GETTING TO YES - AGREEING RESEARCH PROJECT MARKS WITHOUT TEARS

Kathryn Davies
Unitec New Zealand
and
and Roger Birchmore
Unitec New Zealand

Abstract

One of the challenges facing the management of undergraduate research projects is achieving and maintaining consistency in the marking process. High staff turnover, the introduction of new academics to supervisory teams and the desire to benchmark internationally exacerbate the challenge.

The current assessment process within the Unitec Bachelor of Construction programme requires the student project first to be marked by the student's supervisor. This is then followed by second, independent marking of all the student assignments by an external academic. When significant variation of marks occurs and post-marking negotiation between the markers cannot achieve agreement, a third independent marker is utilised.

This paper outlines the development of an assessment rubric intended to provide clear standards and goals for both students and supervisors. The introduction of a rubric is intended to reduce the number of times significant variation in marks is experienced between markers. In cases where variation still occurs, the use of a rubric serves to define the problem and clarify the marking expectations. This assists with the negotiation process between the first and second markers, ideally removing the need for a third marker in most instances. In the most difficult of cases, negotiation will be required between three markers. Again, the use of the rubric allows a clear statement of the issues under discussion and the areas of divergence, allowing the participants to focus on reaching a satisfactory outcome.

Keywords: assessment, supervision, research project, rubric

Introduction

Unitec has its historical roots as a polytechnic and has grown quickly to the point where it delivers 21 undergraduate degrees. In order to successfully deliver degrees, Unitec's research culture has also had to develop rapidly. The School for the Built Environment (SoBE) has been successful in this process, with the current staff profile comprising 10 PhDs and 9 Masters, and with research activity and publication being transparently recognised in the workload model.

- The Bachelor of Construction offered by SoBE is a three year undergraduate degree. It comprises ten courses per year, each of twelve credits. The research project, seen as the culmination of a student's study, is twice the value of a standard course, with 24 credits. This course requires students to develop and justify a research question, analyse relevant literature, develop and defend a methodology, collect data, and analyse and document the results. The principal output of the course is a substantial written report which contributes 80% of the course assessment. Success in this course is considered to be a key demonstration of a student's suitability for graduation, and the course is a major contributor to the programme being internationally benchmarkable against programmes of longer duration and those that have honours components.

The BCons is well supported by students, with an increase in student numbers of 33% over the past 7 years. The growing number of students has produced an increasing supervisory load for the research project. The school staff has limited supervisory experience compared to the typical staff profile of a traditional university, and the team has had to confront a steep learning curve.

The research project course is intended to be student driven, but with supervisor guidance and direction assisting along the way. Supervisors mark the reports submitted by their own students. An external examiner provides the second mark for all of the reports, providing an independent view of the quality of the reports overall. Where the two marks differ by 10% or less, an average of the two is taken as the final mark for the student. Where variation greater than 10% occurs, the supervisor and external examiner are given the opportunity to discuss their marks and negotiate an agreed mark. If post-marking negotiation between the markers cannot achieve agreement, a third independent marker is utilised.

The Issues

From discussions with supervisors and markers at Unitec, a set of concerns has been identified around how the marking and negotiation process is taking place in the Research Project course. Some of these issues are focused on the initial marking stage, while others are more to do with the negotiation that takes place following a difference in marks. Some concerns fall into both categories.

Variability of marks

Abeysekera and Nummy (2006) outlines some of the issues surrounding the marking process at Unitec, particularly focusing on the variability of marks. In this work, a report reviewed by nine supervisors was given marks ranging from 41% to 91%. More recent marks in the Research Project course have shown differences less extreme than this example, but still significant. Last year, for example, 8 out of 26 report marks had a variation between the supervisor and external examiner's marks of greater than 20%. Only 12 out of the 26 report marks were within 10% of each other.

The likelihood of inconsistency in assessment increases as more supervisors become involved in the course (Pathirage et al, 2007: 272). Thus, the increased number of students undertaking the Unitec course, and the subsequent increase in the number of supervisors required, only serve to exacerbate the problem.

One of the conclusions of Abeysekera and Nummy (2006) was that further work should be conducted to clarify course expectations with the supervisors/markers, and to make marking requirements more explicit. As expressed by Pathirage et al (2007), "criteria designed carefully and used with clear procedures can reduce inconsistency in assessment"; while established criteria were in use at Unitec, they were considered unclear, and were applied inconsistently with markers assessing "against their own expectations." (Abeysekera & Nummy, 2006).

