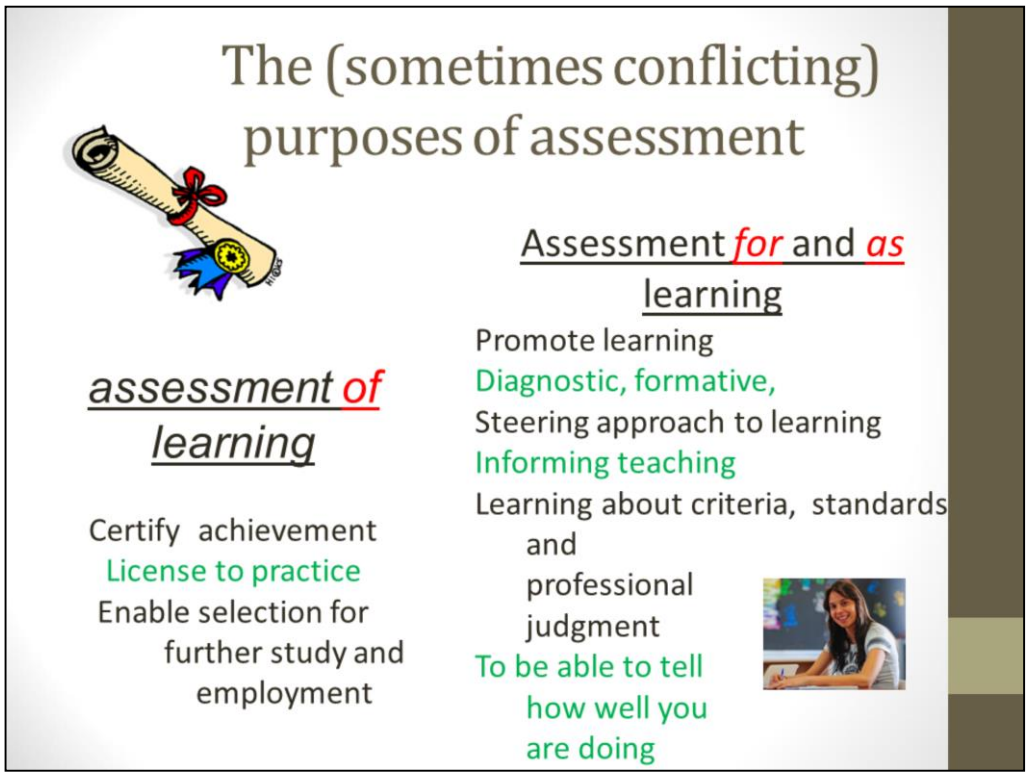
It has been established for decades that assessment defines what students regard as important, how they spend their time and how they come to see themselves as students and then as graduates (for example, Boud 1988; Brown, Bull, and Pendlebury 1997)

There have been multiple and sustained efforts to improve HE assessment over the last 20 years and large sums of public money has been spent but, as a JISC report stated in 2013 it remains 'stubbornly resistant to change' (JISC 2012). Ferrell, J (2012) A view of the Assessment and Feedback Landscape: baseline analysis of policy and practice from the JISC Assessment & Feedback programme, JISC. http://www.jisc.ac.uk/media/documents/programmes/elearning/Assessment/JISCAFBaselineReportMay2012.pdf [accessed 4 April 2014]

Before we look at the problems and potential solutions, let's consider what we are trying to do with assessment – its purposes.



Let's briefly look at these broad purposes of assessment. As I've shown in this slide, assessment must serve two main masters – (as David Boud says – it has to do 'Double Duty'). Traditionally we recognise that assessment has an important purpose in selection and certification – testing students have met course requirements, ranking them, assuring that they deserve a license to practice in professional fields, enabling selection for jobs, scholarships, etc. This is assessment of learning.

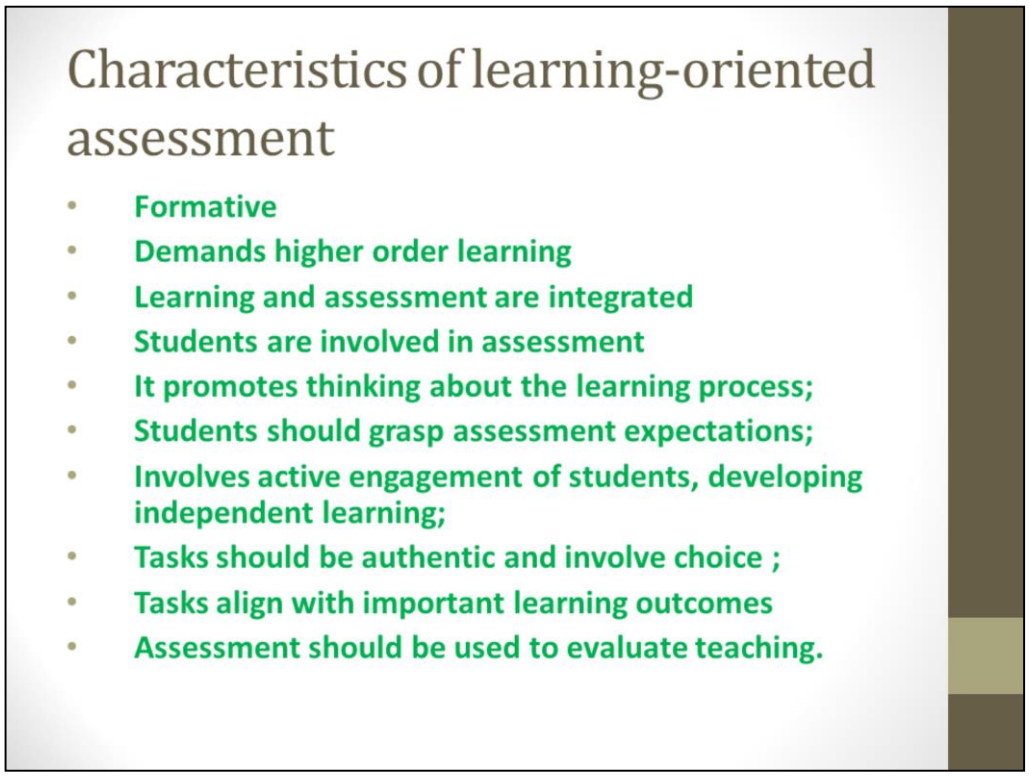
but assessment also has an important role in promoting effective learning (assessment *for* and *as* learning). Assessment *for* learning is designed to be formative and diagnostic, providing information about student achievement to both teachers and learners which allows teachers to respond to the needs of the learner. It recognises the overwhelming evidence that assessment and feedback at a formative stage can improve student achievement and inform better teaching.

