COMMUNICATING ASSESSMENT CRITERIA IS NOT SUFFICIENT FOR INFLUENCING STUDENTS' APPROACHES TO ASSESSMENT TASKS – PERSPECTIVES FROM A DIFFERENTIAL EQUATIONS CLASS

Dann G. Mallet
Faculty of Science and Technology, Queensland University of Technology
dg.mallet@qut.edu.au

Jennifer Flegg
Faculty of Science and Technology, Queensland University of Technology
and Oxford Centre for Collaborative Applied Mathematics, University of Oxford
j.flegg@qut.edu.au

This report presents the findings of an exploratory study into the perceptions held by students regarding the use of criterion-referenced assessment in an undergraduate differential equations class. Students in the class were largely unaware of the concept of criterion referencing and of the various interpretations that this concept has among mathematics educators. Our primary goal was to investigate whether explicitly presenting assessment criteria to students was useful to them and guided them in responding to assessment tasks. Quantitative data and qualitative feedback from students indicates that while students found the criteria easy to understand and useful in informing them as to how they would be graded, the manner in which they actually approached the assessment activity was not altered as a result of the use of explicitly communicated grading criteria.

Key words: differential equations, assessment experiment, criterion-referenced assessment

Introduction

Criterion-referenced assessment (CRA) is assessment that is constructed with the intent to measure student performance that can be explained with reference to clearly delineated learning tasks (Linn & Gronlund, 2000). It involves identifying the extent of a learner's achievement of predetermined goals or criteria and fundamentally involves assessing a student without reference to the performance of others (Brown, 1988; Harvey, 2004; TEDI, 2006). When CRA is used it requires an underlying set of course learning outcomes, an assessment program designed to gather information about a student's performance in relation to those learning outcomes, and importantly the communication of criteria and standards (a two dimensional view) which inform students how they will be judged and to provide directions for assessors.

Generally speaking, Australian Universities impose or very strongly encourage the use of criterion referenced or standards based assessment in the courses that they offer. At the authors' home institution this is no different with the University's Manual of Policies and Procedures stating that the University "has adopted a criterion-referenced approach to assessment where assessment is based on pre-determined and clearly articulated criteria and associated standards of knowledge, skills, competencies and/or capabilities" (QUT, 2011). Furthermore, the policy states that assessment is "clearly communicated to students" and used as "a strategy to support student learning" (QUT, 2011). In essence, the implication is that a particular method of assessment, CRA, is imposed on all teaching academics so as to ensure students are aware of how they are

being assessed and because it is useful in supporting students in their learning. In this study, we challenge this perception by appealing to the thoughts of students themselves.

Until recently, the directive to employ CRA has been largely ignored in the context of many quantitative courses such as those in mathematics and the sciences. Lecturers regularly justify this; claiming that quantitative studies involve assessment responses that are either right or wrong and that right/wrong are sufficient. In this study we report on the successful implementation of elements of criterion referenced assessment into a Differential Equations course that goes beyond simple "right-wrong" criteria while maintaining the mathematical integrity of the assessment program. Specifically, we unpack the usual collection of "right" and "wrong" judgments and collect them into groups related to the learning outcomes of the course, thereby providing assessment criteria. We present findings based on quantitative and qualitative feedback from students regarding their perceptions of criterion referencing and how it is used in guiding their learning throughout the course.

It is important to place this study in context by comparing the assessment experiment with the methods previously used to assess students in the course. Over approximately the past 10 years, the course has been taught by four different lecturers, however the assessment strategy has essentially been to employ 1-2 assignments (problem solving tasks with a 2-4 week completion timeframe) and a mid-semester and final examination. These tasks generally contribute 30-40% (assignment) and 60-70% (examination) of the final grade for the course, respectively. Assessment of students on all of these tasks has been carried out using what we refer to as the "traditional method" for mathematics assessment and *not* criterion-referenced assessment. That is, the academic responsible for assessment writes an examination or assignment, along with his or her own set of "correct" solutions. The set of correct solutions is annotated with points or marks throughout the solutions where points correspond with reaching certain points in the solution process. Assessment using the traditional method involves making judgments as to whether a student is right or wrong at various points in a solution procedure and makes no explicit reference to the learning outcomes that the academic intends students to obtain as a result of undertaking the course of study.

