Does a well-defined scoring rubric lead to an improvement in student results?

Dragi Klimovski[a].
Swinburne University of Technology[a]
Corresponding Author Email: dklimovski@swin.edu.au

Structured abstract

BACKGROUND
In order to achieve an accurate and reliable measurement of students' performance in a subject, good quality assessment and evaluation tools must be used. One such tool is a detailed and well-defined rubric (Arter & McTighe, 2001). A rubric is a clear and unambiguous indication of what is expected of students in order to achieve the various grade levels for a piece of assessment (Rubrics, 2012; Moskal, 2000). Rubrics are often used to aid academics in measuring the ability of students to use and apply factual, conceptual and procedural knowledge (Bloom, 1956; Anderson & Krathwohl, 2001). Littlefair (2008) has presented work on the use of rubrics to identify the key areas of assessment and has shown these methods to reduce the impact of subjectiveness in the marking of engineering projects.

PURPOSE
The focus of this project is to investigate the effect on student results of providing a detailed, well-defined marking rubric in a final year design and development project subject.

DESIGN/METHOD
Over a decade (2003-2013) a final year design and development project subject has been run with the aim of allowing students to showcase the knowledge gained throughout their course and their ability to communicate via the three fundamental modes; oral, visual and written. In 2008 a well-defined scoring rubric was developed to make the marking requirements for the students' clearer and ensure consistent marking amongst the different academics. The research question is evaluated both by comparing the results of the student cohorts who used the rubric and those who did not. An anonymous survey was conducted to determine if the students actually used the rubric in completing their required tasks.

RESULTS
The comparison of results for the students' performance in both assessment components before and after the rubric do not differ significantly as the trend lines for both sets of graphs are very close to flat which indicates very little variation between the results before and after the rubric was introduced. The student survey results reveal that 92% of the students are aware of the scoring rubric for the poster and report requirement for this subject. 82% of the students used the rubric when creating their poster and 87% used the rubric when writing their report. Of the students who used the rubric, 69% believed it to be a well-defined rubric.

CONCLUSIONS
The results indicate that the use of a scoring rubric has not greatly impacted on the subject average results of the poster and design report components over the 10 years of results. However, a positive outcome is that a very high proportion of students have used the rubric.

KEYWORDS
Scoring Rubric, consistency, engineering project subject

Introduction

Evaluating student learning has been a challenge for as long as mankind has been sharing knowledge and ideas through teaching. Over the centuries since Socrates invented the teaching practice of pedagogy (Boghossian, 2003 Vol 23 Book 2), many approaches to assess the level of knowledge attained have been developed, tried and tested. The evolution of the transference of knowledge and the assessment of the level of understanding has led to modern day assessment styles which are needed to deal with the great abundance of information available through the Internet. In order to achieve an accurate and reliable measurement of students' performance in a subject, good quality assessment and evaluation tools must be used. A quite detailed yet brief review of the literature on assessment in the general classroom can be found in Brookhart (1999). The work by Brookhart (1999) references many different tools, one of which is a detailed and well-defined scoring rubric which is clearly illustrated in both (Brookhart, 1999) and Arter (2001). A rubric is a clear and unambiguous indication of what is expected of students in order to achieve the various grade levels for a piece of assessment (Rubrics, 2012). Rubrics are often used to aid academics in measuring the ability of students to use and apply factual, conceptual and procedural knowledge (Bloom, 1956; Anderson & Krathwohl, 2001). A number of authors have stated that criterion referenced assessment is forming the heart of the modern university structure, with rubrics a key part of it (Dunn, Parry, & Margan, 2002). Littlefair (2008) has presented work on the use of scoring rubrics in identifying key areas of assessment and has shown these methods to reduce the impact of subjectiveness in the marking of final year engineering projects. Using this line of logic, this research aims to explore the validity of the above claim and evaluate the impact a rubric can have on student performance.

In this paper the background for using rubrics in assessment is presented, followed by an outline of the approach used in gathering the data needed to undertake the evaluation of the research question posed. A discussion of the results shown in graphs is presented as well as the rationale to conduct a student survey. Lastly the conclusions and recommendations are presented.

