**Creating and Using Good Rubrics**

**Introduction**

Creating and using good rubrics can simplify the grading process for instructors and help provide general feedback on class performance on an assignment. Rubrics can also clearly outline to students what is expected for each assignment and satisfy them that their grades are being assigned objectively.

Rubrics essentially detail how marks/scores should be distributed based on the quality of each student’s completed assignment. They can be broken down into sub-sections for each assignment, but to be useful they must be detailed yet easy to understand and follow, so that different individuals using a rubric will award the same marks/scores when they grade the same student assignment.

Holistic rubrics require graders to assess the learning process as a whole without judging individual components on their own, whereas analytic rubrics operate in the opposite way; they require graders to score individual components of a student’s work on their own (e.g. different questions on an assignment) and then sum the total scores to provide one final grade1.

Holistic rubrics may be suitable for some writing assignments if you are happy for students to make errors in individual components providing their final product is still of high quality (e.g. perhaps a few grammatical errors are tolerable when the main learning objective is to research the literature and present a content-heavy essay that is supported by the literature).

Generally, analytic rubrics are preferred when a relatively focused response is required (e.g. when you want to assess student writing ability based on grammar, punctuation and mechanics, structure, content, logic, and use of sources, or if there are many individual tasks that students need to complete in one assignment). Whether you use a holistic or analytic rubric should depend on the assignment and associated learning goals.2

**Creating a Rubric**

Each rubric will differ based on the assignment you ask your students to complete, but you should focus on the specific learning objectives you wish students to develop3. Think about the observable attributes you want to see from them, as well as those that you don’t.

If you are creating a rubric for an existing assignment that has been used in past offerings of the course, it may be useful to re-read some of those assignments to gain a perspective of the typical spectrum of answers that students provided; it is helpful to predict the kind of answers you expect to see when designing a rubric to cover as many of these as possible, as well as keeping in mind what students must typically show in their assignments to achieve low, middle and high grades.

When creating rubrics (particularly if they are holistic), it is helpful to divide each criterion into levels, such that you create categories that relate to different responses that reflect the progression from novice to expert-like writing (e.g. a score of 1 might represent *emerging* ability, whereas a score of 5 might represent the quality of an *expert*.

Defining these categories can be useful for students who wish to monitor their learning progress, especially because the scores/marks will not represent linear progression (e.g. a student who scores 4 out of 5 for logical development does not mean that student has twice as much ability as one who scores 2 out of 5).

If you are creating a rubric for a new assignment, you should keep the following tips in mind:

1. First divide the total marks for the assignment into different sections (e.g. five marks for the depth of content, five for the quality of sources used, five for integration of these sources, five for quality of argument, and five for paragraph structure and transitions). Note that most rubrics focus on six to seven criteria4, but the absolute number should depend on what the assignment5 asks of students; in most cases, limiting the number of criteria is more practical, but sufficient scope should exist within these criteria to distinguish the full range of likely student abilities.
2. Provide a detailed explanation of what an answer would need to show to be awarded any mark/score within the range available for each section (e.g. if you have allocated up to five marks for the depth of content, you should clearly state what a student must cover to gain one, two, three, four and five marks).
3. Do not use potentially ambiguous explanations. For example, do not propose a scale of 0 – 3 marks where *weak, fair, good* and *very good* are the descriptors used to differentiate between scores of 0, 1, 2, and 3 because different graders will likely differ on their interpretation of what is weak, fair, good or very good. Instead, try to provide objective definitions (e.g. *less than one primary source = 0, one or two primary sources = 1…).* If your rubric is holistic rather than analytic, you should provide detailed summaries (with examples) to clearly distinguish between marks/scores.
4. Add a qualitative description to each of these marks/scores if you plan to share the rubrics with students at any stage, so that they can assess their own learning development based on the grade they receive for the assignment.
5. Share the rubric with colleagues and TAs and ask for feedback. Expect to make changes based on this feedback.
6. If you plan to share the rubrics with students, ask these same students whether it is clear to them before the assignment begins. If they misinterpret any part of it, this may be a sign that you should make changes.