Non-expert supervisors

Each supervisor is able to supervise only a limited number of students. Where many students choose a similar area of interest (the current passion is around sustainability and Green Star rated buildings in particular), there are not enough supervisors with research interests aligned with this focus. Other staff will be drawn in to supervise projects that fall outside of their areas of expertise. This is not a problem confined to Unitec; Rowley & Slack (2004: p179) identify that "supervisors quite often find themselves supervising students working with topics in relation to which they can not offer specialist expertise."

In this situation, the role of the supervisor is to guide the student in the research process, and to assist them in finding appropriate sources of topic-specific advice, rather than to attempt to be the source of topic knowledge themselves. This has been a criticism repeated by the external examiner, that supervisors working outside their area of expertise struggle to give students appropriate guidance.

Informed supervisor/uninformed external examiner

Conversely to the above situation, there are occasions where the external examiner is operating from a position of relative ignorance about the topic whilst the supervisor has a greater level of

involvement and knowledge. It is impossible for one person to be an expert in the wide variety of topics pursued by students, but this is effectively what is asked of the external examiner.

Taking it personally

The process of negotiating with an "outsider", the second or third marker, is often seen as threatening by supervisors. A disagreement between parties as to the mark for a project may be seen by the supervisor as a personal attack. After eight months of working with the student, the supervisor sometimes feels that this is 'their' project that is being criticised. Alternatively, the view is that it is their ability as a supervisor or marker that is under the microscope.

Another personal aspect of the process is the relationship that develops between supervisor and student over the year. Supervisors in most cases invest a great deal of time and energy into assisting 'their' students through the process to a successful conclusion. In the marking and subsequent negotiation process they may see their role as getting the best mark possible for their student. In some instances the opposite is true, where a student-supervisor relationship has been particularly difficult or challenging and as a consequence the supervisor has a negative view of the resulting report.

It is interesting to note that of 26 reports last year, 20 were marked higher by the supervisor than by the external examiner. Only 3 were marked lower, and 3 reports were marked the same by the two markers. This weighting of higher marks by supervisors indicates an element of bias towards the student.

Process or performance vs. product

Related to the issue of the personal relationship between supervisor and student is the question of what exactly is being marked. The external examiner does not meet the students and sees none of the process they go through in developing their project. Thus it is quite clear that he is marking the product, i.e. the final report.

The supervisor, on the other hand, gains an appreciation of how the student tackles the project, the effort they put into it and their engagement with the course. It can be difficult to view the product separately from the process and the student's performance as a whole. This is illustrated by one supervisor who noted that the student's willingness to travel around the country to collect data, and to participate in related activities at their own expense, contributed to awarding a better mark that might otherwise be warranted.

Novice vs. expert

Supervisors in the research project course have experience in supervision from their own Masters or PhD study, and some have completed a course in Research Supervision through the Graduate Diploma of Higher Education offered by Unitec. On the whole however, the majority have had no instruction or guidance in the requirements of supervision, and are left to pick it up as they go along.

Supervisors who have little or no experience of supervision or marking often defer, either tacitly or explicitly, to the expertise of the second marker. Instead of defending their own judgement and negotiating an agreement, they accept the mark given. For example, a comment from a first time supervisor last year explicitly stated that due to their lack of experience in the process, they were happy to defer to the mark given by the supervisor. This means that the student is effectively marked by only one person, negating the moderating effect of the double marking process. In contrast, more experienced supervisors are prepared to enter into robust negotiation and their students generally end up with a higher mark.

Marking Process Development

Original marking schedule

The marking schedule previously used in the Unitec research project assessment, as reported by Abeysekera and Nummy (2006), is an analytical model framework (see Appendix A). This is a

prescriptive system using a set of general headings and sub headings to allocate marks in blocks of between 10 and 25 marks.

Because of the breakdown of marks in the different categories, markers could be within 1 mark on each category and still end up with more than a 10% variation in the final mark.

A major flaw in this marking schedule was the degree of overlap between the headings, with markers tending to reconsider weaknesses under several sections; for example, a flaw in the methodology would be penalised not only in that section, but also in the data and conclusions sections. The sub headings for the sections do not clarify what is to be considered against each heading, so the examiners are forced to apply their own interpretation.

