



THE HONG KONG
POLYTECHNIC UNIVERSITY

香港理工大學

Learning and Teaching Committee



Guidelines *for* **Implementation of** **Criterion-Referenced Assessment**

1. Preambles

- 1.1 In its 38th meeting dated 18 June 2004, the Senate approved the proposal to adopt, with effect from 2005/06, the criterion-referenced approach for assessment and the corresponding removal of the “guideline distribution” for grading students’ overall performance in subjects.
- 1.2 This paper describes the meaning and rationale of criterion-referenced assessment, and suggests guidelines for implementing criterion-referenced assessment effectively in PolyU.

2. Criterion-Referenced Assessment versus Norm-Referenced Assessment

- 2.1 Criterion-referenced assessment is an approach of assessing students according to the extent to which a particular student has achieved the intended learning outcomes of a subject or programme. Students’ performance is assessed on the basis of *pre-determined* criteria and performance standards, which are clearly defined and explicitly communicated to the students *before* teaching and assessment take place. Where a grade is assigned, it will be assigned in accordance with the *absolute performance standard* the student has demonstrated on the criteria, independent of the performance of other students in the same cohort. There is *no* pre-set distribution of grades for the assessment. That is, it is possible in theory for all students to obtain a distinction or failure grade in the assessment although in practice, a distribution across grades is expected as students’ performance will inevitably vary.
- 2.2 Norm-referenced assessment, on the other hand, refers to an assessment approach in which judgments are made about a particular student by comparing his/her performance with that of the other students. In other words, students’ performance is assessed by reference to the achievements of other students in the cohort. Where a grade is assigned, it will be assigned in accordance with the student’s *relative achievement* on a pre-set grading distribution, normally determined by the institution. This assessment approach is commonly known as “grading on the curve”, and only a fixed proportion (within a certain limit) of students can be awarded a certain grade.

3. Why Adopt Criterion-Referenced Assessment?

- 3.1 In its recent curriculum revision exercise for the triennium 2005-08, PolyU is moving towards an outcome-oriented curriculum, which emphasises a clear delineation of the intended learning outcomes of all programmes and subjects, and a constructive alignment between intended learning outcomes, teaching, learning and assessment. Criterion-referenced assessment, with its focus on the assessment of students’ performance standard in relation to the intended learning outcomes, is more in line with the underpinning philosophy of an outcome-oriented curriculum. It is therefore appropriate for PolyU to change its assessment approach from a norm-referenced one to a criterion-referenced one.

- 3.2 Criterion-referenced assessment has the following added advantages:
- The establishment of detailed and transparent criteria and standards of assessment helps to set clear learning goals for students which, in turn, enhances student motivation and learning.
 - By focusing on students' achievement of learning outcomes, it emphasises what students can do with the knowledge rather than retention and recall of "book knowledge".
 - It encourages the use of real-life, contextualized and authentic assessment methods to evaluate various learning competencies and skills of students, as many of the intended learning outcomes cannot be assessed by the conventional objective tests or unseen three-hour written examination.
 - The process of setting explicit and clear criteria and standards contributes to improved validity and reliability of the assessment.
 - It enables formative feedback be made and given to students with respect to their performance levels with respect to each of the criteria or learning outcomes.
 - It provides useful information to teachers on how well their students have attained the various intended learning outcomes of the programme or subject that they set out to teach. The assessment results can be used to inform curriculum and teaching improvements.

4 How to Implement Criterion-Referenced Assessment Properly and Effectively?

- 4.1 Criterion-referenced assessment is somewhat more difficult to implement than norm-referenced assessment because it requires a vigorous and systematic attempt to derive criteria and performance standards that are appropriate to the intended learning outcomes for the particular group of students. A diagrammatic representation of the steps for a proper and effective implementation of criterion-referenced assessment is shown on the next page and explained further in the subsequent sections. Relevant examples or resources, where available, are also indicated for reference.