By following the above steps, you should design rubrics that are straightforward and objective to use. However, when more than one person will grade the assignments, it is important for all graders (e.g. instructors and TAs) to meet to troubleshoot any issues at an early stage. Ensuring your graders can use your rubric objectively is a crucial part of the design process, which is why the following steps are very important before you grade the assignments from an entire class:

1. Try to select a range of assignments from students of varying abilities and then take turns to grade each one using the rubric.
2. Compare the grades you each assigned the same assignments and calculate your **inter-rater reliability (see below)**.
3. If you have a low reliability (e.g. you have awarded different grades to the same assignments), this indicates issues with the rubric or the way it is being interpreted.
4. In such a scenario, work through each section and pinpoint where – and why – graders have awarded different scores based on the rubric, and then rephrase the rubric to make it more objective.
5. Grade a new set of assignments in the same way before comparing grades until you are satisfied that your inter-rater reliability is sufficiently high, and you are confident that you are all using the rubric in the same way.

**Inter-rater Reliability (IRR)**

Calculating inter-rater reliability (IRR) provides an estimate of the degree of agreement between different graders using the same rubric. A well designed, objective rubric should result in a high IRR (approaching 1), whereas a poorly designed, ambiguous one will result in a low IRR (approaching 0 or –1, depending on your method).

All graders who will be grading assignments using your rubric should take part in IRR assessments; if they do not, the IRR estimate you obtain will not encompass data from everyone whose interpretations will provide the final grades. As a result, the rubric may not be effectively assessed before it is used to grade the assignments of a whole class.

There are various techniques for providing IRR estimates, and the best one to use depends on the situation6. When you obtain data from three or more coders, it is generally best to use an extension of Scott’s Pi statistic7, or compute the arithmetic mean of kappa8 (a statistic used in IRR analysis9). There are no cast-iron guidelines for an acceptable level of agreement, but popular benchmarks for high agreement using kappa are 0.7510 and 0.811. Hallgren6 provides a detailed overview of these procedures.

**Using a Rubric to Provide Feedback**

You can save time when providing assignment feedback to large classes by referring to the rubric when highlighting example student answers, and then explaining how/why they were scored/marked as they were. By showing a spectrum of answers, you can also indicate to students how they must approach similar questions in future to attain high scores/marks.

Depending on the situation, you may also wish to provide a marked rubric with each student’s assignment, so that they can see how they have been assessed on each question12. This can be very useful for students, and can again save you time in providing feedback in large classes as it can reduce the need to write explanations to justify scores/marks throughout.

**Sample Rubrics and Useful Links**

Some useful links to rubric templates and examples are shown below:

* <http://uwf.edu/offices/cutla/supporting-pages/examples-of-rubrics/>
* <https://www.ncsu.edu/midlink/ho.html>
* <http://uwf.edu/media/university-of-west-florida/offices/cutla/documents/peer-review-guidelines.pdf>

**References**

1: Nitko AJ. Educational assessment of students. 3rd ed. Upper Saddle River, NJ: Merrill. 2001.

2: De Leeuw J. Rubrics and Exemplars in Writing Assessment. In: Scott S, Scott DE, Webber CF, editors. Leadership of Assessment, Inclusion, and Learning. Switzerland: Springer International Publishing; 2016. p.89-110.

3: Mertler CA. Designing Scoring Rubrics for Your Classroom. Pract Assess, Res & Eval. 2001; 7(25).

4: Andrade HG, Wang X, Du Y and Akawi RL. Rubric-referenced self-assessment and self-efficacy for writing. J Educ Res. 2009; 102:287-301.

5: Covill, AE. College Students’ Use of a Writing Rubric: Effect on Quality of Writing, Self-Efficacy, and Writing Practices. J Writ Assess. 2012; 5(1).

6: Hallgren KA. Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. Tutor Quant Methods Psychol. 2012; 8(1):23-34.

7: Scott WA. Reliability of content analysis: The case of nominal scale coding. Public Opinion Quarterly. 1955; 19(3):321-325.

8: Davies M, Fleiss JL. Measuring agreement for multinomial data. Biometrics. 1982; 38(4):1047-1051.

9: Cohen J. A coefficient of agreement for nominal scales. Educ and Psych Measurement. 1960; 20(1): 37-46.

10: Fleiss JL. Statistical methods for rates and proportions. 2nd ed. New York: John Wiley. 1981.

11: Altman D. Practical statistics for medical research. Boca Raton, FL: CRC Press. 1991.

12: Huba ME, Freed JE. Learner-Centred Assessment on College Campuses. Boston: Allyn & Bacon. 2000.