These multiple purposes for assessment are reflected in the basic principles for assessment design which you will find in any recent textbook on the subject and around which I am going to structure my discussion about how we might improve our practices.

Assessment is critical to learning

assessment is critical in the learning process (Kearny, Perkins & Kennedy-Clark, 2015) because

- Its backwash effect on all teaching and learning activities (Boud, 2007; Havnes, 2004; Watkins, Dahlin, & Ekholm, 2005).
- The anticipation of assessment has a strong influence on what and how students study. (Boud, 2007; 2010),
- it drives and models the learning process (Jensen, McDaniel, Woodard, & Kummer, 2014; Vu & Dall'Alba, 2014).
- It can encourage or inhibit plagiarism (http://www.itlal.org/index.php?q=node/155)

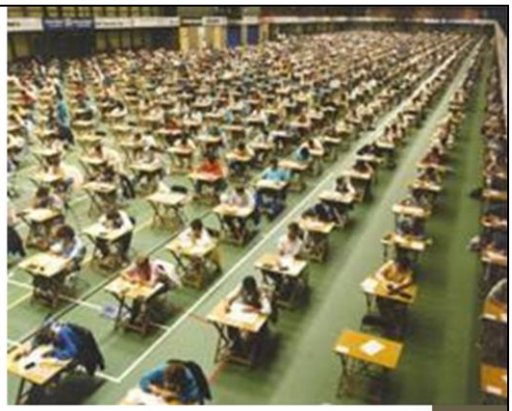


So a first important reason is to make sure that our assessment is valid in assessing the learning outcomes for a programme, including modern learning outcomes that prepare students for life and employment in a 21st century world.

Secondly, we need to consider how assessment will encourage learning, not just test what has been learnt. Research suggests that the characteristics of learning-oriented assessment are:

- Assessment should have a formative function, providing 'feedforward' for future learning which can be acted upon. There is opportunity and a safe context for students to expose problems with their study and get help; their should be an opportunity for dialogue about students' work;
- Tasks should be challenging, demanding higher order learning and integration of knowledge learned in both the university and other contexts;
- Learning and assessment should be integrated, assessment should not come at the end of learning but should be part of the learning process;
- Students are involved in self assessment and reflection on their learning, they are involved in judging performance;
- Assessment should encourage metacognition, promoting thinking about the learning process not just the learning outcomes;
- Assessment should have a formative function, providing 'feedforward' for future learning which can be acted upon. There is opportunity and a safe context for students to expose problems with their study and get help; their should be an opportunity for dialogue about students' work;
- Students should grasp assessment expectations;
- Tasks should involve the active engagement of students developing the capacity to find things out for themselves and learn independently;
- Tasks should be authentic; worthwhile, relevant and offering students some level of control over their work;
- Tasks are fit for purpose and align with important learning outcomes
- Assessment should be used to evaluate teaching as well as student learning

Exams

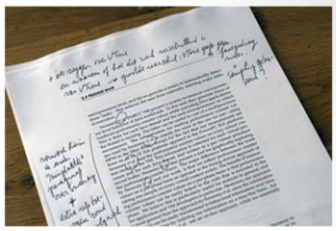


- They come at the end of a course,
- They are not integrated into the learning,
- The criteria are oblique,
- They rarely result in useful feedback,
- Time limits mean they struggle to test complex higher order learning involving use of new and complicated evidence,

- They systematically disadvantage students for whom English is a second language.
- They require 'special arrangements' for increasing numbers of students.

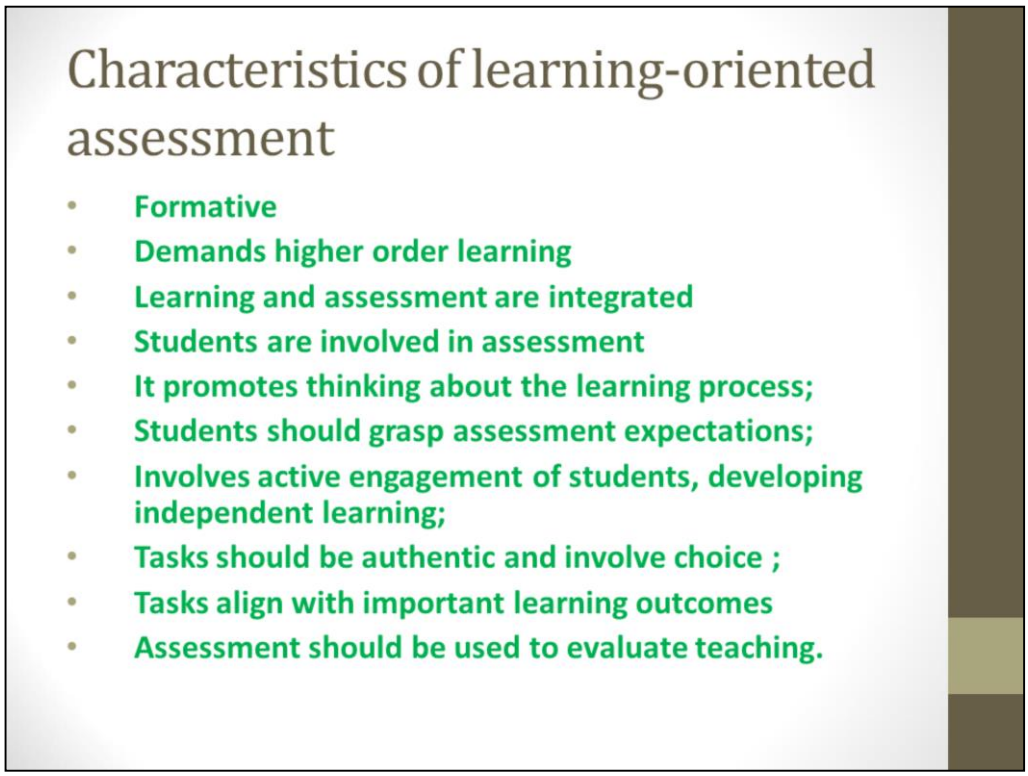
In addition to the failure to support learning, they have some other serious disadvantages – see shaded box

