First-Class Marks

David Young, February 1999

Background

I brought a very brief paper on the subject of the use of marks of 70% and over to COGSAS in November 1997 ("A-Grades and Percentage Marks"). That was intended as a basis for discussion; COGSAS decided that the issue was one which needed to be addressed University-wide, and reported this in its minutes. However, higher committees did not take the matter up, and no advice was received. Subsequent meetings of COGSAS and the COGS Graduate Steering Committee decided that we should therefore make policy within the School.

My first paper was partly motivated by difficulties in the transition from letter grades to percentage marks, then underway. Letter grades have now almost vanished, but the problem of how to mark excellent work has not. This paper present a fuller view of the difficulty; it is rather long because I believe the issues are difficult and stem from the lack of a public rationale for our marking scale. As there is a general push towards using the upper part of the marking scale, it is important to look at the consequences of awarding very high marks.

This paper considers the undergraduate classification rules, but similar problems exist for postgraduate taught courses.

Why first-class marks are a problem

First-class marks are allocated 30% of the marking scale. Other classes of degree correspond to bands of only 10%. This means that first-class marks, when averaged into a mean, can have a dramatic effect on final classification. For a given assessment unit, the difference between getting a bare first and a very good first is the same as the difference between a poor third and a good upper-second. For our best students, it matters a great deal whether examiners give 75%, 85% or 95% to work that really impresses them.

Some examiners are clearly uncertain about how to use the wide top range. The only guidance in the BA/LLB Handbook is that "It is essential that the full marking scale should be used". Yes, but how and when? There is nothing in the BEng/BSc Handbook. Programme and course-specific marking criteria are published, but these often do not discriminate between different parts of the first-class range. The CSAI Handbook does include an attempt: "Tutors should not be reluctant to award marks in the 80s or even 90s in the case of really excellent work, although grades in the 90s should be reserved for work deemed to be outstanding". So marks in the 70s and 80s are for work that is excellent but not outstanding — a nice distinction.

Different kinds of work imply different styles of marking. Examiners are traditionally very reluctant to give essays marks over 80%, but tests in areas such as programming can have a "right answer" and can generate an automatic mark of 100%.

Different areas use marks differently. BA students receive a first-class degree if their overall mean is at least 67% and they have first-class marks for a certain proportion of their work. BSc students in general need an average of 70% or more, without regard for the distribution of marks.

These inconsistencies affect COGS in particular. The BA programme in AI shares courses with various BSc programmes. A mark of 95% for an AI major means a different thing from the same mark for a CSAI major, and tutors cannot reasonably be asked to distinguish the different groups of students when marking. However, the issues are by no means confined to COGS, since all examiners need at some time to decide what marks in the 80s and 90s are supposed to mean.

Why do we have the wide first-class band?

Presumably, the committees that determined the canonical marks boundaries had reasons for their choices. However, whilst the decisions have been passed down to examiners and tutors, the rationale has not, as far as I know. One can only attempt to reconstruct how this division was supposed to work.

The central idea appears to be one of rewarding excellence, even if that excellence is patchy: a student who does something outstandingly good, even at the expense of other work, should have it recognised. Thus (to take an extreme) someone who consistently gets 69% receives a 2i, but someone who scrapes by with 45% in half their units but gets 95% in the other half receives a first. In general, a mark in the 80s or 90s cancels out quite weak performance in other units regardless of the student's final mean. Of course excellent students often perform well all round — but there are always many borderline candidates for whom these biases in the marking scale can be highly significant.

There are at least two problems with this principle. The first is that the implications for different subjects have not been debated, as far as I know. Perhaps in some subjects it is right that a demonstration of excellence in a few topics — student-chosen topics at that — is what matters. In other subjects, however, one might feel that mediocre performance in certain areas should count against the student however well he or she performs elsewhere.

The second problem is that the Arts area has a further mechanism for rewarding sporadic excellence. Roughly, if a student gets first-class marks on one third or more of his or her units, then the overall threshold for a first drops from 70% to 67%. (This rule is frequently invoked.) A good first-class mark can therefore doubly help a BA candidate who does less well in other areas: it has a disproportionate effect on the grand mean because of the wide first-class mark band, and it also may help to push down the threshold the mean has to reach. This is the converse of double jeopardy.