In the assessment experiment reported on in this paper, we have attempted to maintain the previously employed assessment program as much as possible. In particular, we maintained progressive, non-examination assessment of 40% and used mid-semester and final examination contributing 60% of the students' final grades. However, we implemented a criterion-referenced method of grading students in the assignment tasks completed during semester. This involved presenting students with a set of criteria and definition of standards in addition to the actual problems to be solved. Rather than simply implying that students would be marked right or wrong up to some number of points as is the case in the tradition method, students were provided with details of exactly how responses to the mathematical problems would be graded and how translation between the mathematics and the standards and criteria would be carried out.

Our goals in conducting this small-scale experiment fall into two main areas: to gauge students' perceptions regarding criterion referenced assessment and its usefulness, and to a lesser extent, evaluating the motivation for effecting culture change among mathematics academics. With regard to students' perceptions, we investigated how students viewed the understandability and the usefulness of criterion referencing and how they employed the additional information provided to them via the criteria and standards definitions in directing their learning and assessment responses. Implicitly, we believe that such an investigation and its results can then be

used to effect culture change among mathematics teachers at universities by changing the way they view criterion referenced assessment, taking CRA from a directive imposed by administrators to a useful tool for mathematics learning.

This paper presents the analysis and implications obtained from a small-scale, mixed methods study of the use of criterion-referenced assessment in an undergraduate differential equations class. While the study has focused on a single course, the authors expect that the findings of the study could generally be carried over to other similar courses in applied mathematics, particularly in the Australian context. A similar study is currently being undertaken in a more advanced partial differential equations course and a comparison of findings in the two (albeit somewhat similar) contexts is currently under preparation.

Literature Review

While there is extensive practical experience and significant literature relating to the use of criterion-referenced assessment for mathematics at the school level, literature that describes the use of CRA in university level mathematics classrooms is close to nonexistent. Furthermore, at the time of undertaking this research, the authors were unable to find any published research discussing student perceptions regarding CRA and its impact on their learning process.

Niss (1998, in Pegg 2003, p.228) notes that mathematics assessment identifies and appraises the knowledge, insight, understanding, skill and performance of a student. Pegg however points out that this is not in fact the reality of assessment in mathematics and that rather, it is most often concerned with reproduction of facts and computational skills or algorithms (Pegg 2003). It is our contention that this is how previous years' assessment programs for the course under investigation have been presented to students. In the assessment experiment discussed in this report, we attempt to explicitly link the subtasks of the assessment activities with the learning outcomes of the course, which include such concepts as knowledge, insight and understanding in addition to skills. In this way we believe that our assessment becomes more of an educational tool for students than it has been in previous versions of the course, and that it allows for a more "constructive alignment" (in the sense of Biggs, 1996) of the content, pedagogy and assessment.

Criterion referenced assessment involves determining the extent to which a learner achieves certain predetermined goals or criteria, importantly, without reference to the performance of others (Brown, 1988; Harvey, 2004; TEDI, 2006). The implementation of CRA involves the design or statement of a set of learning outcomes for a course, design of a program of assessment to obtain information about a student's performance in relation to the learning outcomes, and the presentation of a criteria set and definition of standards which serves to both inform students how their performance will be judged and to provide directions for assessors.

Pegg (2003) notes that while the movement towards assessment based on outcomes and standards (rather than individual comparison) did initially have some basis in research regarding student learning, the links remain tenuous. As such, there is debate among teachers and academics alike as to whether the claims regarding the benefits of criterion referenced assessment are supported by strong research. Through research such as that presented in this study, we attempt to provide a research base that advocates the benefits and warns of the pitfalls of criterion-referenced assessment in the undergraduate mathematics classroom.

Conceptual Framework

In this study we carry out descriptive research related to questions around student perceptions and criterion referenced assessment. This descriptive research involves statistical and textual

analysis/synthesis of data collected from a student population undertaking a course in differential equations in an attempt to understand student perceptions and provide guidance for academic staff in undertaking more useful assessment in mathematics courses. The context of an undergraduate differential equations course, described in the next section, was chosen due to its representativeness of typical applied mathematics courses and hence the potential for maximum transfer of findings across applied mathematics teaching and learning.

