequently you are not concerned about the differences in	It is true that much judgement of complex work in HE is holistic. That does not mean that criteria are irrelevant. Ideally the holistic judgements are taking into account the same sort of qualities, and the judgements of the criteria should be reasonably similar.
----------	---	---

	judgements against the criteria and have confidence in the consistency of judgement.	Are these final marks the same more by luck than judgement? Discussion between markers should reveal these differences and influence their judgements. Occasional meetings of all markers can 'recalibrate' what is important in judgements.
B	Criteria are so imprecise and overlap with one another that it is impossible for markers to use them consistently. You are happy to ignore these differences.	This is a controversial position to take given the emphasis given to criteria to structure judgements, to guide students and to frame feedback to students. However, it is true that criteria are imprecise and overlap, which is why they should be only the starting point for sharing understandings of academic standards with colleagues and students.
C	You can see that there is a technical problem here but decide that dissertation marking is such a 'can of worms' that you will focus on where there are major discrepancies rather than where there is agreement.	This is quite a nuanced issue and it is quite possible that there are more fundamental problems that need to be addressed in relation to academic standards. In looking at more major discrepancies in judgements it would be useful to consider whether improvements to processes might help to develop shared standards, e.g. requiring a joint feedback to be produced by the marking pairs.
D	You are concerned about the problems this will cause for feedback to the students. At the moment they stand to get two versions of feedback. You recommend that there is one version of feedback produced that is agreed between the two markers.	This recommendation may prompt some closer examination by the markers of the rationales for their judgements and so help align their judgements.

Dilemma – Unexpected student work

The module leader has provided details of the assignment task, guidance given to the student and a marking rubric as well as the model answer used in the marking process. In the sample of work, you notice one piece of work that differs markedly from the others and looks to be an imaginative and thoughtful response to the task using a lot of visual images instead of relying exclusively on text. This has been given a fail grade.

Optional responses

Comments on the options

A	Looking at the student information you can see that the work conforms, albeit differently, to the brief. So you question the fail grade.	While an assignment brief cannot be expected to identify all possible interpretations and specify those that are excluded, it does need to make as clear as possible the parameters of the work. You need to ensure fairness for this student and all other students. It
----------	--	--

		appears that this work does comply with the parameters set, but it would be wise to discuss the issue with the module leader to see if there is further relevant information and why the fail was deemed appropriate.
B	You can see that the use of visual images in the work requires interpretation and some visual literacy on the part of the marker. You suspect that the markers are unsure about the quality of the work as they operate in a largely text-based subject, and have included the work in the sample for that reason. You decide it is necessary to discuss the quality of the work to ensure that the student is not treated unfairly.	A marker's personal preference for a particular style(s) can mean that they are 'biased' against work that is outside that style. If it is a personal bias, this may be unfair on the student whose perspective is different or who has sought to develop a more original response to a task. A discussion about the work may help to reveal its qualities in relation to the rubric and may lead to a revision of the mark.
C	You think this fail has been awarded because the marking relied entirely on the model answer as the only good answer. You decide to discuss the dangers of using model answers for complex tasks.	Model answers are built on an assumption that there is only one good response to the assignment and judgements are made according to how close the work is to this 'ideal'. Where the task is sufficiently complex it is likely that there are several ways to respond well to the task. Work that has taken a different approach is in danger of being penalised unfairly. A collection of exemplars could illustrate a range of possible answers.
D	The task specifies that the work should be prepared for a specific audience. You assume that this audience is likely to be well educated, but that they would normally expect information and ideas to be text based. You think the student should have taken this into account and so you agree with the fail.	You need to recognise the assumption you are making. Check what the assignment brief says about addressing this particular audience. If this student is to be penalised for a less than traditional approach it is important to check whether the work of other students has been assessed as rigorously in respect of addressing the audience.

Dilemma – Maverick student work

The assignment brief consists of an essay title, marking criteria (knowledge, analysis, synthesis and evaluation) and the university's generic level descriptors. The essay title does not specify a particular topic although all but one essay in the sample you have is on a particular topic. The one exception has been given a fail grade. It is an exceptionally good essay on a topic within the module. In the module report the module leader has explained that the student was absent from the class when the essay was set and the topic made clear.

Optional responses		Comments on the options
A	You can see that this essay does address the essay title, and the criteria and level descriptors do not stipulate the topic. You are torn between failing the work and making an exception for such a good essay. You ask about the student's attendance record and mark profile so you can decide what is fair.	It is not clear how attendance or mark profile will help you. If the student's attendance is poor should they be penalised for that (especially if there may be good reason for it)? If the student's marks are generally good does this mean they should be given the benefit of the doubt, or that one poor mark does not matter? You may well need more contextual information and you also have to bear in mind whether the student has benefitted from dealing with another topic in a way that other students could have done but did not get the chance.
B	It seems straightforward. Based on the task instructions and published criteria the essay should be awarded a high mark and this is what you will recommend.	One of the criticisms of criteria is that they are content free, i.e. they do not capture the discipline knowledge that is central to student learning. So to rely solely on the simple one-word criteria to make a judgement is not appropriate. If the topic was from a completely different subject or discipline then it would not be acceptable, so is it appropriate if it is a different topic within that subject? This is not straightforward.
C	You are concerned about the student's motives for writing about a different topic. You think one solution could be to agree with the fail and offer a resit on this module, but suggest the student is counselled that the work they have done may help in the other modules.	This is quite a muddled compromise solution. Has the student been fairly treated? Has the student reached the threshold standard for this module, taking into account the module learning outcomes and the task brief?
D	You agree that the work has probably got to fail although there should be a resit opportunity. Your concern is that it should not happen again. You will recommend that the topic is made clear in the brief.	The module leader has explained that the students were briefed in class about the assignment. It is good practice to do this as it allows students to ask questions and clarify explanations. However you cannot rely on this alone as, in this case the student was absent. So your advice to make the topic clear in the brief is sound. The decision to fail the student is not a straightforward one (see above).

Dilemma – Master's-level complexity

The sample of work for a master's module includes examples for all grades. Each one fits clearly into its band, aligning well with the level descriptors of the

rubric. The pieces of work have similar structure and have taken a very similar approach to the task.

