

Improving student assessment: creating valid, fair and transparent practices in higher education

‘The assessment of students’ learning outcomes is a pressing change frontier for higher education’. (Coates, 2016)

Prof. Sue Bloxham
S.Bloxham@cumbria.ac.uk

Extract from Luth Report, 2010

“further progress requires consideration of the purposes of grading, and how grading relates to more global issues of assessment for learning, assessment of learning, alignment of assessment with course goals and objectives – indeed, with the quality of our teaching. This report is the beginning of a **conversation in the academy**, not the end of one.”

The (sometimes conflicting) purposes of assessment



assessment *of* learning

Certify achievement
License to practice
Enable selection for
further study and
employment

Assessment *for* and *as* learning

Promote learning

Diagnostic, formative,

Steering approach to learning

Informing teaching

Learning about criteria, standards
and
professional
judgment

To be able to tell
how well you
are doing



Assessment principles

- Validity
- Effectiveness (learning)
- Equity
- Transparency
- Reliability
- Comparability and consistency



Principles of assessment - UAPPOL

- Assessment should be integrated into and aligned with the learning experiences and stated objectives/ outcomes of a course and program;
- Policy..... does not limit the development of other, additional, innovative forms of effective assessment;
- General assessment methods and grading standards must be communicated clearly to students at the beginning of the course or program of study;
- Clear and transparent assessment criteria should be provided to students throughout the course;
- Provide reliable and valid information in which students, prospective employers and accrediting bodies can have confidence;
- Assessment should be varied and timely. Student achievement and performance should be assessed in a formative manner during a course and in a summative manner both during and at the end of a course and program;
- In the design, delivery and reporting of summative assessments, the University is committed to open, accountable and equitable processes.

Principles of assessment - UAPPOL

- Assessment should be integrated into and aligned with the learning experiences and stated objectives/ outcomes of a course and program;

Validity

- Policy..... does not limit the development of other, additional, innovative forms of effective assessment;

Validity/ Learning

- General assessment methods and grading standards must be communicated clearly to students at the beginning of the course or program of study;

Transparency

- Clear and transparent assessment criteria should be provided to students throughout the course;

Transparency

- Provide reliable and valid information in which students, prospective employers and accrediting bodies can have confidence;

Reliability/ Validity

- Assessment should be varied and timely. Student achievement and performance should be assessed in a formative manner during a course and in a summative manner both during and at the end of a course and program;

Learning

- In the design, delivery and reporting of summative assessments, the University is committed to open, accountable and equitable processes.