Background

Over many years a final year design and development project subject has been used as a vehicle to allow Biomedical, Electrical and Telecommunication Engineering students to showcase the knowledge they have gained throughout their course. This subject also allows the students to demonstrate their ability to communicate this knowledge via the three fundamental modes; oral, visual and written. Many approaches have been trialled to ensure a fair and consistent assessment of the student's knowledge. A number of the methods used were self-assessment (Boud, 1995) and peer assessment (Boud & Falchikov, 1989). The self-assessment method allows students to be involved in the assessment exercise and encourages them to partake in self-reflection and critical analysis of their own work. In addition when involved in peer assessment, students better understand the assessment criteria. This enables them to improve their judgement on "how-well" their peers have met the set criteria. To allow students to engage in these forms of assessment it is very important to make the criteria for marking as clear as possible. These methods, of self-assessment and peer assessment, were implemented in this subject in an unplanned manner over the period 2003-2007. The lack of consistency over this period was primarily due to the frequent changing of the subject convener, where each convener implemented their preferred tools for assessing the students. Research shows that reliably assessing student knowledge can be improved through the use of rubrics (Jonsson & Svingby, 2007). The same authors also draw the following conclusion from their research, "*rubrics seem to have the potential of promoting learning and/or improve instruction".* The scoring rubric is valuable to both the student and instructor because it clearly conveys to each party what is considered important and what the assessors are looking for (Arter & McTighe, 2001; Busching, 1998; Brookhart, 1999). Given the positive research findings on scoring rubrics and the subject panel's concerns of

consistency in marking between project supervisors, it was decided to introduce a scoring rubric in 2008. There are 4 broad categories of scoring rubrics; holistic, analytic, general and task-specific (Mullinix, 2009). The type of scoring rubric chosen to be used in this final year project subject was task-specific, which concentrated on evaluating the major assessment components, the poster and the design proposal report.

Purpose

The purpose of this research is to determine the effect of using a detailed, well-defined scoring rubric on student performance. The student performance is gauged through their report and poster marks, in a final year design and development project subject. In addition to the impact on student results, the research attempts to determine whether the scoring rubric is clear and well-defined through a student survey.

Approach

A final year design and development subject has been run for the Biomedical, Electrical and Telecommunication Engineering students for many years. This subject has been used as a vehicle to allow students to showcase the knowledge gained throughout their course and their ability to communicate via the three fundamental modes; oral, visual and written. These communication skills are in strong demand in the job market for the current engineering profession. The students are required to present a poster and report of their design proposal.

The research conducted here concentrates over the period 2003 to 2013. Over this decade a number of issues were raised concerning the best method to assess students; ensure that students' have a good understanding of the criteria used for their assessment; and deliver consistency in the project marking by the various project supervisors involved. In 2008 a well-defined scoring rubric was developed using some of the lessons learnt in (Popham, 1997) and (Tierney & Simon, 2004) e.g. clear indicators, consistent performance criteria and to provide clearly worded criteria. A number of good examples of criteria, levels of performance, scores and descriptors (University) were also consulted in the creation of the scoring rubric.

Every year from 2008 onwards the students were advised of the existence of the scoring rubric and its possible benefit when used. The students were strongly encouraged to use the scoring rubric, but note that no strict measures were put in place to mandate its use i.e. penalties or meeting hurdle requirements. The two main aims of the scoring rubric were; to meet the challenges of clearly stating the marking requirements for the students; and ensuring marking consistency between the different academics, who were supervising the various projects. The research question is evaluated by comparing the average marks of the poster and the design proposal report between student cohorts who used the scoring rubric and those who did not. The period chosen, 2003-2013, provides an equal number of results of data both before and after the introduction of the scoring rubric.

To ensure consistency of the delivery of the subject the following parameters were kept as constant as possible over the period 2008 to 2013; lecturer(2 academics with a detailed handover at the change); assessment task; student population (always final year students) and course delivery.

It is worthwhile noting that the provision of a scoring rubric does not guarantee its use. In an effort to obtain honest feedback from the students, the current cohort of students was surveyed anonymously to ascertain whether they actually used the scoring rubric in completing their tasks. The survey also asks the students to judge the scoring rubric on the two criteria; how well-defined the requirements were and how unambiguous the statements presented were. The sample size for the anonymous survey was 48 students.

Item	percentage	Fail	Pass level	Credit level	Distinction level	High Distinction
Oral presentation	10%	All members do not participate in the presentation. Unclear outline of the problem.	Gives a brief outline of the problem. Some evidence of structure to the presentation, but limited detail.	Pass level plus engages the audience and provides a well structured presentation with logical flow. Covers detail of problem well.	Credit level plus identifies related issues to the problem	Distinction level plus shows extensive knowledge about the current situation and refers to possible future developments
Handling question and answer session	5%	Unable to answer questions.	Limited answers to questions, or unable to answer some questions.	Attempt to answer all questions, but only some questions answered fully.	Answers questions fully but only some team members able to answer	Gives full and substantial answers to all questions with all team members participating.
Content and format of poster	20%	Poorly structured poster. Lack of clear outline of suggested approach.	Poster laid out clearly with a basic overview of suggested approach. Approach is not new.	Pass level, with topic fully covered, but little or no detailed explanation. Some evidence of innovation in the approach used.	Credit level with detailed explanations. Only one form of presentation technique used (e.g. text only)	Distinction level plus layout very clear. Information flows well. Information is comprehensive, yet succinct. Good use of a range of different presentation techniques.

