



Queensland University of Technology
Brisbane Australia

This is the author's version of a work that was submitted/accepted for publication in the following source:

Mallet, Dann G. & Flegg, Jennifer (2011) Towards a classification of criterion referenced assessment models in mathematics courses : student and academic perspectives. In Barton, S. (Ed.) *Proceedings of Global Learn Asia Pacific 2011*, AACE, Melbourne, Australia, pp. 479-482.

This file was downloaded from: <http://eprints.qut.edu.au/41328/>

© Copyright 2011 Please consult the authors.

Notice: *Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source:*

Towards a Classification of Criterion Referenced Assessment Models in Mathematics Courses – Student and Academic Perspectives

Dann Mallet and Jennifer Flegg
Queensland University of Technology
Australia
dg.mallet@qut.edu.au
j.flegg@qut.edu.au

Abstract: We present the findings of a study into the implementation of explicitly criterion-referenced assessment in undergraduate courses in mathematics. We discuss students' concepts of criterion referencing and also the various interpretations that this concept has among mathematics educators. Our primary goal was to move towards a classification of criterion referencing models in quantitative courses. A secondary goal was to investigate whether explicitly presenting assessment criteria to students was useful to them and guided them in responding to assessment tasks. The data and 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, it did not alter the way the actually approached the assessment activity.

Introduction

In the Australian university context, it is generally the case that the use of criterion referenced or standards based assessment is imposed or very strongly encouraged in the courses on offer. At the authors' home institution, the situation is no different with official policy 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." The view among academics, and indeed students themselves, in many quantitative disciplines such as mathematics and the sciences is that criterion-referencing amounts to assessment overkill, providing no further valid information compared with traditional assessment methods and simply adds to already high workloads. Directives to the effect of imposing criterion referencing in these disciplines are sometimes ignored, or at best dealt with cursorily with the claim at right answers against wrong answers is criteria enough in such fields of study.

In this paper, we report on a study involving the successful implementation of a number of explicitly criterion referenced assessment items in undergraduate mathematics courses at a large public university in Australia. First we discuss the successful implementation of this relatively uncomplicated model of criterion referenced assessment (CRA) that goes beyond simple "right-wrong" criteria while maintaining the mathematical integrity of the assessment program for assessment items of problem solving type in mathematics courses. That is, the assessment items are mathematics problems and not reports, and they do not contain or require long blocks of written text. Next we discuss 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. Finally, we present an introductory discussion regarding a classification of different models of criterion-referenced assessment in mathematics courses arising out of debate and discussion with mathematics academics.

Context

The courses in which this study was embedded were both mathematics courses at a large Australian public university. One course, *Differential Equations* (DEs), is nominally at second year level while the other, *Partial Differential Equations* (PDEs), is third year level. The DEs course had 56 enrolments (52 undergraduate and 4 postgraduate) while the PDEs course had 52 enrolments (48 undergraduate and 4 postgraduate). Both courses have 3 hours of lectures per week and 1 hour of small group or multiple tutor tutorial classes. The structure of the assessment program for both courses is shown in the following table.

Assessment item	Weighting	Due
-----------------	-----------	-----

Quiz 1	5%	End week 2
Problem solving task 1	15%	End week 4
Quiz 2	5%	End week 5
Problem solving task 2	15%	End week 8
Core exam	30%	Week 10 & Repeated end of semester
Advanced exam	30%	End of semester

Table 1: Assessment items, weighting and due dates for the assessment program of both courses

It is important to place this study in context by comparing the assessment experiment with the methods previously used to assess students in the courses. Over approximately the past 10 years, the courses have been taught by a number of people, 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 courses, respectively. Grading of all tasks has been quite traditional in the sense that the academic responsible for assessment writes their own set of “correct” solutions and assigns points or marks throughout the solutions corresponding with reaching certain points in the solution process.

In the assessment experiment reported on in this paper, we have attempted to maintain the character of the previously employed assessment program as much as possible. This was achieved by maintaining progressive, non-examination assessment of 40% and by using mid-semester and final examinations contributing 60% of the students’ final grades. However, we implemented an explicitly criterion referenced method of grading students in the assignment (or Problem Solving) tasks completed during semester. This explicit criterion referencing involved presenting students with a set of criteria and standard definitions in addition to the actual problems to be solved. 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.