Equity/ Transparency

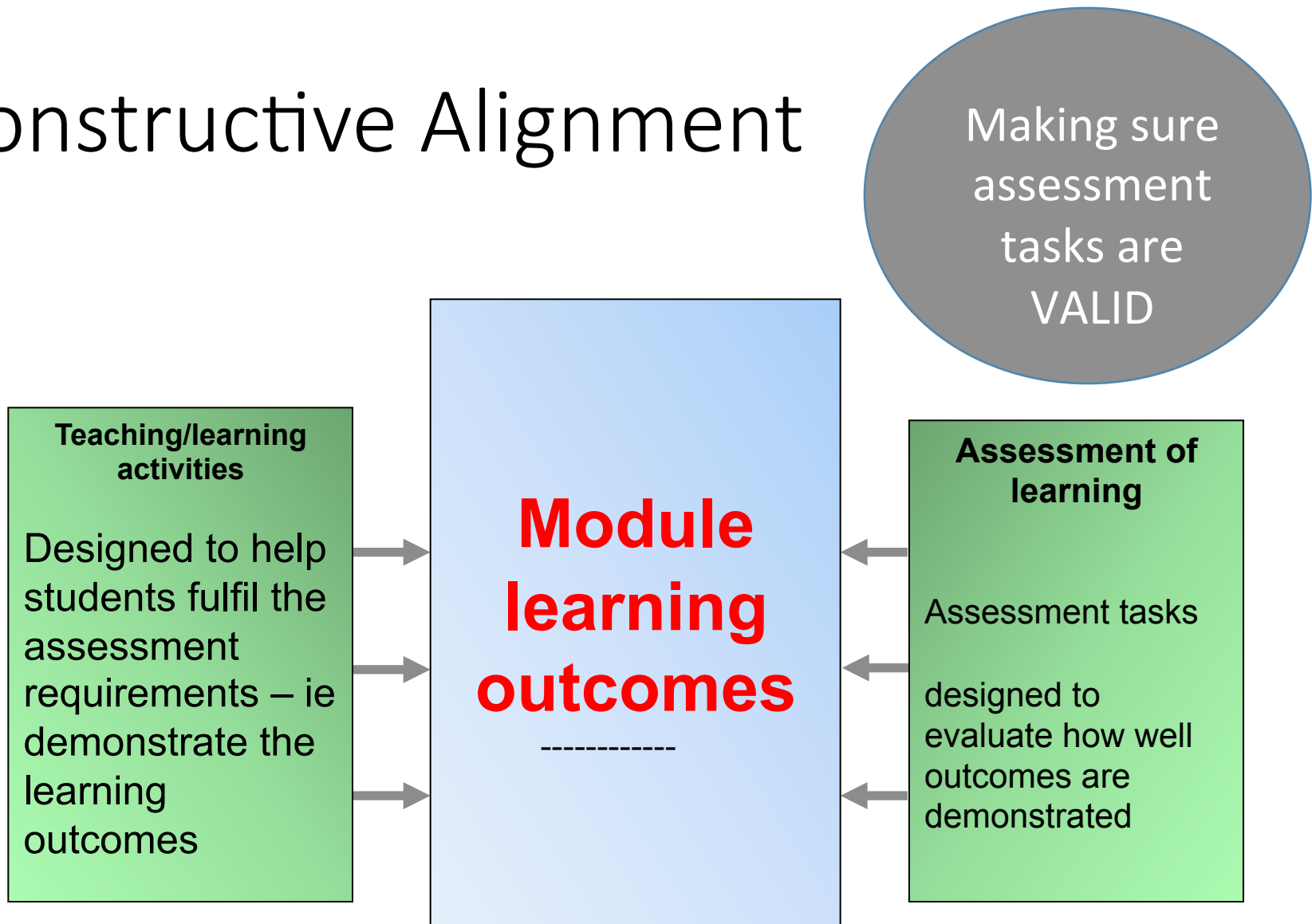
Approaches to curriculum

- Traditionally curriculum design has started with the content of the course, which is then used to decide the teaching method (i.e. how the content will be covered) and subsequently how it will be assessed
- The 'outcome-based' approach has emerged over the last few decades where the focus is on, not the content that the student should have accumulated, but **what they are able to do** by the end of the programme or module.

It is student-learning centred rather than content/ teacher centred.



Constructive Alignment



Adapted from Biggs & Tang (2011)

21st Century higher education

- Skills for the ‘knowledge economy’*
 - Critical thinking and problem-solving
 - Collaboration across networks and leading by influence
 - Agility and adaptability
 - Initiative and entrepreneurialism
 - Effective oral and written communication
 - Accessing and analysing information
 - Curiosity and imagination

* ‘Must have’ skills for the future to tackle the ‘Global achievement gap’ (Wagner, 2008)

Do these attributes figure in Program/course outcomes and do assessment tasks foster and test these broader skills and capacities?



The Taxonomy Table

(Anderson and Krathwohl, 2001)

Knowledge Dimension	The cognitive Process Dimension					
	Remember	Understand	Apply	Analyse	Evaluate	Create
Factual	MCQ test	MCQ test	A problem sheet of equations	Design & conduct experiment	Research project	Phd Thesis
Conceptual	MCQ test	Essay exam	Present a legal argument			
Procedural						
Metacognitive				Self/peer assess	A reflective log of practicum learning	

(Adapted from Anderson, 2003:29)

Discussion

- Consider your existing assessment methods and identify the challenges for achieving really valid assessment of course learning outcomes;
- How well do assessment methods match learning outcomes for 21st century graduates.



