

Mark Scaling: A Brief Overview and Synthesis of Sector Perspectives

Professor John G. Sharp

Head of the Lincoln Higher Education Research Institute

Report

This report focuses on mark scaling as a practice within some UK universities. Understanding mark scaling as a component part of assessment and quality assurance processes, including its range of application and limitations, is considered essential ahead of adoption.

In essence, mark scaling, or the systematic and arithmetic adjustment of academic outcomes after formal marking and moderation procedures have concluded, is the process by which the grades of an entire group or cohort of undergraduates or postgraduate students are moved up or down in order to better reflect performance and achievement against learning outcomes or assessment criteria and to maintain current standards or the trend of standards across years. Courting controversy for obvious reasons (e.g. rationale, imprecision and subjectivity), and certainly less common now than it once was (as far as can be determined), mark scaling remains an accepted and regulated practice at certain Russell Group and now disbanded 1994 Group members at the very least. In the absence of any immediately obvious and traceable research literature on the subject, and from a simple Google search of key words alone, these include Edinburgh, Newcastle, Southampton, Durham, Bath, Newcastle, Exeter and Cardiff. Each provides online access to scaling policies which exhibit both common ground and individual variation.

The basis for mark scaling and its application

The decision to scale marks is acknowledged as a matter of professional judgement. This would be undertaken most frequently when the measured performance of an individual module component, an individual module, a number of modules on a single programme or an entire cohort of undergraduates or postgraduates differs markedly and outside of the anticipated range either from within the current year or in comparison with previous years (e.g. anomalous or unexpected outcomes, atypical mean values, distributions or trends and unusually high pass or fail rates). Other things to be considered include the student profile or demographic which should remain constant in order to be ruled out as a causal factor. Mark scaling is therefore undertaken only in exceptional circumstances rather than routinely and repeatedly, and only used when all other options have been exhausted or unlikely to resolve the issue at hand (e.g. the remarking or further moderation of work).

Among the many common reasons provided for its initiation, though by no means a definitive or necessarily causal list, mark scaling may be deemed necessary by those in authority (e.g. module leaders, programme leaders or Heads of School/Department) in the following situations:

- where students were taught by different lecturers or inexperienced or visiting staff
- where assessment components were marked by different individuals who misinterpreted guidance or awarded marks somewhat spuriously
- where assessment irregularities, including invigilation incidents, were known to have occurred
- where the amount of time required to complete work was deemed insufficient
- where the content of what was taught was at odds with the assessment type or marking criteria.
- where the form of assessment was considered too easy, too hard or contained ambiguous or misleading materials (e.g. in wording, graphs, figures or tables)
- where an identified group of students was considered particularly disadvantaged

In all instances where scaling flags a teaching, assessment or procedural problem or issue at module, programme or disciplinary level, remediation as a normal part of quality assurance procedures and arrangements should be attended to ahead of subsequent assessment and marking phases and certainly before the next academic session as appropriate.

Interestingly, guidance on the scaling of optional modules, modules involving professional or work-based placements, modules in, for example, the first year of study which may not contribute to degree awards and modules exhibiting other individual characteristics remains far from clear.

Scaling practices

From the various institutional documents available in the public domain, scaling would seem to be applied most frequently with the involvement and approval of external examiners after the completion of marking and moderation, and when all other associated quality assurance procedures have been satisfied as indicated, but before marks are ratified at Examination Boards after which time they are fixed. Ideally, all students affected should be informed at the earliest opportunity and ahead of summative announcements. All discussions and decisions associated with mark scaling, including the reasons for its adoption and the statistical basis and mechanism upon which any changes were made, should be entirely transparent and reported by those responsible also at Examination Boards, where all matters should be formally recorded and documented in minutes.

In terms of mark scaling itself, a variety of linear and non-linear strategies and algorithms have been considered and applied either manually or employing an approved software programme to reduce human error. These include:

- the equal and uniform addition (or subtraction) of a single mark to raise (or lower) averages
- multiplication by a simple and common factor to raise (or lower) averages
- multiplication by a complex or differential factor applied to particular areas of a mark profile in the case of spreading marks
- applying a bespoke adjustment to generate an acceptable comparison with the marks from other modules or previous years

In all instances, the range of adjusted marks must never exceed 0-100% in the first instance. Following scaling, due care and attention should also be directed towards changes in pass/fail ratios, earlier fail grades becoming pass grades and where the cumulative marks of individuals result in the transgression of degree categories. The original rank order of individuals should always remain the same. Further adjustments for impaired student performance or mitigation should be considered after scaling and scaling should never be used where a zero mark was awarded for non-submission or some other academic violation.

It is also worth noting that scaling may be difficult, if not impossible, to achieve with any degree of accuracy and confidence where the number of students involved is small (e.g. fewer than 10-15 as a rule of thumb for statistical purposes). Scaling is also considered a criterion-referenced adjustment and should not be used with norm-referencing in mind (e.g. attempting to create a normal distribution with the express purpose of defining a specific number of degree classes).

Conclusion and recommendations

Without a detailed evaluation of all marking and assessment policies across the sector, evidence suggests that the scaling of student marks following assessment and moderation is an accepted and regulated practice at a small number of prominent UK universities at the very least. As far as can be determined, the processes and procedures associated with mark scaling appear absent from within the research literature as might be expected. While it would be entirely within its own power to do so, any UK university could, if it so desired, introduce mark scaling into its own assessment regulations and quality procedures. Such an introduction would, however, require careful consideration of intended purpose and how the practice might be monitored over time.