Figure 1: Scoring Rubric for the Poster Component

Figure 1 shows a sample of the scoring rubric used over the last 6 years. The sample provided is for the poster component, testing the student's ability to convey accurate and informative information both orally and visually.

Item	mark	Fail	Pass level	Credit level	Distinction level	High Distinction
Report presentation (structure)	10%	No evidence of minimum structure requirements (i.e. page numbering, main topic areas, reference list as shown in report example in BB). Considerable spelling and grammatical errors.	Meets minimum structure requirements; shows that proper care was taken in the final submission. May have some spelling and grammatical errors in text.	Pass level plus appropriate and logic breakdown of main areas into subtopics. No spelling errors. Text understandable but may have some grammatical errors	Credit level plus tables and figures to support text, appropriately labelled and embedded in report. Limited references used but use is appropriate. Grammatical structures generally correct but English used may be limited.	Distinction level plus detailed fluent explanations. Appropriate and comprehensive referencing. Good grammatical structure with comprehensive use of English.
Technical content	30%	Problem not clearly described or no way of approaching the problem given.	Gives an overview of the problem with aims and objectives clearly stated. May suggest only one way of approaching the problem. No evidence of research.	Pass level plus describes problem in detail Suggested multiple ways of approaching the problem. Evidence of applying prior knowledge/background in achieving objectives. Limited evidence of research into the area.	Credit level plus fully outlines the approach chosen. Some evidence of innovation in the approach chosen. Gives some evidence of research into the area and the reasons why the approach was chosen.	Distinction level plus gives comprehensive coverage of the problem, outlines the approach chosen, which is new and innovative and supports the reasons for the approach with substantial research.

Figure 2: Scoring Rubric for the Design Proposal Report

Figures 2 and 3 show samples of the scoring rubric used over the last 6 years for the design proposal report which tests the student's ability to convey information in a written format. This component tests their ability to convey a clear definition of the problem at hand as well as the rationale for its solution. In addition their capability to structure a report and communicate technical content is assessed.

Design rational, results analysis (for research project), discussion and lessons learnt	15%	No explanation given for design approach (or obtained results)	Gives some explanation but does not provide insights into the chosen approach/obtained results. No consideration of the environment and sustainability issues.	Provides detailed explanation and insights into the approach/results. Shows evidence of effort/method taken to obtain those insights. Takes in account environment and sustainability issues.	Credit level with useful discussion about lessons learnt from the discussion in relation to achieving the objectives of the project. Fully aware of the outcome of the project in terms of environment and sustainability.	Distinction level with insights that lead to an innovative approach and new ways of achieving the objectives (either in designing the product or conducting research).
Planning and Timeline	10%	No evidence of plan, or plan very sketchy.	Plan shows major aspects of the project listed with associated timeline.	Pass level plus deliverables included and some evidence of a breakdown of major tasks.	Credit level with comprehensive breakdown of tasks. Some evidence of resource allocation included and some attempt to keep plan up to date.	Distinction level with full allocation of resources shown. Evidence of plan kept fully up to date over time.

