

Marks Should Not Be the Focus of Assessment — But How Can Change Be Achieved?

Darrall G. Thompson

Faculty of Design, Architecture and Building, University of Technology, Sydney
darrall.thompson@uts.edu.au

ABSTRACT: This paper attempts to address the possibility of real change after a hundred years of exam-based assessments that produce a single mark or grade as feedback on students' progress and abilities. It uses visual feedback and analysis of graduate attribute assessment to foreground the diversity of aspects of a student's performance across subject boundaries. The use of assessment as the key driver and data source requires the refocusing of staff and student attention away from marks and towards the development of knowledge literacies, conceptual frameworks, and a broad range of personal qualities and skills. It is suggested that the use of student self-assessment with visual feedback on progress in important attributes can engage students with assessment as part of their personal and professional development. Web-based software (REVIEW) developed by the author and refined for more than a decade to facilitate this approach retains categorized and colour-coded student progress data through the day-to-day criterion-referenced marking of assignments and exams. It has been practically applied in four Australian universities with pilot schemes now beginning at the secondary/high school level. This paper examines a typical application at the author's institution with background research and insights into the challenges involved.

Keywords: Assessment, self-assessment, graduate attributes, assessment software.

1 INTRODUCTION

The predominance of high-stakes exam-based assessment throughout the school years encourages both parents and children to focus on marks and grades as the determiner of value and progress. The single mark given conflates the reasons for differentiation between student performances and obscures the diversity of orientation that may inform future study and employment.

In Australia, universities are complicit in this focus on marks, with entry requirements almost exclusively reliant on a single ranking number calculated from Higher School Certificate (HSC) results. As this number is based mostly on examinations, one could facetiously argue that university entrance foregrounds the ability to regurgitate memorized information in handwritten form within a time constraint.

It is unsurprising that the focus on marks for the majority of children's education may cause students to be strategic in their approach to the system when they enter university level education:

(2016). Marks should not be the focus of assessment—But how can change be achieved? *Journal of Learning Analytics*, 3(2), 193–212.
<http://dx.doi.org/10.18608/jla.2016.32.9>

I hate to say it, but what you've got to do is have a list of the facts, you write down ten important points and memorize those, then you'll do alright in the test... If you can give a bit of factual information — so and so did that, and concluded that, for two sides of writing, then you'll get a good mark. (Ramsden, 1984, p. 144)

This quote from Ramsden's research is from a student who received a first-class honours degree (ironically in Psychology), and reveals a strategic/surface approach to learning even though they "hate to say it." This comment implies that the student is uncomfortable with achieving success through a memory-based formulaic approach that results in good marks every time. It is interesting to reflect that the focus on marks may encourage the less able students to resort to a more plagiaristic strategy than the one adopted in this example.

In essence, the assessment method itself is adversely affecting students' approach to, and value for, their study. It is clear that without careful regard for assessment constructs, other curriculum initiatives are unlikely to succeed (Boud, 1988). Rowntree (1987) highlights the importance of assessment to teaching and learning in number one of his seventeen recommendations:

Here is what I believe we need to do, both to teach better and to provide reports that are in our students' (and the community's) best interest. Let us:

1. Articulate as clearly as possible the criteria by which we assess — the aims and objectives we espouse, what qualities we look for in students, in general and in individual cases; let us strive to become more aware of our implicit assessment constructs, and constantly question why we value the qualities we do. (Rowntree, 1987 p. 240)

Elliot Eisner spoke of the need for broad systemic change in his John Dewey Lecture for 2002 at Stanford University:

I am not talking about the implementation of isolated curriculum activities, but rather, the creation of a new culture of schooling that has as much to do with the cultivation of dispositions as with the acquisition of skills ... The public's perception of the purpose of education supports the current paradigm. We need to sail against the tide.

Eisner's term "the cultivation of dispositions" as an addition to the "acquisition of skills" points to the need to broaden the attributes that we are attempting to develop in students. There are important initiatives attempting to gather data to support a new culture, such as the Browser of Student and Course Objects (BoSCO) developed by researchers at Michigan State University. They suggest that access to gathered Institutional Analytics, Developmental Analytics, and Learning Analytics "allows for a bridge to be built between the analytics space and the course/curriculum design environments" (Dunbar, Dingel, Prat-Resina, 2014, p. 231). Institutional Analytics is defined as "helping to run the business of higher education," Developmental Analytics relates to understanding the personal and emotional development of students, and Learning Analytics in their model has to do with understanding student learning behaviours from course management and student information systems. Whilst their stated goal

(2016). Marks should not be the focus of assessment—But how can change be achieved? *Journal of Learning Analytics*, 3(2), 193–212.
<http://dx.doi.org/10.18608/jla.2016.32.9>

in all these is “student success and retention,” the aim of BoSCO as a tool is to feed that information back into the improvement of curriculum design. The REVIEW system described in this paper uses an approach to change focus specifically on assessment processes as the key fulcrum of engagement for both staff and students. Whilst the BoSCO goal is much broader, the data from ongoing day-to-day criterion-referenced assessment of student work focused on broad attribute development in REVIEW could potentially contribute data to both their Developmental Analytics and Learning Analytics dimensions.

In the field of psychology, Howard Gardner at Harvard also advocated broadening the spectrum:

It is of the utmost importance that we recognize and nurture all of the varied human intelligences.... If we can mobilize the spectrum of human abilities, not only will people feel better about themselves and more competent; it is even possible that they will also feel more engaged and better able to join the rest of the world community in working for the broader good. (1993, p. 12)

It was in the context of such research that I embarked in 2002 upon my own research with a range of initiatives about the development of attribute-based education with a particular focus on the facilitation of assessment processes through software design and development.

My background in visual design, design thinking, and engagement in educational research focused my attention on the visual feedback systems often used in instructional design to foreground attribute development. The resultant REVIEW online criteria-based assessment system is now owned by and commercialized through the University of Technology Sydney. Some of its functions are briefly described with screenshots in this paper.

2 THE GRADUATE ATTRIBUTE ASSESSMENT SPACE

The use of a single percentage mark or grade to describe student performance might be convenient for certification but does not accommodate the diversity required to be useful for feedback on the development of graduate attributes. Most assessment marks and feedback relate to a specific subject or unit of study with minimal feedback developmental across subjects and over the years of a course of study. Figure 1 positions the REVIEW software in an assessment space where both summative marks and formative feedback come together to provide longitudinal feedback about the development of graduate attributes across subject boundaries. As Whitelock argues in her work on activating assessment and Web 2.0 tools, we need to be “moving towards an assessment for learning agenda which provides students with advice for action that will assist them on their paths of lifelong learning” (2001, p. 320).

(2016). Marks should not be the focus of assessment—But how can change be achieved? *Journal of Learning Analytics*, 3(2), 193–212.
<http://dx.doi.org/10.18608/jla.2016.32.9>

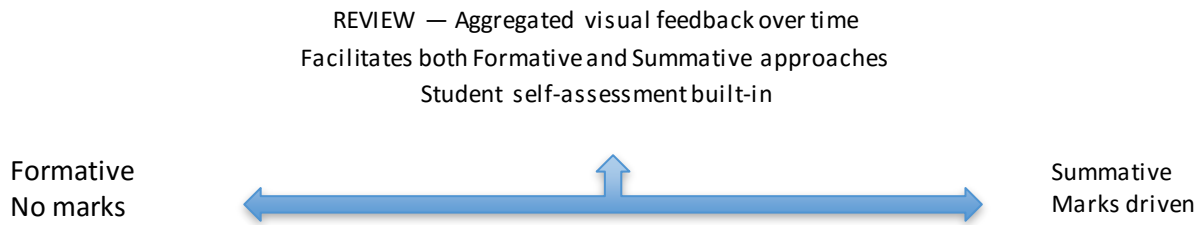


Figure 1: Positioning of the REVIEW web-based system regarding formative and summative assessment approaches.

