

Achievement Matters

External Calibration and Peer Review of Accounting Learning Standards

Mark Freeman - University of Sydney

Phil Hancock - University of Western Australia

HECQ Conference – Melbourne – 10 November

Definitions

Learning outcome

Learner knows, understands and can do..... independently, on demand and in a range of contexts (Sadler, 2013)

Learning standard

“a definite level of excellence or attainment..... so established by authority, custom, or consensus” (Sadler, 2013)



Context – Australian legislation (2011)

- 1.1 The course of study meets the Qualification Standards
- 1.2 Robust internal processes for design and approval of the course of study which *take account of external standards* and requirements, e.g. *published discipline standards*, professional accreditation, input from relevant external stakeholders, and comparable standards at other higher education providers
- 5.5 Academic standards intended to be achieved by students and the outcomes actually achieved by students in the course of study are benchmarked against similar accredited courses of study

Challenge

- Regular external referencing including grading of students' achievement of learning outcomes

Standard 5.4.3, HES Framework (2015)

Problem – Interrater reliability



Study of 6 experienced external examiners (EE) in 4 disciplines given 5 students work (QAA, 2013)

- Only 1 jointly highest (of 5) by all 6 EE in discipline
- 9 of 20 ranked both best and worst (of 5)



“The idea that a single external examiner could make a comparative judgement on the national, and indeed international, standard of a programme has always been flawed” (HEA, 2012)

Literature

- Variation due to differing experience and knowledge
- Assessors vary in the marks they award to the same student
- Marks differ in their distribution
- Reliability is low for essays & problem style questions but higher for knowledge recall style questions
- Not clear that use of assessment criteria, grade descriptors and marking rubrics increase reliability as markers may not agree with the benchmarks

Solution – Calibrate before assessing

“Assessment is largely dependent upon professional judgement and confidence in such judgement requires the establishment of appropriate forums for the development and sharing of standards within and between disciplinary and professional communities”

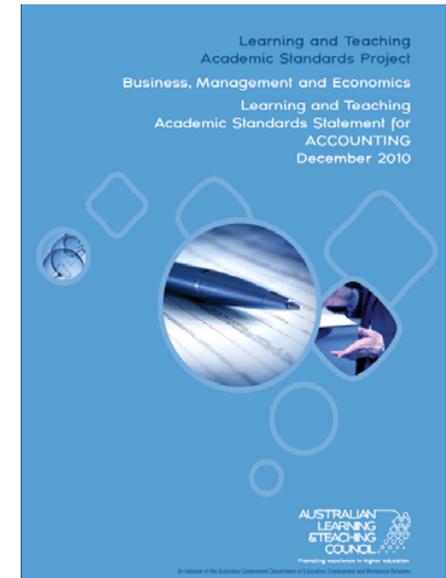
Tenet 6: Price et al (2008)

United Kingdom

- Calibration workshops trialled UK subject centre (Rust, 2009)
- Major review recommends sector wide calibration (HEFCE, 2015)
- ‘Degree Standards’ project (£1.8m ~ A\$3.0m) to engage all 15,000 external examiners over 2016-18 in calibration (HEFCE, 2016)

Accounting learning standards

- Collaboratively developed 2010
 - ALTC commissioned
 - 3 cycles of consultation
 - 2,100 participants nationally
 - Deans council endorsed
- Collaboratively assessed 2011-2016
 - Deans council commissioned with professional bodies
 - 2 academics from each institution
 - 10 cycles of calibration includes practitioners
 - 5 cycles review (double blind; 2-3 external + 1 internal)
- Collaboratively revised 2016
 - Academic association commissioned



