Supplementary advice regarding scaling of marks

This document should be read in conjunction with the Examination and Assessment Framework which details the University's policy relating to scaling of marks.

Definition of scaling

Scaling is the systematic adjustment of marks for an entire cohort, carried out on a piece of assessed work in order to ensure the final marks better reflect the achievement of the students as defined by the University's standard scale for the expression of agreed final marks i.e. the qualitative marks descriptors.

Scaling is not a mechanistic process but one which requires examiners to use their academic judgement to determine (i) whether scaling is appropriate and (ii) if the outcome of the scaling process has resulted in marks that fairly reflect student achievement.

Determining whether scaling is appropriate

Boards aim to set questions and mark schemes which will generate a spread of marks that fairly reflects the student cohort's performance compared with the University's standard scale for the expression of agreed final marks (i.e. the qualitative marks descriptors) as set out in the course examination conventions. However, it is recognised that despite the very best efforts at the examination setting stage, an examination may not generate the expected spread of raw marks. Scaling, with academic judgment via qualitative checks, may then be needed to translate raw marks to marks that are a fair reflection of the performance of candidates on the University scale.

Agreed scenarios when scaling can be applied

Scaling of marks on a paper is only appropriate when the examiners can supply evidence for at least one of the following scenarios:

(a) A paper was more/less difficult than in previous years

(b) An optional paper was more/less difficult than other optional papers taken by students in a particular year

(c) A paper has generated a spread of marks which is not a fair reflection of student performance on the University's standard scale for the expression of agreed final marks, i.e. the marks do not reflect the qualitative marks descriptors

Evidence requirements

Examiners will need to demonstrate there is more than one source of evidence that one or more scenarios has been exemplified by the paper in question. Examples of the types of evidence which could be considered include:

- A paper has a higher/lower median/mean mark relative to previous years;
- The examiners' academic evaluation of the performance of the candidates (possibly guided by qualitative descriptors of each class);
- A comparison with the questions set in previous years' papers;
- An analysis of the spread of candidates' performance in compulsory papers compared to their performance in the paper in question;
- A paper does not produce the expected spread of raw marks for the cohort.

Scaling practice

What scaling should be:

- Transparent – the scaling methods (for example an algorithm) should be made publicly available to examiners and students.
- Justifiable – the rules of any algorithm used must be fully consistent with the examination conventions, the reason(s) for scaling should be included in the examiner's report and the external examiners should report on the application of the scaling.
- Fair – the scaling should be applied to all candidates, not just to problematic subsets, for example failures or high passes.
- Reasonable – the scaling should not be applied to candidates if this takes their mark below 0 or over the total marks available for the script.

Scaling should not be:

- Used mechanistically to fit the spread of classes on a paper to historical norms (i.e. norm referencing).
- Applied if the number of students being considered is small, unless the examiners are confident that the statistical analysis is a reliable tool.
- Considered on only one piece of evidence; for example, the differences in mean or median scores of students taking different optional papers could simply be the result of natural variation in ability within the cohort of students.

Examples of scaling algorithm methods

- Simple addition – a fixed number of marks is added to either (i) a particular assessment component of a script, or (ii) the final script mark, as long as no scaled marks are then greater than the total marks available for the script.
- Simple subtraction – a fixed number of marks are subtracted to either (i) a particular assessment component of a script, or (ii) the final script mark, as long as no scaled marks are less than 0.
- Multiplication by a factor – all marks on a script are multiplied by a particular factor (for example 0.96).

Examiner process for applying scaling

Examiners who intend to scale should:

1. Document that there is sufficient evidence to support that the paper to be scaled meets one or more of the scenarios (this will need to be included in the examiner's report).
2. Ensure the paper to be scaled (i) was taken in the current academic year (scaling cannot be carried out in later academic years) (ii) has been moderated
3. Consider the scaling method to be applied and ensure (i) the rules of any algorithms are consistent with the examination conventions (iii) advice has been sought if there are any concerns over the construction and/or operation of the algorithm.
4. Apply the scaling
5. Review the outcome of the scaling process by (i) considering a sample of papers either side of the classification borderlines to ensure that the outcome of scaling is

consistent with academic views of what constitutes a paper in each class (ii) asking the external examiner(s) to consider the application of the scaling and report on this stage of the process.

6. Agree if outcome of scaling should be applied.
7. Document in the examiner's report (i) why scaling was necessary (ii) how scaling was applied. This same information should also be supplied to students.

Points to bear in mind

1. Examiners may discover very early on in the marking of papers that the paper was too hard. Therefore, the Board of Examiners can decide to adjust the mark scheme prior to continuing with the marking instead of applying a wholesale rescaling at the end.
2. If the scaling alters the marks by more than half a class band, the Board of Examiners should consult the External Examiner to review the process.
3. Examiners should consider that addition/subtraction affects only the mean/median and not the spread of marks, whereas multiplication affects both the mean and the spread. The approach chosen by the Board should be guided by the issue they are trying to correct, this could involve one or the other or a combination of two.