Context

The course that was used for the experiment described in this study was a second year undergraduate ordinary differential equations course. Content covered in the course included first and *n*th order linear equations, series solution methods, Laplace transform solutions, linear systems of differential equations, phase portraits, Bessel, Legendre and Cauchy-Euler equations and Fourier series solutions. While officially the prerequisite knowledge required for students to enter the course included advanced calculus or linear algebra (at the second year level), the actual requirement was only understanding first order ordinary differential equations as typically covered in first year level courses.

The cohort included 52 undergraduate students and 4 additional coursework postgraduate students (although the course content was second year undergraduate level). Teaching activities involved 3 hours per week of lectures presented in one 2 hour block and one 1 hour block to the entire groups as well as 1 hour per week of smaller group "tutorial" sessions with additional teaching assistance. The course ran for 13 weeks with a one-week mid-term break.

The official course outline lists the following learning outcomes for the differential equations course:

1. Engage your critical thinking skills to understand the principles of and develop theoretical knowledge regarding differential equations.
2. Draw on a range of your thinking skills to identify, define and solve real world and purely mathematical problems using existing knowledge and knowledge developed in this unit.
3. Communicate your theoretical understanding and problem solving attempts in methods appropriate to the context of this unit.
4. Demonstrate independence and self-reliance in retrieving and evaluating relevant information and in advancing your learning.

These rather broad objectives can be summarized as an intention to facilitate students developing critical thinking skills and theoretical knowledge, retrieving and evaluation relevant information, developing ability to identify, define and solve problems, and communicate results.

The assessment package included a 30% end semester exam, a 30% in semester exam (week 10), 2 problem solving tasks (weeks 4 and 8) totaling 30% and 2 short multiple choice quizzes (weeks 2 and 5) contributing 10% to the student's grade. It is the problem solving tasks, worth 15% each, that are specifically of interest in this study as these were the items assessed using explicitly communicated criterion referenced assessment. In particular, the marks allocated using a "traditional method" of assessment were analyzed and grouped into categories related to the learning outcomes of the course. The standards for each criterion were then determined by weighting with regard to the marks achieved in each category. We refer to this method of CRA as the "frequency-based standard allocation" and note for the reader's reference that taxonomy of standard allocations in applied mathematics CRA is the topic of other research (in preparation) by the authors.

Learning Outcome	Criterion	7 HD	6 D	5 C	4 P	3 MF	2 F	1 LF
2,4	*[20%] Correct identification of equation types and appropriate solution procedures*	Always or nearly always	Occasional errors	More than half time	About half the time	Less than half the time	Rarely	Never
1,2,4	*[70%] Apply existing and newly developed knowledge/skills to correctly arrive at solutions/results*	Consistent, effective and efficient throughout	Consistent and effective in most questions	Effective more than half the time	Around half the time	Less than half the time	Rarely correctly applied	Never or very rarely applies knowledge correctly
3	*[10%] Communication of mathematics/problem solving attempts*	Generally clear, logical	Minor/few errors and/or inconsistencies	Some errors and inconsistencies	Regular errors and inconsistencies	Many errors and logical inconsistencies	Rarely clear or logical	Meaningless or no attempt at communicating

Mapping of marks to questions and criteria Criteria	Q1	Q2	Q3	Q4	Q5
Correct identification of equation types and appropriate solution procedures	2	1	1	1	1
Apply existing and newly developed knowledge/skills to correctly arrive at solutions/results	4	3	5	4	5
Communication of mathematics/problem solving attempts	1	0.5	0.5	0.5	0.5

Figure 1: Criteria and standards along with a mapping of marks to criteria for one of the two problem solving tasks assessed using CRA in this study.