Objectives

Our objectives in conducting this study fall into two main areas: to gauge students’ perceptions regarding criterion referenced assessment and its usefulness and to initiate a classification system for models of criterion referencing in mathematics, and quantitative disciplines in general, that has deep connection with standards expected by the discipline as opposed to external policy makers. 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. Debate and discussion with university-level mathematics educators allowed for an investigation of existing perceptions of and models of criterion referencing in mathematics, and to move towards a system of classification of such models.

Literature review and positioning of this research

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 courses 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 courses, 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 courses, 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). It is assessment designed to measure performance that is interpretable in terms of a clearly defined and delimited domain of learning tasks (Linn & Gronlund, 2000). 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.

Initially, the movement towards outcomes and standards based, rather than norm referenced, assessment did have grounding in research regarding student learning. However Pegg (2003) notes that the links remain tenuous at this time. This is especially the case in university-level, rather than school-level, education. 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. It is our intention through research such as that presented in this study; to provide a research base that advocates the benefits and warns of the pitfalls of criterion referenced assessment in the university-level mathematics classroom. Through careful classroom experiment, and discussion guided by teaching experts, we intend to develop a classification framework for models of criterion referencing in university-level mathematics and their impacts on learning.

Theoretical perspective/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.

Methods

We have used three primary data sources in this study, one quantitative and two qualitative, in an attempt to address our research goals regarding criterion referenced assessment. One quantitative and one qualitative source were delivered via means of online survey. The first source was a survey allowing free-text responses on two questions of interest, while the second was a 10-item survey using a 5 level Likert scale. Both surveys were conducted at the end of the courses of study, following the provision of feedback to students on the criterion-referenced items and also following the post mid-semester exam feedback sessions. All 56 students in the DEs cohort and 52 students in the PDEs cohort were offered the chance to respond, with 46% and 39% response rates achieved, respectively. Another source of data that will be commented upon, although to a lesser extent, are the assessment responses themselves. Numerical and statistical analyses of the Likert-survey were conducted, while textual analysis and synthesis was carried out on the free-text responses. The second of the qualitative data sources involved a number of sessions of discussion and debate with mathematics educators from around the region of the authors' home institution (although, not restricted to that same institution). These discussions were used to frame the preliminary classification of CRA models.

Results

The data collected on the surveys of student perceptions indicates that while students found assessment criteria easy to understand and useful in informing them as to how they would be graded, it did not alter the way the actually approached the assessment activity. 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 tasks 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.

The discussions with mathematics educators led to a preliminary classification of models of criterion referencing in mathematics courses. This classification is shown in (Tab. 2).

Model	Brief description
Frequency	Standards of achievement are based on frequency of demonstrating tasks associated with criteria
Compounding/Complexity	Standards are based on demonstrating increasingly complex levels of tasks associated with a criteria
Completion	Standards reflect levels of completion of tasks of various types.
Exemplar	Standards are based on model solutions or exemplars (provided <i>a priori</i> by the teacher/lecturer)

Table 2: A preliminary classification of models of criterion referencing in mathematics courses.

Implications/Applications

Unlike humanities-based disciplines and disciplines that employ assessment programs involving a high level of written reporting, mathematics and similar quantitative disciplines often do not employ explicitly criterion referenced assessment strategies. Reflecting this, the qualitative and quantitative data collected as part of this study 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. In particular, guiding them 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.

The construction of the criteria and standards for mathematical courses is by no means straight forward – but it is important, because these are exactly the types of judgements we are normally making in an implicit, content-centred manner. Academic staff need to be closely guided in the development of these elements of any criterion referenced assessment strategy. While this study involved a frequency-based criterion referenced assessment model (see Tab. 2) where standards simply reflect the frequency of a student demonstrating a particular criteria, other models of CRA described briefly in (Tab. 2) would require different criteria setups – it is essential that the criteria and standards are chosen appropriately to reflect the assessment philosophy of the teacher and to provide an accurate measurement of the achievement of the student.

References

- Biggs, J. B. (1996). Enhancing teaching through constructive alignment, *Higher Education*, 32, 347-364.
- Brown, S. (1988). Criterion referenced assessment: what role for research? in Black, H.D. & Dockerell, W.D., *New developments in educational assessment*, British Journal of Educational Psychology, Monograph series no. 3, 1-14.
- Harvey, L. (2004). *Analytic Quality Glossary* <http://www.qualityresearchinternational.com/glossary/#c>. Accessed June 12, 2010.
- Linn, R.L., & 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 J.M. Royer (Ed.), *Mathematical Cognition*, 227-259, Information Age Publishing.
- 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.