Bringing graduate attributes into the traditional assessment space requires a brief discussion of terms. The use of the term “graduate attributes” has become the standard term in Australian universities. However, there are a number of terms used in different countries and levels of education, such as *key skills* (Drew, Thorpe, & Bannister, 2002), *generic attributes* (Wright, 1995), *key competences* (Mayer, 1992), *transferable skills* (Assiter, 1995), and the terms *employability skills* and *soft skills*, which are increasingly popular in the business sector (BIHEC, 2007). Considering these terms listed as “skills” seems too narrow to embrace personal, ethical, and social issues. The word “generic” can also be problematic as it tends to imply independence from content when educational research asserts that attributes need to be developed within the context of a field of study rather than “bolted on” (Barrie, 2004).

In Australia, the term graduate attributes was defined as early as 1992: “The skills, personal attributes and values which should be acquired by all graduates regardless of their discipline or field of study” (Higher Education Council, 1992, p. 20). However, despite a number of Australian Government initiatives, graduate attributes have tended to remain as motherhood statements in documentation and the reporting of attribute development in assessment has gained minimal adoption by Australian universities (Thompson, 2006). The Higher Education Council (1992) definition implies separation from discipline content or a field of study and therefore follows this thrust of educational research. For the purposes of this paper, I will use my own definition of graduate attributes, published in 2009, designed to include the learning and application of discipline knowledge:

Graduate attributes are the skills we want students’ to develop, the qualities we want them to acquire *and* the knowledge literacies and conceptual frameworks we want them to construct, through a progressive program of discipline-based assessment tasks. (Thompson, 2009, p. 402)

Given this holistic definition, there needed to be a thoroughly researched formulation of Graduate Attribute (GA) categories grouping together attributes that include a broad range of student learning and development. For this categorization to be accepted, it needed to accommodate all assessment criteria for assessing student work to potentially give ongoing feedback to students about their development in each area. Whilst it is recognized that many exams are poor vehicles for attribute development, if they are marked using assessment criteria they can be validated for their contribution to student development through the linking of criteria to some of the graduate attribute categories.

(2016). Marks should not be the focus of assessment—But how can change be achieved? *Journal of Learning Analytics*, 3(2), 193–212.
<http://dx.doi.org/10.18608/jla.2016.32.9>

Adopting a visual feedback system would also require colour-coding plus symbols to allow for the participation of students and staff with colour blindness.

3 DEVELOPING GRADUATE ATTRIBUTE CATEGORIES

Assessment and grading do not take place in a vacuum. Professional judgements about the quality of student work together with interpretations of such judgements are always made against some background framework or information. (Sadler, 2005, p. 177)

I am suggesting that unless a framework is defined, understood, and valued by students, tutors, and academic coordinators, the judgements and their interpretations are likely to be flawed. Each university has its own approach to the development of graduate attribute frameworks. Some are mandated at university level or faculty and school levels. At UTS, there has been a general statement at the university level with each faculty left to develop their own.

The process of developing a framework in the form of Faculty Graduate Attribute categories was eventually based on my own research over an extended period prior to the university's intention to foreground graduate attribute development in degree courses. The research aimed to define categories that captured sufficient variation in student performance to enable categorization of assessment criteria without ambiguity. This included my own experience as a young designer in London studying colour theory and colour-coding using the five psychological primary colours (Foster, 1891, p. 921), together with approaches to personal development that used these colour-codings (Ive, 1998, p. 10). The literature surveyed whilst developing the REVIEW software in Australia was extensive, including multiple intelligences, graduate employability, and psychometric frameworks.

For example, the Co-Intelligence Institute¹ compiled a list of 24 intelligences with definitions derived from a variety of authors, including Gardner's (1993) nine intelligences. Whilst the REVIEW software does not specify or restrict the number or range of attribute categories or sub-categories, the research summarized below may be useful to others intending to develop or simplify their approach to defining graduate attributes or capabilities.

Furnham and Petrides (2004) used five intelligences in conducting a number of surveys: 1) general, 2) emotional, 3) analytic, 4) creative, and 5) practical (p. 13).

Volansky & Friedman (2003, p. 78) proposed five intelligences as a self-organizing system:

1. **Reflective intelligence:** includes task-oriented behaviours, professionalism, and improvement
2. **Strategic intelligence:** involves formulating a mission statement, defining objectives, and deciding how to carry them out
3. **Contextual intelligence:** includes community-orientation, thinking big, and a system-oriented perspective

¹ <http://www.co-intelligence.org/multiIntelligence.html>

(2016). Marks should not be the focus of assessment —But how can change be achieved? *Journal of Learning Analytics*, 3(2), 193–212.
<http://dx.doi.org/10.18608/jla.2016.32.9>

- 4. **Collegial intelligence:** involves shared management within a school, co-operation with parents and the community, support, mutual assistance, and giving credit where it is due
- 5. **Values intelligence:** based on humanistic, spiritual, and ethical values

In “5 Minds for the Future,” Gardner (2009) describes the disciplined mind, the synthesizing mind, the creating mind, the respectful mind, and the ethical mind, the descriptions being similar to the previous five intelligence models outlined. Gardner reflected on his previous research into multiple intelligences, reducing them to five minds that he predicted would become increasingly important:

Why these five particular minds? Could the list be readily changed or extended? My brief answer is this: the five minds just introduced are the kinds of minds that are particularly at a premium in the world of today and will be even more so tomorrow. They span both the cognitive spectrum and the human enterprise — in that sense they are comprehensive, global. (2009, p. 3)

In extending this research to psychometric testing I investigated the “Big Five Inventory” (Gosling, Rentfrow, & Swann, 2003) from the field of business psychometric testing. This seemed to provide a manageable and simple range of attribute categories and related well to the five intelligences and minds mentioned. The inventory, also known as the “Big Five Traits,” is a well-supported taxonomy in psychometric research and is used extensively in business contexts. Table 1 shows some keywords describing the traits correlated with my initial categorization of graduate attributes (the keywords in brackets were derived from a number of lists and descriptions available in different texts).

Table 1: Initial categorization of Graduate Attributes relating to the psychometric Big Five Traits and with academic references into multiple intelligences: 1) Furnham & Petrides, 2004 p. 13; 2) Gardner, 2009, p. 3; 3) Volansky & Friedman, 2003, p. 78.