To maximize the feedback provided to students, and hence maximize support of student learning, students were provided with both the traditional feedback of "points" (with ticks and crosses) annotated upon their submissions and an annotated version of the criteria map shown in Figure 1. To fully elucidate the concept of frequency-based standards allocation employed in this study, we provide an example with reference to Figure 1. At the university where this study was undertaken a grade of 7 generally corresponds with a score of 85%-100% while a grade of 6 with a score of 75%-84%. Consider only criterion 1 and suppose a student scored 1 out of 2 points for question 1, and then 1 out of 1 point for each of the remaining questions. This gives 5/6 points or 83.3% for criterion 1 over the entire assessment item. Using the frequency-based standard allocation method this student would score a grade of 6 for criterion 1 on the problem solving task. The idea here is that students are providing not only with the tradition feedback of points, ticks and crosses, but also an explicit mapping of these ticks and crosses to criteria related to the learning outcomes of the course, with a view to providing them with deeper information regarding where they are progressing well and where they are struggling in the course.

Methods

We have used two primary data sources, one quantitative and one qualitative, in an attempt to address our research goals regarding student perceptions of criterion referenced assessment. The quantitative source involved a Likert scale survey while the qualitative tool comprised two questions to which students were requested to respond in free-text format. For maximum flexibility, both data collection tools were deployed online using SurveyMonkey. In order to gauge the full impact of the assessment experiment, surveys were conducted at the end of the course of study, following the provision of feedback to students on all criterion-referenced items and also following the post mid-semester exam feedback sessions. All 56 enrolled students were

given the opportunity to take part in the survey and a 50% response rate was achieved at survey closure.

The first data collection tool was a 10-item survey using a 5 level Likert scale. Students were presented with the 10 items and asked to choose which response – strongly agree, agree, neutral, disagree or strongly disagree – best described their feeling regarding each statement in turn. The data collected is presented in full in Table 1. Numerical and statistical analyses of the Likert-survey were conducted with findings presented in the remaining two sections of this paper.

The second data collection tool was a survey allowing free-text responses on two questions of interest. The first question given to responders was

"*I would say that the impact of having an assessment task being marked by criterion referenced assessment on my approach to learning was ...*"

while the second was

"*I see the educational benefits of using criterion references assessment as: ...*".

The intention of the first question being to elucidate the student's perception of how CRA impacted on their own learning and their approach to assessment items completed *after* the problem solving tasks assessed using CRA. The second question was intended to obtain the student's wider view regarding the benefits of CRA. Textual analysis and synthesis was carried out on the free-text responses and again, findings and discussions are presented in the remaining sections of this work.

Results

The quantitative data collected via the first of the student surveys is presented in summary form in Table 1. The data indicates that while students found assessment criteria easy to understand and useful in informing them as to how they would be graded (items 1 and 10, 5 and 8), it did not alter the way they actually approached the assessment activity (item 2). Interestingly, on the whole it did not seem that students felt strongly that CRA provided more useful feedback to them than the traditional method in terms of preparing for future assessment (item 9). There was a similar almost uniform spread of responses regarding whether students found CRA useful at all (item 6).

	SA	**A**	**N**	**D**	**SD**
I feel that the details of the criterion referenced assessment (CRA) guidelines were made clear to me early in the semester.	**42.9% (12)**	39.3% (11)	14.3% (4)	3.6% (1)	0.0% (0)
I found that the way I approached completing the assessment task was different, given that I had the CRA sheet describing exactly how I would be graded.	14.3% (4)	14.3% (4)	10.7% (3)	25.0% (7)	**35.7% (10)**
Being assessed with a CRA sheet seems to me to be the best way that mathematics assessment can be graded.	3.6% (1)	17.9% (5)	**39.3% (11)**	14.3% (4)	25.0% (7)
I am so used to being assessed using CRA sheets that being assessed in mathematics by this method did not concern me.	0.0% (0)	21.4% (6)	**46.4% (13)**	17.9% (5)	14.3% (4)
The information provided by the CRA sheet made it clearer to me what was	17.9% (5)	**32.1% (9)**	25.0% (7)	17.9% (5)	7.1% (2)

expected of me in order to get a particular grade.					
I found the CRA sheets useful/helpful.	10.7% (3)	21.4% (6)	**32.1% (9)**	21.4% (6)	14.3% (4)
I believe I understand the educational benefits of CRA.	7.1% (2)	**42.9% (12)**	32.1% (9)	10.7% (3)	7.1% (2)
I feel to some extent that the CRA sheet demystified the way that marks are allocated in the assessment piece.	10.7% (3)	**42.9% (12)**	21.4% (6)	14.3% (4)	10.7% (3)
I feel that the CRA grade provided me with more feedback on how I had performed in the assessment task and where I could improve in the future than a mark out of a total does.	14.3% (4)	**25.0% (7)**	21.4% (6)	17.9% (5)	21.4% (6)
I found the categories on the CRA sheet (ie, communication, problem solving) easy to understand.	21.4% (6)	**67.9% (19)**	10.7% (3)	0.0% (0)	0.0% (0)

