Experience and Reflection on the use of Rubrics for Senior Capstone Design Course Assessment

R.W. Brennan, T. Freiheit, T.I. Khan, and S. Spiewak
Department of Mechanical and Manufacturing Engineering
University of Calgary, 2500 University Dr. N.W., Calgary, AB, T2N 1N4
Email: rbrennan@ucalgary.ca

Abstract

In this paper we report on our experience with the use of rubrics for the assessment of senior capstone design reports and presentations. The motivation for this approach to assessment was to create tools that have clear requirements, are clearly linked to the course objectives, are flexible, and are fair. The paper provides a summary of the main rubrics used for assessment as well as an analysis of their use both in the form of student feedback and the authors’ general impressions.

1. Introduction

Two key aspects of senior capstone design courses in CEAB (Canadian Engineering Accreditation Board) accredited engineering programs are teamwork and project management; closely related to this is the need for student design teams to be able to effectively communicate the analysis, interpretation and decision-making process followed during the course of an open-ended design project to faculty, sponsors, and their extended team [4]. As a result, oral and written communication is paramount to successful student design projects and is central to student team assessment.

This central role of communication in the capstone design course sets it apart from other courses in the engineering curriculum, but also presents a number of challenges to teaching faculty, the foremost being assessment.

One might ask why assess project work at all? For example, in most capstone design courses the project advisor and in some cases the project sponsor or “customer” serves as the examiners. Each of these individuals typically has intimate knowledge of the project and the team’s performance and could make a good argument that it is a foregone conclusion.

As Powell notes [8], there are a number of substantial challenges to this line of thinking. In particular, it is tricky being facilitator and judge at the same time. To address this, it is common to use a report or series of reports and design reviews for assessment [5]. The key advantage to this approach (over the “foregone conclusion”) is that the student team has the opportunity to defend their decisions and learn from their mistakes.

The challenge is to create assessment tools that, (1) have clear requirements, (2) are clearly linked to the course objectives, (3) are flexible, and (4) are fair. First and foremost, student teams must have a clear understanding of the expectations of the assessment. Well-formulated requirements for team reports and presentations can help here, though they must be linked with the course objectives and provide enough flexibility to accommodate a wide range of projects to be truly effective. For example, if the assessment tools are not clearly linked to the course objectives, students may feel that the course is more about writing and presentations than about actual “design”. As well, they may feel “boxed-in” by the requirements if they are not flexible enough to allow the team to explore a creative solution to their problem.

In a capstone design course it is particularly important that the assessment is clearly linked to the course objectives. As noted by Hansen [4], students tend to develop their study strategy according to the type of assessment they are facing. As a result, a “hidden curriculum” may result: i.e., a set of unspoken rules that students follow to get higher grades. As design instructors, we hope that our assessments point to the intended competencies of a well-structured design curriculum (e.g., [1]), rather than just those competencies necessary for passing an exam [4].

Given a well-formulated set of requirements for student assessments the question still remains, how are these requirements translated into a grade that is perceived as being fair and can be defended? This is particularly problematic for engineering faculty given the diverse and subjective nature of assessment.
Collaborative Learning Scoring Rubric

4 - Thorough Understanding
- Consistently and actively works toward group goals.
- Is sensitive to the feelings and learning needs of all group members.
- Willingly accepts and fulfills individual role within the group.
- Consistently and actively contributes knowledge, opinions, and skills.
- Values the knowledge, opinion and skills of all group members and encourages their contribution.
- Helps group identify necessary changes and encourages group action for change.

3 - Good Understanding
- Works toward group goals without prompting.
- Accepts and fulfills individual role within the group.
- Contributes knowledge, opinions, and skills without prompting.
- Shows sensitivity to the feelings of others.
- Willingly participates in needed changes.

2 - Satisfactory Understanding
- Works toward group goals with occasional prompting.
- Contributes to the group with occasional prompting.
- Shows sensitivity to the feelings of others.
- Participates in needed changes, with occasional prompting.

1 - Needs Improvement
- Works toward group goals only when prompted.
- Contributes to the group only when prompted.
- Needs occasional reminders to be sensitive to the feelings of others.
- Participates in needed changes when prompted and encouraged.

