

Using rubrics to improve marking reliability and to clarify good performance

Catherine Hack

University of Ulster
Coleraine BT52 1SA
cj.hack@ulster.ac.uk
<http://www.ulster.ac.uk>

Abstract

A quantitative review of marks achieved by students in two assignments on a distance learning module indicated that students were not acting on the feedback they received from the first assignment. Student assignments were marked by one of five tutors; the average mark in each tutor group varied by over 15%, whilst the standard deviation within groups varied from 3.1 to 17.6, this indicated that tutors were marking across different grade ranges. The introduction of rubrics both raised the marks students achieved in their second assignment and improved marker reliability, reducing inter-group variability to less than 10% and increasing the average standard deviation within groups to 17.7. Feedback from students indicated that they found the rubrics very helpful in clarifying performance and promoting self-assessment, whilst the tutors felt that it was a time efficient and informative method of providing feedback.

Keywords

Rubric, feedback, marker reliability

Introduction

Rubrics are an assessment tool which can be used to evaluate student performance over a range of criteria. A rubric normally comprises three main features (Reddy and Andrade, 2010):

Evaluation criteria: which are usually mapped to the learning outcomes or competencies that are to be measured;

Quality criteria: qualitative descriptions of what is expected for a given grade or mark;

Scoring system: Grade ranges or degree classifications mapped to the quality description (Figure 1).

Rubrics have been used in a wide range of subject areas, and have been particularly effective in providing student guidance on specific skills and competencies such as reflective writing (Koole et al, 2012; Wald et al, 2012), discussion board contributions (Giacumo, Smith and Savenye, 2012, and Penny and Murphy, 2009), scientific reasoning (Timmerman, et al, 2011) and oral presentations (Welch, Suri and Duran, 2009, Lunney, and Sammarco, 2009). Rubrics have also been shown to improve rater reliability (Oakleaf, 2009 Thaler, Kazemi, and Huscher, 2009, Timmerman, et al, 2011), and are therefore particularly useful for assessing large cohorts and where multiple or inexperienced assessors are employed or for peer assessment. Students can be provided with a rubric prior to submission of an assignment, which helps to clarify what is required for good performance. Rubrics have also been shown to promote self-reflection and critical assessment of a student's own work. The provision of feedback via a rubric can help students identify the specific elements of their work that require improvement.

Criteria	Fail	Pass	Merit	Distinction
Introduction/Conclusions	<p>7 to 9 points</p> <p>The introduction provides background to a topic. Some evaluation and synthesis of key issues and material presented in conclusions.</p>	<p>10 to 11 points</p> <p>The introduction provides a brief background to a challenging topic e.g. why is it relevant, and why does it raise ethical issues. The conclusion provides critical evaluation and synthesis of key issues and material.</p>	<p>12 to 13 points</p> <p>The introduction provides a brief background to a challenging topic e.g. why is it relevant, and why does it raise ethical issues. The conclusion provides a critical evaluation and synthesis of complex issues and material. Both intro & conc should demonstrate an original and reflective approach.</p>	<p>14 to 20 points</p> <p>The introduction provides a brief background to a complex and challenging topic e.g. why is it relevant, and why does it raise ethical issues? The conclusion provides a critical insightful evaluation and synthesis of presented evidence.</p>
Literature Review	<p>7 to 9 points</p> <p>Limited evidence of reading. Limited referencing and bibliography. Missing either scientific or ethical evidence, or not balanced arguments.</p>	<p>10 to 11 points</p> <p>Evidence of reading supplementary sources. Adequate referencing and bibliography for both scientific and ethical aspects of review. Papers that support either side of argument.</p>	<p>12 to 13 points</p> <p>Evidence of extensive reading of supplementary sources. Comprehensive referencing and bibliography for both scientific and ethical aspects of review. Papers that support either side of argument.</p>	<p>14 to 20 points</p> <p>Extensive evidence of integrating supplementary sources. Outstanding referencing and bibliography for both scientific and ethical aspects of review. Papers that support either side of argument.</p>
Scientific basis	<p>7 to 9 points</p> <p>Basic knowledge and depth of understanding of key principles and/or concepts only.</p>	<p>10 to 11 points</p> <p>Demonstrates understanding of the specialist topic that underpins the ethical issue in question.</p>	<p>12 to 13 points</p> <p>Presents a clear understanding of specialist topic that underpins the ethical issue in question.</p>	<p>14 to 20 points</p> <p>Presents a systematic understanding of the specialist topic that underpins the ethical issue in question.</p>