Optional responses	Comments on the options
<p>A It is clear to you that this marking is straightforward and is aligned with the local academic standards, so you are happy to agree the marks.</p>	<p>There may well be no issue here, but perhaps you might consider whether the task is complex enough for master's level if the marking is so straightforward. Does it move beyond requiring 'right' answers? You might also wonder how decisions are made at the borders of grade boundaries. Is the marking quite so clear cut? You might want to ask that in the future the sample includes some borderline examples where judgement is often more difficult.</p>
<p>B In your experience, judgement about complex work is quite messy and you find it rather odd that all the work in the sample fits so clearly into each grade. You wonder how the samples are selected as you think this sample might be contrived. You ask to see a few more examples, specifically at the borderline of the distinction grade.</p>	<p>It is important to know how the samples of work are selected and by whom as it may well change the picture you get. Many institutions make all the work available electronically so that you can select your own sample. Decisions at the borderlines often focus what is really valued by markers and the academic standards being used. If you have the time to look at borderlines this will help you judge whether you can have confidence in the marks as the sample suggests.</p>
<p>C The consistency and precision of the marking suggests to you that a highly structured marking scheme may have been used. You think you should engage in discussion about the nature of master's level work and the need to allow students to develop their ideas in their own way, which must be reflected in the assessment judgements.</p>	<p>It is possible that here the drive towards reliability in marking has been privileged, providing a neatness to a usually imprecise process of judgements. This is more appropriate in some subjects than others but it is important to consider whether the task is complex enough for master's level. Whether or not there is a marking scheme, you might want to talk to the module leader about the task, its purpose and how the students are prepared for the assessment.</p>
<p>D The similarity of student work seems inappropriate for master's' level. It is not clear whether this similarity arises from detailed briefing in class on the approach expected, and perhaps the provision of one exemplar followed by the cohort, or the cohort has independently discussed the approach to take. You need to know more about this assessment, including the rationale for the task, expectations of the work to be produced, extent of guidance given to</p>	<p>Your worry that this 'perfectly marked' sample may be disguising other issues with the module are worth thinking about and a brief discussion may indicate problems or reassure you.</p>

students and time allowed to do the task.	
---	--

Dilemma – Overlapping assessment

The three modules that have run in the most recent semester have assessments in which you can see considerable overlap. Although the assignments are ostensibly different (problem-solving task, report writing and preparing an informative PowerPoint presentation for a particular audience) you can see that students are using the same or a very similar research base, theoretical models and analysis to prepare each piece of work.

Optional responses

Comments on the options

A	You are concerned that students are not being exposed to a wide enough range of ideas or theories to prepare them for the next stage of the course. You ask the course team to look collectively at the assessments being used and how they fit together both concurrently and in terms of progression through the programme.	The situation may have arisen because of disjointed planning or 'drift' due to incremental changes to module assessment. There needs to be thought given to the student experience of assessment and the progression of challenge and academic standards through the programme. Your advice is good. As part of their discussion the team should resolve whether they think this overlap is appropriate.
B	It seems to you that students are getting credit for the same work three times. This could have serious implications for final results. You ask to look at any noticeable changes in students' overall result after these marks have been added to their profile.	If the similarity and overlap is so great then this could make a difference to students either negatively or positively, but only if the overlap in the work has resulted in similar marks being awarded for each module. Taking a numbers-based approach to examining this problem may not be possible as there are so many factors that affect marking and academic standards being used.
C	As the assignments require students to do different things with the theory and analysis it helps them develop the broader graduate attributes relating to communication. With this practice they should be able to replicate these skills for other topics. It is not possible to assess student on all the topics in your discipline. So you do not worry about the overlap.	Assessment has to meet many requirements including providing evidence of graduate attributes. However it may be preferable for students to demonstrate these attributes while also developing a wider knowledge of the subject. It would still be a good idea to point out the issue to the team and ask what they could do to resolve it while still developing attributes.
D	Given this overlap you see an opportunity for cross-module	This suggestion would provide an opportunity to move towards programme-based

assessment. You suggest that the module leaders collaborate to develop a more complex task that would cover learning outcomes from each of the modules, and give credit to be counted in all three modules. You are not sure what reaction you will get but go ahead and suggest it anyway.

assessment, i.e. focused on the programme learning outcomes rather than individual module learning outcomes. This could potentially reduce marking load and acknowledge the commonality that already exists. This may be an idea that has not occurred to the team.

Session 7: Social moderation and calibration of standards

This session aims to demonstrate a range of approaches to social moderation and calibration. It should enable you to recognise that calibration methods can be designed to suit a subject or discipline and that external examiners should take a lead in demonstrating the value of calibration.

Task 7: Social moderation

Purpose

- To demonstrate social moderation as a process for obtaining consistency in understanding and applying academic standards, based on discussion of concrete examples and drawing on relevant reference points, such as assessment criteria.
- To demonstrate the difference between local moderation of grades and broader calibration.
- To recognise the importance of calibration for undertaking the role of external examiner.

What you should get out of it

- You should recognise the effects of calibration to mitigate variability in standards and to create stronger external examiner and subject communities.

Relevance to external examiners

- Calibration of standards used by external examiners will provide greater confidence in the external examiner system and maintain a peer-based system.

Description of task

The task is to consider and compare two examples of social moderation (video of examiners marking a trumpet recital and written case study concerning dissertation marking). Once you have watched the video and read the case study on the following page, you are asked to discuss the following questions in your groups:

- What do you see as the main differences, if any, between the post-marking moderation shown in the video and the written case study?
- How do these social moderation methods compare with reliance on criteria to provide consistent judgements?
- To what extent does each method contribute to a common understanding of academic standards?

Case study of a social moderation process

The following case study describes measures taken by a university department to improve the reliability of their assessment of dissertations. Despite having established explicit assessment criteria, there was still concern that this on its own had not established a common set of standards. The case study describes the two workshops devised and the social moderation processes involved in establishing greater common understanding of the criteria and greater reliability in the standards applied.

Case study: Summary of a social-moderation process to improve the reliability of dissertation assessment in a university department

The teaching team decided to use an assessment process based upon nine sub areas of assessment for which they subsequently devised criteria. These sub areas were grouped into three assessment areas.