Essays



have the potential to meet many characteristics but:

- Often no formative element
- Questions may ask students to 'evaluate' or 'critically assess' a topic but if students can pass adequately by regurgitating others' evaluation or criticism (from lectures or reading), they may avoid higher order learning.
- Students not involved in assessment or thinking about learning
- Rarely authentic
- Criteria often oblique
- Easier to plagiarise



- So what can we do about it?

- How can we make our assessment more learning oriented in a practical way that doesn't overburden us with work?

- I haven't got time today to talk about all these characteristics although you will see that they are hard to separate in good assessment/ teaching design.

- However, I'll start with the most important - formative assessment

- Much formative assessment in HE rests on the idea that feedback on summative assessment will be used to improve future work. There are lots of problems with that assumption. Let's start with the difficulty with our way of doing feedback.

Feedback – signal and response

Room too cold

Thermostat senses low temperature

Sends signal to boiler

Room heated to comfortable temperature

Let’s think of this in the context of feedback as part of a monitoring system

Feedback in higher education?

Signal is ‘dangling data’ – no response required or checked for – we often don’t know if the student has grasped what you say or made any adjustment to their understanding.

Engaging students in formative assessment – key requirements

- It clearly feeds into summative assessment tasks;
- The students must submit it in some way (bring to class, post on line, hand it in, do in class) and action is taken if they don't;
- Students receive useful feedback on it;
- It is not contaminated by summative purposes.

Formative assessment integrated into teaching

- http://www.youtube.com/watch?v=wont2v_LZ1E

The assessment of what students know is seamlessly blended with teaching. The students learn whether they understand, the lecturer knows if more 'teaching' is needed.

Eric Mazur, teaching physics, Harvard

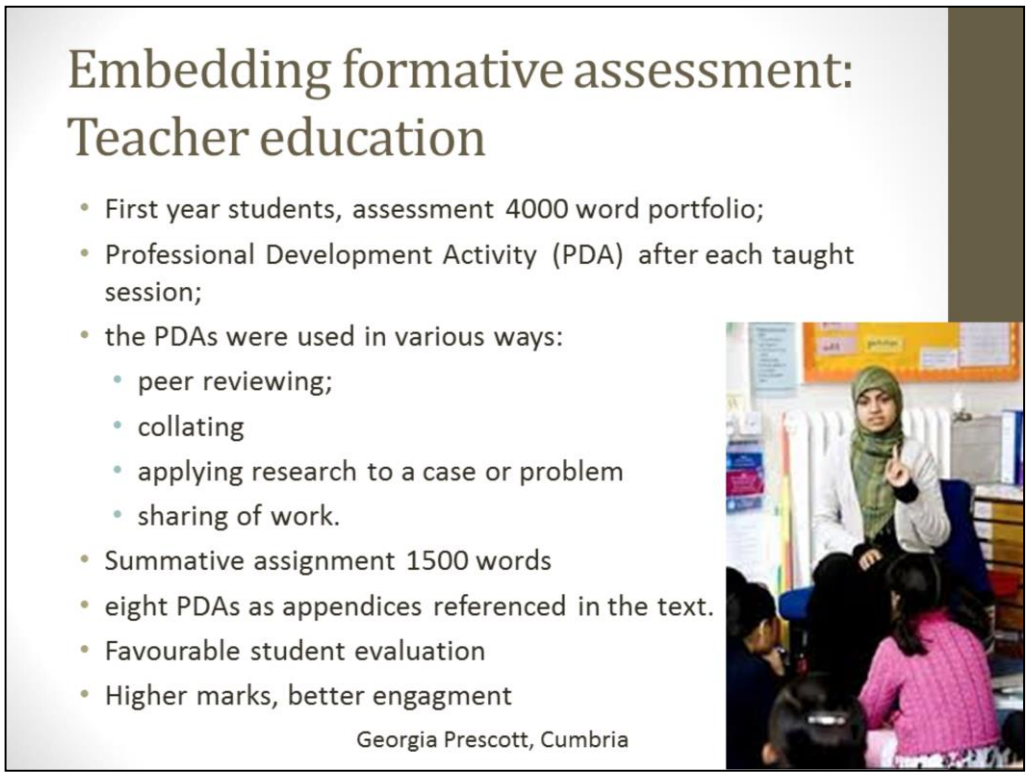
Note that learning to drive and the PhD embody much of what we are now trying to achieve for other students: The learning and the assessment are one and the same thing – fully integrated - although we don't call it assessment most of the time because it is largely formative and low stakes. The teacher is constantly assessing the student's progress and providing immediate feedback. The student will normally act on the feedback because it is seen as useful as it comes at a point where they can use it to improve

Note the opportunity in the video for students to:

develop learning through explaining and listening;
Get immediate feedback on whether they understand something.