Figure 3: Scoring Rubric for the Design Proposal Report cont'd

Actual Outcomes
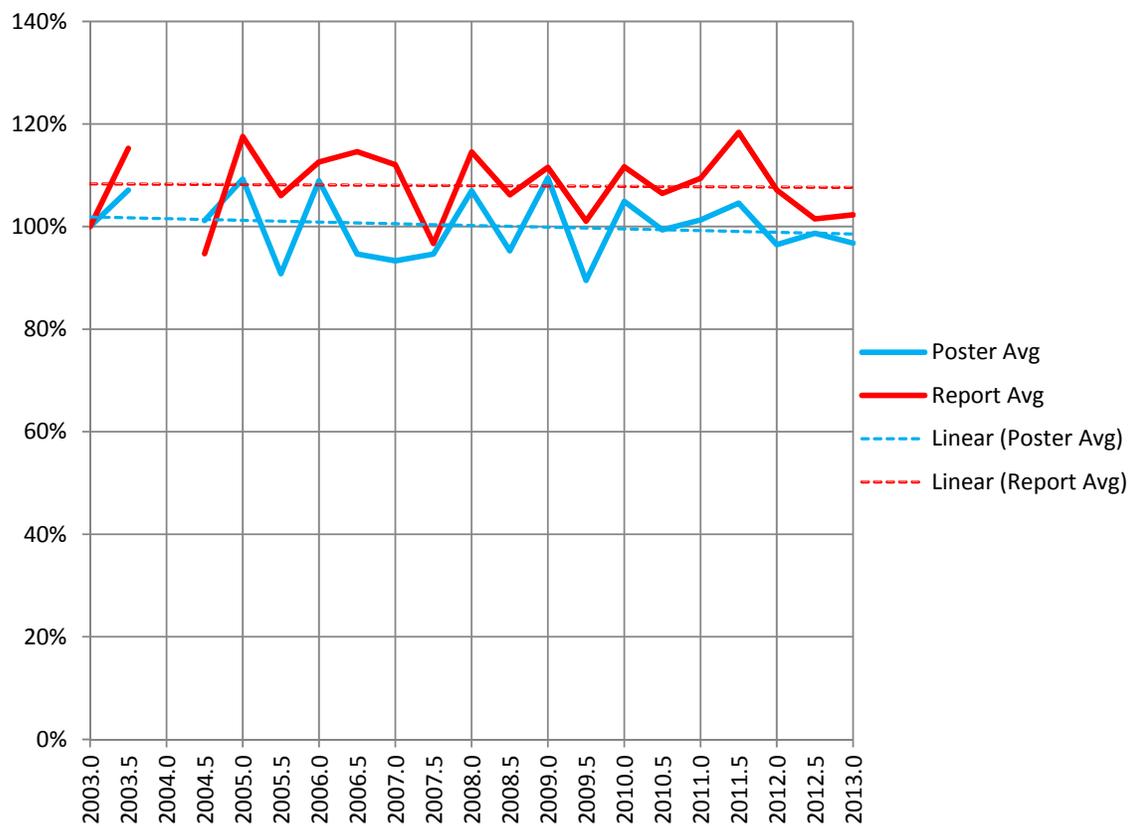
Figure 4: Avg Results for Poster and Report

The students' performance for both the poster component and the design proposal report component of the assessment are graphed and presented in figure 4 per semester. Note these results have been normalised so that they are relative to the 2003 results. This has been achieved by dividing the current year result by the corresponding 2003 result. Also note

that the result for the 2004 average poster and report components has been lost and hence cannot be displayed.

The comparison of average semester results for the students' performance, in both assessment components, before and after the application of the scoring rubric do not differ significantly. It is possible to explain the observed results as natural variations in student results based on the variability of students' skills and knowledge on a year-to-year basis.

The trend lines for both sets of graphs have negligible gradients which indicate very little variation between the results before and after the scoring rubric was introduced. In fact if anything, both trend lines have continued in the same direction after the introduction of the scoring rubric and have a slight negative gradient. This indicates that the scoring rubric has had minimal impact on the results. The expectation was that when the results from 2008 to 2013 were included that both trend lines, as a result of the expected positive step input in the results, would yield gradients which were positive thereby showing an improvement in outcomes.

The graphed average poster result has a negligible but larger negative gradient of -0.04, whilst the average report result has a negative gradient of -0.01 implying that the scoring rubric has not had the expected positive effect. Based on other researchers' findings (Jonsson & Svingby, 2007) and (Arter & McTighe, 2001) that rubrics can have a positive input to student learning, it was expected that the use of a scoring rubric would lead to higher average results. It was expected that once a detailed structure was provided to students outlining how they were going to be assessed and what was expected from their submissions, that the students would attain a higher average mark. This expectation was based on other literature findings of taking out the guesswork about their submissions. In fact there is an increase in 2008 but the results go down the semester, hence it is difficult to attribute the improvement to the introduction of the scoring rubric. A possible explanation of why higher average results were not attained could be that the students in this subject are at the final stage of their engineering degree. Hence they should already have a high level of maturity and drive to ensure they understand the criteria used to mark their assessable work. Possible future work with scoring rubrics could be trialled in the earlier year subjects where the level of maturity is not as high and students could benefit from the direction a scoring rubric provides.