1. Task definition and approach
2. Literature review, findings and evaluation
3. Communication of ideas and presentation

Although the procedure and criteria were felt to be robust, the appointment of new lecturers, and a desire within the dissertation team to ensure that comparability of standards was being maintained between all lecturers involved in assessing dissertations, prompted two workshops to examine the actual use of the criteria.

The initial workshop

Before the initial workshop, lecturers were given a complete copy of a dissertation from another higher education institution. By using an unknown dissertation they ensured that assessment was not biased by knowledge of the student's previous performance. Copies of the assessment procedure, criteria and a pro forma for recording comments were also supplied. Lecturers were asked to assess independently the dissertation using these criteria before the workshop.

As each lecturer arrived at the workshop they recorded their 'marks' and comments for each sub area of assessment, using a number line on a flip chart. This enabled a clear visual representation of the marks to be obtained. A range of marks was recorded for each sub area of assessment. Assessments of the dissertation were then debated in detail. Criteria were reconsidered and other possible reasons for the differences in assessment explored. Comments made by lecturers proved helpful here. The discussion did not alter the criteria but made them clearer. The guidance notes on each of the sub areas of assessment (issued to both lecturers and students) were modified to reflect this discussion.

The second workshop

A second workshop was held at which the same lecturers were asked to reassess the same dissertation based on their more detailed understanding of the criteria. Results were recorded in the same manner as before. Informal discussion between lecturers indicated that they were now far more confident in using the criteria to determine the overall grade, due to their shared understanding. Ownership of the criteria and the assessment process appeared to have increased compared to the first workshop, with lecturers usually referring to them as 'our' rather than 'the' criteria. A decline in the standard deviation for each of the assessment sub areas from the assessments in the first workshop suggested an increase in uniformity of assessment between lecturers. However, it should be noted that there was still some variation in 'marks'.

Implications for good practice

The first implication for good practice, based upon this experience, is that joint development of criteria by those assessing the work provides a useful start for ensuring that each lecturer understands them in the same way.

Our second implication for good practice relates to the continued consistent understanding of the criteria among the assessment team. The workshop experience suggests that, over time, understanding and application of criteria does alter. As a consequence, it is important that the criteria are periodically revisited and debated by lecturers. Such debates are also helpful in engendering ownership and enabling new colleagues to obtain a clearer understanding which is

consistent with those expressed by the rest of the team. Workshops at which criteria are debated can also serve to remind experienced lecturers of the required standard, reducing the dangers of standards becoming too demanding or less stringent over time.

Summary of: Mark N.K. Saunders, Susan M. Davis, (1998) "The use of assessment criteria to ensure consistency of marking: some implications for good practice", *Quality Assurance in Education*, Vol. 6 Issue: 3, pp.162-171,

Comparing video and case study

- What do you see as the main differences, if any, between the marking process shown in the RNCM video and the written case study?
- How do these social moderation methods compare with reliance on criteria to provide consistent judgements?
- To what extent does each method contribute to a common understanding of academic standards?

AdvanceHE

Social moderation to calibration

Moderation is generally focused on agreeing marks rather than sharing standards as in our discussion of the Part 1 assignments.

Calibration is about shared knowledge of standards, often achieved through social moderation processes as in the case study, in order to create 'calibrated' academics.



'calibrated' academics ... are able to make grading judgements consistent with those which similarly calibrated colleagues would make, but without constant engagement in moderation. The overall aims are to achieve comparability of standards across institutions and stability of standards over time. (Sadler 2012)

AdvanceHE

Calibration of standards

AdvanceHE

Effective calibration

- A collegial process within subject communities
- Working with de-identified exemplars – taking focus away from institution or marker;
- Multistage process – individual judgement > small group > large group;
- Robust but respectful conversation;
- Well-facilitated discussion allowing diverse opinions.
- Development of enhanced assessment literacy;

Where might you find opportunities for calibration, even in your own department/ institution?

AdvanceHE

Implications for external examining

- Programme teams expect you to be able to say whether they are using appropriate and comparable standards for their assessment
- You, as external examiner need to be confident that you represent the standards in your subject area
- That confidence can be established through calibration with other external examiners in your subject community
- Hence this project involves a parallel initiative piloting approaches to external examiner calibration



AdvanceHE

Getting involved in Calibration

How can you calibrate your standards with other external examiners in their field?

The HEA Degree Standards Project is piloting various approaches and have produced calibration toolkits and case studies.

Examples of calibration initiatives: the Royal Geographical Society, Royal Society of Chemists, Veterinary Schools Council Education Committee and Conservatoires UK, a consortium of NW HEIs

For further information on calibration, visit the Degree Standards project website:
<https://www.heacademy.ac.uk/hefce-degree-standards>

AdvanceHE

Summary log

Session 7 summary log: Social moderation and calibration

Please comment on the distinction between social moderation and calibration. Use examples to illustrate the differences

Make notes on your experience of engagement in social moderation and calibration activities. AND Identify new ways in which you could develop yourself as a calibrated academic.

Example of calibration: a project in accounting education in Australia

In this extract from their paper, Watty et al (2013) describe the method used in their project to calibrate assessment standards in accounting education in Australia.

In this section, we describe the social moderation process that we have developed for deriving calibrated peer assessments (ethics approval to conduct the research had been obtained from each of the author's institutions). In this project, the focus is on one part of

one of the five threshold learning standards for accounting, namely written communication skills at the undergraduate level. The reason for this choice of focus was that communication skills is one of the skill sets most frequently cited by employers as lacking in accounting graduates. Our initial choice of written rather than oral communication skills was driven by the data on written-based assessment tasks being far more readily available. The discipline standards define communication and teamwork as follows:

Bachelor graduates in Accounting will be able to:

Justify and communicate accounting advice and ideas in straightforward collaborative contexts involving both accountants and non-accountants.