Table 1: Students were asked to read each item and select the response which best described how they feel about the statement. SA=strongly agree, A=agree, N=neutral, D=disagree and SD=strongly disagree.

Qualitative feedback from almost 100% of respondents indicated that in general the criteria provided were not used to determine how a student would approach individual questions or the assessment task as a whole. Interestingly, a similar percentage of students stated that they found CRA beneficial as it made the process of allocating scores by graders much clearer. A small percentage of students indicated that they did refer to the criteria sheets after the tasks were graded in order to get a different, higher level representation of where they had made errors in their responses.

Implications For Future Teaching Practice

The analysis of the data collected during this study indicates that while the concept and practice of CRA was clearly explained, and CRA sheets provided better guidance as to what was expected for different grade levels/marks, immediate and post-feedback learning approaches were not greatly altered. This research study has opened up new questions for future research. For example, we are now considering the impact on graders/academics and the usefulness they perceive in employing criterion referenced assessment.

With regard to application in the classroom in the future, both the qualitative and quantitative data indicate that students and graders alike, need to be explicitly informed exactly why they are provided with criteria and how they can be used to assist learning. Only 11 of 28 responded that they agreed in any way that CRA was more useful than traditional assessment for the purposes of preparing for future assessment items. Guiding students in their response attempts (showing them what the grader will deem to be “important”) and also aiding them in understanding the feedback they receive following the grading of their work are important benefits of CRA that should be communicated to students so that they may best use the feedback provided to them.

The actual construction of the criteria and standards is by no means straightforward. In the free-text responses, students indicated that in a general educational context they see CRA as

providing better feedback, more guidance about how to approach a solution and an element of grading transparency. Clearly then, the process of constructing the criteria and standards is important, because these are where students gather this additional information and transparency. The criteria and standards must be carefully designed and worded so that they are exactly the types of judgments the grader is using while assessing students' work. Academic staff need to be closely guided in the development of these elements of any criterion referenced assessment strategy.

Finally, we return to the point made in the introduction of this paper, namely that CRA is imposed on all teaching academics so as to ensure students are aware of how they are being assessed and because it is useful in supporting students in their learning. In this small-scale study, we have shed light on the fact that while CRA may be useful in raising student awareness about the assessment process, it is not sufficient in itself as the assessment method of choice to support students in the learning process. In fact, it may be more important to educate students regarding "how" to use feedback at all as a way to assist in their learning process, rather than to rely solely on the method itself.

References

Biggs, J. B. (1996) Enhancing teaching through constructive alignment, Higher Education, 32, pp. 347-364.

Brown, S. (1988) Criterion referenced assessment: what role for research? in Black, H.D. and Dockerell, W.D., New developments in educational assessment, British Journal of Educational Psychology, Monograph series no. 3, pp. 1-14.

Harvey, L. (2004). Analytic Quality Glossary http://www.qualityresearchinternational.com/glossary/#c. Accessed June 12, 2010.

Linn, R.L. and Gronlund, N.E. (2000). Measurement and Assessment in Teaching, (8th Ed). Upper Saddle River, NJ: Prentice Hall.

Pegg, J. (2003) Assessment in mathematics: A developmental approach, in Mathematical Cognition, J.M. Royer (Ed.), p. 227-259, Information Age Publishing.

Queensland University of Technology (QUT). (2011) QUT Manual of Policies and Procedures C/5.1 Assessment http://www.mopp.qut.edu.au/C/C_05_01.jsp#C_05_01.05.mdoc. Accessed March 24, 2011.

Teaching and Educational Development Institute (TEDI). (2006) Teaching and Learning Support: Criterion referenced assessment http://www.tedi.uq.edu.au/teaching/assessment/cra.html. Accessed June 15, 2010.