The student survey results revealed a number of very positive outcomes as well one outcome which is a cause for some concern. For the students who were enrolled in the project subject in semester 1 of 2013 the survey results can be summarised as follows:

- 92% are aware of the scoring rubric for the poster and design proposal report components.
- 82% used the scoring rubric when creating their poster.
- 87% used the scoring rubric when writing their design proposal report.
- 69% believed it to be a well-defined, clear and unambiguous scoring rubric.
- 67% judged it to be one of the best aspects of the subject

The high percentage of students being aware of and using the rubric is very pleasing. The fact that only 69% of the students believed the scoring rubric to be clear and unambiguous is disappointing. More work needs to be done to ensure that a higher percentage of students find it clear and easy to understand. The finding that 69% found the rubric to have good clarity aligns well with the results that 67% of the survey respondents thought the scoring rubric was one of the "best" aspects of their final year project subject. This indicates that the students do not fully comprehend the value of a scoring rubric and is an area which needs to be better communicated.

Conclusions

A scoring rubric is a good tool to ensure the expectations of students' learning is clear and unambiguous. However, the outcomes of this study showed that the use of scoring rubrics, in a final year engineering project subject, did not produce an improvement in the average final semester results. Research shows that there is benefit in the use of scoring rubrics and perhaps their use should be trialled in earlier year subjects, where the structured marking guidance would be of greater benefit to newer students in the University system than in final year subjects.

A pleasing outcome is that a very high proportion of students have used the rubric, but further work is required in order to ensure that the definition of the rubric is improved so that all students can easily understand the requirements to achieve the various grade levels.

References

Anderson, L., & Krathwohl, D. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.* New York: Longman.

Arter, J., & McTighe, J. (2001). *Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance.* Thousand Oaks, CA: Corwin Press.

Bloom, B. S. (1956). Taxonomy of educational objectives. In B. Bloom, *Handbook 1. The Cognitive Domain.* New York: David McKay Co Inc.

Boghossian, P. (2003 Vol 23 Book 2). How Socratic Pedagogy Works. *Informal Logic - Reasoning and Argumentation in Theory and Practice (Teaching Supplement)*, pp17-25. Retrieved from http://ojs.uwindsor.ca.

Boud, D. (1995). *Enhancing Learning through Self Assessment.* London: Kogan Page.

Boud, D., & Falchikov, N. (1989). Quantitative studies of student self assessment in higher education: a critical analysis of findings. *Higher Education, 18*, 529-549.

Brookhart, S. M. (1999). *The Art and Science of Classroom Assessment: The Missing Part of Pedagogy.* The George Washington University, Graduate School of Education and Human Development. Washington DC: ASHE-ERIC Higher Education Report.

Busching, B. (1998). Grading inquiry projects. New Directions for Teaching and Learning. 89–96.

Dunn, J., Parry, S., & Margan, C. (2002, October 28). *Seeking quality in criterion referenced assessment.* Retrieved June 2nd, 2013, from Education-Line: http://www.leeds.ac.uk/educol/documents/00002257.htm

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 2*(2), 130-144.

Lister, R. (December 2000). On blooming first year programming, and its blooming assessment. *Proceedings of the Australasian conference on Computing education*, (pp. 158-162). Melbourne, Australia.

Littlefair, G., & Gossman, P. (2008). BE (Hons) final year project assessment–leaving out the subjectiveness. *Proceedings of the Annual Conference for the Australasian Association for Engineering Education - Industry and Beyond.* Brisbane.

Moskal, B. M. (2000, 3 29). *Scoring rubrics: what, when and how?* Retrieved August 15th, 2013, from Practical Assessment, Research & Evaluation, 7(3): http://PAREonline.net/getvn.asp?v=7&n=3

Mullinix, B. B. (2009, June 16). *Rubrics.* Retrieved August 13, 2013, from The TLT Group: Teaching, Learning and Technology A Not for Profit Organisation: http://www.tltgroup.org/RubricsExamples_files/MullinixRubrics.htm

Popham, W. J. (1997). What's Wrong—and What's Right—with Rubrics. *Conference on Educational Leadership - Schools as Safe Havens*, *55*, pp. 72-75.

Rubrics. (2012, September 25). *Rubrics.* Retrieved June 5th, 2013, from Institutional Assessment and Studies at the Univeristy of Virginia: http://avillage.web.virginia.edu/iaas/assess/tools/rubrics.shtm

Tierney, R., & Simon, M. (2004, 1 28). *What's wrong with rubrics: Focusing on the consistency of performance criteria across scale levels.* Retrieved 08 13, 2013, from Practical Assessment, Research and Evaluation: http://pareonline.net/getvn.asp?v=9&n=2

University, N. I. (n.d.). *Rubrics for Assessment.* Retrieved 08 13, 2013, from Northern Illinois University, Faculty Development and Instructional Design Center: http://www.niu.edu/facdev/resources/guide/assessment/rubrics_for_assessment.pdf

Copyright statement