Sample of assessors

To operationalise the consensus (social) moderation process and form our group of peer calibrated assessors, we obtained volunteer accounting academics and practitioners totalling over 30 people. The inclusion of practitioners in our assessor group reflected the close collaboration between the research project team and the professional accounting bodies who were directly involved in the development of the accounting threshold learning standards. Many of the practitioners had experience in similar moderation processes associated with professional bodies' accounting examinations. Ten of Australia's 40 universities participated in this study. A sample of 10 universities was accepted as representative of the Australian accounting education environment, due to the broad cross-section of business schools which participated. We used random, anonymous samples of final-year work, chosen by the higher education providers, as evidence of the relevant learning standard. The assessors participated in workshops in three locations: first in Darwin (20 assessors), then Melbourne (19 assessors) and finally Adelaide (24 assessors).

The same group of academic assessors from the 10 participating universities attended each workshop, with only slight variations where attendance was not possible. Practitioner assessors were only able to attend workshops in the city in which they were domiciled. Assessors were selected on the basis that they were senior accounting academics, and two came from each of the participating universities.

The peer review calibration process

To operationalise the calibration process, selected universities agreed to provide final-year students' assessment data which aligned with the selected threshold learning standards of undergraduate written 'communication skills'. The calibration process required all assessors to participate in three stages: a pre-workshop calibration exercise assessing sample student work; then a face-to-face workshop; and finally, each assessor was required to reassess each student's submission following the workshop (within two weeks).

The calibration process entailed the assessors rating student sample work against each assessor's understanding of the standards reflected in the national threshold learning standards. Each piece of work was graded by a peer assessor along a continuum out of 100 of 'not meeting' and 'meeting' the threshold learning standards. Scores of 50 and above were deemed as 'meeting' the threshold learning standards. In addition, each assessor had to provide written reasons as to why they had awarded the grade that they had, and to suggest how work could have been approved to obtain a higher score.

In the first stage of the calibration process, individual assessors were required to record their assessment of student submissions using an electronic repository known as the Self and Peer Assessment Resource Kit (SPARK). SPARK helps facilitate anonymity in ratings and comments, ensures confidentiality while reducing the risk of influence by other reviewers

and minimises the burden of participating for any assessor. SPARK was also used to aggregate, track and disseminate assessments.

Prior to each workshop, but after the completion of the pre-workshop individual assessments, the whole set of responses were anonymously revealed to each assessor for their own analysis and interrogation. Assessors could interrogate the set to clarify why others may have rated a piece of work differently, or even similarly but for different reasons. These two reflective steps assisted assessors to construct their understandings of the 'national standard'.

In the second stage of the calibration process, all reviewers met in person at the workshops and shared their views. The face-to-face workshops varied between three to six hours in duration. The process encouraged dialogue around the pre-workshop reviews; however, reviews remained entirely anonymous. The workshops provided a forum to foster a shared understanding of what constitutes the concept of a 'fair' and agreed assessment. The workshops commenced with an analysis of the pre-workshop individual ratings given by reviewers on the set assessment task (for task validity) and then a review of each piece of student work, before the reviewers separated into small groups to discuss the strengths and weaknesses of particular sample student assessments. The aim of this was to develop a shared understanding among reviewers, resulting in a reduction in the standard deviation of the reviewer's post-workshop evaluation scores. Finally, the whole group was involved in a vigorous discussion to arrive at a consensus regarding understanding of the threshold learning standards.

In the third stage, once assessors had returned to their home institution, they were again asked to reassess the sample student work using SPARK. This provided an opportunity to explore whether there was persistence in the shared understandings that had been reached during the workshops. Assessors were also encouraged to add comments on their understanding of the threshold learning standards, and the process of developing shared understanding. These evaluations remain anonymous to all participants and members of the research team, though the research team is able to review aggregated comments and assessments from the group.

Watty, K, Freeman, M., Howieson, B., Hancock, P., O'Connell, B., de Lange, P., and Abraham, A., 2013. Social moderation, assessment and assuring standards for accounting graduates. *Assessment and Evaluation in Higher Education*, 39(4), pp.461-78.

Session 8: Conclusion

This session reinforces learning points from the course.

Task 8: Completing the summary log and identifying key points

Purpose

To engage in reflection on their learning and clarifying key messages to take away.

What you should get out of it

You should achieve clarification of key messages, and identify potential change in your approach to external examining.

Relevance to external examiners

External examiners need to review their approach to the role to ensure they meet expectations of stakeholders.

Description of task

Use your notes from previous sessions to identify key points from the course for the summary log.

In your own time, compare your final list, plus other sections of your summary log, with a completed version (supplied as a handout).

Task 8 Completing summary log and key points

AdvanceHE

Completing summary log and key points



Review your summary log and reflect on key learning points for you from the course (part 1 and 2). Do you have any remaining questions? What might you change about your external examining?

In your own time, complete your summary log (key learning points) and compare it to the already completed version available to take away

AdvanceHE

Summary log

Session 8 summary log: Review

Please fill in key points from the course

Follow-up activities

This section provides follow-up activities which may be useful to you. They include:

1. how as an aspiring examiner you can find external examiner vacancies;
2. continuing professional development as an external examiner;
3. thinking about your external examining practice and collecting evidence;

1. How as an aspiring examiner you can find external examiner vacancies

Vacancies are advertised on a regular basis on the JISC external examiners mail base. Just search for 'JISC external examiners mail base' online or go to <https://www.jiscmail.ac.uk/cgi-bin/webadmin?AO=EXTERNAL-EXAMINERS> where you can subscribe to the list for free.

2. Continuing professional development as an external examiner

Here are some suggestions for continuing your professional development as an external examiner.

- Look again at the further reading and sources at the end of the briefing paper on assessment scholarship (Part 1 Activity 3).
- Complete your summary log and compare this to the already completed version distributed during the course.
- You might like to contact your subject community body about calibration exercises. At the earliest opportunity, take up any calibration opportunities they might offer.

3. Thinking about your external examining practice and collecting evidence

If you are undertaking a PG Cert in Academic Practice or thinking of applying for fellowship recognised by the HEA, consider what evidence you could produce about your understanding of the role of external examiner, ideally with practical examples of undertaking that role. One way of doing this is to undertake a critical incident analysis, to help you reflect on your practice and collect evidence for a fellowship application. Here are some suggestions as to how you might do that, depending on your experience of external examining. See below for guidance on critical incident analysis.

