Is implementing criterion-referenced assessment worth the effort with gen Y?

Moira Cordiner
Centre for the Advancement of Learning and Teaching
University of Tasmania, moira.cordiner@utas.edu.au

Dr Deb Stenzel
Faculty of Science
Queensland University of Technology, d.stenzel@qut.edu.au

A/Prof Louise Hafner
Faculty of Science
Queensland University of Technology, l.hafner@qut.edu.au

Well implemented criterion-referenced assessment (CRA) requires dedicated time and effort, especially in describing realistic expectations of evidence of achievement to students in the form of criteria sheets (or grading rubrics). It is also takes time out of delivering content to teach students how to judge their own work using criteria sheets. In 2007, to engage third year Microbiology students in using criteria sheets for the first time in their degree, we devised an innovative assessment tutorial supported by online resources. We were sceptical of much of the literature that reported 'agreed' characteristics of our predominantly gen Y cohort, because of the older ages of the majority of authors. These authors claim gen Y has a propensity for digital media, overconfidence in their own abilities and a collaborative orientation. We rejected this stereotype when developing the tutorial. Evaluations by students were positive and there was no dramatic change to grades for the unit. These results are similar to those in the literature for non gen Y cohorts. This lends support to our claim that giving students control over their own learning, irrespective of their generational label, is worth the time and effort.

Keywords: *criteria sheets, students judging grades, role play*

Background

Most students currently do not use criteria sheets in their Science units at our university or know how academics make judgments about students' assessment work. Queensland University of Technology requires that academics base their judgments about grades for tasks and units (subjects) on how well students meet criteria and described standards. That is, academics are to implement criterion-referenced assessment (CRA). Many find this very challenging and time consuming as they are more comfortable with the practice of 'norming', or allocating grades based on a preset formula that relies on a bell curve. To implement CRA well requires dedicated time and effort from busy academic staff, with many competing demands on their time. For example, rethinking unit design, improving the alignment between objectives, graduate attributes, course outcomes and assessment criteria, describing in detail the standard of evidence for five levels of achievement and redesigning assessment tasks.

The time and effort spent on these activities does not, however, guarantee that students know the

'rules of the game' (Orrell, 2005) such as what to do with criteria sheets and how academics grade students responses to assessment tasks and give overall unit grades. Teaching students the 'rules of the game' is a necessary part of helping them develop control over their own learning. Tan, for example, advocates that students need to be able to make judgments independently about their own learning (or self-assess, self-monitor, self-evaluate), because it is a vital component for lifelong learning.

> … the most critical need for students to meet their own future learning needs is their capacity to judge what their own learning needs are and how they can go about meeting these needs. (Tan, 2008, p. 27)

Sadler maintains that, in a criteria and standards system, students can develop the skills of *self*-evaluation and therefore, control the quality of their own work. In a norm-referenced system, students' work is compared to the work of others and students develop the skills of evaluating the achievement of *other* students to ensure they maintain (or improve) their relative position (Sadler, 1998). The research literature reports various ways of teaching students the 'rules of the game' in relation to criteria sheets. All agree that, by themselves, criteria sheets, no matter how explicit or whether accompanied by definitions, are still just words on a page (or online) and have little effect on students learning (Rust, Price & O'Donovan, 2003; Carlson, MacDonald, Gorely, Hanrahan, & Burgess-Limerick 2000; O'Donovan, Price & Rust, 2004). Simply providing criteria sheets to students does not guarantee they will have the same understanding of the written expectations as the assessor (O'Donovan et al. 2004). Students should be actively engaged in the discussion and application of criteria and standards (Woolf, 2004) and this should be via a combination of knowledge transfer processes (O'Donovan et al. 2004).

The cohorts we currently teach in this unit are predominately generation Y (or gen Y) students. That is they were born between 1976 and 1991 (Williamson, 2008), although exact years differ between authors. Donnison comprehensively reviewed the literature on gen Y and found almost all of it was written by older non-gen Y authors (2007). These authors, while not all agreeing on the 'location' of the generation (in terms of time frame), tended to agree on gen Y characteristics. That is, these students are considered to have a propensity for digital media, are overconfident in their own abilities and a have a collaborative orientation (2007). Like Donnison (2007), we were sceptical of these 'agreed' characteristics. Much of the literature on gen Y is written from the perspective of how to attract and retain them in the workplace. Williamson for example, says that employers should 'strive for transparency of processes', give gen Y staff 'freedom to manage their own workloads' and 'infuse work with purpose and a sense of fun' (2008:61). We maintain these strategies are applicable to any generation and to university teaching. Hence, we agreed not to treat these students as fitting a predictable stereotype or to spend extra time tailoring our approach specifically to gen Y. Rather, we wanted to develop an approach that would be considered by them as good teaching, regardless of generational type casting.