Amended marking schedule

An interim step in the development of a revised marking schedule took the original schedule and added discussion and comments on what aspects should be considered in each section (See Appendix B). The overall sections and allocations of marks remained the same, but overlaps were made explicit, and more direction was given to markers on how marks should be allocated within sections. This revised marking schedule was not used formally for marking research reports, but was circulated to all supervisors. It was referred to in discussions about the marking process last year, but the marking schedule used was the same as in previous years.

Adoption of rubrics

The use of rubrics for improving assessment was emphasised by Maier (2006) in a presentation given at Unitec as part of a Teaching and Learning Seminar for the School of the Built Environment. This stimulated discussion and development about how they could be used to define and develop the marking regime for the Research Project course.

Rubrics were first introduced into the course last year, for marking interim presentations. Students have to present their work at three stages during the course of the project. At that time, each presentation was worth 5% of the final research project mark. Simple rubrics were introduced to improve consistency between panels of supervisors who were marking groups of presentations in isolation from each other. The rubrics were made available for students' reference before the presentation, and were used by the supervisors as marking sheets.

The rubrics contributed to an improvement in the quality of presentations and the thinking that went into them. Comments from both supervisors and students indicated that criticisms from panel of supervisors became more constructive as they were focused on the elements of the rubric. Students were conversely less defensive and more open to reflection on the comments given. Other changes in the presentation format were implemented at the same time so it is difficult to know how much to credit the effect of the rubric alone; however, students have been very positive in their feedback regarding how they use rubrics to focus their thinking and achieve good outcomes in their presentations.

Following the encouraging results from the presentation rubrics, a more extensive version was prepared for the report marking process (see Appendix C). This drew on aspects of the presentation rubrics as well as the interim marking schedule that had previously been used.

Woolf (2004) identifies three applications of the term 'assessment criteria' – to identify the parts of a work that are to be assessed; to establish the levels of performance that are required for different grades; and to specify the minimum acceptable standard for a pass. These correspond with the three essential components of a rubric identified by Popham (2000): evaluative criteria, quality definitions, and a scoring strategy.

The rubric developed here combines all three elements. It is first broken down into a set of six areas (Literature review and analysis, Research question, Research design/methodology, Data collection, Data analysis, discussion and conclusions, and Final report). These areas are then further defined in terms of what aspects of each are to be assessed. Marks are allocated to each component. Two different standards are given for each component: 'minimum', describing the least that a student can do for that section and still receive a pass mark; and 'excellent', to describe the level of work required for an A grade.

Popham (2000) describes two possible scoring strategies when using a rubric. One approach is to consider that the rubric provides guidance for a holistic marking, where the marker considers all of the aspects described in the rubric but does not use it directly to assign a mark, instead forming an impression of the student's work overall. Alternatively, a rubric can be used prescriptively to assign marks to each element identified, with the total of these giving the student's final mark. The latter approach has been adopted for the report marking described here, in order to maintain links with the previous system used. For the interim presentations described above, a holistic approach is used. In both instances, academics being what they are, there are individuals who will use the alternate approach. An example of this is seen in Abeysekera and Nummy (2006) where, despite the prescriptive nature of the marking schedule, one supervisor used the holistic approach to marking the report. Conversely, with the holistic rubric used in the interim presentations, one or two supervisors allocate points based on each criterion and use these to determine the final mark. Despite these variations in application, a consistent and detailed marking schedule should still serve to limit variation in marks.

Getting To Yes

The rubric provides improved clarity in expectations and specific descriptions of characteristics of a report to be marked. As a result, this should reduce the variability in marks coming out of the initial marking process. It should also help in issues concerning what is being marked, by focusing the markers' attention firmly on the product, rather than on the process.

The rubric is not just a tool to reduce variability in the making process. One advantage of using the rubric as a reference point throughout the course is to firmly focus the attention of the supervisor on the process aspects of the project, which form the core intentions of the research project course. While there will always be a need for some engagement with the specific topic chosen by the student, the supervisor's role falls largely in the monitoring and guidance in the research process, rather than the ins and outs of the specific research question and the data collected. The course is intended to judge the student's ability to carry out an independent piece of research, not on the value or otherwise of the findings that result. The rubric makes this clear, and serves to guide the contribution of the supervisor throughout the supervision process. This particularly supports new or inexperienced supervisors by making explicit the content and standards expected in the report.