And for the lecturer to:

Get immediate feedback on whether the students have understood;
Create engaged students.



A Level 4 undergraduate module in teacher education, assessed through a 4000 word 'portfolio', was redesigned to increase the involvement of students throughout the module and spread their workload whilst reducing the marking. Students completed a Professional Development Activity (PDA) in their own time after each taught session in preparation for the following week. The PDAs were an extension or application of learning from the session. The following week, the PDAs were used in various ways, such as peer reviewing; collating or applying research; or sharing of work. One week, in groups of four, students peer reviewed a mini essay, commenting in turn on content; academic conventions and writing, and the use of description /analysis. At the end, they wrote how they would improve their writing in the final assignment. They found this formative assessment very useful both as a giver and recipient of feedback.

Because the PDAs involved a lot of work, the final summative assignment was reduced to 1500 words. Students had to submit the eight PDAs as appendices with the expectation that they would make some explicit references to them in their work.

In module evaluations the students overwhelmingly commented very favourably about the module and its approach to assessment. They liked the spread of workload and the formative assessment feedback they received. Engagement with the PDAs increased because these were always used in the following session, thus creating a sense of real purpose. If they weren't completed, students might let someone else down, or squander an opportunity for personal feedback.

The PDA follow up activities brought them closer together as a group because they involved sharing personal perspectives, or because they had to work collaboratively to create a joint product. I felt they were more involved and engaged in their learning because they had made a greater prior investment into the sessions.

My marking load was reduced. I did scan through the PDAs but did not use them to grade the assignment. I commented explicitly when I saw that they had clearly improved their final piece of work in response to previous formative feedback, which many of them had.

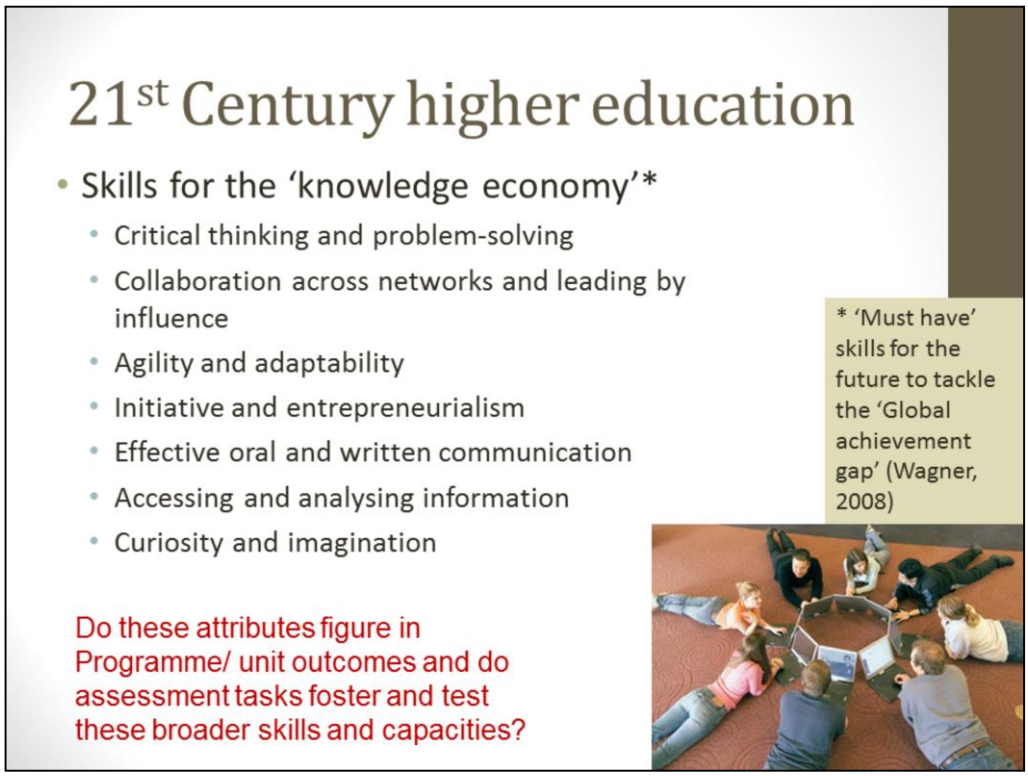
Group based formative assessment: Psychology

- 560 students in groups of 6-7;
- 3 week cycle culminating in 700-800 word essay

  e.g. 'Assess the strengths and weaknesses of Freud's and Eysenck's theories of personality. Are the theories incompatible?
- Guidance provided for tackling the question and working in a group;
- Best definitions & essays posted on VLE as feedback;
- Students used familiar language to discuss academic concepts – Dialogue and explanation.