Figure 1. An example of a collaborative learning scoring rubric [10]

In order to address these challenges, the authors supplemented the assessment requirements documents used in their senior capstone design course with rubrics. A rubric can be thought of as a scoring guide that gives specific criteria for evaluating student performance. This tool is particularly useful for subjective assessments, allowing the instructor to explicitly link expectations with grades. As well, given that rubrics are typically available to students at the same time as the requirements documents, they set expectations in advance and consequently reduce “surprises” when grades are returned.

In this paper, we report on our experience with the use of rubrics for the assessment of senior capstone design reports and presentations. The paper begins with background on the use of rubrics for assessment. In section 3 we summarize the assessment structure used in the authors’ course. Next, we provide a description of some of the key rubrics used in the course and our experience with their use in section 4.

The paper concludes with a critique of the use of rubrics in senior capstone design courses as well as our thoughts on next steps.

2. What is a “rubric”?

A rubric is a “descriptive scoring scheme developed … to guide the analysis of the products or processes of students’ efforts” [2]. In other words, a rubric is a scoring guide, useful to both students and the graders, that gives specific criteria for evaluating student performance. In this section, we look more closely at rubrics in the context of design project assessment and contrast them with other forms of assessment.

2.1 Using rubrics for assessment

Figure 1 provides an example of a scoring rubric used for collaborative learning. This rubric uses four levels to guide both the student and the grader. In this case, each level describes a characteristic of the student’s collaborative work that would receive a respective score.

It should be noted that rubrics provide benefits to both the student and grader throughout the assessment process. For example, the rubric would typically be prepared in advance of the assignment’s due date, thereby providing students with a better idea of expectations. As well, at the time of grading there is a higher likelihood that multiple graders – typical in engineering design courses – will assign the same score to a given student. In other words, the “rater reliability” is increased by the use of a rubric [6].

Once the grading has been completed, the rubric also has the advantage of providing clear feedback to the student. For example, if the grader selected a “3 – Good Understanding”, the student can see immediately how her collaborative work compares with a “4 – Thorough Understanding”.

For written assignments, student feedback can also be provided with hand-written notes within the assignment. When coupled with the rubric, more individualized feedback is possible than just using the rubric by itself. This approach also has an added advantage over written feedback alone. In particular, the written feedback can now be linked with the original expectations as well as the resulting grade.
2.2 Alternate forms of assessment

Two alternative forms of assessment that are commonly used in engineering design courses are checklists and numerical weights. A checklist is quite similar to a scoring rubric in that a list of criteria to be met by the student is given. Unlike a scoring rubric however, checklists do not support the evaluation of the extent to which these criteria have been met [6].

This may be satisfactory in some cases; however a more common approach is to assign a numerical weight to each aspect of the assignment. For example, the “executive summary” of a written report is given a weight of 10/100. The difficulty with this approach is that a student who receives a 6/10, for example, may not know how to improve her work on the next assignment [6]. A rubric implicitly addresses this by giving the student a description of her grade as well as descriptions of higher grade levels. These descriptions alone and/or in combination with written comments may be used by the student to infer how her grade may be improved the next time.

In summary, rubrics provide two main benefits over standard forms of assessment [6]: (1) they support the examination of the extent to which the specified criteria has been reached, and (2) they provide feedback to students concerning how to improve their performance.

2.3 Types of rubrics

Depending on the purpose and characteristics of a given assessment, one may choose from several types of rubrics. In this section, we focus on four main types of scoring rubrics that have been used in classroom settings from K-12 to post-secondary.

The first distinction is between holistic scoring rubrics and analytic scoring rubrics. Figure 1 gives an example of a holistic scoring rubric. In this case, the entire performance is evaluated and scored as a whole. In other words, the assessment is not separated into independent factors and, as a result, there is overlap between criteria. For example, each category in Figure 1 encompasses factors such as “work towards team goals”, “sensitivity to others”, and “contribution”.