Accounting learning standards – Bachelor & Masters

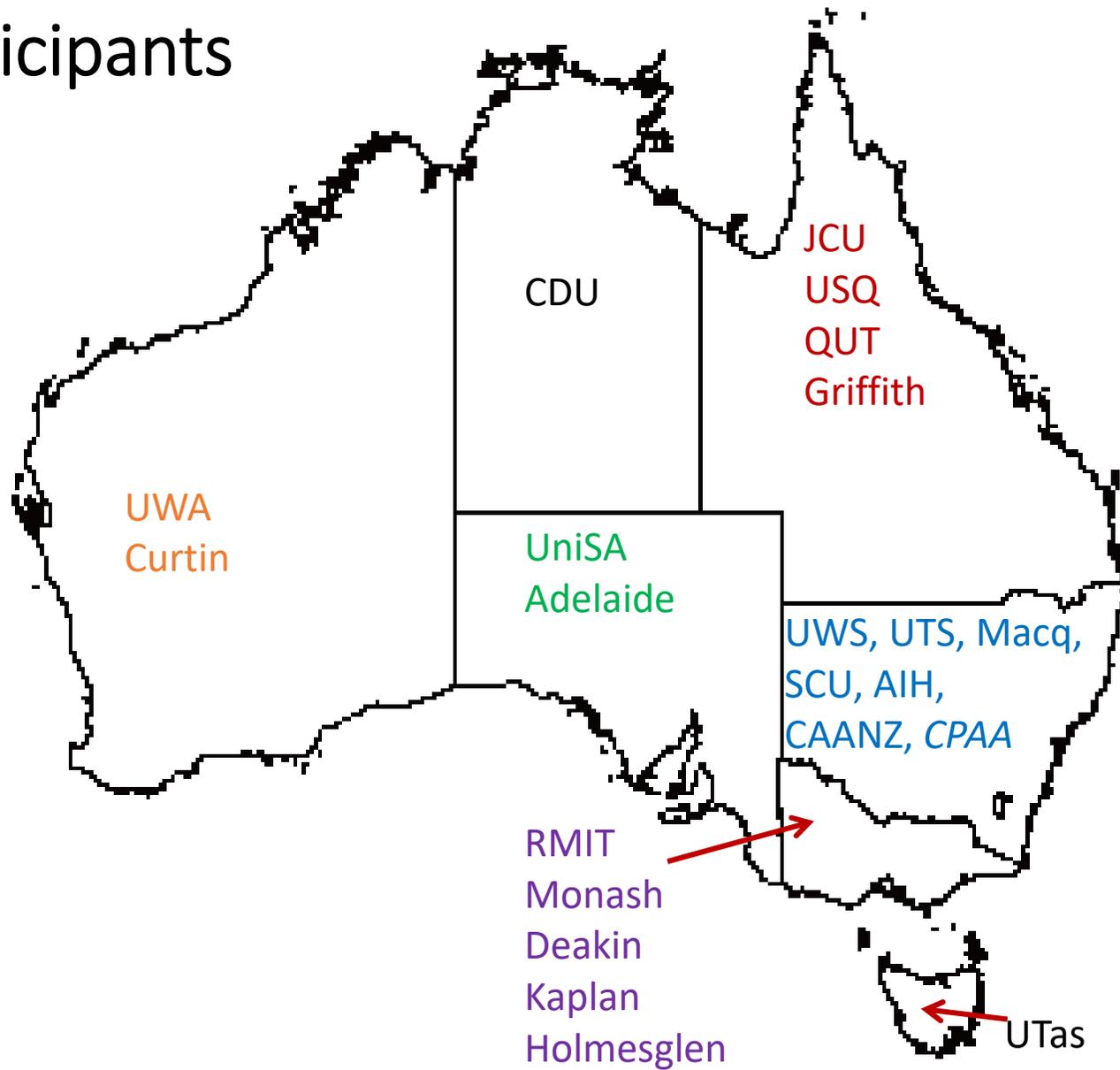


Accounting learning standards – Revised 2016



advanced
↓
Critically apply theoretical and technical accounting knowledge and skills to provide possible solutions to ~~routine~~ business issues
↑
emerging and/or advanced