The Big Five Traits	The author’s Initial Graduate Attribute Categories* with linked comparative research references
Openness to Experience (creative, versatile, divergent, lateral)	*Creativity and Innovation 1. Creative Intelligence 2. Creative Mind 3. Strategic Intelligence
Agreeableness (warmth, empathy, sensitive communication)	*Communication and Interpersonal Skills 1. Emotional Intelligence 2. Respectful Mind 3. Collegial Intelligence
Conscientiousness (principled, careful, good at self-regulation/self-discipline)	*Attitudes and Values 1. General Intelligence 2. Ethical Mind 3. Values Intelligence
Extraversion (active, outgoing, practical)	*Practical and Professional Skills 1. Practical Intelligence 2. Disciplined Mind 3. Contextual Intelligence
Emotional Stability (emotions do not impact behaviour; research based)	*Critical Thinking and Research Skills 1. Analytic Intelligence 2. Synthesizing Mind 3. Reflective Intelligence

(2016). Marks should not be the focus of assessment—But how can change be achieved? *Journal of Learning Analytics*, 3(2), 193–212.
<http://dx.doi.org/10.18608/jla.2016.32.9>

These five initial graduate attribute categories formulated in 2007 for the UTS School of Design (Thompson, 2009) carried descriptions that related well to the academic research and added confidence to the debate about categories that might be more broadly adopted. However, the broad adoption of such a system in the author's faculty across 700 subjects in three diverse schools needed buy-in from hundreds of academics and thousands of students if it was to be effective.

Together with research on appropriate categories of attribute development, evidence to support the change was to be found in critical business and industry reports on graduate capabilities (BIHEC, 2007). This criticism did not foreground that employers wanted graduates who can remember the content from their university studies, and such critiques have been consistent for more than a decade. For example, a study of design engineering graduates and their employers in the UK (Garner & Duckworth, 2000) exposed specific flaws in graduate profiles. The employers' criticisms included the following reflections about the attributes lacking in university graduates:

- They need greater ability to take other people's ideas on board.
- They have a lack of resilience to criticism.
- They have a weak ability to muster a reasoned defence of their contribution.
- They need to improve listening skills.
- They need higher-quality written, graphic, and verbal communication.
- They need to be able to be critical of their own work and contributions.

(Garner & Duckworth, 2000, p. 208)

Such evidences together with a University initiative to promote graduate attribute integration were used as levers to foreground a change in assessment practices through use of the REVIEW software. Critical to adoption of the changes was a university requirement for Courses to report on graduate attribute development and a subsequent Faculty Board decision that REVIEW would be the validated system through which evidence could be provided. However, from many years of experience as an academic, it is one thing to inform an academic of a requirement and another to inspire a deeper engagement beyond a box-ticking response. The challenge of encouraging course teams to engage required an acceptance of complexity by allowing each course to develop its own discipline-specific language in describing the intended learning outcomes. Also, prior to the practical adoption and approval of Graduate Attribute categories for assessment, student focus groups, an industry advisory group, and a course directors' faculty retreat were conducted.

The sub-categories of the five GA categories termed Course Intended Learning Outcomes at UTS were then developed through six planning workshops with 14-degree course teams. The one-day workshops were conducted in school groupings with very large pre-prepared colour printouts of each course/subject structure and a facilitated process of writing outcomes relevant to subjects in each progressive year of the course. This was a challenge for some degree courses that were modular in construction but gradually over a period of work with course directors, the clarity of the five categories resulted in three or four discipline-specific sub-categories under each.

(2016). Marks should not be the focus of assessment—But how can change be achieved? *Journal of Learning Analytics*, 3(2), 193–212.
<http://dx.doi.org/10.18608/jla.2016.32.9>

Although the general differentiation of the initial categories was accepted, a major challenge was the introduction of the system in course documentation together with staff and student recognition of the attribute categories. There needed to be a memorable acronym, colours, and symbols for each category in order to foreground the system in the common parlance of staff and student interaction. The acronym CAPRI came from the course directors' faculty retreat and caused slight changes in the wording of the GA category titles together with their one-paragraph descriptions.

The following colour-coded CAPRI categories and descriptions gave guidance for course teams to word their own Course Intended Learning Outcomes as sub-categories under each of the five Graduate Attribute categories. The idea being to use the marking and feedback software (REVIEW) to collect marks in the background from the day-to-day marking of assessment criteria linked to both Course Intended Learning Outcomes and the five CAPRI categories.

C – Communication and Groupwork (Colour – Yellow; Symbol – Pentagon). This category concerns the quality and clarity of such things as oral presentations, written essays, explanations, and visual presentations. In addition, the development of communication, in-group interactions and various team roles.

A – Attitudes and Values (Colour – Blue; Symbol – Circle). This category is to do with respect for one's own work and the work of others, including ethics. Developing care, understanding, and patience, with consideration for others' points of view, including indigenous and multicultural perspectives.

P – Practical and Professional (Colour – Red; Symbol – Square). This category of development involves technical skills and operational techniques together with the methods and experience/knowledge required to function as a professional in a broad range of environments.

R – Research and Critique (Colour – White/Grey; Symbol – Triangle). This category of development involves fact-finding, literature surveys, research methods and the ability to think analytically. Also developing the ability to make informed criticism of one's own work and the work of others.

I – Innovation and Creativity (Colour – Green; Symbol – Star). This category has to do with inventiveness, versatility, thoughtful risk-taking, imagination, creative concepts, innovative problem-solving, natural curiosity, creative experimentation, and the innovative application of technologies and processes.

In order to promote this new approach to assessment, video/animations explaining the system were produced.²

² Explanation for Staff (2min 51secs): <https://www.youtube.com/watch?v=wLIFcwTae7A>

Explanation for Students (1min 39secs): <https://www.youtube.com/watch?v=ZN1lojufSTA>

(2016). Marks should not be the focus of assessment—But how can change be achieved? *Journal of Learning Analytics*, 3(2), 193–212.
<http://dx.doi.org/10.18608/jla.2016.32.9>

4 THE REVIEW ONLINE CRITERIA-BASED ASSESSMENT SYSTEM

Students normally pass subjects and courses with a percentage mark or grade so it is not surprising that they focus on the mark rather than the criteria assessed to arrive at that result. In many cases, students do not have feedback about the ways in which assessment criteria relate to their development of important attributes, such as those made explicit in the CAPRI categories outlined in the previous section.

The design and implementation of the REVIEW system was intended to make explicit those criteria links. The following screenshots from the REVIEW software version 3.3.5 illustrate the practical implementation of this approach.

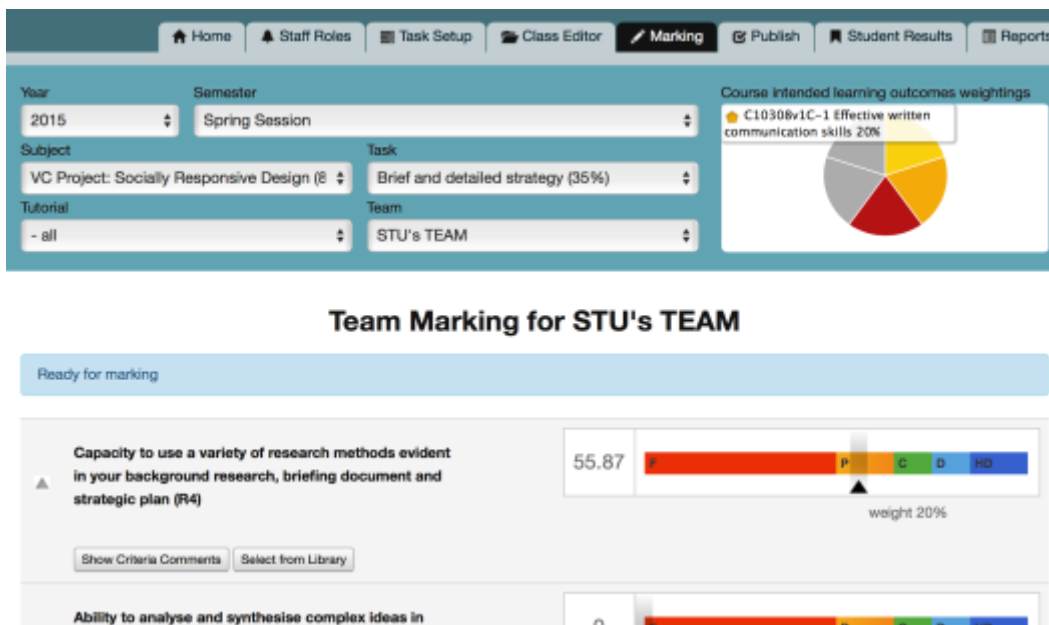


Figure 2: Staff view showing marking sliders and a pie chart with a communication goal highlighted.