Alternatively, with an analytic scoring rubric, the student’s performance is evaluated and scored on several distinct criteria. Much like a checklist, an analytic scoring rubric allows for the separate evaluation of multiple criteria, each of which is scored on a different descriptive scale [2]. An example of an analytic scoring rubric developed by the authors will be described in section 4.

The second way in which scoring rubrics differ is in their approach to the learning task being assessed. More specifically, a rubric may be task specific or generic. A task specific scoring rubric is designed to evaluate student performance on a single assessed event [6]. For example, a scoring rubric that is developed to assess student teams’ written report for the conceptual design phase of a design project would fall into this category.

Alternatively, a generic scoring rubric may be applied to multiple assessed events. For example, a scoring rubric that is developed to assess student teams’ management of a design review meeting may be used for the design requirements, conceptual design, embodiment design, and verification reviews.

2.4 Words of caution

Before describing our work on the development of rubrics for capstone design courses, we will conclude this section with a few words of caution concerning the use of rubrics. When developing any assessment tool it is important to first determine the purpose and objectives of the assessment [7]. More specifically, the instructor should ask what competencies will be assessed (its purpose) and how students will demonstrate these competencies (its objectives). Only after the instructor has a clear notion of the assessment’s purpose and objectives can the criteria for the assessment be developed.

Moskal and Leydens [7] suggest a three step process, summarized in Figure 2 that concludes with an assessment of the final rubric. Of course, this process would typically be iterative in nature, and potentially extend past the implementation of the assessment itself (i.e., to improvements for the next time the course is run).

As noted in step 3 of this process, it is important to reflect on whether or not any of the scoring criteria are unrelated to the objectives. This is stressed by Stiggins [9], who notes that to develop sound performance criteria it is important “to describe levels of quality, not merely judge them”.

![Figure 2. Steps toward developing valid rubrics (adapted from [7])](image-url)
Finally, a detailed description of the development of rubrics is beyond the scope of this paper, however one may consult [6] and [11] for more comprehensive descriptions of the process.

3. Assessment of a senior capstone design course

In the remainder of this paper we describe our experiences using scoring rubrics to assess a final-year, capstone design course in mechanical and manufacturing engineering. In this section, we provide a brief overview of the assessment approach used in this course. For the design course, assessment rested primarily on student teams’ performance on preparing and defending a final design report. In order to split the workload over the fall and winter terms, the final design report was assessed in four interim reports with associated design reviews:

- Design Requirements Report (week 6),
- Conceptual Design Report (week 10),
- Design Embodiment Report (week 18), and
- Verification / Prototyping Report (i.e., “Final Design Report”) (week 26).

These interim reports were intended to summarize the project team’s work on the main stages of the product development process (i.e., planning, concept development, system-level and detailed design, and testing and refinement respectively). The deadlines associated with these interim reports, as well as the associated design reviews can be thought of as “gates” between the main stages of the product development process where student teams were encouraged to use the feedback from their reports and presentations to not only help in revising to their report, but also to help guide their work in the subsequent stages of the product development process. The relationship between the generic product development process and the main course deliverables is illustrated in Figure 3.

4. Applying rubrics in a senior capstone design course

In order to assess the various aspects of the capstone design course a variety of analytic scoring rubrics were used by the authors. As summarized in Table 1, both task specific and generic rubrics were developed, however the majority of the rubrics tended to be generic in nature.

As will be discussed in the next section, the rubrics evolved over the course of the fall and winter sessions. Of particular interest, a specialized form of the analytic scoring rubric was developed that included more detailed descriptions of the criteria used for assessment. At first, this modification was used to make it easier for the grader to obtain more information about each criteria (rather than referring to the requirements documents associated with each assessment), however as will be discussed later, this