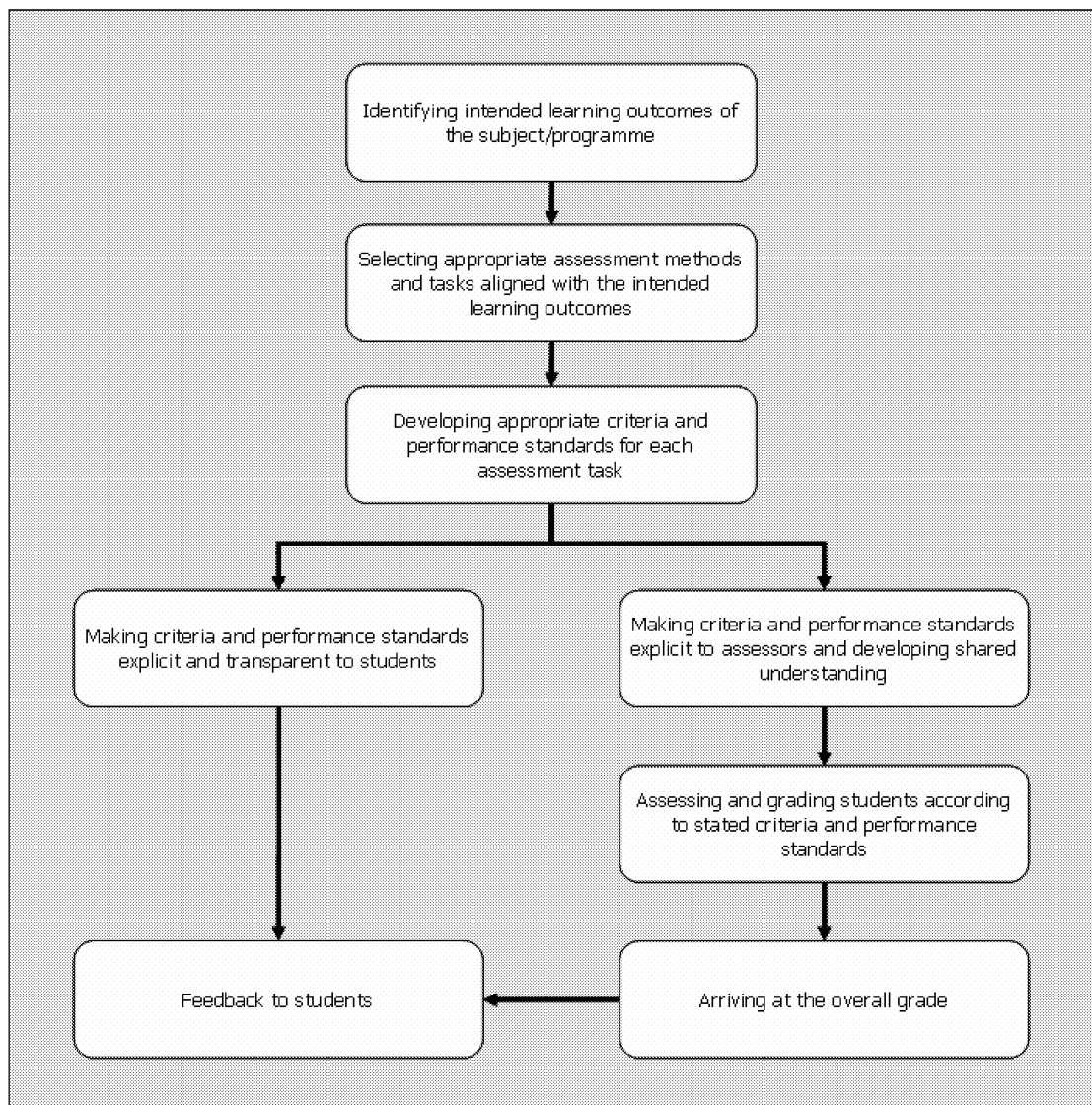
4.2 Identifying the intended learning outcomes of the subject/programme

Learning outcomes refer to the intellectual abilities, knowledge, skills and attitudes that are to be learnt and performed by the students. They specify the kinds of skills and competencies we expect our students to be able to possess on completing the subject/programme and hence, should be the starting point for designing a criterion-referenced assessment. They are different from content topics in the traditional syllabi.

4.3 Selecting assessment methods or tasks that are aligned with the intended learning outcomes

To validly assess students' performance with respect to the intended learning outcomes, it is important and critical that we select appropriate assessment methods and tasks that enable students to demonstrate their learning and truly reflect how well they have achieved the intended learning outcomes.

It should be noted that different learning outcomes may require different assessment methods and tasks, and the conventional objective tests or written examinations are



not appropriate for assessing some of the important learning outcomes. Authentic assessment tasks in real-life contexts may be needed to evaluate students' development of generic and professional skills and competencies. For example:

- The use of practicum or work-integrated learning for assessing knowledge and skills in dealing with professional problems in the workplace environment,
- The use of seminars and presentations for assessing students' presentation and communication skills.
- Using case studies, problem-solving exercises and/or student projects for assessing application and creative problem solving.
- The use of reflective journals for assessing application and reflection.