Figure 1: Part of rubric for assignment 2 indicating evaluation and quality criteria with scoring system

The importance of high quality feedback for student learning is acknowledged by the National Student Survey (<http://unistats.direct.gov.uk>.) At the University of Ulster, seven principles of feedback and assessment were identified to support students and staff in providing effective feedback (University of Ulster, 2011) which are to:

1. Clarify good performance
2. Encourage time and effort on task
3. Provide timely high quality feedback

4. Provide an opportunity to act on feedback
5. Encourage positive motivational beliefs
6. Develop self-assessment and reflection skills
7. Encourage interaction and dialogue

These principles informed the assessment strategy for a post-graduate module, which was designed to provide prompt feedback on a short piece of work, which the students should be able to implement in a more substantive second piece of work. The module was delivered to over one hundred distance-learning (DL) students and to 10-20 on-campus students. The DL students are organised into groups comprising 20-25 students with a tutor who is responsible for assessment and feedback. A quantitative analysis of marks indicated that students who failed or presented 'adequate' or 'acceptable' quality of work in their first submission, did not improve significantly in the second assignment. This was confirmed by a paired student t-test, in which it was shown that in 3 out of the five tutor groups there was a significant decrease in the average mark between assignments. The group mark average for assignment 1 varied from 58% to 73%, with a similar variation in assignment 2. The standard deviation varied from 3.1 to 17.6, indicating that the tutors were marking across very different grade ranges. To ensure there was fair and consistent marking across the module the marks had to be moderated by the module coordinator; this required a significant amount of work and presented delays in giving feedback to the student. This reduced the opportunity for the students to act on the feedback for their second submission. It was also noted that there was wide variability in the quantity and quality of feedback given to students. The aim of this study was to introduce rubrics into the module and determine their effect on the consistency of marking across tutor groups and whether they can improve performance of students as they progress from assignment 1 to assignment 2.

Method

The rubrics were designed by using generic assessment criteria for qualitative based work at level 7 (University of Ulster, 2012). These were then contextualised within the specific requirements for the pieces of coursework. There are many on-line tools available for creating rubrics; for this assignment the rubric tool within our Virtual learning Environment (VLE), Blackboard Learn (BBL) was used which provided a streamline process for marking, moderating and providing feedback and marks. Students were presented with the rubric in BBL alongside the assignment requirements, and they received their feedback via the rubric grid. Feedback from the students was obtained via an on-line questionnaire. The rubric was piloted with the on-campus cohort; minor modifications were made to the quality criteria prior to implementation with the DL cohort.