For academics who are already external examiners: critical incident analysis

Think of a critical incident that has arisen in your own practice as an external examiner. This is likely to be similar to the scenarios that were used on the professional development course. Use this incident to analyse what happened and to reflect critically on your actions in the light of (selected) concepts and frameworks introduced on the course, as well as in relation to the UKPSF. Write 500 - 700 words.

For academics who are new to external examining, but have worked with external examiners appointed to the programmes they are involved in

Think of a critical incident that has arisen in your encounters with an external examiner. This is likely to be similar to the scenarios that were used on this professional development course. Use this incident to analyse what happened and to reflect critically on your own and the external examiner's actions in the light of (selected) concepts and frameworks introduced on the course, as well as in relation to the UKPSF. If you can't think of a critical incident from your own experience, you might seek out someone who has worked with or been an external examiner and interview them about a critical incident. Write 500 - 700 words.

For academics who are new to external examining and have not yet got an experience of working with external examiners appointed to the programmes they are involved in

Choose one of the scenarios in this Handbook that describe critical incidents which you may encounter in the role of external examiner. Use this incident to analyse the situation and make suggestions for possible courses of action in the light of (selected) concepts and frameworks introduced on the course. Alternatively, you might seek out someone who has worked with or been an external examiner and interview them about a critical incident, using the guidance below. Write 500 - 700 words.

If you are thinking of applying for Senior Fellowship, think of an example of your activities as an external examiner in which your actions demonstrated leadership and impact on the practices of others. For instance, your interactions with the programme team you were an external examiner for may have led to changes in assessment practices of modules or programmes, or the insights gained from your external examining activities may have influenced changes made in your home institution. Write this up as example evidence to use in your application for Senior Fellowship.

Consider the way in which your activities as an external examiner demonstrate 'successful co-ordination, support, supervision, management and/or mentoring of others (whether individuals and/or teams) in relation to teaching and learning' (UKPSF Descriptor 3 VII.). Use selected concepts and frameworks from the literature and policies introduced on the course to underpin your reflections and offer a rationale for your actions.

Consider developing a critical incident analysis of an aspect of your external examiner experience which demonstrates leadership or impact on the practices of others. Use the guidance on writing a critical incident set out below.

4. Guidance for critical incident analysis

The following questions are intended to guide you through the critical incident analysis.

1. What was the incident and what were (might have been) the events leading up to it?
2. What happened (could happen) because of the incident?
3. What was (might be) your response to the incident and what actions did (might) you take?
4. What occurred (might occur) subsequently because of your actions?
5. Reflecting on the incident, do you think that your original actions were appropriate and effective? Would you now act in the same way or prefer to have intervened in a different way?
6. Which policies might be relevant and offer principles and guidance for appropriate actions?
7. Which concepts from the literature and research on external examining might throw light on this critical incident?
8. What have you learned from the incident for your future practice as an external examiner?

Criteria

Please use these criteria to think about and develop the quality of your critical incident analysis.

- Quality of critical analysis and evaluation, taking account of different perspectives.
- Clearly articulates links between policy, theory and practice.
- Convincing rationale for conclusions drawn.
- Clear communication and argument, accurately referenced.
- Links made to relevant dimensions of practice of the UKPSF.

Glossary of terms

Assessment criteria

The defined basis on which students' submissions are judged. They are used by staff in assessing students' work and are usually provided to students to guide their work.

Academic standards

The standards of student attainment set by degree-awarding bodies for their courses (programmes and modules) and expected for their awards.

Assessment for Learning

Assessment for Learning is the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there (Assessment Reform Group, UK 2002). In the higher education context, it can also mean designing assessment tasks that specifically integrate learning.

Assessment as learning

Assessment as learning is a subset of assessment for learning and sees student involvement in assessment, using feedback, participating in peer assessment, and self-monitoring of progress, as moments of learning themselves. Students come to have a better understanding of the subject matter and their own learning through their close involvement with assessment.

Assessment literacy

This includes an appreciation of the relationship between assessment and learning, knowledge of the principles of sound assessment including the related terminology, understanding of the nature of standards and criteria, skills in self and peer assessment, recognition of varied purposes of assessment methods, skills and techniques, and ability to select and apply appropriate approaches and techniques to assessment tasks (summarised from Price et al, 2012).

Award algorithm

The method of calculation that determine a student's degree classification. It typically involves some element of weighting of marks from different levels of the programme and the averaging of marks. It may also involve rules regarding omission of some module marks (e.g. first year or lowest scoring modules), condonement and compensation, and/or elements of discretionary judgement at the boundaries of each classification.

Blind double marking (BDM)

Two assessors grade the work independently without prior sight of the other's grades or comments.

Calibration

Calibration is a process of peer review carried out by members of a disciplinary and/or professional community who discuss, review and compare student work in order to reach a shared understanding of the academic standard which such work needs to meet. Its key purpose is to set and assure standards as well as ensuring that the tasks used are valid assessments of key graduate learning outcomes in the discipline. Calibration is based on the assumption that standards are socially constructed and it therefore involves dialogue, negotiation and joint decision making.

Compensation

Definitions of compensation vary between higher education providers and are sometimes the same as the definition of condonement used elsewhere. Commonly, compensation is defined as the process by which an assessment board may decide that a strong performance by a student in one

part of the curriculum may be used as the basis for the award of credit in respect of a failed performance elsewhere.

Condonement

Definitions of condonement vary between higher education providers and are sometimes the same as the definition of compensation used elsewhere. Commonly, condonement is defined as the process by which an assessment board, in consideration of the overall performance of a student, decides that without incurring a penalty, a part of the programme that has been failed does not need to be redeemed.

Constructive alignment

Constructive alignment has two aspects. The 'constructive' aspect refers to the idea that students construct meaning through relevant learning activities. That is, meaning is not something imparted or transmitted from teacher to learner, but is something learners must create for themselves. Teaching is simply a catalyst for learning:

'If students are to learn desired outcomes in a reasonably effective manner, then the teacher's fundamental task is to get students to engage in learning activities that are likely to result in their achieving those outcomes ... It is helpful to remember that what the student does is actually more important in determining what is learned than what the teacher does.' (Shuell, 1986: 429)

The 'alignment' aspect refers to what the teacher does, which is to set up a learning environment that supports the learning activities appropriate to achieving the desired learning outcomes. The key is that the components in the teaching system, especially the teaching methods used and the assessment tasks, are aligned with the learning activities assumed in the intended outcomes. The learner is in a sense 'trapped', and finds it difficult to escape without learning what he or she is intended to learn.' From Biggs, J. *Aligning teaching for constructive learning*. York: The Higher Education Academy available

at https://www.heacademy.ac.uk/system/files/resources/id477_aligning_teaching_for_constructing_learning.pdf

Criterion-referenced assessment

Means that learners are assessed against pre-defined criteria or standards. An example of this is the assessment of learners against specific competencies in some professional area such as nursing.