Participants



External peer review 2011-16

- Peers
 - 2 academics from 18 universities + 3 other
 - Submit de-identified assessment data online
 - Participate as reviewer only after calibration (often next day)
 - Double blind review task. If valid, review sample student work
- ‘Editor’
 - Chooses standards for review & random numbers for student work
 - Facilitates calibration workshops + invites practitioners
 - Allocates 2 external (3rd if disagree) + 1 home reviewer
 - Collates report to department head
- More: <http://achievementmatters.com.au>



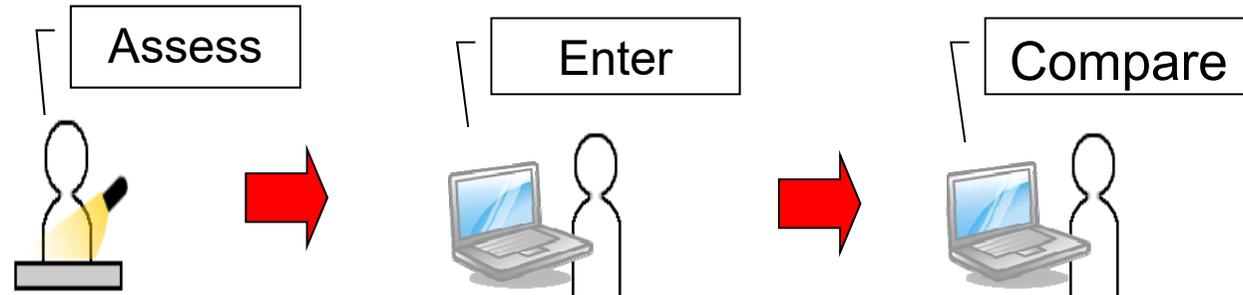
Martin Morris 20/2/12

<http://www.flickr.com/photos/martin55/6929721749/>

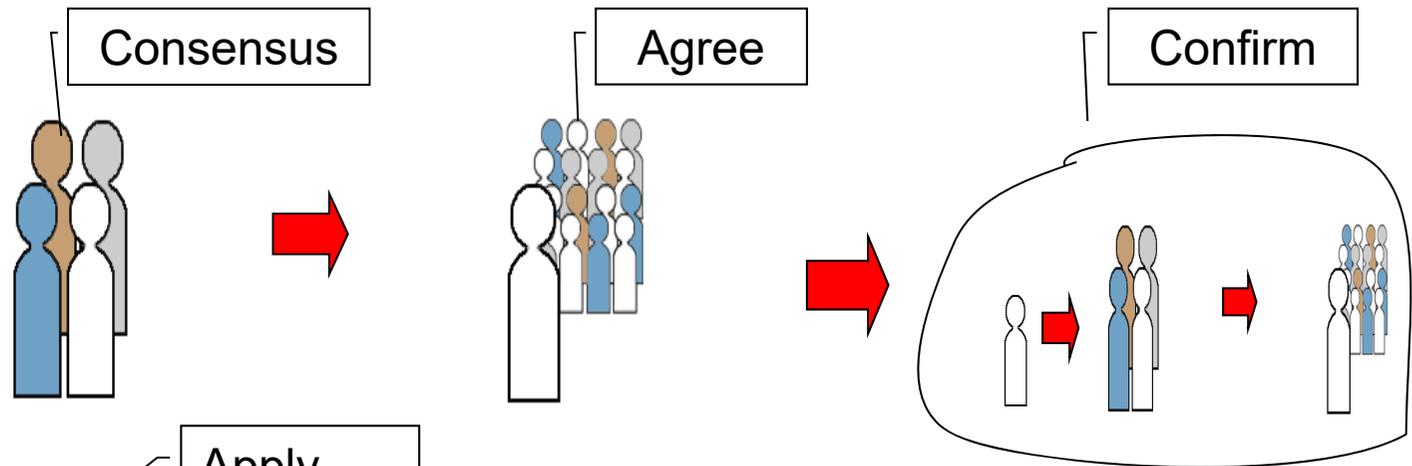
Date	Location	Learning standard	Level	External calibration	External review
Jul 2011	Darwin	<ul style="list-style-type: none"> Written Communication 	Bachelor	✓	x
Sep 2011	Melbourne	<ul style="list-style-type: none"> Written Communication 	Master	✓	x
Feb 2012	Adelaide	<ul style="list-style-type: none"> Written Communication 	Bachelor	✓	x
Sep 2012	Sydney	<ul style="list-style-type: none"> Knowledge Written Communication 	Master	✓	Delay
Feb 2013	Adelaide	<ul style="list-style-type: none"> Knowledge Oral Communication 	Master	✓	Next day
Jul 2013	Perth	<ul style="list-style-type: none"> Application Judgement Oral Communication 	Bachelor	✓	Next day
Feb 2014	Adelaide	<ul style="list-style-type: none"> Application Judgement Teamwork 	Master	✓	Next day
Jul 2014	Sydney	<ul style="list-style-type: none"> Written Communication Teamwork Self management 	Bachelor	✓	Next day
July 2015	Hobart	<ul style="list-style-type: none"> Written communication Knowledge Self management 	Masters	✓	Delay
July 2016	Brisbane	<ul style="list-style-type: none"> Written communication Knowledge Application 	Masters	✓	Delay