Nicol 2009

University of Strathclyde Glasgow



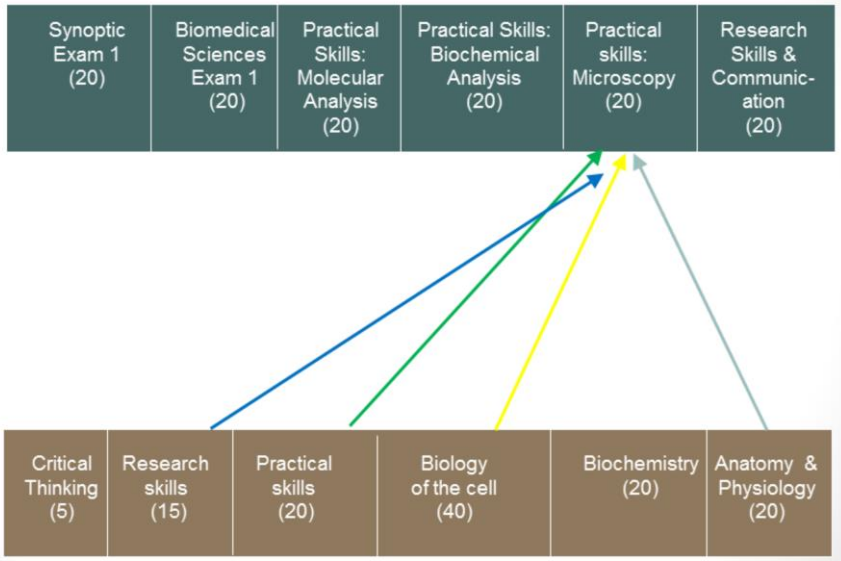
A third characteristic of learning oriented assessment is that it links to important learning outcomes. Let's just think This issue of validity in assessment is important as the knowledge we expect students to acquire in a modern university programme. We have now come to understand that factual and conceptual knowledge is not enough for an effective graduate. For example, there is a perception, particularly amongst employers, that higher education is not 'fit for purpose'. It is not providing graduates with the skills and attributes they require to deal successfully with a complex and rapidly changing world; a world which needs graduates to be creative, capable of learning independently and taking risks, knowledgeable about the work environment, flexible and responsive. Whilst we might argue with this view about the shortcomings of higher education, we can certainly improve how our assessment fosters and encourages those qualities in our graduates.

To become a good professional in any field of study, it is not only necessary to master the knowledge and disciplinary or technical skills, professional competences are also required (Guzzomi, Male & Miller, 2015). On response is that we need to think about learning and assessment activities which are more authentic, holistic and collaborative (Jackson, 2013). Professional competences are not separate skills that are learned in isolation (Jackson, 2013), they need to be contextualized and problematized to be discussed together with others. For example, the ability to ask questions, share ideas with other colleagues, validate decision-making within work teams, find meaning in experiences, incorporate ideas from others, and reflect on professional practices (Smith & Trede, 2013).

Integrated programme assessment – Bio Science at Brunel

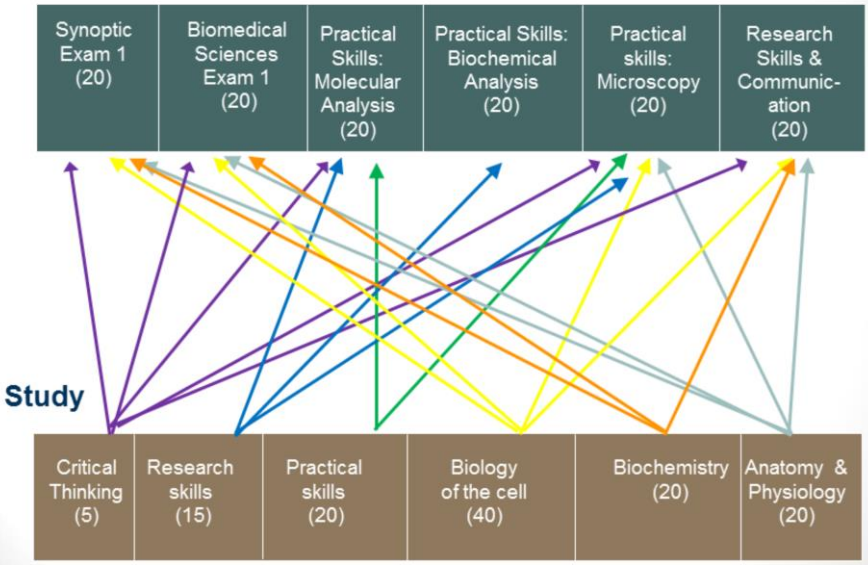
- Uni regs require 120 credits of **assessment** per year/level, not 120 credits of modules, freeing up the separation of teaching from assessment.
- Staff identified key learning outcomes for each level and created a range of methods to assess these LOs in an **integrated** and **holistic** way.

IPA - Level 4


Synoptic Exam 1 (20)
Biomedical Sciences Exam 1 (20)
Practical Skills: Molecular Analysis (20)
Practical Skills: Biochemical Analysis (20)
Practical skills: Microscopy (20)
Research Skills & Communic-ation (20)
Critical Thinking (5)
Research skills (15)
Practical skills (20)
Biology of the cell (40)
Biochemistry (20)
Anatomy & Physiology (20)