Feedback

Includes information to students on their achievements, the purpose of which is to help learners to improve their performance. Feedback can be part of a formal assessment system (when it might be called formative assessment) or it can be informal and available from multiple sources.

Finch Report

Review of external examining arrangements in universities and colleges in the UK: Final report and recommendations. This review and associated report was undertaken by Universities UK and GuildHE, and chaired by Professor Dame Janet Finch, former Vice-Chancellor of Keele University. Available at: <http://www.universitiesuk.ac.uk/policy-and-analysis/reports/Pages/external-examining-arrangements-review.aspx>

Formative assessment

Formative assessment provides feedback to the student so that they can improve and self-regulate their work, and to the lecturer or tutor so that they may adjust their teaching.

Framework for Higher Education Qualifications (FHEQ)

A published formal structure that identifies a hierarchy of national qualification levels and describes the general achievement expected of holders of the main qualification types at each level.

Grade-point average (GPA)

GPA is a measure of student achievement used both during and on completion of a programme of study. There are different approaches to and scales for GPA, but the core North American model accords letter grades to assessments, with implied descriptors, such as A = excellent, B = good / above average, and so on. Typically with GPA, student work is graded A to D, or F (a fail), and then converted to a grade point (where A = 4.00 and D = 1.00). In North America, a GPA is then calculated by averaging the grade points for every module from all years of study (i.e. a 'straight average'). A GPA system gives students access on a continuous basis to a cumulative average as well as their receiving an end of programme GPA score.

Higher Education Achievement Report (HEAR)

The HEAR has been developed to provide a rich record of student achievement at university including curricular and accredited co-curricular and extra-curricular achievements. The HEAR is being used by an increasing number of universities; it encompasses the degree transcript, and provides information and content which meets the requirements of the European Diploma Supplement. The HEAR can, in principle, include any summative judgement of student academic achievement including either a degree classification or a GPA, or both. See <http://www.hear.ac.uk/> for more information.

Higher Education Quality Council (HEQC)

The forerunner of the QAA.

Marking scheme

This is an outline of the response expected from students to an assessment task, with an indication of the division of marks between sections or topics. It is used by staff when marking work.

Model answer

This is a precise marking scheme where the response from students can be specified exactly. For example, this may be the case with some mathematical problems or other questions with a clearly defined solution.

Norm-referenced assessment

This type of assessment produces a mark or grade based on each student's achievement in relation to the rest of the group. For example, the best ten performers may receive a Grade A, the next ten Grade B and so on.

Peer assessment

This is where students are involved in assessing the performance of their fellow students. As with self-assessment, it may involve helping to devise criteria, giving feedback or allocating marks.

Professional, statutory and regulatory bodies (PSRB)

Organisations that set the standards for, and regulate entry into, particular profession(s) and are authorised to accredit, approve or recognise specific programmes leading to the relevant professional qualification(s) for which they may have a statutory or regulatory responsibility.

Programme

The overall curriculum followed by an individual student, normally comprising a specified set of modules or option choices.

Programme assessment strategy

A strategic approach to planning the assessment across a programme, based on enabling the students to demonstrate the programme learning outcomes. Such a strategy is also designed to provide for progression, assessment for learning and employability.

Reference point

A national, institutional or professional framework, guidance, standard or regulation which is relevant to designing and evaluating the programme of study. For example, the UK Quality Code for Higher Education, the Framework for Higher Education Qualifications, Professional body standards and university regulations.

Quality Assurance Agency (QAA)

The QAA is an independent body entrusted with monitoring and advising on standards and quality in UK higher education. It has been responsible, with the sector, for developing the UK Quality Code for Higher Education, subject benchmark statements and frameworks for higher education qualifications. It currently works on behalf of the UK Standing Committee for Quality Assessment, in consultation with the higher education sector.

Quality standards

Standards in the assessment cycle, such as task setting, marking, feedback, moderation and examination board procedures.

Reliability

Assessment tasks should generate comparable marks across time, across markers and across methods. Reliability is demonstrated, for example, when different markers make the same judgements about an assignment or when one marker makes consistent judgements about a piece of work at different times.

Second marking, sample second marking

Checking of marking by a second tutor, involving only a selection of work in the case of sample second marking.

Self-assessment

This is the involvement of the student in the assessment process, by, for example, engaging students in helping to devise the criteria by which their work will be assessed. It may involve students in evaluating the strengths and weaknesses of what they have achieved or in attributing marks to their work.

Social moderation

Also known as consensus moderation, this is a process for obtaining consistency in understanding and applying academic standards based on discussion of concrete examples and drawing on relevant reference points such as assessment criteria or standards. Social moderation may be carried out to decide the appropriate mark for student work or to resolve differences in markers' judgements.

Subject benchmark statement

A published statement that sets out the knowledge, understanding, abilities and skills expected of those graduating in each of the main subject areas (mostly applying to bachelor degrees), explaining what gives that particular discipline its coherence and identity.

Summative assessment

Assessment which counts towards, or constitutes a final grade for a module or course, or where a pass is required for progression by the student.

Threshold (academic) standard

The minimum acceptable level of achievement that a student must demonstrate to be eligible for an academic award.

UK Quality Code for Higher Education

The UK Quality Code (2018) enables providers to understand what is expected of them and what to expect from each other. It has been developed by QAA on behalf of the UK Standing Committee for Quality Assessment, in consultation with the higher education sector. <https://www.qaa.ac.uk/quality-code>

UKSQA

UK Standing Committee for Quality Assessment. It provides sector-led oversight of higher education quality assessment arrangements that continue to be shared across the UK. The committee brings together partners from across the higher education sector, drawing on academic, student and regulatory expertise. <https://ukscqa.org.uk>

Validity

The requirement for assessment tasks to assess the stated learning outcomes; that is, they are assessing what they purport to be assessing.