Results

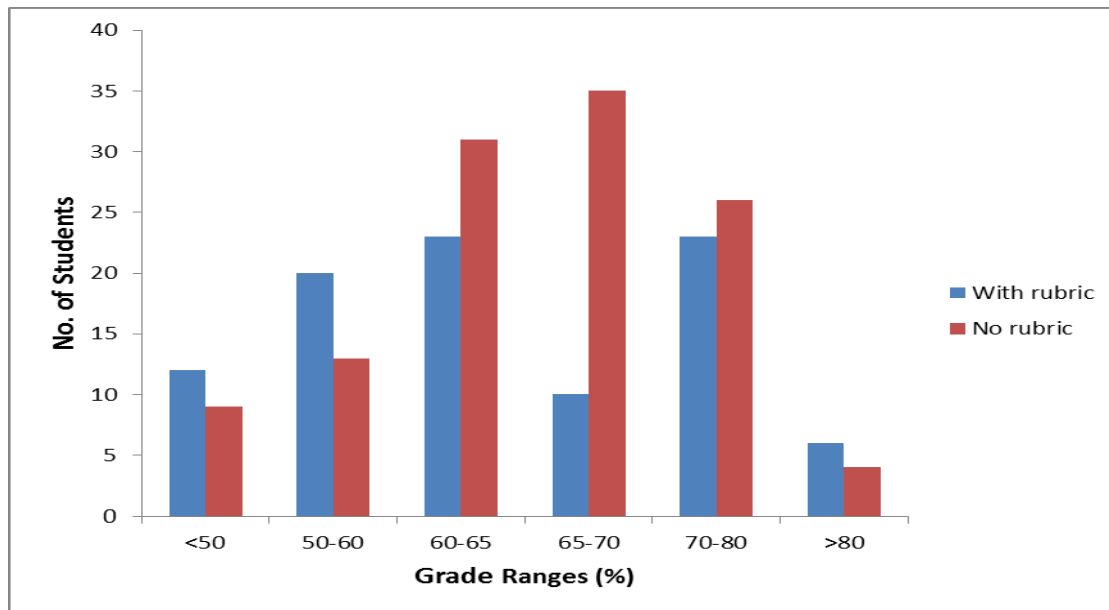


Figure 2: Marks achieved by DL students pre and post introduction of rubrics

The introduction of rubrics to the pilot group of on-campus students increased the average student mark by over 5% from assignment 1 to assignment 2. A paired t-test indicated that there was a significant improvement in student marks. This was a very encouraging result as the second assignment was a more substantive piece of work than the first, and required that the students reached many more learning outcomes. Interestingly, students who were scoring grades in the Merit range (60-70%), had the biggest increase in their grades, achieving a Distinction (>70%) in the second assignment. Following the successful outcome of the pilot, the rubrics were introduced into assignment 1 of the DL module. The difference in the group average decreased from 15% to less than 10% after the introduction of the rubrics, whilst the standard deviation for each group was between 13 and 20 (Figures 2 and 3).

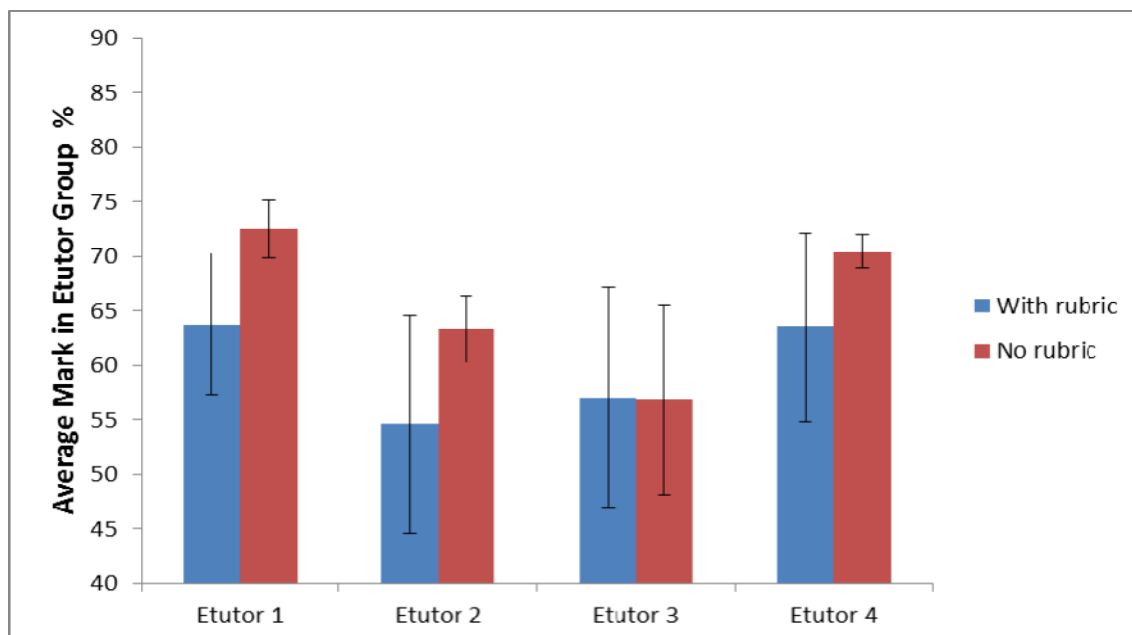


Figure 3: Average mark within each tutor group before and after the introduction of rubrics