IPA - Level 4

Assessment



Benefits

Staff

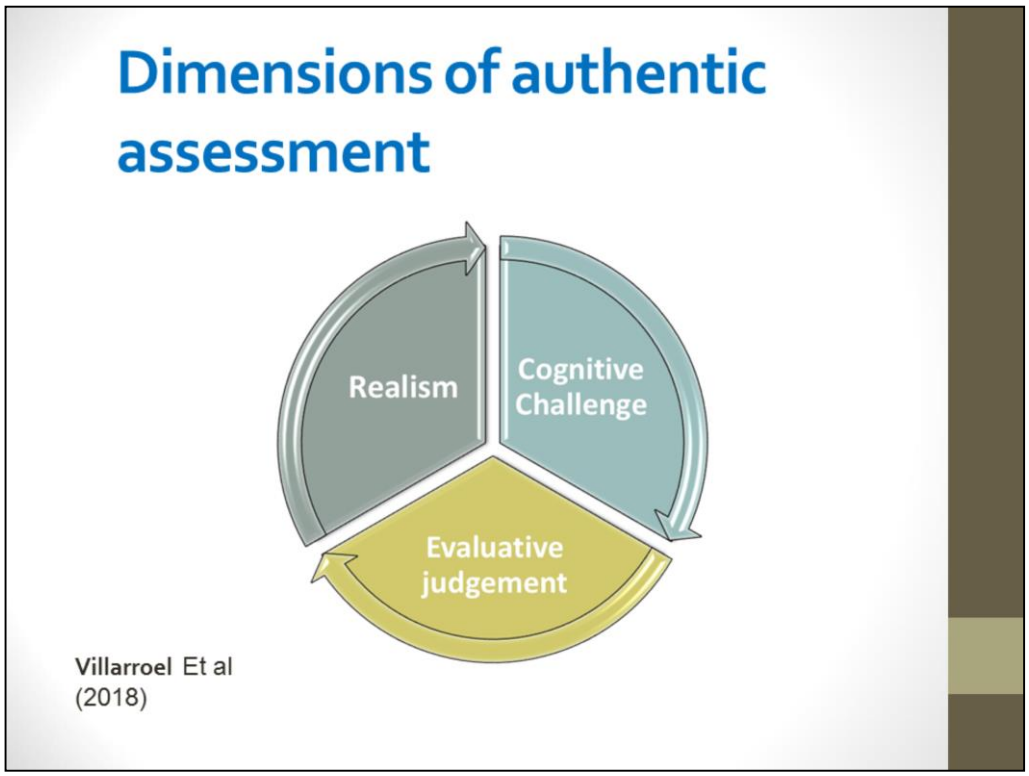
- Assessment burden is reduced
- Marking is shared
- Teaching has become a 'community property'

Students

- Fewer more interesting assessments
- Formative activities supports learning
- Recognise graduate attributes

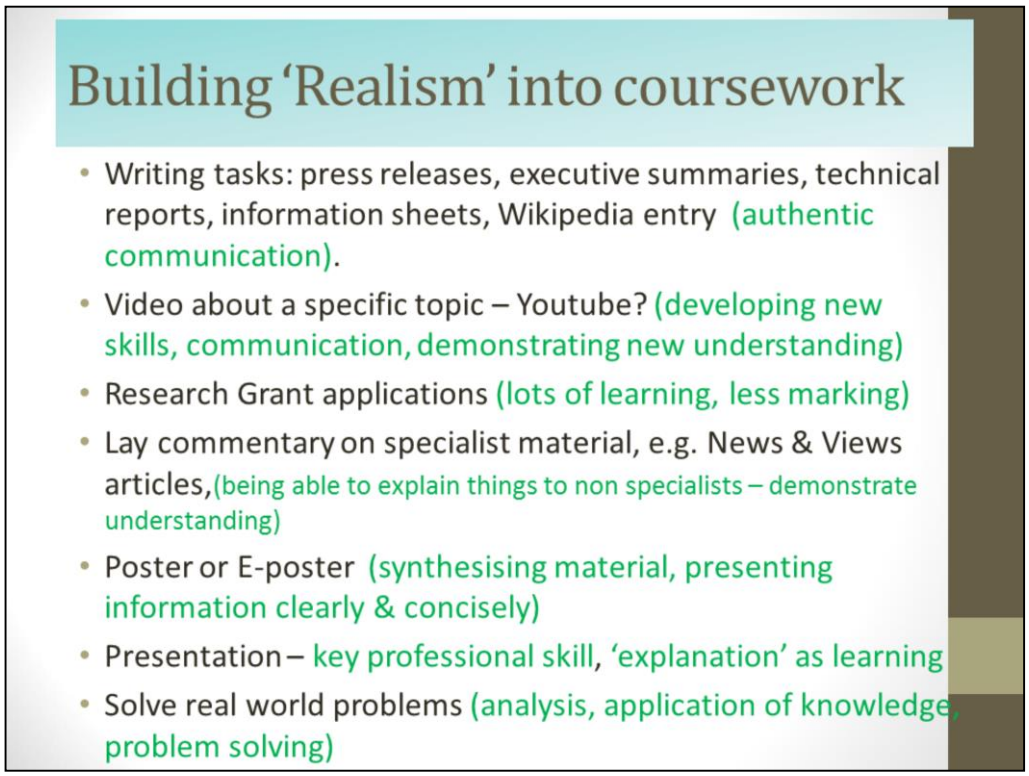
- Better students outcomes
- Better preparation for employment
- Increased student satisfaction
- Highlighted as good practice by professional bodies

19



Recent research by Veronica **Villarroel** Et al (2018) suggests that AA is not based on a well developed theoretical idea – it is not a robust concept. However we can draw on their work to understand what it means in practice; what we should be considering if we want to make our assessments authentic., Their review of recent literature on the topic identifies 3 components to authentic assessment which occur in many of the publications.

I will spend more time on this in my workshop later today but note that these components (although worded slightly differently) are also ones related to supporting student learning from assessment – authentic, higher order and being able to judge whether your work meets expectations – make your assessment more authentic and you are likely to increase its learning potential for students.



There are many other alternative assessment methods that can be used which provide for realism involving developing subject knowledge and skills but also, developing the skills and knowledge for 21st century graduates and post graduates.