The screen shot in Figure 2 shows the teacher’s view of a marking screen for a fictitious student (Stu Dent) who at UTS is automatically added to every subject for the purposes of practice and demonstration. In this case, Stu is in a team assignment and the pie chart (top right) is generated automatically from criteria weightings ascribed to the five Faculty Graduate Attribute categories. These five categories are sub-categorized with related program goals referred to at UTS as Course Intended Learning Outcomes. In this case, the task only covers three of the five CAPRI categories, but staff and students can hover the cursor over the pie chart to view the Course Intended Learning Outcome sub-categories within each coloured section. The marking of student work happens by clicking and dragging the marker along the sliding scale to produce automatically calculated percentage marks (the system does not show students their marks and unit coordinators are discouraged from doing so). The tabs at

(2016). Marks should not be the focus of assessment—But how can change be achieved? *Journal of Learning Analytics*, 3(2), 193–212.
<http://dx.doi.org/10.18608/jla.2016.32.9>

the top of the screenshot are those viewed by subject or unit coordinators with a reduced tab set for tutors who do not need particular functions.

The screenshot in Figure 3 shows a staff view of the REVIEW marking screen for a different task after the staff member has saved their gradings against each criterion. The student’s self-assessments are now visible as blue arrows on the top edge of each grading slider. This shows that the fictitious Stu Dent has overrated his performance on the first criterion but underrated on the second.

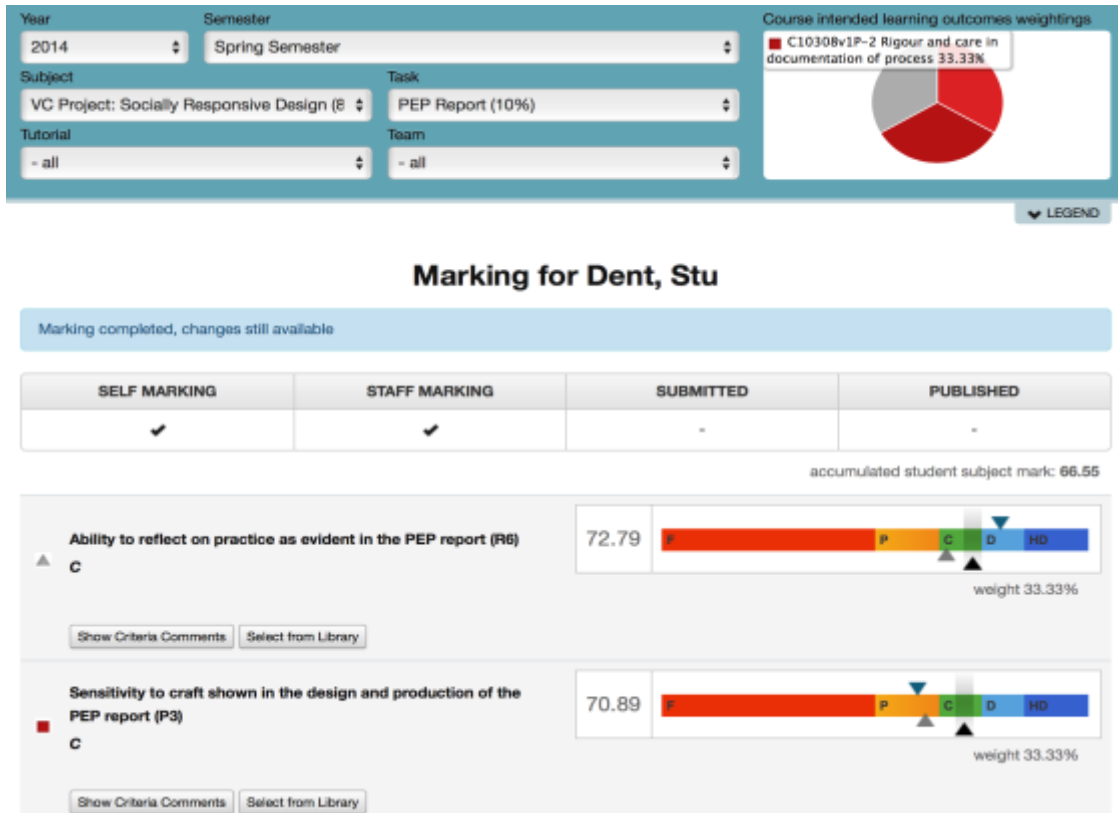


Figure 3: Staff view with student self-assessment now visible (blue arrows above the scale).

In this staff view, the numbers are calculated for each criterion and the total when the black triangle marker under each grading scale is dragged by the tutor or marking staff. Staff can alternatively type in a number in the box to the left of the scale or click anywhere on the scale. The student’s own self-assessment triangle marker is only shown above each grading scale after staff have entered and saved their own grading judgements.

(2016). Marks should not be the focus of assessment—But how can change be achieved? *Journal of Learning Analytics*, 3(2), 193–212.
<http://dx.doi.org/10.18608/jla.2016.32.9>

Educational research based on four semesters of data from REVIEW using voluntary student self-assessment has shown benefits to student learning and the calibration of their judgement over time (Boud, Lawson, & Thompson, 2013). Due to the benefits of this self-reflective process for students, a number of ways to incentivize student engagement have been attempted, including rewarding or penalizing with marks for accuracy or engagement, adding a criterion to assess the quality of students’ gradings, and promoting the activity through animations and video introductions. In my experience, the most successful method has been an introduction by the unit coordinator in combination with tutors who genuinely value the student gradings and demonstrate this feature by marking a piece of work in a large lecture context. Involving students in this live marking activity engages both them and the tutors in further understanding the criteria. In some cases, academics have offered to give extra feedback comments where students’ self-gradings are very different from their own as official markers. The screenshot in Figure 4 shows Stu Dent’s view of the marking screen during his own self-assessment against the criteria before a staff member has marked his work. The paragraph of text under the pie chart area explains self-assessment and the linkage of criteria to attribute categories with a diagram showing the types of triangle markers that will be displayed on the grading scale.

