10 Steps for Creating Marking Rubrics

Defining Rubrics.

The term 'rubric' has taken on a number of different meanings for educators. Dawson (2017)[1] provides an overview of these interpretations and offers a detailed description of 14 design elements that can differ between separate rubrics. Despite these variations, Dawson (2017) offers the following broad definition:

"A rubric is a tool used in the process of assessing student work that usually includes Popham's (1997) three essential features: evaluative criteria, quality definitions for those criteria at particular levels and a scoring strategy."[2]

The figure below depicts the structural features described in this definition. Point 6 below covers the concept of a scoring strategy.

	e.g. Fail	e.g. 3rd class	e.g. 2.2	e.g. 2.1	e.g. 1st class
Evaluative criterion #1	Quality definition	Quality definition	Quality definition	Quality definition	Quality definition
Evaluative criterion #2	Quality definition	Quality definition	Quality definition	Quality definition	Quality definition
Evaluative criterion #3	Quality definition	Quality definition	Quality definition	Quality definition	Quality definition
Evaluative criterion #4	Quality definition	Quality definition	Quality definition	Quality definition	Quality definition
Evaluative criterion #5	Quality definition	Quality definition	Quality definition	Quality definition	Quality definition

Particular Levels (columns) — *List of Evaluative Criteria* (rows)

1. **Decide what is needed - a common rubric or separate rubric(s).**

    Where multiple assessments within a course have a similar structure or focus, it may be possible to use the same marking rubric. Conversely, where assessments have very different formats or test highly divergent skills/knowledge, it is likely that separate rubrics are needed. For example, it would be difficult to use the same rubric to assess an

---

[1] Dawson, P. (2017). Assessment rubrics: towards clearer and more replicable design, research and practice. Assessment & Evaluation in Higher Education, 42(3), 347-360.
[2] Popham, W. J. (1997). What's wrong-and what's right-with rubrics. Educational Leadership, 55, 72-75.

OSCE (objective structured clinical examination) and a literature review assignment. The principle is that the linkage between the rubric and any assessment should be strong enough to offer staff/students meaningful guidance on how to complete the task(s) and interpret marks/feedback.

Even where a series of rubrics are used across a programme, there is value in achieving as much consistency as possible in terms of appearance; the kind of feedback fields used; and choice of terminology (e.g. 'patients' vs. 'clients' vs. 'service users'). This will help the cohort (and teaching team) progressively build a consistent vocabulary and become more familiar with how to work with rubrics.

2. **Identify around five key evaluative criteria.**

An assignment could, for example, ask participants to devise a new treatment programme. The task could be broken down into i) critical exploration of available evidence/theories/concepts; ii) mapping of proposals to prevailing professional/regulatory standards; iii) detailed specification of the treatment protocol; iv) consideration of the challenges of change management; and iv) general quality of communication/presentation.

Whilst most criteria will be assessed across a continuum of grading levels (e.g. Fail, Pass, Merit, Distinction at Masters level), some criteria may be 'pass/fail' only. One example might be a criterion relating to 'demonstration of safe clinical knowledge'. In cases like this, where an individual criterion is of critical importance, the rubric might command that a fail judgment for this criterion results in immediate failure for the assignment overall. Such criteria can be referred to as 'red lines'. Where any red lines are in operation they must be clearly flagged to students embarking on the assignment.

3. **Consult available descriptors to establish minimum expectations for the stage of study.**

For different stages of academic study[3] SEEC descriptors and QAA descriptors identify the expected level of students' knowledge/skill in relation to matters such as conducting enquiries; critiquing information; applying understanding to different contexts; working autonomously; collaborating with others etc. For example, at Masters level students are expected to show greater ability than undergraduates in terms of identifying new methods of investigation; applying understanding in novel contexts; and appreciating inter-relationships between multiple domains of knowledge. Careful reading of these

---

[3] The National Qualification Framework (NQF) is used; Bachelors degrees span tiers 4-6 and Masters degree are at tier 7.

frameworks can, therefore, generate various ideas as to how expected performance in each evaluative criteria can be differentiated across the grade levels.

Whereas SEEC/QAA descriptors are generic, there is guidance on the discipline-specific knowledge and skills that students should accrue by the end of given Bachelors/Masters courses. This can be particularly helpful when writing rubrics for the final stage of a programme. The QAA provides dedicated benchmark statements for many established disciplines. Furthermore, PSRBs (professional, regulatory, or statutory bodies), such as the Nursing and Midwifery Council or the Institution of Engineering and Technology, issue detailed standards for the programmes that they accredit. Whilst it is important to consult these sources, typically they do not provide specify expectations for individual years of study as is the case for SEEC/QAA descriptors.

4. **Start by defining the expected performance to achieve a 'pass'.**

Given that SEEC/QAA descriptors convey the expected level of knowledge/skill to *pass* a given academic stage, it makes sense to start by formulating the quality definition for passing each of the evaluative criteria. One can then work progressively upwards and downwards from this datum.