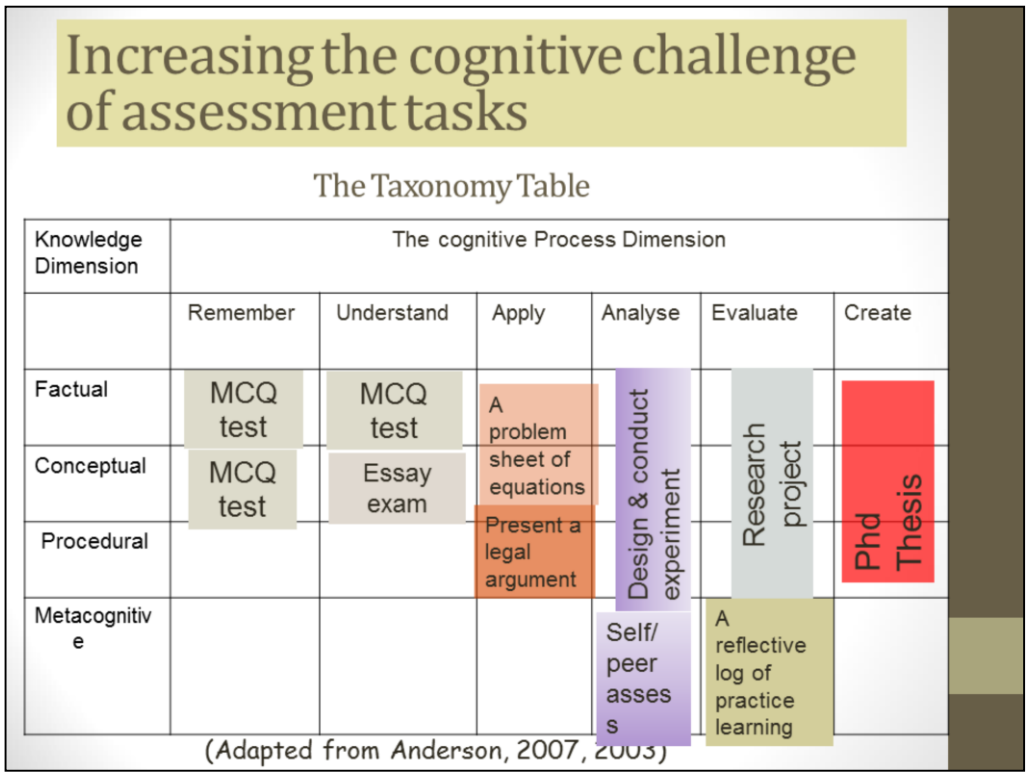
For example, we can imagine a wide range of writing tasks which are more authentic than an exam or essay. These are the types of writing required in work – often clear and concise. Less marking for the teachers.

Asking students to articulate concepts or information in their own words helps them develop their understanding of those ideas. There is growing evidence of the power of explanation in learning – having to explain something helps you understand it. So presentations, posters, Wikipedia entries, lay commentaries, are all good examples of this. E.g. Poster or E-poster (presenting information clearly & concisely) e.g. plant science Dr Elizabeth Sheffield, University of Manchester Group eposters -

We can also capitalise on students interest with social media and video. For example making a Youtube video to demonstrate a skills or illustrate an idea can be an authentic and interesting task. Modern students have the technology for this in their pocket and my experience is that they are motivated to do the reading and thinking for something that they are presenting in a new medium or two a third party audience.

Authentic tasks – for example a research grant application for trainee medical staff or case study analysis for business students – can help develop and assesses required skills.

Various assessment tasks can help develop the skills of metacognition such as diaries, learning logs and reflective journals.



If we try to plot assessment methods on this matrix in terms of the type of knowledge and thinking they are assessing, we can see that some methods tend to assess lower level knowledge of facts and concepts whereas others can assess more challenging cognitive processes which often align with the expected learning outcomes of a university level education. In addition, some common assessment methods do not really assess the procedural and metacognitive knowledge dimensions of learning so essential to Wagner's 'must have' skills for the 21st century.

MCQ and other tests are a good example – whilst some questions can demand higher order thinking, – for example application and cognitive problem solving, logical thinking - these are extremely difficult to write (give example of £1000 per question for medical stuff) and we often don't have the resources to create really challenging questions. And for the most part written or online tests cannot really measure procedural and metacognitive knowledge

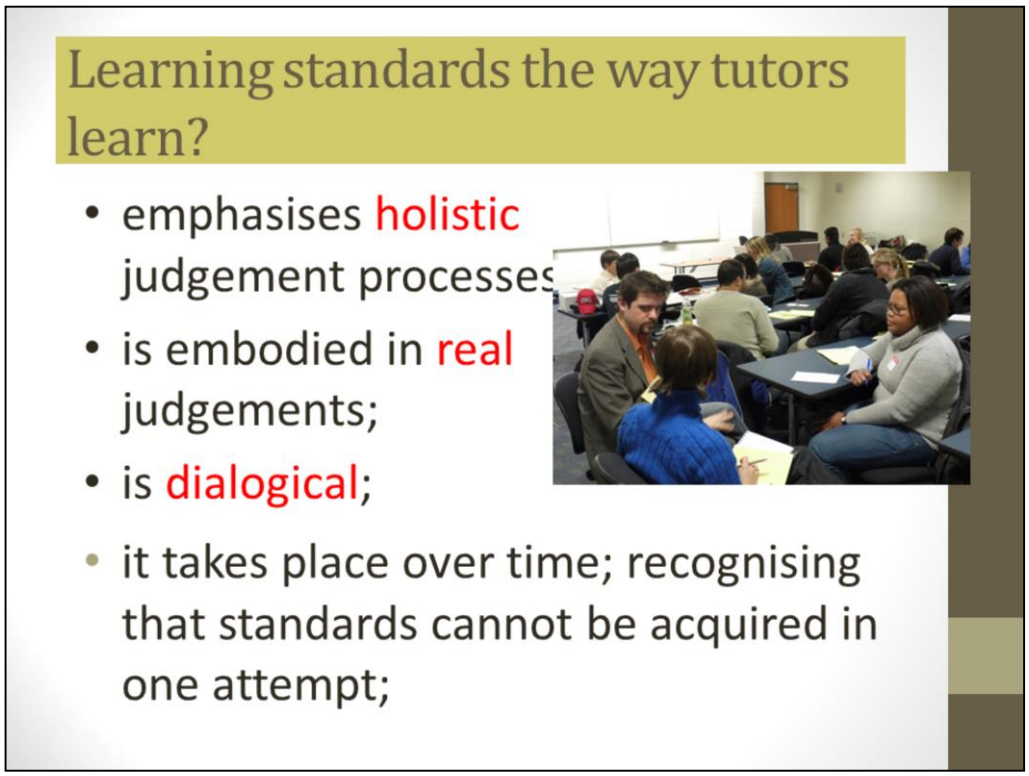
Also, memory based assessment (ie done in controlled conditions without access to information) can have a negative backwash on student learning, encouraging short term memorisation rather than learning with understanding which students are likely to retain

Other types of assessment can be more valid in authentically assessing the full range

of knowledge dimensions and higher order thinking by students.