An additional benefit, as seen in the adoption of rubrics for the interim presentations, is that the rubric provides guidance to the students as to how to approach the task and where to concentrate their efforts.

Despite this range of anticipated benefits, adopting the rubric can not be expected to solve all of the problems encountered in the marking process. Circumstances will still arise where differences between supervisor and external examiner marks will require negotiation.

Getting to Yes: Negotiating an agreement without giving in (Fisher, Ury & Patton, 1991) is a well respected "how to" guide built on experience at Harvard. Its overall tenet is to define the negotiation process into four components:

> People: Separate the people from the problem.
> Interest: Focus on interests not positions.
> Options: Generate a variety of possibilities before deciding what to do.
> Criteria: Insist that the result be based on some objective standard.
>
> (Fisher, Ury & Patton, 1991: p11)

The use of a rubric for marking the research projects assists in the negotiation process in the first and fourth of these components.

Adopting a rubric helps to remove the focus from the people, and refocus on the problem. The problem is identified quite clearly in the elements of the rubric as being the research report; it is not the student, the supervision process or the student-supervisor relationship. The rubric helps to deflect any criticism of supervisor shortcomings (or perceptions of criticism) by making clear the terms of reference for the marking and negotiating exercise.

This connects with the fourth component of the Getting to Yes approach, using an objective standard as the basis for negotiation. The rubric makes explicit the standard expected for a minimum pass and an excellence (A grade). Thus any differences in marks between supervisor and external examiner can be explored in the context of the descriptions provided.

For inexperienced supervisors, on in cases where either marker has limited expertise in the topic of the report under scrutiny, the definition of objective standards supports the negotiation process, and diffuses challenges to the knowledge or experience of the markers.

Future Work

The definitions of terms contained in the rubric have yet to be discussed in detail. Webster et. al. (2000) identifies that one of the crucial factors in using published criteria is to ensure that terms used are objective and unambiguous. While this has been kept in mind during the development of the rubric, it is a difficult thing to achieve. Expressions such as “interesting and engaging description of research”, “persuasive explanations” or “appropriate language” describe the essence of what is required but unfortunately have many potentially different connotations or interpretations. To reduce the length of time spent in writing and agreeing every phrase used in the rubric, a discussion is scheduled with all markers, including the external examiner, at a mid-point of the course. This will allow such terms to be debated and defined, to promote a common understanding of any that may cause confusion.

As well as promoting involvement of all participants in the development of marking criteria, Saunders and Davis (1998) goes further to recommend that this be repeated over time. They suggest that this is important both to involve new supervisors in the process, but also to benchmark understanding and application of the criteria to ensure they are not altering over time. While staff participation in research project course workshops have always been required, this has previously focused on student presentations and activities. This marking development exercise has highlighted the need for ongoing staff workshops to maintain standards and skills in the supervision and marking process.

Conclusions

As a marking schedule introduced for use this year, the rubric described here has not yet been tested. Some of the benefits expected may not eventuate. The terms and descriptions used may not prove as objective as intended, or terms may require further discussion and clarification.

It is likely that further refinement will be necessary to establish the rubric as a useful tool for marking the research project. It is evident, however, that the discussion it has prompted around how to approach the marking exercise has helped to clarify and make explicit many assumptions and expectations that have not previously been explored or challenged. In this respect the development rubric has already contributed to improvement of the course for supervisors and students.

References

Abeysekera, V. & Nummy, D. (2006). A critical incident approach for the improvement of an assessment regime for undergraduate research reports. In G. Runeson & R. Best (Eds) *Proceedings of the 31st Australasian University Building Educators Conference*, University of Technology Sydney, 12-14 July, Sydney, Australia.

Fisher, R., Ury, W. & Patton, B. (1991). *Getting to Yes: negotiating an agreement without giving in, 2nd edition.* London: Century Business.

Maier, H. (2006, August 25). *Some features of good assessment.* Presentation to School of Built Environment, Unitec New Zealand.

Pathirage, C., Haigh, R., Amaratunga, D. & Baldry, D. (2007). Enhancing the quality and consistency of undergraduate dissertation assessment. *Quality Assurance in Education, 15*(3), pp271-286.