The definitions above are based in part on the [QAA glossary](#) and on the former HEFCE's glossary. In addition, we have drawn on definitions from HEA (2012) and Brian Kjaer Andreasen, Aalborg, 8 March 2001 Assessment Glossary.

References

- Assessment Reform Group, 2002. *Assessment for Learning: 10 Principles* Assessment Reform Group. https://www.researchgate.net/publication/271849158_Assessment_for_Learning_Research-based_principles_to_guide_classroom_practice_Assessment_for_Learning
- Baume, D., et.al., 2004. What is happening when we assess, and how can we use our understanding of this to improve assessment? *Assessment and Evaluation in Higher Education*, 29 (4), pp.451–477.
- Bell, A., Mladenovic, R. and Price, M., 2013. Students' perceptions of the usefulness of marking guides, grade descriptors and annotated exemplars. *Assessment and Evaluation in Higher Education*, 38(7), pp.769–788.
- Bloxham, S., Boyd, P. and Orr, S., 2011. Mark my words: the role of assessment criteria in UK higher education grading practices. *Studies in Higher Education*, 36(6), pp.655–670.
- Bloxham, S with J. Hudson, B. den Outer, M. Price, (2015) External peer review of assessment: an effective approach to verifying standards? *Higher Education Research and Development*, 34 (6) pp. 1069-1082.
- Broad, B., 2003. *What We Really Value: Beyond rubrics in teaching and assessing writing*. Logan, Utah: Utah State University Press.
- Brooks, V. 2012. "Marking as Judgement." *Research Papers in Education* 27 (1): 63-80.
- Colley, H., and Silver, H., 2005. *External examiners and the benchmarking of standards*. York: Higher Education Academy.
- Crisp, V., 2008. 'Exploring the nature of examiner thinking during the process of examination marking', *Cambridge Journal of Education*, 38(2), pp. 247–64.
- Dobson, S., 2008. Applying a validity argument model to three examples of the viva. *Nordisk Pedagogik*, 28, pp.332–44.
- Dressel, P.L., 1957. Facts and fancy in assigning grades. *Basic College Quarterly*, 2, pp.6–12.
- Elander, J. and Hardman, D., 2002. An application of judgement analysis to examination marking in psychology. *British Journal of Psychology*, 93, pp.303–328.
- Felton, J. and Koper, P.T., 2005. Nominal GPA and real GPA: A simple adjustment that compensated for grade inflation. *Assessment and Evaluation in Higher Education*, 30(6), pp.561–69.
- Gibbs, G., 1999. Using Assessment Strategically to change the way students learn. In S. Brown and A. Glasner, A., eds. 1999. *Assessment Matters in Higher Education: Choosing and Using Diverse Approaches*. Buckingham: SRHE / Open University Press. pp.41–53.
- Grainger, P., Purnell, K. and Zipf, R., 2008. Judging quality through substantive conversations between markers. *Assessment and Evaluation in Higher Education*, 33(2), pp.133–42.
- Greatorex, J., 2000. Is the glass half full or half empty? What examiners really think of candidates' achievement. In: BERA (British Educational Research Association), British Educational Research Association Conference, Cardiff, United Kingdom, 7–10 September 2000.
- Hand, L. and Clewes, D., 2000. Marking the Difference: An Investigation of the Criteria Used for Assessing Undergraduate Dissertations in a Business School. *Assessment and Evaluation in Higher Education*, 25(1), pp.5–21.
- Handley, K. and Williams, L., 2011. From copying to learning: using exemplars to engage students with assessment criteria and feedback. *Assessment and Evaluation in Higher Education*, 36(1), pp.95–108.

- Hanlon, J., Jefferson, M., Molan, M. and Mitchell, B., 2004. *An examination of the incidence of 'error variation' in the grading of law assessments*. United Kingdom Centre for Legal Education. Available at: http://www.law.uwa.edu.au/_data/assets/pdf_file/0006/1888611/Hanlon.pdf [Accessed December 2016].
- Hartog, P. and Rhodes, E.C., 1935. *An examination of examination*. London: Macmillan.
- Hawe, E., 2002. Assessment in a pre-service teacher education programme: the rhetoric and the practice of standards-based assessment. *Asia Pacific Journal of Teacher Education*, 30, pp.93–106.
- Higher Education Academy, 2015. *A review of external examining arrangements across the UK*. Available at: <http://www.hefce.ac.uk/pubs/rereports/Year/2015/externalexam/Title,104316,en.html> [Accessed December 2016].
- Higher Education Academy, 2015. *A Marked Improvement: Transforming assessment in higher education*. Available at: https://www.heacademy.ac.uk/system/files/a_marked_improvement.pdf [Accessed January 2017].
- Higher Education Quality Council, 1997. *Assessment in higher education and the role of 'graduateness'*. London: HEQC.
- Hugh-Jones S., Waterman MG., Wilson I., 2009. Accessing and understanding the tacit dimensions of assessment. *Psychology Learning and Teaching*, 8(2), pp.7–15.
- Hunter, K. & Docherty, P., 2011, 'Reducing variation in the assessment of student writing', *Assessment and Evaluation in Higher Education*, 36(1), pp. 109–24.
- Jawitz, J., 2009. Learning in the academic workplace: the harmonization of the collective and individual habitus. *Studies in Higher Education*, 34(6), pp.601–14.
- Kelly, G.A. (1991). *The psychology of personal constructs: Volume 1: A theory of personality*. London, UK: Routledge.
- Laming, D., 1990. The reliability of a certain university examination, compared with the precision of absolute judgements. *Quarterly Journal of Experimental Psychology*, 42, pp.239–254.
- Leach, L. et al. 2001. Assessment and empowerment: some critical questions. *Assessment and Evaluation in Higher Education*, 26(4), pp.293–305.
- Miller, G.A., 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63, pp.81–97.
- Moss, P.A. and Schutz, A. 2001. Educational standards, assessment and the search for consensus. *American Educational Research Journal*, 38(1), pp.37–70.
- Newstead, S.E. 2002. Examining the examiners; why are we so bad at assessing students? *Psychology Learning and Teaching*, 2, 70-75.
- Newstead, S.E. and Dennis, I., 1994. Examiners examined: The reliability of exam marking in psychology. *The Psychologist*, 7, pp.216–219.
- Nonaka, I., 1991. The knowledge-creating company. *The Harvard Business Review*, November-December, pp.96–104.
- Norton, L. 2004, Using assessment criteria as learning criteria: a case study in psychology, *Assessment and Evaluation in Higher Education*, 29 (6) pp. 687-702.
- O'Donovan, B., Price, M. and Rust, C., 2004. Know what I mean? Enhancing student understanding of assessment standards and criteria. *Teaching in Higher Education*, 9(3), pp.145–158.