Process

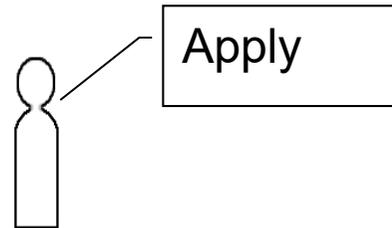
Pre-F2F



F2F
calibration



Post-F2F



ACHIEVEMENT MATTERS



User Manual / Help

2.0.0 RC20

Hi CI-Mark,

Due date: 1 Feb 2013 11:00pm

Instructor: CI-Mark Freeman

Period: Assessment

1. Drag the sliders for each of the criteria
2. Enter feedback if necessary
3. Click Save

Rating Key:

NA = Not Acceptable

A = Acceptable

**Remotely assess
(Rate, justify, advise)**

SELECT ACTIVITY:

CAS5 - Calibrating Academic Standards - Adelaide, 2013, Feb

SELECT A GROUP:

1

SELECT TASK:

1. Peer review task validity

CI-Mark Freeman

Save

TASK REQUIREMENTS - ORAL COMMUNICATION - MASTERS ENTRY GRADUATES

NA | A

The assessment task requirements are suitable for each student to independently and rigorously demonstrate that they can meet the national learning standard for oral communication.

'Master (Entry) graduates in Accounting will be able to identify and communicate accounting advice and ideas in diverse collaborative contexts involving both accountants and non-accountants.'

NB. See full document (<http://abdc.edu.au/download.php?id=325154,282,1>) also for examples and definitions (eg. 'diverse' contexts for Master entry graduates versus 'straightforward' for Bachelor graduates)

Please justify your rating. If rated 'not acceptable', suggest improvements so it would. Even if 'acceptable', suggest possible improvements. At least 20 words are needed to successfully submit your rating. (1953 words left)

There are two possible improvements to make this task acceptable to assess potential graduates against the national threshold standard for oral communication:

1. Ensure each student has an individual role (eg. external analyst; internal accountant) in the group and takes part in the presentation in that role.



User Manual / Help

2.0.0 RC20

Hi Mark,

Due date: 1 Feb 2013 11:00pm

Instructor: CI-Mark Freeman

Period: Published

WELL DONE!

The period for rating is over.