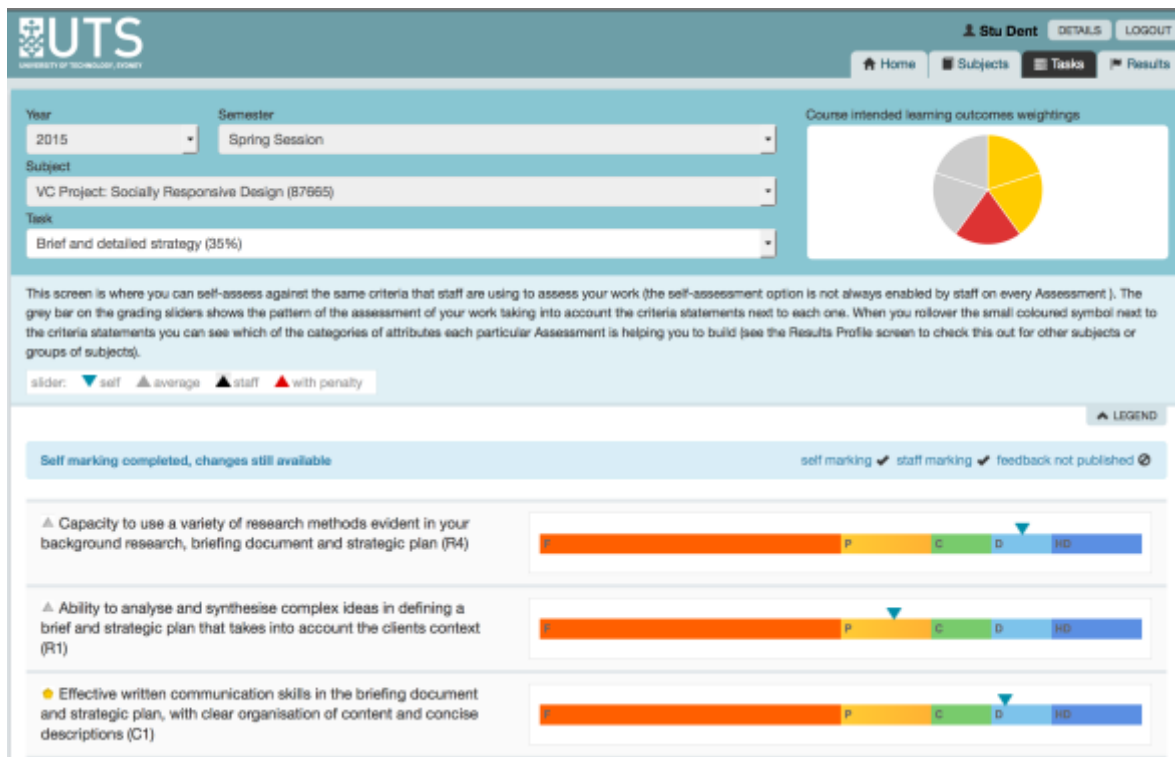


Figure 4: Student view of the task assessment screen prior to staff marking showing the student’s own self-assessment gradings against some of the task criteria.

Figure 5 shows Stu Dent’s view of the marking screen for a task after a tutor has marked his work. The most important aspect of the screen’s design is that there is no percentage mark shown for Stu’s result. This has the effect of focusing student attention on the gradings against the criteria and the position of their own self-assessment when compared to the tutors’ gradings. As Stu has significantly overrated his

(2016). Marks should not be the focus of assessment—But how can change be achieved? *Journal of Learning Analytics*, 3(2), 193–212.
<http://dx.doi.org/10.18608/jla.2016.32.9>

performance on one of the criteria and underrated another, the tutor has given further explanation with a comment under each of the criteria descriptions. Staff may decide to enter a comment under a particular criterion as well as an overall comment at the bottom of the list of criteria.

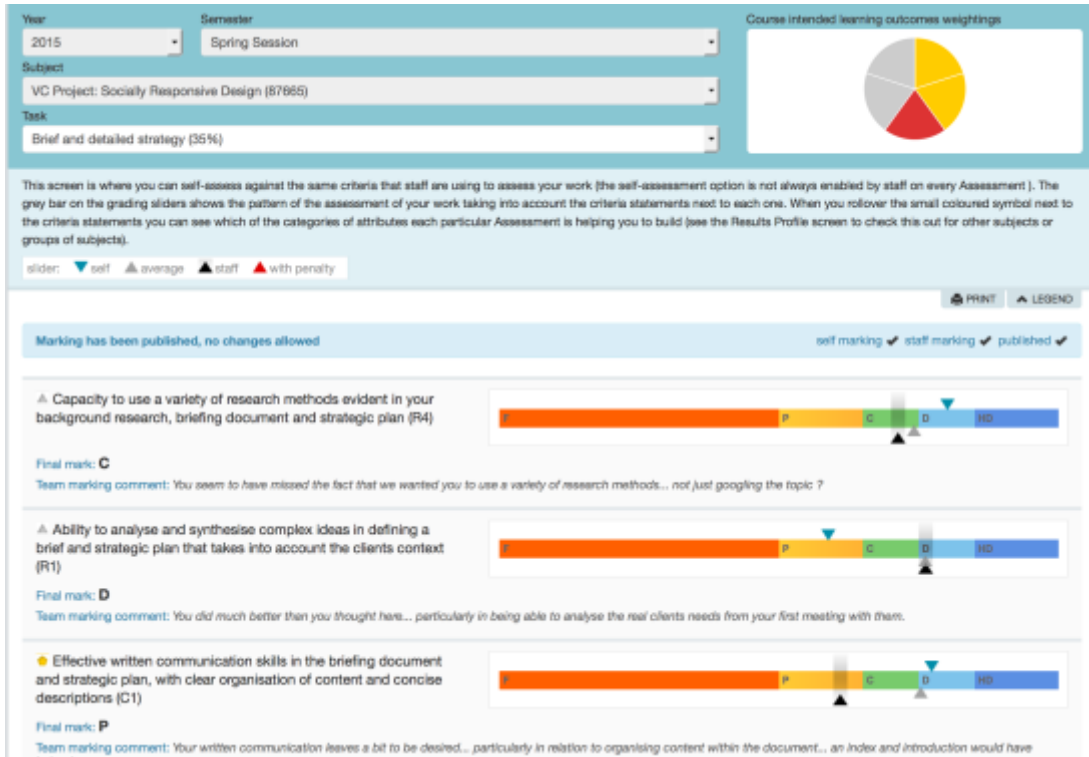


Figure 5: Student view of the marking screen for a task after a staff member has marked and published their own gradings and comments.

Figure 5 also shows a grey triangle slider below the grading scale on each criterion. This is an average of the marks given to all the students concurrently taking this task and gives the student a very approximate idea of whether their score for a criterion is above or below the cohort's average. Unit coordinators can monitor these averages and export a range of data reports on self-assessment and tutor marking. Spreadsheets can be exported and imported to accommodate exam, test, or essay marking that may be happening outside REVIEW but still relate to program learning goals and attribute categories.

Figures 6 and 7 show student views of screens intended to inform them about their progress in the five CAPRI categories with sub-categories for Course Intended Learning Outcomes indicated by lined sections within the five colours on the pie chart.

(2016). Marks should not be the focus of assessment—But how can change be achieved? *Journal of Learning Analytics*, 3(2), 193–212. <http://dx.doi.org/10.18608/jla.2016.32.9>



Figure 6: Student screen showing their assessment results in the five CAPRI attribute categories in one subject or unit of study.