An alternative rationale for the wide band might be that there is somehow a three times wider range of performance between a bare first and the best possible work, than between a bare 2i and a bare first, say. This suggests that the mark of 100% is to correspond to some ideal, and the 30% margin is necessary to reflect the distance by which excellent work may still fall short of perfection. Whether this makes sense again depends on the subject and on the nature of a particular exercise, but in any case this view of the scale does little to help the individual examiner who has to decide whether 95% is or is not attainable by a real-life student who has written a computer program or a report on a lab class.

Possible strategies for assigning first-class marks

Marking on a curve

Adjusting marks to fit a distribution (i.e. the top X% of students in a given assignment always get marks of Y or more, where the relationship between X and Y can be tabulated and is fixed) would cut through the problem. This would represent a radical departure from the position that we sustain absolute standards in assessment.

At the moment, formally marking on a curve is not really an option, because it is too much of a departure, and because there are real differences between cohorts of students which would not be visible. Nonetheless, it is the case that our standards are influenced by our experience of students: we do and must take into account what we can reasonably expect. The BSc examination boards often have mark distributions in front of them when classifying, and BA boards have requested such distributions.

Guidance about first-class marks given in distribution form would at least be readily understandable. It might say, for example: “Only one student in 100 is expected to be capable of marks over 90% on any given assignment.” The general objection to marking on a curve — that it negates any attempt to sustain absolute standards — has less force when only a small number of bright students are being considered.

I suspect that many examiners at present use some informal notion of “exceptional work” when giving very high marks. This implies some sense of frequency of occurrence as a criterion, but this is not explicitly quantified.

Working back from classification

One way to finesse the problem of assigning marks is to assume that examiners all know what degree classes mean. Then most marks have easy interpretations: for example 65% means work of 2i standard. Very high marks have a more complex meaning, but their consequences for classification can still be used to derive it. A mark of 75% in this interpretation corresponds to work that by itself would merit a first class degree, but marks of 85% and 95% mean that the work is so good that it compensates for much weaker work in other units.

One problem is that the degree of compensation varies between BSc and BA degrees, with particularly complex effects under the BA rules. Another is that it may be difficult to formulate simple and generally agreed guidance in these terms, particularly as the effects of very high marks on classification are probably accidental rather than considered.

In general, examiners who are not on examination boards are not familiar with the classification rules for the BA. In general, the implications for classification are probably not considered explicitly when awarding very high marks.

Verbal descriptions of the qualities required

Descriptions of the qualities required for the 10% bands above 70% could be given as they are for the lower bands. The problem with this is that “outstanding”, “exceptional”, “very original” and so on will mean different things to different examiners, and both formulating and interpreting such descriptions would be even more difficult than it is for the existing bands.

Subject-specific conventions

Within a subject, there are often conventions which may be applicable. For example, in many subjects essays never receive 100%, and high marks are used very sparingly. In others marks for tests may simply reflect the proportion of the course material that the student has learnt, and to get 100% just needs extra effort. These conventions do not explicitly take account of the implications for classification.

Conclusions

The use of marks above 80% causes difficulties. These are largely due to the "pulling up" effect of high marks: the degree to which such marks can compensate for weak work in other areas may not be justified. The general case for strongly rewarding sporadic excellence has not been convincingly made.

There is no obvious way in which to tell examiners when to award very high marks. Strategies used in practice may lead to unforseen consequences at classification.

Differences between the BA and BSc classification rules make it particularly difficult to achieve consistent standards in COGS.

Recommendations

If marks in the 80-100% range are to be used with any significant frequency, then:

1. The classification rules for BA degrees should be reexamined. COGS's external examiner has already commented adversely on the rules that can lower the class thresholds by 3%. The double advantage conferred by a few very high marks needs to be looked at carefully.

2. The meanings of the 70-80%, 80-90% and 90-100% bands need to be defined University-wide in the same way as are the bands from 30% to 70%. A description in terms of frequency would be the easiest to understand and apply.

3. It is impossible at present for COGS to produce a policy on high marks that will be fair to all students. Either AI majors will be at a disadvantage compared to other BA students, or they will be at an advantage compared to CS and CSAI majors taking some of the same courses. Given this, I propose that within COGS:

   a) The trend to encourage the use of marks over 80% should be resisted. Whilst high marks on individual assignments may sometimes be appropriate, the distribution of average marks for a given course should have very small numbers in the 80s and 90s.

   b) Examiners should be made aware of the consequences of high marks at classification.