When deciding expected performance for higher grade levels it can be helpful to consult the SEEC/QAA descriptors for the next academic tier. For example, according to descriptors the ability to propose new investigative methods is a hallmark of Masters performance (NQF 7). Whilst it would be unreasonable to make this a pre-requisite for passing a final year Bachelors module (NQF 6), it could be used to exemplify the kind of performance expected to achieve a 2.1 or 1st grade.

For a given evaluative criterion, a rubric will typically provide a separate quality definition for each grade level (e.g. fail, 3rd class, 2.2, 2.1., 1st class). A historical quirk of UK HE is that the upper and lower grade levels have a large span. In Bachelors programmes, for example, 'fail' spans from 0-40% and '1st class' spans 70-100%. Some markers have decided that these levels should be further subdivided, with rubrics offering distinct quality definitions. For example, failure at undergraduate level could be subdivided into 0-19% (i.e. a 'poor fail') and 20-39% (i.e. a 'marginal fail'), whilst 1st class performance could be subdivided into 70-84% and 85-100% (i.e. a 'strong 1st class'). Making this effort to expand the range of quality definitions gives students further clarity regarding the expected level of performance, and markers the confidence to use the full marking range.

5. **Avoid subjective language in favour of explicit guidance.**

   As one moves across the quality definitions for a given criterion it is sometimes apparent that the author has used progressively more complimentary adjectives. For example: 'poor use of literature' > 'acceptable use of literature' > 'good use of literature' > 'excellent use of literature'. Problematically, such wording does not provide the student with any explicit guidance on how to perform well and leaves the marking process subjective and opaque. For each evaluative criterion, it is better if the associated quality definitions provide clear manifestations of performance at different levels. For instance, in relation to the example here – using literature – a rubric might differentiate performance between levels according to whether the student incorporated *contemporary* literature; used literature relating both to theory and empirical evidence; critiqued the methodological rigour of the literature; appraised the *relative* authority of sources; identified significant gaps in the current literature base; and so on.

   Some commentators have suggested that defining such explicit requirements is damaging; students may mechanically 'tick off' the sub-tasks on the rubric without exercising much original thought (effectively completing 'assessment by numbers'). This interpretation assumes that well-defined expectations are easy to accomplish, which is simplistic. Clearly signaling to students that they must, for example, 'propose alternative investigative methods' or 'apply knowledge to an unfamiliar context' does not negate the individual effort required to achieve each milestone, or the ingenuity/creativity that participants can exercise in the process.

   An overarching principle is that all language used within the rubric should be as clear and concise as possible.

6. **Decide the scoring strategy for generating a final mark.**

   Where an educator feels strongly that certain evaluative criteria are more important than others, they can choose to reflect this through weighting (e.g. criteria 1-3 each contribute 30% towards the final grade, whereas criteria 4-5 only contribute 5% each). If adopting this approach care is required, as it may provoke students to concentrate fixedly on performance in particular evaluative criteria and their sub-elements. An alternative approach is to highlight which grade level the work achieved for each evaluative criteria but without applying any weighting. An overall verdict can then be offered. For example, where work from a Masters student meets the quality definition of a 'Merit' in 3 evaluative criteria and a 'Distinction in 2 evaluative criteria the marker might choose to award a percentage commensurate with a 'high Merit' (e.g. 68%).

   Whilst one of the fundamental purposes of a rubric is to enhance accuracy of assessment, excessive efforts to systematize the marking process fail to recognize that

with rare exceptions (e.g. some mathematical calculations) most work produced by students remains open to some subjective interpretation. Consequently, alongside well-conceived rubrics there is need for a mature conversation between staff and students about how, courtesy of the numerous elements of assessment built into programmes and processes such as moderation, it is highly unlikely that marking variance in any single assignment will impact on an individual's final award classification.

On a related topic, teams may consider adopting a policy of 'categorical' marking. This involves restricting markers to, for example, three marks within each grade level, which are commensurate with 'lower', 'medium' and 'higher' performance (e.g. 52%, 55%, 58%). This approach recognizes that whilst markers can make broad judgements about the quality of work oftentimes they cannot operate at any finer level of granularity (e.g. defending why one assignment attained 65% whilst another achieved 66%). The use of categorical marking also dissuades the awarding of marks very close to grade level thresholds (e.g. 69%), which may provoke strong student reactions.

7. **Decide how the rubric will be used to provide feedback.**

Before writing any free text, markers can use rubrics to convey valuable feedback. For each evaluative criterion, the appropriate quality definition can be shaded or highlighted to inform a student where his/her performance lay. The student can then easily read the quality definition for the neighbouring grade level(s) to establish how a higher mark may have been achieved. This use of rubrics might be particularly beneficial during formative assessment, by explicitly conveying to students the mark they are on course to achieve whilst there remains time to respond.