The aims of this paper are to:

1. Report Science students' reactions to an innovative approach in the use of criteria sheets
2. Advocate that this approach can benefit all students regardless of their generational grouping or discipline

Method

At the start of a third year (exit) unit in Microbiology, we run an 'assessment' tutorial to teach students the 'rules of the game'. This involves immersing students in the language of assessment, helping them to demystify criteria sheets, making expectations and methods of judgments explicit and transparent and giving them the opportunity to develop self-evaluative expertise. Students evaluate the tutorial and at the end of the unit, their use of criteria sheets. The tutorial requires students to match anonymous samples of responses to assessment tasks (from previous students), to descriptors on three criteria sheets (one written assignment, a theory exam and a practical exam). These samples range in quality from highest to lowest standard. Students then decide on a grade, in discussion with their peers and orally support their judgments. We then orally validate the grades they give to the samples. This exercise is repeated so that students can fine-tune their judgments based on this feedback and learn to use the language of assessment. Students are also let in on the 'secrets' of making judgments for whole tasks and units as they practice using hypothetical profiles and a set of judgment (or grading) rules. These rules show how grades awarded for achievement in each criterion are combined according to task weightings. The exam criteria sheets are supported by online practice exams with answers and comments, plus graded examples of previous students' responses (to the written task). These materials remain online for the duration of the semester.

A role play is then conducted to determine how well students have engaged with the preceding activities. In our team, one of us acts as a 'difficult student', one is the 'supportive academic', while the other is the 'controller of the action'. The 'difficult student' asks a range of questions of the 'supportive academic' and part-way through the 'academic's' answers, the 'controller' freezes the 'action' and selects a student to 'finish the answer'. This action is repeated and varied so that different students play the role of the 'difficult student' or the 'supportive academic'. The script is devised to be free flowing, but also to incorporate some complex issues about assessment. Much humour and discussion results, especially when students are shown the incomplete and the less than remarkable assessment profile of the 'difficult student' (who, in reality, is a highly regarded academic and one of their lecturers).

Our approach to revealing the 'rules of the game' differs from those reported in the literature in two ways:

1. The requirement for students to use judgment or grading rules to decide on overall grades for tasks and units and not just grade examples in terms of individual criteria
2. The use of a role play to check whether students could apply their new skills to challenging, but realistic assessment scenarios

Our approach shares similarities to those reported in the following sample of literature, e.g. Forbes & Spence, (1991), Stefani, (1992), Carlson, MacDonald, Gorely, Hanrahan, & Burgess-Limerick, (2000), Elwood and Klenowski, (2002), O'Donovan, Price & Rust, (2004), Hughes, Hinchy, & Cappa, (2007) in the following three ways:

1. We use a series of marking (or grading) exercises that require matching examples to descriptors on criteria sheets
2. Students are required to discuss the grades they award with each other and justify them to the larger group
3. Models, examples and student discussion are involved

Results and discussion

Despite many not liking the early morning winter start in 2007, 68% of the 44 students attended the 8am tutorial. Tables 1 and 2 present quantitative data from two student surveys: one of the tutorial (held at the start of the unit) and the second of their use of the criteria sheets (data collected at the end of the unit). The number of 'no answer' responses is not shown in Table 1.

Table 1: Quantitative data from student survey #1: tutorial on using criteria sheets to guide responses to assessment tasks

Question	yes	no
Q1. Did this tutorial help you to understand how to use criteria sheets?	23	4
Q2. Did this tutorial help you to understand how grades are awarded for tasks and units?	28	1
Q3. Was early in the semester the right time to have this tutorial?	27	1
Q4. Was anything missing from this tutorial that would have improved your understanding of the main concepts?	7	20

Table 1 (n=30) shows that student evaluation of the tutorial was overwhelmingly positive and of those who attended, 90% said they thought it was at the right time in the semester. In 2008, 64% of the cohort of 55 attended the tutorial and their positive evaluations are almost identical to the 2007 data, despite the unfortunate timing of the tutorial (7-9pm).

Table 2: Quantitative data from student survey #2: how you used criteria sheets to guide responses to assessment tasks

Question	**yes**	**no**				no answer
Q1. Did you use the criteria sheets in planning how to respond to the assessment tasks?	**32**	**11**				1
Q2. We would like your feedback about the criteria sheets. (i) Were the criteria sheets the right length?	**40**	**1**				3
(ii). Were the descriptors sufficiently detailed?	**37**	**5**				2
(iii). Did you understand what was expected of you if you were aiming for a passing grade (standard 4) for each task?	**39**	**0**				5
(iv) Which one of the criteria sheets needs the most improvement?	none	prac exam	assignment	theory exam	all of them	
	19	**3**	**13**	**1**	**1**	7
Q3. Did you refer to the online examples of student responses in preparing for your assessment?	yes	no				
	37	**2**				**5**
Q5. Which was more helpful – the online examples or the criteria sheets?	onlin	criteria sheets	neither	both		
	24	**6**	**4**	**4**		**6**

