A Primer on Performance standards, Cut scores and Weights

(Dr Zita Lysaght, IoE, DCU)

Performance standards

Rubrics are often accompanied by scoring guides that specify the distribution of marks for different elements of the assessment under review. When designing scoring guides and rubrics to support criterion-referenced assessment, the first step is to decide on the number of performance standards to be used (e.g., *First Class Honours, Second Class Honours, Grade I, Second Class Honours, Grade II etc.)* and then write descriptions for each level. In some cases, particularly when the aim of the assessment is to make a dichotomous judgement about students' mastery of specific content, the range of levels may be reduced to two options (e.g., *pass/fail* or *satisfactory/unsatisfactory*), without compromising on standards necessarily. For instance, the pass or satisfactory level may be set very high at 80%, 90% or even 100% (as is the case in some medical exams and driving tests where high-level competence is essential), or relatively low, perhaps at 40%, where the stakes are lower and basic competence is acceptable.

The statement in Section 6.2.8 of DCU's *Marks and Standards* document (https://www.dcu.ie/sites/default/files/ovpaa/marks_and_standards_2019.1.1.pdf) that "Student performance in a module can be marked on a scale 0 - 100 or allocated a Pass / Fail grade as described in the approved module descriptor" underscores that decisions of this kind regarding performance standards rest largely on individual assessor's/lecturer's professional judgement.

Cut scores

One of the most difficult decisions in using performance standards arises when assessing 'border-line' student work. This is because, while performance level descriptors indicate what students *should know* and *be able to do*, cut scores determine the number of score points students must attain in order to demonstrate that their work is of a particular standard, i.e., cut scores are used to distinguish between performance levels.

The following points, based on research in this area, may be of use when trying to make decisions about cut scores:

1. All judgements about cut scores contain a degree of subjectivity reflecting an assessor's judgement about different levels of competence in a specific domain or module
2. By definition, this means that there is no objectively perfect or definitive cut score
3. The validity of inferences made about students' learning, based on performance standards and cut scores, is directly related to the competence and expertise of those who create them

4. Decisions about cut scores should involve a team of people with discipline knowledge and experience of judging students' performance at different levels, particularly in the case of high-stakes assessment
5. There are many approaches that can be taken to determine performance levels and/or to set cut scores. All should be iteratively reviewed to ensure fitness for purpose and, where possible, multiple-methods of assessment employed to provide complementary and comparable evidence of student learning.

Weights

A key question to ask when devising an assessment is: *will this assessment, or combination of assessments, give me the information I need to judge my students' performance in the relevant area/domain?* As lecturers, we have a range of assessment options from which to choose. For example, one might elect to use a performance assessment rather than an exam because it would allow students to demonstrate their learning more effectively. Or a combination of assessments might be required to ensure that the range of knowledge, skills and concepts taught in a module is reflected (construct validity). In this case, an essay (to assess students' knowledge of a topic), followed by a performance assessment such as a debate (to assess students' skills to apply that knowledge in real life), might be used. When there are multiple assessments within modules (as is customary in composite modules, for example), the issue of weighting arises.

Weighting refers to the allocation of marks to components of an assessment and, in general, one of two approaches is used: the percentage correct method or the total points method. (An 'eye-balling' approach can also be used, which involves making an informed guess based on perusal of assessments and results, but this has clear limitations).

Example 1: Combining assessment results using the percentage correct method

Let's assume that a student has completed three assessments for a module. The weighting is 30%, 30% and 40%, respectively and the student has been awarded the following marks:

Assessment 1 = 85%

Assessment 2 = 55%

Assessment 3 = 90%.

To obtain the weighted percentage correct score, each score is multiplied by the relevant weight (expressed as a decimal):

Assessment 1 = 85% x .3 = 25.5%

Assessment 2 = 55% x .3 = 16.5%

Assessment 3 = 90% x .4 = 36%

The weighted scores are then totalled to calculate the overall composite score.

Composite score: 25.5% + 16.5% + 36% = 78%.

These calculations are straightforward because each of the component assessments uses a common percentage correct scale. Where different scales are used that give different types of scores and grades, these must be converted to the same scale and weighted before calculating the overall mark.

Percentage correct calculators can be helpful (see https://percentagecalculator.net/) .

Example 2: Combining assessment results using the total points method

The total points method is based on allocating points to each contributing assignment in a manner that reflects its weighting. Using the previous example, if, taken together, the total number of points available for the module assessment was 500, then, based on the weightings given, the following points would be have been allocated to the student for each component assessment:

Assessment 1 = 150 points

Assessment 2 = 150 points

Assessment 3 = 200 points.

In this case, if the student achieved the percentage correct results listed in Example 1, s/he would be awarded the following point ratings and composite score:

Assessment 1 = 127.5/150 points

Assessment 2 = 82.5/150 points

Assessment 3 = 189/200 points

Composite weighted score: 127.5 + 82.5 + 180 = 390/500 = 78%.

Related documents of potential interest include:

A Primer on Differences between Norm-reference based and Criterion-referenced Assessment

A Primer on Criterion-Referenced Assessments and Rubrics

A Primer on Norm-reference based Assessment and Grading on the Curve

An Example of a Weighted Rubric

A PowerPoint on Rubrics.