If each quality definition within the rubric is 'indexed' using letters/numbers the marker can, at opportune points, use shorthand to annotate an assignment, highlighting links with the rubric (e.g. '1M' could refer to criterion 1, definition for Merit). If even more immediate impact is sought, each quality definition might be sliced into short statements, so forming a 'comment bank' that can be used to annotate student assignments via 'cut and paste'. This represents a time saving, compared with manually writing annotations.

When concluding the marking process it is likely that colleagues will want to provide some open comments. To help structure feedback some markers provide a succinct statement against each evaluative criterion, regarding how future work might be enhanced. The rubric can be expressly formatted for such an approach by inserting a column of feedback 'boxes' adjacent to the criteria. Naturally, the volume and depth of feedback that the marker provides for each criterion may depend on the grade level that the student achieved.

Instead of such a system of multiple feedback boxes, some markers prefer a single field. Again, this can be incorporated onto the same file/piece of paper as the shaded/highlighted rubric, such that they can be read in conjunction. Where this approach is taken there is value in ensuring that feedback is succinct and focused clearly on how to enhance future performance. For example, markers may present feedback at 'top three' or 'top five' recommendations for enhancing future performance, employing a series of bullet points.

8. **Work as a team to master the rubric.**

   At the start of the process any current members of the marking team should have sight of the rubric and check, via discussion, their shared understanding of it. This can occur even if one individual has led the initial process of devising the rubric. Where the rubric has been devised by an individual for his/her exclusive use there is still value in sharing it with colleagues as a way of gauging its quality and clarity.

   Once shared understanding is achieved, teams are advised to use the rubric as a central element of a 'calibration exercise'. For example, shortly after a submission deadline the team may mark the same subsample of student work. They can then convene to compare marks, using the rubric to resolve any disparities and so forge further understanding. Team members can then proceed to mark the remaining submissions independently. The subsample of work, agreed marks, and feedback/notes made by the team, should be retained. These could be shared with future student cohorts as agreed exemplars of work in different grade levels.[4] Furthermore, they could act as a valuable source of guidance to newly recruited staff who did not have opportunity to engage in initial team discussions but are being asked to contribute to marking.

   It is also important to acknowledge that a rubric is not a static document. Rather, the need for changes or enhancements might arise from markers' experiences of using it on the 'frontline'; student feedback; modifications to the associated assessment; or developments in the local context or wider discipline. Consequently, it is important that teams revisit the rubric periodically (e.g. on an annual basis) to check that it remains fit for purpose.

9. **Ensure that students have early opportunities to engage with the rubric.**

   Even the most thoughtfully-devised rubric will have little positive impact on students' performance unless they use it diligently to inform their work. Empirical studies have demonstrated improvement to students' grades following opportunities within the

---

[4] Permission should be sought from the students who produced the work.

formal curriculum to build experience and understanding of rubrics.[5, 6] Amongst a range of possible activities, students can work independently or in groups to mark exemplar assignments using the rubric, prior to the lecturer revealing and explaining the received mark. Alternatively, students can be asked to systematically assess their own work or that of a peer using the rubric, either as an informal activity or a formal part of the submission process. Activities like these are said to enhance students 'assessment literacy'. Encouragingly, evidence indicates that following participation students continue to achieve higher marks on subsequent modules, where such activities are not embedded. One explanation is that early exposure to assessment literacy activities inculcates students with awareness about the pivotal role of rubrics and the skills for deconstructing them.

Delivery teams should take a panoramic view of their programmes to identify early opportunities in core modules, where assessment literacy activities might be embedded. In modules that are not earmarked for this purpose, the leader should still ensure that sufficient time is allowed for informal discussion about the assessment and corresponding rubric.

10. Consider using rubrics as a vehicle for student engagement.

There is a growing discourse around the importance of giving students opportunities to help make decisions about their education. This is reflected in the addition during 2017 of a new bank of questions to the National Student Survey (NSS), relating to the 'student voice'. Traditional methods for using student input are not collaborative; teaching staff survey student opinion and use this to make executive decisions. Under a different, 'partnership' approach, staff and students work together to resolve challenges and plan future aspects of the programme.

The production of marking rubrics could be an ideal activity to realise student engagement. Whereas changing some aspects of the curriculum can be protracted owing to quality assurance rules (e.g. learning outcomes), assessment rubrics can be modified on a flexible basis. Furthermore, on programmes comprising multiple years students can be invited to work on a rubric from the next academic year. This addresses one of the fundamental challenges attending traditional systems of student engagement – that individuals rarely benefit themselves from the input they provide.

---

[5] Rust, C., Price, M., & O'Donovan, B. (2003). Improving students' learning by developing their understanding of assessment criteria and processes. Assessment & Evaluation in Higher Education, 28(2), 147-164.

[6] Jones, L., Allen, B., Dunn, P., & Brooker, L. (2017). Demystifying the rubric: a five-step pedagogy to improve student understanding and utilization of marking criteria. Higher Education Research & Development, 36(1), 129-142.