Why evaluative judgement: Learning

- An essential step in improving performance is knowing how well you are doing and what is needed to improve.
- So preparation for life after HE means being able to judge your own performance against the required standard and to take action to improve.
- This is crucial to (post)graduates being **safe and effective** workers - to know whether their work is up to standard and, importantly, when they need to ask for help from others or undertake further learning



In attempting to achieve this, we need to draw on our understanding of informal social learning as suggested in theories of social practice (Eraut 2004, Lave 1996, Lave and Wenger 1991 from Javitz p 601). Such activities provide important 'access to the knowledge needed to sustain those practices' (e.g. knowledge used in marking) and therefore we need to think how we create the same *social* learning for students in order to develop our guidance and feedback to help students grasp a sense of standards; to 'stand in the shoes' (Handley et al) of their tutors and perceive standards in the same way. The brief discussion above of the informal, repetitious, and social processes underlying tutor learning about standards and the socially situated nature of marking may help us agree some key principles for developing more effective guidance and feedback to help students both develop a *standards' framework* closer to our own and better understand professional judgement. Key principles might be that the process:

emphasises holistic judgement processes; - getting students thinking about the whole meaning and purpose of their assessments, not just individual criteria.

is embodied in real judgements rather than disembodied theoretical descriptions of judgement; - using exemplars, peer and self assessment.

is dialogical; - opportunity for dialogue with tutors and students – using class time for more of this and technology for the delivery of content.

it takes place over time; recognising that standards cannot be acquired in one inoculation; - getting students doing this at different stages in their course, reflecting on their performance

Knowledge as contestable.

Selling peer assessment

Evidence shows students find their peers a useful and more approachable source of help with assignments but we need to stress the main value in peer assessment is standing in the shoes of the assessor – not being assessed – because:

- learning about standards – absolutely crucial to making progress and understanding feedback
- Seeing other ways of going about the task – develops strategies for taking action
- Key employability skill – being able to judge own performance and assessing and giving feedback to others
- More opportunity for dialogue
- Chance for more, rich, formative feedback

Peer assessment needs to become a regular feature of programmes so that it is taken seriously and taken for granted as part of learning at this level.

Assessment to support transition for diverse students

- **Ensure plenty of formative assessment opportunity**
- Help students understand 'the rules of the game'
- Resist the temptation to 'spoon-feed' students – keep to higher order learning
- Help students develop academic and library skills, giving them plenty of opportunity for practice and feedback
- Capitalise on the potential of students to help one another

Conclusion

Developing assessment can help:

- Balance the different purposes of assessment;
- Integrate assessment into teaching and involve ongoing formative assessment opportunities
- Make assessment more valid for a 21st century curriculum;
- Ensure that students' study efforts are directed towards meaningful, programme level, learning;
- Increase achievement;
- Support the retention and achievement of diverse learners;
- Reduce plagiarism.

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.

Bell, A., Mladenovic, R. and Price, M., 2013. "Students' perceptions of the usefulness of marking guides, grade descriptors and annotated exemplars." *Assessment & Evaluation in Higher Education* 38(7), pp.769-788.

Bloxham, S (2014) Assessing assessment (**case study by Georgia Prescott**), in H. Fry et al (Eds) *A handbook for teaching and learning in higher education* (4th Edn) London: Routledge.

Boud, D. 2007. "Reframing Assessment as If Learning Were Important." In *Rethinking Assessment in Higher Education; Learning for the Longer Term,* edited by D. Boud and N. Falchikov, 14-25. London: Routledge.

Boud, D. (2010). Sustainable assessment: rethinking assessment for the learning society. *Studies in Continuing Education*, 22 (2), 151-167, doi: 10.1080/713695728.

Havnes, A. 2004. "Examination and Learning: An Activity-theoretical Analysis of the Relationship between Assessment and Educational Practice." *Assessment and Evaluation in Higher Education* 29 (2): 159-176.

Jensen, J., McDaniel, M., Woodard, S., & Kummer, T. (2014). Teaching to the test... or testing to teach: exams requiring higher order thinking skills encourage greater conceptual understanding. Educ. Psychol. Rev, 26, 306-329, doi: 10.1007/s10648-013-9248-9

Con't

Nicol, D (2009) Assessment for learner self-regulation: enhancing achievement in the first year using learning technologies *Assessment & Evaluation in Higher Education,* Vol. 34, No. 3, June 2009, 335–352 Villarroel, V., Bloxham, S. Bruna, D., Bruna, C. & Herrera-Seda, C (2017): Authentic assessment: creating a blueprint for course design, Assessment & Evaluation in higher Education, DOI: 10.1080/02602938.2017.14123969

Vu, T. & Dall´Alba, G. (2014). Authentic Assessment for Student Learning: An ontological conceptualization, Educational Philosophy and Theory, 46(7), 778-791, doi: 10.1080/00131857.2013.795110.

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.

Watkins, D., Dahlin, B., & Ekholm, M. (2005). Awareness of backwash effect of assessment: A phenomenographic study of the views of Hong Kong and Swedish lectures. Instructional Science, 33, 283-309, doi: 10.1007/s11251-005-3002-8.