Reliability

Assessment tasks should be generating similar marks across time, across markers and across methods – consistent, dependable, reproducible

When

- ✓ different markers make the same judgement about an assignment
- ✓ one marker makes consistent judgements about a piece of work at different times;
- ✓ Students receive similar grades when the same learning outcome is assessed in different ways



Unreliable grading

Many studies show poor reliability amongst those assessing complex performance at higher education level (see Bloxham 2012 for a review)

- ✓ Grading habits – markers focus on different aspects of student work
- ✓ Different expectations of standards at different grade levels
- ✓ Ignoring or choosing not to use the criteria
- ✓ Different interpretation of criteria or standards as a result of varying professional knowledge, experience, values,
- ✓ Use personal criteria different to those stated.
- ✓ Importance given to different criteria

The reliability dilemma

The greater the drive for reliability, the more we choose assessment methods which do not allow for **varying**, **unpredictable** and **unique** responses.

Yet higher order, **complex performance** necessarily involves work with these characteristics



Other ways to build trust in grades:

- ✓ Multiple judgements
- ✓ Calibration of markers
- ✓ Moderation of grading

Transparency

What it means:

Information, guidance, rules and regulations on assessment should be clear, accurate, consistent and accessible to all staff and students.

Did students understand what was required of them to succeed in the task set for them?



This is a key element of **fairness**.

There has been lots of effort to improve transparency for students – but there is still a **mismatch between what we say and what we do**

indispensable conditions for improvement (Sadler 1989)

- the student comes to hold a concept of quality roughly similar to that held by the teacher,
- is able to monitor continuously the quality of what is being produced during the act of production itself.
- And has a repertoire of alternative moves or strategies from which to draw

(knows the standard or goal)

(recognises the 'gap' between their work and the standard)

(can take action to fill the 'gap')

What we say and what we do

What our guidance implies	Reality of academic judgement
standards are explicit in criteria and rubrics	Markers' standards are internalised and tacit (difficult to express)
Criterion-referenced grading	Norm-referencing – grading to a curve or mean
There is a 'right' mark	Marker variation

Overall, our judgements can lack transparency for students – it's not surprising that they complain about grading?

Transparency - where final course grades come from

- Students should know how grades earned during a course combine to make up the final grade.
 - Are some assignments and exams weighted more than others;
 - are some assignments not counted;
 - are there marks for attendance and participation for which the criteria/ grading method is not explicit.



Improving transparency: course information

Clear information in the course syllabus on:



- ✓ the assessment methods;
- ✓ assessment dates;
- ✓ assessment criteria;
- ✓ how different marks will be combined to create the final grade;
- ✓ How class attendance/ participation grades are calculated.

Improving transparency:

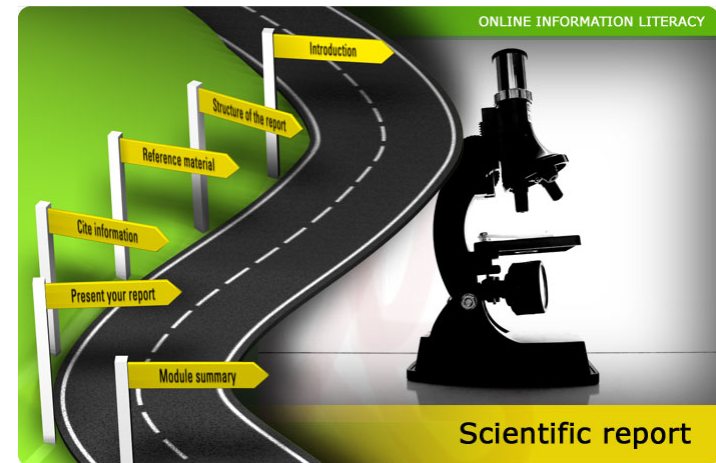
It is easier to convey the required standards for assessment through real examples rather than through descriptions of standards:

- Practice grading
- Exemplars
- Peer review



Transparency through exemplars: Scientific reports

- Biological sciences
 - Students mark two exemplars of a scientific report – one high & one medium standard
 - Student discussion about marks given and justification
 - Tutor tells and explains actual marks awarded
 - Followed by students peer reviewing each others' draft assignments
 - Extremely positively evaluated by students – increased confidence & reduction in anxiety
- Yucel et al (2014)



General benefits of exemplars

- Students & staff very positive
- Enable students to better grasp overall quality, structure, language, style than lists of individual criteria;
- Contributes to confidence to tackle assessments;
- Evidence of improved marks across different ability levels;
- Helps students recognise different 'expressions of quality'.



Discussion

- Consider how transparent the assessment guidance, regulations, requirements and criteria are for your students?
- Can you share any ways you help students to understand assessment requirements more clearly?



Comparability and consistency

Comparability

- students studying for the same courses and awards have a comparable experience, with similar standards and workload applying across sections and courses.
- Also there are comparable approaches to the summative assessment requirements of awards of the same level across an institution.

Consistency

- Assessment standards and outcomes are maintained across cohorts and years.

Above all, comparability and consistency are related to assuring quality standards and fairness

Grading to a curve / mean: upside

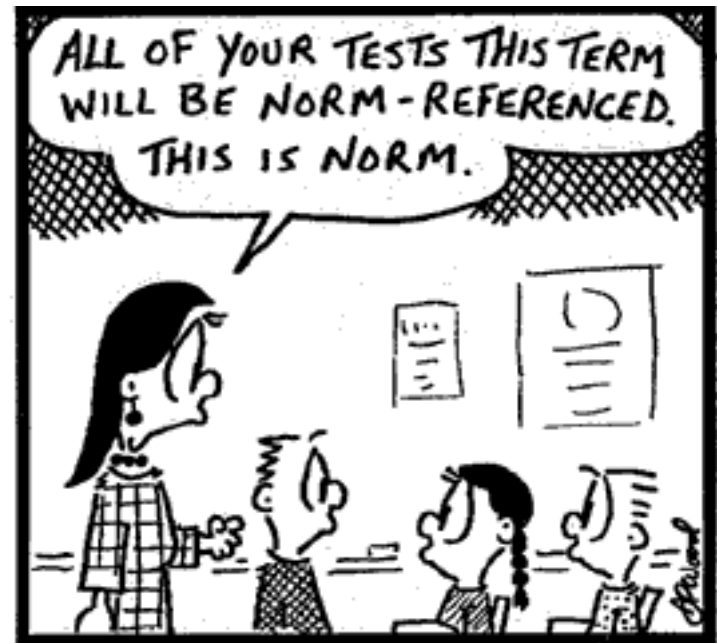
- Offers students the protection that they won't be disadvantaged by 'hard' tests or assignments;
- Gives the impression of comparable achievement across sections, courses and programs;
- Can prevent grade inflation;
- Can prevent students being disadvantaged by poor course design, teaching or assessment methods.

However, none of these is a good pedagogical reason for grading to a curve or mean.



Reference to other students' performance does 'not provide a stable basis for maintaining standards because as a basis for grade decisions they are unpredictable and highly variable'.

Orrell (2008, p. 259)



Grading to a curve or mean: downside

- Students achievement is controlled by ability and effort of peers as much as their own – penalizes students in strong cohorts;
- Students may not know what is needed to achieve a grade because it isn't fixed;
- Discourages peer collaboration;
- Benefits of good teaching are not recognized;
- Students' actual achievement is not transparent – have they met the stated outcomes for the course?

Grading to a curve or mean is not a fair, valid or transparent way to create comparability and consistency

The power of peer learning

Spanish Computing course

- Students worked on problems in groups, then sat an individual test;
- Student test grades were made dependent on the individual grades obtained by the weakest member of their group.
- Students who did this did significantly better – both individually and collectively – than those not offered this opportunity.
- In this way, students are encouraged to help each other, and their level of commitment is increased.
- Students were positive about the method.

