Abstract

This study set out to ascertain whether percentage mark distributions show evidence of discipline-related marking behaviour and, if so, to consider the implications of this for equity in assessment outcomes. Performance data were obtained for 10 subjects at seven English universities for 1993-94, 1994-95 and 1995-96. The data showed that mark distributions at universities using percentages may be categorised into three Types (A, B and C). 

The types of distributions appear to reflect the relative confidence of assessors in different disciplines in awarding marks. The extremities of the percentage scale are perceived as insecure territory for the assessors of qualitative subject matter. Type A distributions are characteristic of disciplines in which the assessment is qualitative. Type B distributions may reflect the assessment of combined qualitative-quantitative subject matter. Type C distributions are typical of disciplines in which more quantitative matter is assessed. The existence of discipline-related patterns of distribution is a potential cause of inequity of outcomes for students on joint or multidisciplinary programmes. Data derived from institutions using grading scales show that discipline-related patterns are less evident (p. 285).

Introduction

Empirical evidence is presented which indicates differences in subject marking conventions with implications for the outcomes of assessment. Differences include whether full or restricted ranges of percentages are used and the relative ease of gaining high grades in particular subjects.

The study examines the extent to which discipline-related marking behaviour is evident in institutions using traditional percentages and considers the implications for equity and output standards. Marking systems other than those using percentages are also briefly reviewed in relation to the relative equity they provide.
Sources of data
Data were obtained from the home universities of the authors in subjects selected to represent a wide range of disciplines with possible differences in marking behaviour - Biology, Business Studies, Computer Studies, English, Fine Art, French, History, Law, Mathematics, Sociology. Data are presented as bar charts with statistical measures.

Types of distribution
Data derived from two universities were selected for comparison and three types of mark distributions are identified and described (p. 285).

- **Type A distributions (English and History)** are characterised by a steep-sided negatively skewed distribution with a narrow spread. Pronounced microscale variation in the frequency is evident only in the vicinity of the mode between 55-65%.
- **Type B distributions (Biology, Business Studies, Fine Art, French, Law and Sociology)** have a slightly broader spread and display microscale variation in frequency across a wider range of percentiles [40-70%].
- **Type C distributions (Computer Studies and Mathematics)** are characterised by a subdued, slightly negatively skewed distribution with a considerably wider spread. Type C distributions show conspicuous microscale variation in frequency across most of the percentile range.

Further detail is provided in relation to each type of distribution presented.

Interpretation of the patterns
The results confirm previous research that there are disciplinary differences in marking distributions when using a percentage scale. This is attributed to how assessors recognise extremes of performance. For example in Type A subjects such as English, assessors making qualitative judgements tend to avoid the top of the scale with its implications of perfection which they may not be able to defend with confidence. Outstanding performance may therefore be indicated by a mark in the high 70%. As marks are infrequently depressed below 25% the entire spread of marks or percentages is low.

By contrast, in Type C subjects, such as mathematics, the assessor is generally more concerned with the correctness of the students’ work. This offers more secure ground for awarding marks at the top (or bottom) end of the percentage scale. Consequently the spread is wider. However, in Type B subjects, such as Business Studies, the mix of qualitative and quantitative types of assessment judgements results in a mark distribution which is intermediate between Types A and C. Marking distributions are also influenced by traditional grade cut-offs, for example in the UK system, marks are concentrated in the 40% - 70% range as these are the boundaries between a “Pass” and First Class Honours. This can result in circular reasoning if the grade or honours level is determined before (rather than after) the marks are assigned and finer distinctions are made.

The Problem of Unwarranted Precision
A key problem when using percentage scales is the unwarranted precision required for decisions to award 52%, 53% or perhaps 54% with confidence and consistency. The problem is intensified for decisions at the boundaries or cut-offs for particular grades or classifications.
and anecdotal evidence suggests that assessors are encouraged to avoid marks within one or two percentage points of grade boundaries.

**Potential Inequities in Using Percentage Scales**

The use of percentage scales can disadvantage students. This is illustrated through hypothetical comparisons of students undertaking double degrees within disciplines with comparable marking practices (e.g. English and History) with those undertaking degrees across disciplines with different practices (e.g. History and Computer Studies) in which the latter implies a potential disadvantage. The same problem may occur within single subjects in which students may combine courses or subjects with assessment requiring qualitative and quantitative judgements, for example in Business Studies. This has particular significance for students whose overall performance is close to a classificatory borderline (or GPA).

**Marks Distributions at Universities Using Grading Systems**

As some universities used grading scales to overcome the problem of unwarranted precision identified with percentage scales, a comparison of the two approaches was made using distributions for English (Type A), Sociology (Type B) and Computer Studies (Type C). The use of grading scales significantly reduced the differences between the distributions for the different types of disciplines. In particular the use of grading scales expanded the range of the scale used to represent student achievement. The use of grading scales was not without problems however with some aberrations in need of further investigation.

**Conclusions**

The study concludes that there is indeed a discipline effect on marking when percentage scales are used and that this is potentially disadvantageous for students taking particular subjects or subject combinations. It is also suggested that the adoption of grading scales reduces the impact of disciplinary differences. Quality descriptors for individual grades are recommended as minimum requirements for achieving grading consistency.

**References**

The paper provides nine references.

*This series of briefs on assessment topics has been prepared by the UQ Teaching and Educational Development Institute (TEDI) for UQ teaching academics. “UQ ASSESSMENT BRIEFS” of journal articles, book chapters, reviews, websites, reports etc are distributed to Faculty and School Teaching and Learning Chairs in a form designed to encourage wider distribution.*

Please send feedback, requests or suggestions for future “UQ ASSESSMENT BRIEFS” to:

Dr Clair Hughes (clair.hughes@uq.edu.au)
Senior Lecturer in Higher Education
The Teaching and Educational Development Institute (TEDI)
The University of Queensland, St Lucia Qld 4072
Telephone +61 7 336 52456
Facsimile +61 7 3365 1966