The data in Table 2 (n=44; 14 of whom did not attend the tutorial) shows again that student evaluation of the use of criteria sheets was overwhelmingly positive. 84% of students reported that they used the online examples and 54.5% found these more helpful than the criteria sheets. Despite this, 72.7% used the criteria sheets in planning their responses (including 71.4% of those who had not attended the tutorial). The majority of student feedback about the quality of the criteria sheets (Question 2 in Table 2) was positive. Crucially, all students who answered question 2(iii) about understanding the expectations described for the passing grade (standard 4) said 'yes'.

Qualitative data from 2007 revealed that no students had experienced this type of assessment tutorial in their other units. Of the 67 comments made across the four questions in survey #1 (see table 1), 79% were positive and 21% negative. However, the four *negative* comments for question 1 were made by students who said they found the tutorial unnecessary, because they knew how to use criteria sheets from high school. Interestingly for the same question, six of the *positive* comments were made by students who also had used criteria sheets in high school.

Pleasingly, both these groups of students still chose to attend the tutorial, indicating that they all sought more information about their assessment tasks. Table 3 provides some samples of typical positive and negative responses to the tutorial.

Table 3: Examples of qualitative data from student survey #1: tutorial on using criteria sheets to guide responses to assessment tasks

Q1. Did this tutorial help you to understand how to use criteria sheets?
Yes, the nuances of the wording of the criteria were highlighted so that the differences between the grades are easier to understand. *Yes, was enjoyable, given a chance to really look through weighting of tasks, which will help balance out study in relevance to criteria sheet.*
No, used criteria sheets previously.
Q2. Did this tutorial help you to understand how grades are awarded for tasks and units?
Yes, helped me realise how important it is to consider the weightings and do well in areas weighted heavily and I guess it showed me how you can screw up the whole result for a unit by bombing out in one section of one task. *Yes, usually allows me to concentrate on providing accurate assessments – that is as close as you want from us.*
No, I have used criteria sheets in previous study and find them self explanatory.
Q3. Was early in the semester the right time to have this tutorial?
Yes, having it early means that we have something to work towards and know what will be expected in each task and prepare and aim for a grade we want. *Yes, this allows us to concentrate earlier in the semester in providing the best assessable material.* *Yes, glad we finally get to look at what you want*
No, closer to the assessment dates preferred
Q4. Was anything missing from this tutorial that would have improved your understanding of the main concepts?
No, thanks – it is great that criteria assessed tasks are finally being implemented (students aren't mind readers).
Yes, would have been useful to go through the requirements of the assignment rather than such a big focus on criteria and marking.

Qualitative data from the evaluation of how students actually used criteria sheets in the unit (survey #2) was gathered from those who had and those who had not attended the tutorial. Table 4 summarises the number and types of comments made by students.

Table 4: Qualitative data from student survey #2: how you used criteria sheets to guide responses to assessment tasks

	Attended tutorial (n=30)	**Did not attend tutorial (n=14)**	**Whole student cohort (n=44)**
Total number of comments	76	45	121
positive	79%	67%	74.4%
negative	21%	33%	25.6%

Typical student positive comments from student survey #2 include:

- The criteria sheets were good and provided a good overview of what to expect
- They provided detailed information as to what was expected from each piece of assessment
- It was very useful (criteria sheets). Would love to use them for every unit I do
- Very good – made us sure that the things we studied and used in assignments were exactly what the marker was looking for

There are several limitations of this data that restrict its usefulness. Student attendance at the tutorial was voluntary. This resulted in the small size of the 'did not attend tutorial' group, some of whom reported that the timing of the tutorial prohibited their attendance. For equity reasons, a 'matched control' group was not set up. While we can confidently conclude that the majority of the cohort found the criteria sheets useful, we are unable to state whether the tutorial was the main reason for this result. It could have been because students appreciated the quality of the criteria sheets and the associated examples of task responses or the fact that some students were familiar with using criteria sheets in high school. Table 5 shows that in 2007 (the first year of CRA implementation), there were fewer failures and a slight increase in the percentage of highest and passing grades compared to the previous three years.