Rating Key:

NA = Not Acceptable

A = Acceptable

Remotely
compare de-
identified rating
& feedback

SELECT ACTIVITY:

CAS5 Calibrating Academic Standards - Adelaide Feb 2013

SELECT A GROUP:

Team 1

SELECT TASK:

1. Peer review task validity

Mark Freeman

Logout

TASK REQUIREMENTS - ORAL COMMUNICATION - MASTERS ENTRY GRADUATES



The assessment task requirements are suitable for each student to independently and rigorously demonstrate that they can meet the national learning standard for oral communication.

'Master (Entry) graduates in Accounting will be able to justify and communicate accounting advice and ideas in diverse collaborative contexts involving both accountants and non-accountants.'

NB. See full document (<http://abdc.edu.au/download.php?id=325154,282,1>) also for examples and definitions (eg. 'diverse' contexts for Master entry graduates versus 'straightforward' for Bachelor graduates)

Please justify your rating. If rated 'not acceptable', suggest improvements so it would be 'acceptable', suggest possible improvements. At least 20 words are needed to successfully submit your rating.

*University 4 Comments

The audience is not clearly defined in the requirements - the reference to 'a corporate presentation' does not necessarily imply an audience that is sophisticated in its understanding of accounting. In addition, it is not clear whether the students receive marks based on their individual contributions to the presentation.

Assessor 46 Comments

The assessment task requirements would be improved by:

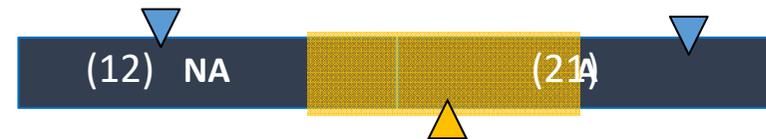
Calibration – Task validity communication

Individual confidence pre-workshop

91%

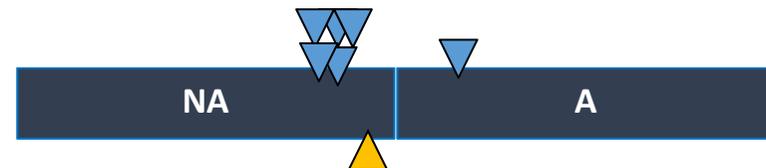
Individual results pre-workshop

- Min & max (n=33)
- Mean ± 1 SD



Group results at workshop

- Small groups (n=6)
- Consensus



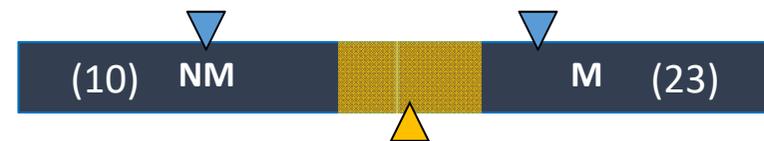
Calibration – Student 1 communication

Individual confidence pre-workshop

85%

Individual results pre-workshop

- Min & max (n=33)
- Mean ± 1 SD



Group results at workshop

- Small groups (n=6)
- Consensus



Results – Reviewer with/without calibration

Student 1 – Application standard – first time assessed				
	Experimental (with calibration)		Control (without calibration)	
	T=1 (Pre)	T=2 (Post)	T=1	T=2
Mean	65.97	61.63	76.80	73.27
SD	14.86	8.07	14.92	13.48
F Statistic		3.3907		1.2251
Prob.		p=0.002		p=0.709

Evaluation

	Agree	Disagree
• Calibration (reaching consensus on pre-workshop + new samples against agreed national learning standard) changed my understanding that might apply locally		
• Knowledge	78%	4%
• Application	74%	4%
• Written Communication	89%	4%
• Confident rating capacity of assessment requirements to allow students to demonstrate agreed learning standard		
• Knowledge	93%	4%
• Application	85%	0%
• Written Communication	100%	0%
• Confident rating students ability benchmarked against agreed learning standard		
• Knowledge	89%	4%
• Application	89%	0%
• Written Communication	100%	0%
• Overall: This project will help establish national agreement on learning standards between accounting degree providers	96%	0%