O'Hagan, S.R. and Wigglesworth, G., 2014. Who's marking my essay? The assessment of non-native-speaker and native-speaker undergraduate essays in an Australian higher education context. *Studies in Higher Education*, 40(9), pp.1729–1747.

OFQUAL, 2016. *Marking consistency metrics*. Coventry: OFQUAL Available at: <https://www.gov.uk/government/publications/markings-consistency-metrics> [Accessed January 2017].

Polanyi, M., 1998. *The tacit dimension*. Reprinted in L. Prusak (Ed) *Knowledge in Organizations*, Boston: Butterworth Heineman.

Price, M., 2005. Assessment Standards: The Role of Communities of Practice and the Scholarship of Assessment. *Assessment and Evaluation in Higher Education*, 30(3), pp.215–230.

Price, M. and Rust, C. 1999. The Experience of Introducing a Common Criteria Assessment Grid Across an Academic Department. *Quality in Higher Education*, 5(2), pp.133–144.

Quality Assurance Agency, 2005. *Outcomes from institutional audits: external examiners and their reports*. Gloucester: Quality Assurance Agency.

Quality Assurance Agency, Higher Education Academy, 2013. *External examiners understanding and use of academic standards*. Available at: <http://www.qaa.ac.uk/en/Publications/Documents/External-Examiners-Report.pdf> [Accessed December 2016]

Quality Assurance Agency, 2015. *Characteristics Statement: Master's Degree*. Gloucester: Quality Assurance Agency.

Quality Assurance Agency for Higher Education, 2016. *The UK Quality Code for Higher Education*. Available at: <http://www.qaa.ac.uk/assuring-standards-and-quality/the-quality-code> [Accessed December 2016].

Read, B., Francis, B and Robson, J., 2005. Gender, bias, assessment and feedback: analyzing the written assessment of undergraduate history essays. *Assessment and Evaluation in Higher Education*, 30(3) pp.241–260.

Ross, V., 2009. External music examiners: micro-macro tasks in quality assurance practices. *Music Education Research*, 11(4), pp.473–484.

Rust, C. et al., 2003. Improving Students' Learning by Developing their understanding of assessment criteria and processes. *Assessment and Evaluation in Higher Education*, 28(2), pp.147–164.

Sadler, D. R., 1987. Specifying and Promulgating Achievement Standards. *Oxford Review of Education*, 13, pp.191–209.

Sadler, D. R., 2009. Grade integrity and the representation of academic achievement. *Studies in Higher Education*, 34(7), pp.807–826.

Sadler, D. R., 2012. Assuring academic achievement standards: from moderation to calibration. *Assessment in Education: Principles, Policy and Practice*, 20(1), pp.5–19.

Scoles, J et al., 2013. No longer exempt from good practice: using exemplars to close the feedback gap for exams. *Assessment and Evaluation in Higher Education*, 38(6), pp.631–645.

Shay, S., 2004, 'The assessment of complex performance: a socially situated interpretive act', *Harvard Educational Review*, 74(3), pp. 307–29; pp. 23.

Smith, E. and Coombe, K., 2006. Quality and qualms in the marking of university assignments by sessional staff: an exploratory study. *Higher Education*, 51(1) pp.45–69.

Snowdon, D, 2002. Complex acts of knowing: paradox and descriptive self-awareness. *Journal of Knowledge Management*, 6 (Special Edition), pp.100–11.

Vygotsky, L. S., 1978. *Mind in society: the development of higher psychological processes*. MA: Harvard University Press.

Watty, K., et al., 2013. Social moderation, assessment and assuring standards for accounting graduates. *Assessment and Evaluation in Higher Education*, 39(4), pp.461–478.

Webster, F., Pepper, D., and Jenkins, A., 2000. Assessing the Undergraduate Dissertation. *Assessment and Evaluation in Higher Education*, 25(1): pp.71–80.

Wolf, A., 1995. *Competence-based assessment*. Buckingham: Open University Press.

Woolf, H., 2004. Assessment criteria: reflections on current practices. *Assessment and Evaluation in Higher Education*, 29(4), pp.479–493.

Yucel, R. et al., 2014. The road to self-assessment: exemplar marking before peer review develops first-year students' capacity to judge the quality of a scientific report. *Assessment and Evaluation in Higher Education*, 39(8), pp.971–986.



Contact us

+44 (0) 3300 416201

external.examining@advance-he.ac.uk

www.advance-he.ac.uk

[in](#) [f](#)@AdvanceHE

(Registered Office)

Innovation Way, York Science Park,
Heslington, York YO10 5BR

-

First Floor, Westminster Tower
3 Albert Embankment, London

-

Peer House, 8–14 Verulam Street
London WC1X 8LZ SE1 7SP

© 2018 Advance HE. All rights reserved.

The views expressed in this publication are those of the author and not necessarily those of Advance HE. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any storage and retrieval system without the written permission of the copyright owner. Such permission will normally be granted for non-commercial, educational purposes provided that due acknowledgement is given.

To request copies of this report in large print or in a different format, please contact the Marketing and Communications Team at Advance HE:
+44 (0) 3300 416201 or external.examining@advance-he.ac.uk

Advance HE is a company limited by guarantee registered in England and Wales no. 04931031. Registered as a charity in England and Wales no. 1101607. Registered as a charity in Scotland no. SC043946. Advance HE words and logo should not be used without our permission. VAT registered no. GB 152 1219 50.