3.1 Student and Tutor Feedback

Both students and tutors were asked how helpful they found the rubrics. The questions were framed in terms of meeting six of the Ulster Principles of Assessment for Feedback and Learning. The majority of both tutors and students agreed that the rubrics were very helpful in meeting the Ulster Principles, with the strongest endorsement of the rubrics for clarifying good performance and developing self-assessment skills. Furthermore the tutors found the rubrics an efficient and informative method of providing feedback.

Discussion

This study supports other works which has demonstrated the value of rubrics for clarifying performance and developing self-assessment skills (Reddy and Andrade, 2010). These preliminary results indicate that student performance can be enhanced by students both using the rubric prior to submission and providing feedback via the rubric. To be effective the rubrics have to:

Map to the learning outcomes of the module

Be informed by the appropriate level assessment criteria

Use clear language that is understood by students and markers

The most noteworthy enhancement in student performance was observed by students in the grade range of 60-70%. This part of the study was only conducted on a small cohort, and there were several factors that could be involved, for example, these students were potentially more capable of interpreting the rubric and the feedback and/or they were more motivated to improve their grade to a distinction. Going forward the rubrics will be used in assignment 2 for the larger cohort which will provide more evidence in this area. The use of rubrics within the large distance learning module was particularly effective at improving grading efficiency and reliability. The tutors provided a wider range of marks when using the rubric, indicating that they were more confident about what was required for each grade boundary, furthermore the tutors stated that they found the rubrics an efficient way of grading assignments and providing high quality feedback.

References

- Giacumo, L.A., Smith, N and Savenye, S (2012) Facilitation prompts and rubrics on higher-order thinking skill performance found in undergraduate asynchronous discussion boards, *British Journal of Educational Technology*, (Early view on-line ISSN: 1467-8535).
- Lunney, M. and Sammarco, A. (2009) Scoring Rubric for Grading Students' Participation in Online Discussions, *CIN-Computers Informatics Nursing*, 27, 26-31
- Oakleaf, M (2009) Using rubrics to assess information literacy: An examination of methodology and interrater reliability, *Journal of the American Society for Information Science and Technology*, 60, 969-983.
- Penny, L and Murphy, E (2009) Rubrics for designing and evaluating online asynchronous discussions, *British Journal of Educational Technology*, 40, 804-820.
- Reddy, Y.M and Andrade, H. (2010) A review of rubric use in higher education *Assessment and Evaluation In Higher Education*, 35 435-448.
- Thaler, N, Kazemi, E. and Huscher, C. (2009) Developing a Rubric to Assess Student Learning Outcomes Using a Class Assignment, *Teaching of Psychology*, 36, 113-116.
- Timmerman, B.E.C., Strickland, D.C., Johnson, R.L., Payne, J.R (2011) Development of a

'universal' rubric for assessing undergraduates' scientific reasoning skills using scientific writing *Assessment & Evaluation In Higher Education*, 36, 509-547 .

Welch, H. Suri, D. and Durant, E. (2009) Rubrics for assessing oral communication in the capstone design experience: Development, application, analysis and refinement. *International Journal of Engineering Education*, 25, 952-961.

University of Ulster (2008) Teaching and Learning Strategy,

<http://www.ulster.ac.uk/academicoffice/download/T&LStrategy.pdf> University of Ulster

(2012) Assessment Handbook, page 136-146,

<http://www.ulster.ac.uk/academicoffice/download/Assessment%20Criteria.pdf>