External calibration

Implications	Principles
↓ groupthink bias	<ul style="list-style-type: none"> • Require pre-workshop remote rating • Use stages – remote judge; workshop consensus team→all
↓ bias	<ul style="list-style-type: none"> • Double blind (ie. de-identify all materials & reviewers) • Small random sample student work (+ new for workshop)
↑ shared understandings of standards	<ul style="list-style-type: none"> • Collaboratively develop explicit learning standards (if n.a) • Actively interrogate explicit learning standards • Prioritise most critical standards • Calibrate task first, then if valid, student work • Collaborate externally & broadly • Consider difference for masters standard • More calibration workshops needed to reduce variability • Engage ≥ 2 permanent, experienced academics • Enfold early career academics
↑ trigger local duplication	<ul style="list-style-type: none"> • Collaborate internally to calibrate understandings • Share good practice exemplars from national workshops
↑ relevance & employability	<ul style="list-style-type: none"> • Include successful practitioners (or graduates) • Include relevant professional bodies (if exist)

External calibration

Implications	Principles
<p>↑ shared assessment resources, literacy and practice</p>	<ul style="list-style-type: none"> • Use 'good' practice exemplars eg. integrative tasks; ↑ self-assessment resources like Writers Diet; ↑ weighting (~30%); ↑ authentic tasks like different audiences; ↓ groupwork • Require minimum level feedback, even for good practice • Use holistic judgements with criteria as feedback prompts • Allow time, and 2 stages at workshop, to reach consensus • Anonymously reveal and challenge good, poor and mismatched feedback • Imitate locally. Extend to early career and casual academics and tutors. Extend to calibrating and assessing progressive achievement. Extend to students.
<p>↑ confidence externals</p>	<ul style="list-style-type: none"> • Use national workshops to identify potential examiner
<p>↑ sustainability</p>	<ul style="list-style-type: none"> • Encourage academic association to lead calibration aspect of 'college of peers' on cost-recovery basis for participants
<p>↓ costs</p>	<ul style="list-style-type: none"> • Use online tools to manage files and reviewer judgements • Schedule workshops with academic association events • Seek sponsorship from workshop host, and if possible, other stakeholders like professional bodies

External peer review/examination

Implications	Principles
<p>↑ inter-rater reliability if more calibrations, less elapsed time since, more involved</p>	<ul style="list-style-type: none"> • Only use reviewers who have participated in calibration • Offer regular calibration opportunities – a single workshop halved standard deviation • Schedule external review close to recent calibration • Use 2 external reviewers, and 3rd if they disagree • Use 1 internal reviewer to assist in local calibration • Develop strategies to evaluate reviewer calibration and optimally use and support (eg. mentor) • Use non-stratified, random, deidentified, clean sample • Compare external with internal review and prior judgement
<p>↓ workload</p>	<ul style="list-style-type: none"> • Don't assess student work if task not valid • Use online tools to manage files and reviewer judgements
<p>↑ sustainability</p>	<ul style="list-style-type: none"> • Align to academic culture practice for research publication • Encourage academic association to lead 'college of peers' for institutions without partnerships and use honoraria then. • Incorporate examining in academic workload for institutions with partnerships around external examining • Develop capstone units of study to exhibit benchmarks

Future of external calibration and peer review

- External calibration workshops
 - Schools pay yearly fee (\$1000?)
 - Professional bodies support
 - Academic association (AFAANZ) manages finances & admin
- External peer review
 - AFAANZ member pay (\$750?) & non-member pay more (\$1000?)
 - 3 calibrated reviewers invited to review and paid \$250 each
 - School receives a summary of the reviews

Key references

AAHE article: [dx.doi.org/10.1080/02602938.2015.1008398](https://doi.org/10.1080/02602938.2015.1008398)

Project report: <http://www.olt.gov.au/project-achievement-matters-external-peer-review-accounting-learning-standards>

Project website: <http://achievementmatters.com.au>

Thank you



mark.freeman@sydney.edu.au