Resources

1. Section 4.3 of the PolyU Curriculum Revision Resource Book at (URL <https://www2.polyu.edu.hk/cr/files/Chapter4.DOC>) provides a range of assessment methods and discusses their appropriateness for assessing different types of intended learning outcomes
2. Table 9.2 of Biggs, J. (2003). *Teaching for Quality Learning at University* (2nd ed.). Philadelphia: Open University Press.

4.4 Developing appropriate criteria and performance standards

Criterion is “a distinguishing characteristic of anything, by which its quality can be judged or estimated” (Sadler, 2005). Standard, on the other hand, refers to “a statement about the degree of quality to be attained” for a particular performance level (Dunn, Parry and Morgan, 2002).

For every assessment task, the assessment criteria have to be clearly set before any judgment can be made on students’ performance. Often, multiple criteria are used for assessing a single piece of student work, e.g., a term paper or a student project. Furthermore, it has been suggested that a good practice in implementing criterion-referenced assessment is to attach expanded verbal statements to the performance standards or levels for each criterion.

The following example illustrates how to design assessment tasks to align with the intended learning outcomes, and to develop appropriate criteria for each assessment task.

Assuming that a subject has the following two intended learning outcomes, requiring students to be able to:

1. Communicate the subject matter effectively to other stakeholders, and
2. Use evidence appropriately in support of an argument.

One appropriate task for assessing students’ attainment of the intended learning outcomes of the subject will be to require students to give a presentation on a topic requiring the use of evidence to support their arguments.

In this case, multiple criteria are needed to assess students’ performance in the assessment task. These may include:

- Organisation of the presentation
- Relevance of evidence used
- Clarity of presentation
- etc.

For each of the criteria, it has been suggested that detailed descriptions can then be provided on the performance standards or levels against which students’ work will be judged. For example, if the criterion is “Organisation”, the levels might include:

- “Organized around a single topic with good transitions”,
- “Connections between ideas are generally good, and writing is about one topic”,
- “Loosely organized, with no central idea”, and
- “Disorganized”.

On the other hand, Gosling and Moon (2001) suggested that if the criterion is “Relevance”, the levels might include:

- “Analytical and clear conclusions well grounded in theory and literature, showing development of new concepts and originality of ideas”,
- “Limited evidence of conclusions and findings as supported by theory and practice”.

Dunn, L., Parry, S., and Morgan, C. (2002). Seeking quality in criterion referenced assessment. Paper presented to the Learning Communities and Assessment Cultures Conference organized by the EARLI Special Interest Group on Assessment and Evaluation, University of Northumbria, 28-30 August.

Gosling, D. and Moon, J. (2001). *How to use learning outcomes and assessment criteria*. London: SEEC.

Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, 30, 175-194. Retrieved December 22, 2004, from Taylor and Francis Online Journal Website:

<http://journalsonline.tandf.co.uk/app/home/main.asp?wasp=n0fdjqeqth5jykbbnq96>

The most difficult part of implementing criterion-referenced assessment is to determine and define the appropriate performance standard against which students' work will be judged with regard to the established criteria, because of the removal of the "comparison group". This is not an easy task, and involves a lot more than awarding grades to students on the basis of some arbitrarily-defined mark ranges. Vigorous attempts have to be made to determine what constitute an "outstanding", "good", "satisfactory", "minimally acceptable" and "unacceptable" level of performance.

Extreme care should be taken when deciding on the criteria and standard of the assessment to avoid the following potential pitfalls:

- "Grade inflation" as a result of setting too low a standard such that students' performance is over-rated,
- "Failing too many students" as a result of setting too high a standard that very few students can reasonably achieve.

Unlike norm-referenced assessment, there is no simple formulaic method of determining the performance standards of students under criterion-referenced assessment. Much informed professional judgment is required in the process to ensure that the standards are appropriate for the target student group and at par with international standard. Inputs and feedback from peers (particularly the external examiners) and/or employers are essential to this endeavour.



Resources

Examples of descriptors can be retrieved via the Assessment Resource of the Hong Kong Polytechnic University at URL
<https://www2.polyu.edu.hk/edc/Staff/assessment/workshop3.pdf>

4.5 Making criteria and performance standards explicit and transparent to students

The design of assessment has strong impact on student learning. It is therefore important that the intended learning outcomes of the subject/programme and how they are to be assessed are clearly and *explicitly* communicated to the students at the beginning of the teaching and learning process, before the actual assessment is conducted. This helps students to develop a clear understanding of the requirements and expectations of the subject/programme, and will have a strong positive effect on students' motivation and learning.