The pie chart shows the assessment weighting of CAPRI categories and their relevant sub-categories. The bar chart, however, shows a student’s actual performance in the subject. In this case, whilst the majority of the criteria are assessing the Communication and Groupwork category, this is the student’s weakest result in this particular subject. The highlighted section in the second (yellow) bar reveals the text of the intended learning outcome C-2 Effective oral communication skills.

However, this attribute-based progress can only be of ongoing longitudinal value to students if the same categories and sub-categories are used throughout their degree program. The adoption of such an approach across subject boundaries yields important data analysis potential for the institution and the students themselves. Students and staff can download more comprehensive PDFs of their performance by year, semester, subject, and task. Figure 7 shows our fictitious student’s progress over three years of study.

(2016). Marks should not be the focus of assessment—But how can change be achieved? *Journal of Learning Analytics*, 3(2), 193–212. <http://dx.doi.org/10.18608/jla.2016.32.9>

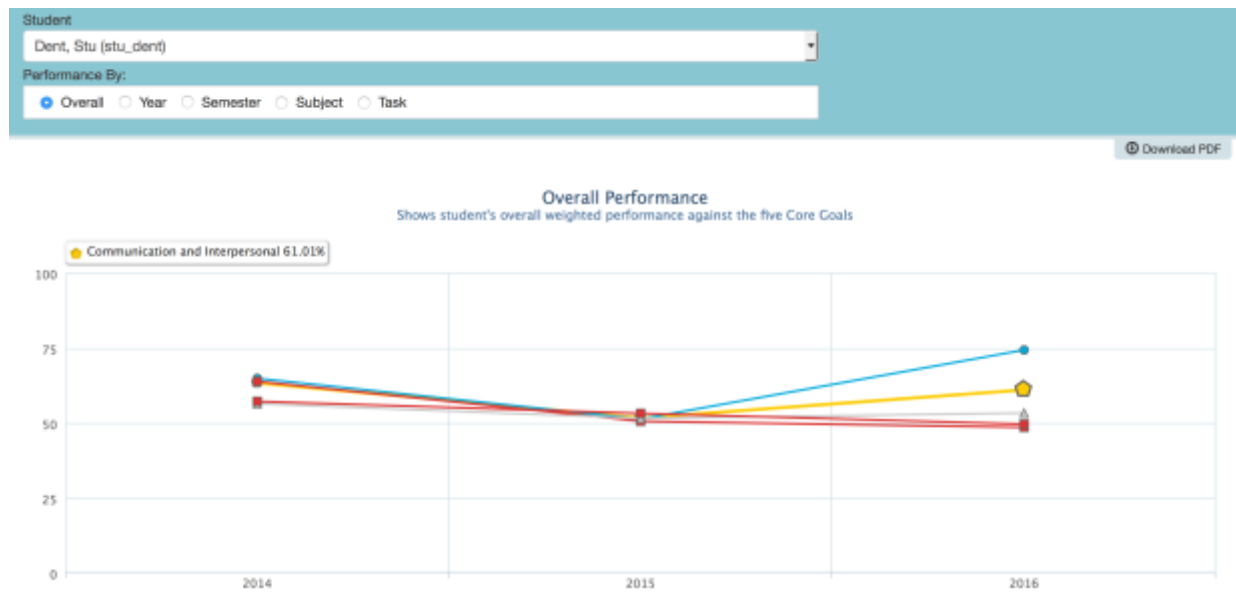


Figure 7: View of an individual student progress overall or by year, semester, subject, or task also available in more detailed PDF reports, including comments.

Apart from the value of visual feedback for students, the system has a large range of reports available for teachers, instructors, coordinators, and administration. Such reports are increasingly valuable for course mapping, tutor benchmarking, institutional accreditation, and the monitoring of subject improvement over time.

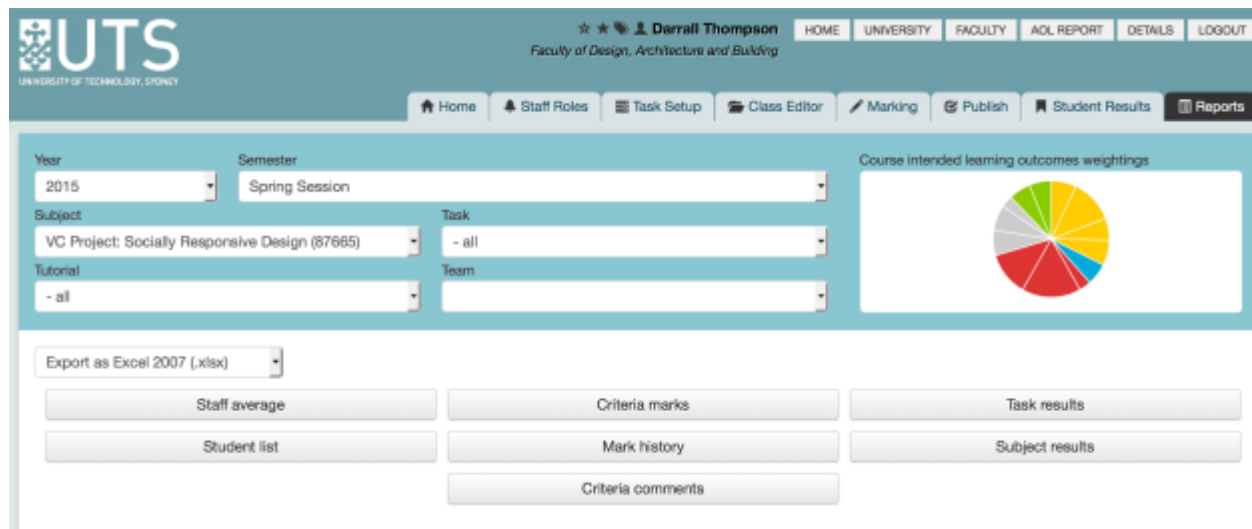


Figure 8: Part of the Reports screen in REVIEW showing seven excel spreadsheet exports available to staff.

(2016). Marks should not be the focus of assessment—But how can change be achieved? *Journal of Learning Analytics*, 3(2), 193–212. <http://dx.doi.org/10.18608/jla.2016.32.9>

Figure 8 shows various reports available for staff, including the Staff average report of tutor marking that highlights in colour any marking significantly different from other tutors in a subject or task. This is used by subject coordinators for tutor benchmarking to monitor and in some cases intervene in task assessment. There is a function available to export marks, make a percentage variation to a particular tutor’s marking, and then re-import the marks. However, this is discouraged in favour of up-front tutor meetings to clarify the marking criteria and standards for assessment tasks.

REVIEW is also used currently to provide data reports for accreditation by various professional and educational bodies. Assurance of Learning (AOL) is a common term in business faculties but other external and internal accreditation processes for other faculties require similar reporting capabilities. The screenshot in Figure 9 shows a highly granular AOL reporting module in REVIEW capable of producing filtered reports designed to accommodate the sophisticated requirements of various international accreditation bodies. The reports can be output as spreadsheets or visual graphs and charts. REVIEW is currently being used to produce official AOL reports at UTS, the University of New South Wales, and Queensland University of Technology.