Popham, W.J. (2000) *Modern educational measurement: practical guidelines for educational leaders, 3rd Edition.* Boston : Allyn and Bacon.
Rowley, J. & Slack, F. (2004). What is the future for undergraduate dissertations? *Education + Training, 46*(4), pp176-181.
Saunders, M.N.K. & Davis, S. M. (1998). The use of assessment criteria to ensure consistency of marking: some implications for good practice. *Quality Assurance in Education, 6*(3), pp162-171.
Webster, F., Pepper, D. & Jenkins, A. (2000). Assessing the undergraduate dissertation. *Assessment & Evaluation in Higher Education 25*(1), pp71-80.
Woolf, H. (2004). Assessment criteria: Reflections on current practices. *Assessment & Evaluation in Higher Education, 29*(4), pp 479-493.

APPENDIX A: Original Marking Schedule

Supervisor:

Unitec NEW ZEALAND

School of the Built Environment
Research Project 2006

RESEARCH PROJECT
MARKING SCHEDULE

NAME _____ **Student ID** _____

No	Assessment event	Mark	Out of
1.	• **Literature Review and Analysis** Relevance Suitability Extent		25
2	• **Research Question** Relationship to Literature Context Currency		10
3	• **Methodology** Relevancy to Research Question Description Defence		20
4	• **Data** Collection Analysis Results		20
5	• **Discussion and Conclusions** Significance of Findings Relationship to Literature		15
6	• **Final Report** Abstract Structure and Appearance Referencing Readability / Logical Development		10
			100 Scale / 80
	• **Oral Presentations** 2nd Block Course - General Description of Research, Methodological Approach, Annotated Bibliography. 3rd Block Course - Research Question, Methodology, Literature Analysis. **Completed** Formal Research Proposal and Ethics Approval 4th Block Course - Research Results and Preliminary Findings • **Research Process, Time management, Block Course Attendance** -defined and met agreed deadlines. [Note 0% will be awarded if you DO NOT attend every Block course		
	Final Mark		100

APPENDIX B: Amended Marking Schedule

Supervisor:

School of the Built Environment
Research Project 2007

RESEARCH PROJECT
MARKING SCHEDULE

NAME _____ Student ID _____

No	Assessment event	Mark	Out of
1.	**Literature Review and Analysis** Extent: a range of approx 20 sources , including seminal texts and some journals as a minimum. Out of 10 Relevance: Sources match the theme being explored 5 Suitability: Sources are credible. (academic texts, journals, reputable websites, Beware popularist books, personal websites . Themes, trends, conflicting views from the literature are made explicit. Students give their view of what the literature says. Avoid penalising the student again for weaknesses in the extent or suitability Out of 10 *(the breakdown of marks into sections prevents a student referencing a wide range of low quality references and still getting a good mark, or failing to analyse)*		25
2	**Research Question** Relationship to Literature: The themes in the literature build logically to a question that is clearly described. Avoid penalising the student again for weaknesses previous section. Out of 4 Context. The question is sensible,. The question is not trivial or filling a gap in the students knowledge that should have been addressed by another course in the programme Out of 3 Currency: Not answered before and worth answering. Out of 3		10
3	**Methodology** Relevancy to Research Question: Quantitative methods are used to answer quantitative questions, qualitative methods are used if opinions are being sought. Out of 5 Description: The description is unambiguous and complete. Out of 10 Defence: There is an explanation of why the method is suitable (been used before by….) Out of 5		20
4	**Data** Collection: Methods are described and appropriate and do not contribute unconscious bias. Ethics procedure complete and followed. Results are collected according to methodology and are complete. Questionnaire or equivalent is included. Sanitised data is included. Out of 10 Results: presented clearly, graphically where appropriate. Out of 5 Analysis: Relevant themes, trends and contradictions in the actual data collected are identified. Avoid penalising student again for weaknesses in methodology Out of 5 *(breakdown recognises a combination of the volume of the effort and the importance of the effort)*		20
5	**Discussion and Conclusions** Significance of Findings: Has the Hypothesis been proved or question answered? Insignificant findings can also be OK if actually identified-requires flaws to be explained and solutions proposed. Out of 10 Relationship to Literature: Does this confirm themes, tends identified previously or contradict them. If so why ? Out of 5		15
6	**Final Report** Abstract :accurate summary includes results and conclusions. Out of 2 Structure and Appearance: Complies with formatting reqts, spell checked grammar checked Out of 3 Referencing. Complete accurate and according to format. Out of 3 Readability / Logical Development: Fluent, reader follows theme Out of 2		10
			100 Scale / 80