See Arevalillo-Herráez & Claver, 2014



Creating comparability and consistency

- ✓ Course learning outcomes scrutinised for match to level of study (cf: Canadian Degree Qualifications Framework);
- ✓ Assessment methods scrutinised for alignment with learning outcomes (and thereby level of study); **i.e assessment validity**
- ✓ All sections aim for same learning outcomes and undergo same summative assessment;
- ✓ Moderation of marking standards for open-ended, complex work.



Moderation/ calibration in the assessment cycle

Stage of assessment	Activities
Design	<ul style="list-style-type: none">• Peer scrutiny of assessment design – what methods are you using, do they test the learning outcomes?
Judgement	<ul style="list-style-type: none">• Pre-teaching discussion of assessment tasks and exemplars to inform teaching and judgement;• Moderation discussion about student work (consensus/social moderation);• Consideration of grade distributions to identify if there are legitimate reasons for differences.
Program evaluation	<ul style="list-style-type: none">• External experts – providing feedback on the standards of student work in ‘terminal stages’ of program.

Achieving meaningful comparability and consistency

Consider replacing the superficial manipulation of grades with a clearer focus on the valid assessment of learning outcomes. This means:

- promoting the enhancement of assessment design, clearly aligned to appropriate learning outcomes;
- considering opportunities for moderation of grading;
- considering opportunities for the calibration of instructors;
- using distribution of grades to identify differences in student achievement – not mask them.

“The institution must have a process to maintain the currency of the program and the quality of its learning outcomes” (CAQC, 2015 p56) – surely evaluation of student assessment outcomes is the most valid process.

Discussion questions

1. Are we happy with existing assessment and grading practices?
2. What influences your decisions on assessment methods and grading?
3. How can we make assessment more valid and transparent?
4. How can we achieve comparability and consistency of grading without merely manipulating the distribution of grades?
5. How can we help students gain a good understanding of what they need to do to achieve?
6. How can we give department chairs confidence that course grades are an accurate reflection of attainment even if they are high or low compared to historical trends.
7. What other questions do we have?

References

- Anderson, L.W., and Krathwohl, D.R. (2001) A taxonomy of learning, teaching and assessing: a revision of Bloom's taxonomy of educational objectives. New York: Longman.
- Anderson, L.W. (2003) Classroom Assessment: Enhancing the Quality of Teacher Decision Making. Mahwah, N.J. Lawrence Erlbaum Associates, Inc.
- Arevalillo-Herráez, M & [Claver](#), J.M. (2014) Impact evaluation of reactive assessment strategies to address social loafing by promoting student cooperation and encouraging mutual support. *Innovations in Education and Teaching International* Volume: 51 Issue: 5 Pages: 523-532
- Biggs, J & Tang, C (2011) Teaching for Quality Learning at University Buckingham, 4th Edn: Open University Press
- Bloxham (2012) 'You can see the quality in front of your eyes': grounding academic standards between rationality and interpretation *Quality in Higher Education*, 18 (2) 185-204
- Campus Alberta Quality Council (2015) Handbook: quality Assessment and Quality Assurance. Edmonton:CAQC
- Coates, H (2016) Assessing student outcomes internationally: insights and frontiers. *Assessment & Evaluation in Higher Education*. 41 (5) 662-676
- Luth, R (2010)[Assessment and grading at the University of Alberta: policies, practices, and possibilities. A report to the Provost and the University by Robert W. Luth, Provost's Fellow 16 June 2010]

- Orrell, J., 2008, 'Assessment beyond belief: the cognitive process of grading', in Havnes, A. & McDowell, L. (Eds.) *Balancing Dilemmas in Assessment and Learning in Contemporary Education* (London, Routledge).
- Sadler, D.R. (1989) Formative assessment and the design of instructional systems, *Instructional Science*, 18 (2):119-144.
- Wagner, T. (2008) *The Global Achievement Gap: Why Even Our Best Schools Don't Teach The New Survival Skills Our Children Need—And What We Can We Do About It* Basic Books.
- 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 & evaluation in higher education*, 39 (8) 971-986