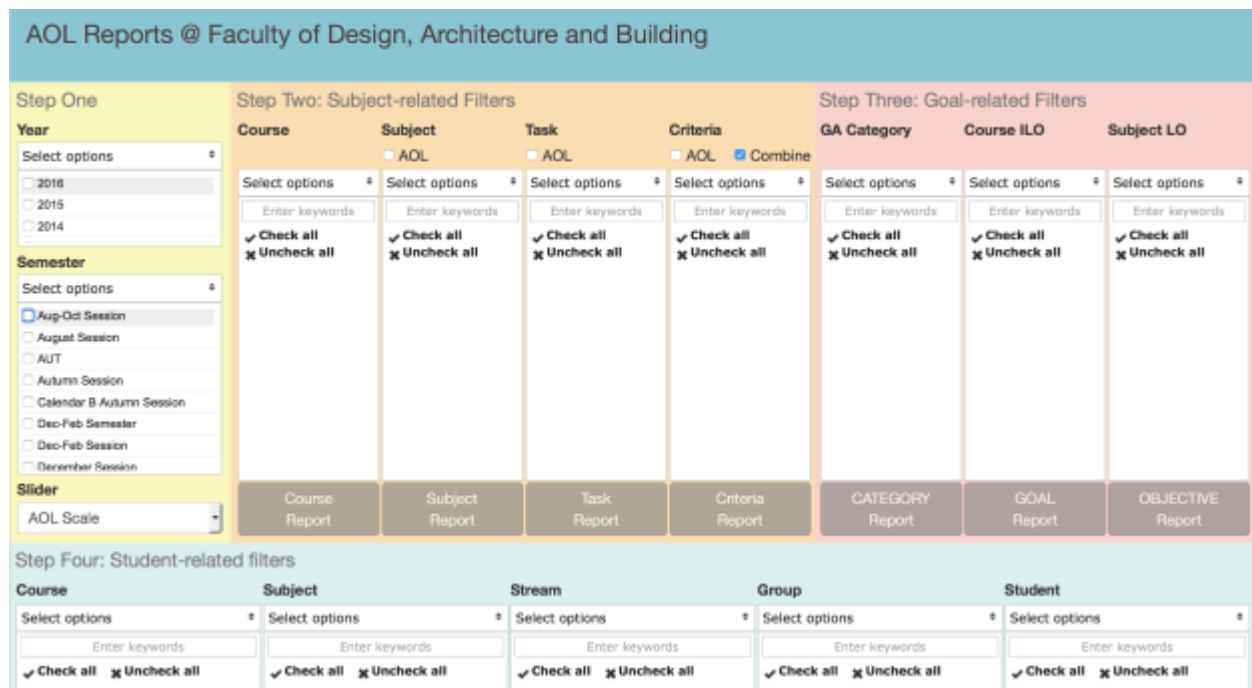


Figure 9: Assurance of Learning (AOL) reporting module in REVIEW that can produce highly granular spreadsheets and visual reports for course improvement, student progression, and institutional accreditation.

REVIEW features are continuously upgrading due to a collaborative funding model that enables universities that require a particular feature to pay for it to be included. For example, the Assurance of Learning reporting system illustrated in Figure 9 was funded by the University of New South Wales (UNSW) because of their requirement for Business School accreditation by the AACSB (Association to

(2016). Marks should not be the focus of assessment—But how can change be achieved? *Journal of Learning Analytics*, 3(2), 193–212.
<http://dx.doi.org/10.18608/jla.2016.32.9>

Advance Collegiate Schools of Business), and EQUIS (European Quality Improvement System). They have used this module in REVIEW extensively for their successful and continuing accreditation processes and maintain that previous methods of collecting and compiling data for these reports was onerous and time-consuming at the most highly pressured times of the year. REVIEW has automated this process with a level of granularity that has assured its adoption across a number of faculties.

The collaborative funding model is a progressive format that enables such Assurance of Learning and other modules to be available for any other user of REVIEW free of charge. Shared or individually funded features are specified, and costs are then estimated by the software developers in Sydney. Extensive modules together with smaller features are implemented with ongoing upgrade versions. There is a REVIEW Users Group (RUG) jointly run by UNSW and UTS as both an academic and technical forum for ideas, feature requests, and upgrade presentations.

5 CONCLUSIONS AND LESSONS LEARNED

The following lessons learned are based on educational research and my experiences as a teaching academic over a ten-year period:

1. All assessment activities, including tests and exams, contribute to the development of attributes. Therefore, all types of assessments can be marked using criteria explicitly linked to a range of attribute categories such as CAPRI. I have found that the conversion of exam and MCQ marking to criterion-referenced assessment linked to attribute categories leads to more scenario-based questions that test a broad range of knowledge and concept application.
2. Many academics seriously consider the teaching aspects of their work, ensuring that the assessment tasks are interesting and challenging, and aligned with the aims and values they personally espouse as part of a degree program. It is therefore essential, particularly when taking on a subject developed by someone else, that the academics' views are respected and valued as the basis for unit development together with the appropriate linking of assessment criteria to Graduate Attribute categories and Course Intended Learning Outcomes.
3. Assessment criteria are a powerful method for communication to students and tutors and form an important “fulcrum of engagement” for both. Academics should therefore be encouraged to develop explicit wording and weighting for assessment criteria that make clear the aspects of students' work that will be assessed. To assist in their understanding of these criteria, students should be encouraged to engage with them through self-assessment prior to handing in their work.
4. In criteria-based or criterion-referenced assessment, two students can receive exactly the same total percentage mark for very different reasons according to the criteria. The single mark gives minimal feedback to students about their learning and development, and therefore strategies should be employed to focus student attention away from marks and onto attributes linked to

(2016). Marks should not be the focus of assessment—But how can change be achieved? *Journal of Learning Analytics*, 3(2), 193–212.
<http://dx.doi.org/10.18608/jla.2016.32.9>

the criteria with visual feedback of attribute progress over time. This approach should be used across all subjects in a course of study for the feedback to be useful for student progression and course accreditation.

5. Top-down directives about graduate attribute integration often involve onerous documentation, alienating busy academics while having minimal impact at the student level. For improvement in feedback to occur, instructors need to be given timesaving strategies and support. Software such as REVIEW must be integrated into the main university system to save time in assessment and reporting processes. The timesaving aspects and ease of use of REVIEW together with its perceived value to staff and students caused it to spread by osmosis, leading to its commercialization by the University of Technology Sydney in 2011.
6. University technology divisions require highly secure systems that do not compromise their existing services. There are a number of approaches for web-based systems hosted internally by each university or externally by a provider. The developer's recommendation is for REVIEW to be externally hosted and undergo rigorous penetration testing with every upgrade release. However, an internally hosted option is available. The configuration of the system and Application Program Interface (API) integration is essential for broad adoption, together with policy approvals by faculty boards, heads of school, and course directors.
7. Assurance of Learning reports for course and school accreditation are easy to produce from the REVIEW system but need careful interpretation if they are to be used in subject or course improvement. Course mapping of graduate attribute assessment has revealed the need to add assessment criteria for some Course Intended Learning Outcomes that were not valued in the assessment of tasks. For example, the course mapping spreadsheets populated from the ongoing day-to-day marking of assessment criteria using REVIEW has shown that some of the Course Intended Learning Outcomes were completely missing from any assessment task in any subject throughout some courses of study. If intended outcomes are not mentioned in assessment criteria, the course team should reflect on whether that outcome is actually valued or needs to be changed or deleted.
8. The gradual implementation of REVIEW in a large range of subjects and the engagement of students in self-assessment has provided statistical evidence of significant impacts on calibrating students' judgement of their own work (Boud et al., 2013). There are early signs in student feedback that the visual display of criteria linked to attribute categories and sub-categories is useful in charting progress and presenting to employers in interview contexts. Employers take these charts seriously because they are derived from actual official assessment criteria from a broad range of subjects over time, rather than badges that may not carry the same weight (Rosewell, 2012). Currently there is no official issue from UTS at student graduation other than a degree certificate and a subject-by-subject testamur of percentage scores. However, there are discussions currently happening about adding a REVIEW graduate attribute report to this range

(2016). Marks should not be the focus of assessment—But how can change be achieved? *Journal of Learning Analytics*, 3(2), 193–212.
<http://dx.doi.org/10.18608/jla.2016.32.9>

of outputs. Students themselves have the option of presenting these to employers but there is only anecdotal evidence of positive employer response to date.