One way to help students "recognise" quality is for teachers to collect sample pieces of work that cover various performance levels of the criteria under consideration and provide opportunities for students to assess and discuss the quality of the sample work according to this performance continuum.



Resources

Sadler, D. R. (2002) Ah!.. so that's quality, in: Schwartz, P. & Webb, G. (Eds) *Assessment: Case studies, experience and practice from higher education*. (London, Kogan Page), 130-136

4.6 **Making assessment criteria and standard explicit to assessors and developing shared understanding**

If more than one assessor is involved in the assessment of students' work, it is critical that *prior to* the assessment exercise, the assessment panel should have agreed on the criteria descriptor for each level of performance and developed shared understanding of the grading scheme. Subject teams should devise and implement parity measures to ensure consistency and fairness of the grading such as internal moderation and/or double marking.

It is also suggested that faculty can assure fairness and validity of grading by periodic review of assessment practices and results. Faculty who teach the same course or use the same criteria should review the accuracy of their assessments by assessing sample student work and comparing their assessments at least annually. If there are disagreements in these assessments, faculty should discuss those differences to clarify the basis for their judgments or to modify criteria or rubrics. Clear, common understanding of the meaning and application of criteria will assure fairness. Individual faculty can validate their own assessments by comparing the assessments of one semester against those of the previous semesters to determine if the criteria are being applied uniformly. By examining over time and across sections of the same course the judgments faculty apply to criteria, faculty can guarantee "fairness" by eliminating significant divergences from common practice as well as criteria that are too unclear to be applied uniformly.

4.7 **Assessing and grading students according to stated criteria and standards**

Students' work should be assessed according to the *pre-determined* criteria and standards. However, teachers or subject teams may need to fine-tune the criteria and standards if they find them to be unreasonably high or low during or after the teaching and assessment process.

4.8 **Arriving at the overall grade**

Where the assessment of a subject involves multiple criteria or assessment tasks, a scheme for combining the grades of the various components into a single overall grade has to be determined.

A common approach to arriving at the overall grade is by a *weighted average* of all the components of the assignment. For example, different weightings can be assigned to different assignment tasks of a subject to reflect their respective importance.

An alternative approach is to report the overall grade based on a *holistic* judgment of the performance of the student in all the assessment tasks against each of the learning outcomes laid down for the subject or programme rather than a simple arithmetic mean. For example, students may be required to obtain at least a pass grade in all components before he/she can be awarded a pass grade in the overall subject grade, to ascertain that all the important learning outcomes of the subject has been achieved.

The table on the next page illustrates how the overall grade will be awarded according to the achievement of intended learning outcomes.

<i>Subject grade</i>	<i>Short description</i>	<i>Revised elaboration on subject grading description</i>
A+	Excellent	The student's work is outstanding. It demonstrates the attainment of all intended learning outcomes and far exceeds the threshold standard required by the subject in all regards.
A		The student's work is excellent. It demonstrates the attainment of all intended learning outcomes and far exceeds the threshold standard required by the subject in nearly all regards.
B+	Good	The student's work is very good. It demonstrates the attainment of all intended learning outcomes and is well above the threshold standard required by the subject in the majority of regards.
B		The student's work is good. It demonstrates the attainment of all intended learning outcomes and is well above the threshold standard required by the subject in some regards.
C+	Satisfactory	The student's work is wholly satisfactory. It meets the threshold standard required by the subject in all of the intended learning outcomes.
C		The student's work is satisfactory. It largely meets the threshold standard required by the subject in essentially all of the intended learning outcomes.
D+	Marginal	The student's work is barely adequate. It fails marginally to meet the threshold standard required by the subject in a few key intended learning outcomes.
D		The student's work is weak. It fails marginally to meet the threshold standard required by the subject in several of the key intended learning outcomes.
F	Failure	The student's work is inadequate. It fails to meet the threshold standard required by the subject in many of the intended learning outcomes OR fails badly to meet some crucial intended learning outcomes (i.e., ones that students cannot fail in order to obtain an overall pass grade.)

4.9 Feedback to students

To help students learn from the assessment process and improve their learning, it is important that in addition to the grade, students should be given detailed feedback on their performance. The following provides some useful guidelines for giving effective feedback.

- Feedback should be provided quickly enough to be useful to students and should be given both often enough and in enough detail.
- Feedback should focus on learning, be understandable for students and linked to the purpose of the tasks and the criteria.
- Student's response to feedback should be taken into consideration.
- Teachers could encourage or require student to engage in peer review and feedback as a form of formative assessment for learning.