Table 5: percentage of students awarded each grade

Year	enrolled students	7 (highest)	6	5	4 (pass)	2 (fail)
2004	32	3.1	9.4	28	60.6	6.2
2005	48	4.2	25	16.7	27.1	6.3
2006	24	12.5	16.7	41.7	29.7	4.7
2007	45	15.6	17.8	40	24.4	2.2

These results are similar to those reported by a number of authors and are not specific to gen Y cohorts: Carlson et al (2000), O'Donovan, Price, & Rust, (2002), McCune & Hounsell, (2005), Defeyter & McPartlin (2007), Cathers (2007), Hughes, Hinchy, & Cappa, (2007). These authors have employed a wide range of strategies to help students understand assessment expectations. Their respective papers demonstrated that these strategies were implemented via excellence in teaching, as defined by Hartman, Moskal & Dziuban (in Oblinger & Oblinger 2005) based on more than half a million student responses about what they thought was excellent teaching at

Central Florida University. Hartman et al found that all students, regardless of their generation, held a similar view. They expected teachers to:

> Facilitate student learning, communicate ideas and information effectively; demonstrate genuine interest in student learning; organise their course effectively; show respect for students and assess students' progress fairly and effectively. (Hartman et al. 2005 6.11).

Conclusions

In the process of implementing CRA, we took time out of delivering content to teach gen Y students, in an innovative way, how to judge their own work using criteria sheets. Evaluation results were similar to those in the literature for non gen Y cohorts. That is, students appreciated our efforts and were more aware of assessment expectations. We advocate our approach as a successful way of giving students control over their learning and the skills of self-assessment, regardless of generational groupings or disciplines.

References

Carlson, T., MacDonald, D., Gorely, T., Hanrahan, S., & Burgess-Limerick, R. (2000). Implementing criterion-referenced assessment within a multi-disciplinary university department. *Higher Education Research and Development*. 19(1), 103-116

Cathers, I. (2007). Case 24. Engaging with the Standards: Using Feed-forward and Feedback. In O'Neill, G., Huntley-Moore, S., & Race (Eds). (2007). *Case Studies of Good Practices in Assessment of Student Learning in Higher Education*. All Ireland Society for Higher Education (AISHE) Readings. 2007(1), 148-150

Defeyter, M., & McPartlin, P. (2007). Helping students understand essay marking criteria and feedback. Psychology Teaching and Review. March 2007, 23-33

Donnison, S. (2007). Unpacking the Millennials: A Cautionary Tale for Teacher Education. *Australian Journal of Teacher Education*. 32(3). (Retrieved 29 May 2008 from) http://ajte.education.ecu.edu.au/issues/PDF/323/Donnison.pdf

Elwood, J. & Klenowski, V. (2002). Creating Communities of Shared Practice: the challenges of assessment use in learning and teaching. *Assessment & Evaluation in Higher Education*, 27 (3), 243 — 256

Forbes, D. & Spence, J. (1991). An experiment in assessment for a large class. In: R. Smith (Ed.) *Innovations in Engineering Education*. London: Ellis Horwood.

Hartman, J., Moskal, P. and Dziuban, C. (2005) Preparing the academy of today for the learner of tomorrow. In D.G. Oblinger & J.L. Oblinger, (eds). (2005). *Educating the Net Generation* [online] Educause e-book available from www.educause.edu/educatingthenetgen/

Hughes, C., Hinchy, T., & Cappa, C. (2007). Case 4. Applying a Criterion and Standards Approach to Assessment by Examination in Law. In O'Neill, G., Huntley-Moore, S., & Race (Eds). (2007). *Case Studies of Good Practices in Assessment of Student Learning in Higher Education.* All Ireland Society for Higher Education (AISHE) Readings. 2007 (1), 31-35

Hughes, C. (2005). Casting some light on a shady practice: towards guidelines for the development of assessment criteria and standards. Paper presented at the *Evaluations and Assessment Conference*, University of Technology, Sydney (30 November – 1 December).

McCune, V. & Hounsell, D. (2005).The development of students' ways of thinking and practising in three final-year biology courses. *Higher Education* 49: 255–289

O'Donovan, B., Price, M., & Rust, C. (2004). Know what I mean? Enhancing student understanding of assessment standards and criteria. Teaching in Higher Education, 9 (3), 325 — 335

O'Donovan, B., Price M., & Rust, C. (2002). Chasing Common Assessment Standards. *Teaching Forum*, 50, Autumn, 12-13

Orrell, J. (2005). Assessment Literacy: A Precursor to Improving the Quality of Assessment. Keynote presentation. *Evaluation and Assessment Conference*, University of Technology, Sydney.

Rust, C., Price, M., & O'Donovan, B. (2003). Improving Students' Learning by Developing their Understanding of Assessment Criteria and Processes. *Assessment and Evaluation in Higher Education*, 28(2), 147-164

Sadler, R. (1998). Letting students into the secret: further steps in making criteria and standards work to improve learning. Paper presented at the *Queensland Board of Secondary School Studies annual conference for State Panel Chairs*.

Tan, K. (2008). Qualitatively different ways of experiencing student self-assessment. Higher Education Research & Development, 27 (1) 15 - 29

Williamson, A. (2008). Gen Y. *Teacher: the National Education Magazine*. ACER. May 2008, 60-61

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