For students to engage with assessment as an important aspect of their development, the relevance of attributes beyond handwriting and memorizing standard answers needs to be made explicit throughout their curriculum and aligned with assessment criteria for all tasks. This paper proposes that one of the most effective ways to achieve change is to introduce assessment systems that foreground and make visible the attribute development evident in student performances over time.

The culture of the single mark and grade impacts students' approach to assessment very early in their educational engagement at high school and increasingly at primary levels. By the university level, marks-driven attitudes are difficult to dislodge in both staff and student approaches. There needs to be earlier intervention if real change to assessment as feedback on student progress and abilities is to be achieved. Whilst their curricula tend to be more prescriptive, it is hoped that this paper can stimulate some pilot schemes at the secondary/high school levels.

REFERENCES

- Assiter, A. (Ed.). (1995). *Transferable skills in higher education*. London: Kogan Page.
- Barrie, S. C. (2004). A research-based approach to generic graduate attributes policy. *Higher Education Research & Development*, 23(3), 2004. <http://dx.doi.org/10.1080/0729436042000235391>
- BIHEC (Business, Industry and Higher Education Collaboration Council). (2007). *Graduate Employability Skills* (Report). Melbourne, Victoria: Precision Consultancy.
- Boud, D., Lawson, R., & Thompson, D. G. (2013). Does student engagement in self-assessment calibrate their judgement over time? *Assessment & Evaluation in Higher Education*, 38(8), 941–956. <http://dx.doi.org/10.1080/02602938.2013.769198>
- Boud, D. (1988). Moving towards autonomy. In D. Boud (Ed.), *Developing student autonomy in learning* (2nd ed), (pp.17–39). London: Kogan Page.
- Drew, S., Thorpe, L., & Bannister, P. (2002). Key skills computerised assessments: Guiding principles. *Assessment and Evaluation in Higher Education*, 27(2), 175–186. <http://dx.doi.org/10.1080/02602930220128742>
- Dunbar, R. L., Dingel, M. J., & Prat-Resina, X. (2014). Connecting analytics and curriculum design: Process and outcomes of building a tool to browse data relevant to course designers. *Journal of Learning Analytics*, 1(3), 223–234. Retrieved from <https://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/3744>
- Eisner, E. W. (2002). *What can education learn from the arts about the practice of education?* John Dewey Lecture delivered at Stanford University. Retrieved from http://www.infed.org/biblio/eisner_arts_and_the_practice_of_education.htm
- Foster, M. (1881-1891). *A text-book of physiology* (5th ed.). London: Macmillan. <http://dx.doi.org/10.5962/bhl.title.1854>

(2016). Marks should not be the focus of assessment—But how can change be achieved? *Journal of Learning Analytics*, 3(2), 193–212. <http://dx.doi.org/10.18608/jla.2016.32.9>

- Furnham, A., & Petrides, K. V. (2004). Parental estimates of five types of intelligence. *Australian Journal of Psychology*, 56(1), 10–17. <http://dx.doi.org/10.1080/00049530410001688074>
- Gardner, H. (1993). *Multiple intelligences: The theory in practice*. New York: Basic Books.
- Gardner, H. (2009). *5 Minds for the Future*. Watertown, MA: Harvard Business Review Press.
- Garner, S., & Duckworth, A. (2000). The employability of design graduates: A study of competencies achieved through undergraduate design education. In C. Swann & E. Young (Eds.), *Conference Proceedings: Re-inventing design education in the university* (pp. 203-210). Perth, Australia: Curtin University of Technology
- Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37(6), 504–528. [http://dx.doi.org/10.1016/S0092-6566\(03\)00046-1](http://dx.doi.org/10.1016/S0092-6566(03)00046-1)
- Higher Education Council. (1992). *Achieving Quality, Higher Education/ Higher Education Council, National Board of Employment, Education and Training*. Canberra: Australian Govt. Pub. Service.
- Ive, P. (Ed.). (1998). *The Penemin: The complete pentagram key to threshold development*. London: Gemstone Press.
- Mayer, E., Australian Education Council. (1992). *Key Competencies. Report of the Committee to Advise the Australian Education Council and Ministers of Vocational Education, Employment and Training on Employment-related Key Competencies for Postcompulsory Education and Training [Mayer Report]*. Canberra: Australian Education Council and Ministers of Vocational Education, Employment and Training. Retrieved from <http://www.voced.edu.au/content/ngv%3A28045>
- Ramsden, P. (1984). The context of learning. In F. Marton, D. Hounsell, & N. Entwistle (Eds.), *The experience of learning: Implications for teaching and studying in higher education* (pp.159–184). Edinburgh, Scotland: Scottish Academic Press.
- Rosewell, J. (2012). A speculation on the possible use of badges for learning at the UK Open University. In *Proceedings of the 25th EADTU Annual Conference: The Role of Open and Flexible Education in European Higher Education Systems for 2020: New Models, New Markets, New Media* (pp.214–228). Cyprus: European Association of Distance Teaching Universities and Authors.
- Rowntree, D. (1987). *Assessing students: How shall we know them?* (2nd ed.). London: Kogan Page.
- Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, 30(2), 175–194. <http://dx.doi.org/10.1080/0260293042000264262>
- Thompson, D. G. (2006). E-assessment: The demise of exams and the rise of generic attribute assessment for improved student learning. In T. S. Roberts (Ed.), *Self, peer and group assessment in e-learning* (pp. 295–322), Hershey, PA: IGI/Information Science Publishing.
- Thompson, D. G. (2009). Successful engagement in graduate attribute assessment using software. *Campus-Wide Information Systems*, 26(5), 400–412. <http://dx.doi.org/10.1108/10650740911004813>
- Volansky, A., & Friedman, I. A. (2003). *School-based management: An international perspective*. Jerusalem: State of Israel, Ministry of Education, Publication Department.

(2016). Marks should not be the focus of assessment—But how can change be achieved? *Journal of Learning Analytics*, 3(2), 193–212.
<http://dx.doi.org/10.18608/jla.2016.32.9>

Whitelock, D. (2011). Activating assessment for learning: Are we on the way with Web 2.0? In M. J. W. Lee & C. McLoughlin (Eds.), *Web 2.0-based-e-learning: Applying social informatics for tertiary teaching* (pp. 319–342). Hershey, PA: IGI Global.

Wright, P. (1995). *What are graduates? Clarifying the attributes of 'graduateness'*. Quality Enhancement Group, Higher Education Quality Council, London.