<table>
<thead>
<tr>
<th>Assessment</th>
<th>No.</th>
<th>Holistic / Analytic</th>
<th>Generic / Task Specific</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reports 1-4 – Style</td>
<td>4</td>
<td>Analytic</td>
<td>Generic</td>
</tr>
<tr>
<td>Report 1 – Content</td>
<td>1</td>
<td>Analytic</td>
<td>Task Specific</td>
</tr>
<tr>
<td>Report 2 – Content</td>
<td>1</td>
<td>Analytic</td>
<td>Task Specific</td>
</tr>
<tr>
<td>Report 3 – Content</td>
<td>1</td>
<td>Analytic</td>
<td>Task Specific</td>
</tr>
<tr>
<td>Report 4 – Content</td>
<td>1</td>
<td>Analytic</td>
<td>Task Specific</td>
</tr>
<tr>
<td>Design Reviews</td>
<td>4</td>
<td>Analytic</td>
<td>Generic</td>
</tr>
<tr>
<td>Logbook</td>
<td>12</td>
<td>Analytic</td>
<td>Task Specific</td>
</tr>
<tr>
<td>Poster</td>
<td>1</td>
<td>Analytic</td>
<td>Task Specific</td>
</tr>
</tbody>
</table>

Figure 3. The main deliverables of the authors’ capstone design course
modification also proved to be useful for our students.

An example of a generic scoring rubric used in our course for assessing design reviews is shown in Figure 4. The modification described above is shown by the arrow in Figure 4. For example, if the grader or student requires more information about the requirements for “meeting documentation” they can click on “meeting documentation” in the main rubric and navigate automatically to a summary of the requirements.

The design review rubric shown in Figure 4 evolved from a generic project presentation rubric developed by one of the authors for teaching smaller classes (20-40 students). In this case, the rubric proved to be very useful for ensuring consistent grading across multiple graders: i.e., the instructor, one teaching assistant, and an additional faculty member. As well, it clearly stated the expectations for what was perceived by students as the most subjective part of their course assessment.

For the capstone design course, we found that this generic presentation rubric did not fit the course objectives for the design review very well. In particular, the design review covered generic presentation skills, but also required other management skills such as meeting preparation, organization, and conduct. The resulting rubric, shown in Figure 4, was still generic to the capstone design course, but unique in terms of its relationship to presentation rubrics used in typical engineering science courses.

5. Experience with the rubrics

The rubrics used in the design course evolved over the year with our experience. At the beginning of the year, the students were given a template explaining the requirement for the four design reports. This template not only provided the format for the report, but also provided content requirements for each progressive report, which culminated in a comprehensive final design report. Each report addressed a different phase of the design and development process, and the template cross-referenced each report to the appropriate material in the textbook. The first rubric had two separate sections covering style and content, with each criteria point weighted equally (with four marks each) on a four point scale – Very Good, Satisfactory, Marginal and Unsatisfactory. It was established prior to the start of class that the style section would have a declining weight with each progressive report.
5.1 Evolution of the design report rubric

The first report rubric had five criteria points each on content and style, and was developed as a means to grade the criteria given in the template, and didn’t provide a lot of detail about the criteria to the grader. It was judged not as successful as hoped, in part as it was expected that the grader (a teaching assistant) would educate himself on the referenced textbook material sufficiently to apply the criteria. As a result, it was determined that additional sub-criteria were needed to both aid the students and the grader in preparing and grading the report.

The second report rubric was expanded from two pages to nine pages, and expanded the content criteria to seven points, but left the style criteria the same (the style criteria remained the same for all subsequent reports). The additional pages covered the expanded sub-criteria, and in some senses was a high-level outline of the material delivered in the textbook and lectures. Some criteria were the same as the previous report, namely the ‘executive summary’ and ‘introduction’, as it was expected that the students would update these sections. The other criteria covered new material. Whereas it was judged that a better assessment of the reports was achieved by the addition of the sub-criteria, it was felt that the rubric was not giving the correct incentives to the students.

The third report rubric was improved by consolidating the ‘executive summary’ and ‘introduction’ criteria, to reduce their overall weight, and a criteria point grading specifically on succinctness was added. From the beginning, students had been requested to limit the new material added to each subsequent report to 10 pages. This was not always achieved, despite requesting that the students maintain a separate design binder where extraneous material and detailed calculations would be stored. The correction of the incentives through these changes was not entirely achieved, since the core of the third report, reporting on the detailed design, still had the same weight as the other criteria.

For the fourth and final report rubric, each criterion used the same four point scale as the earlier three reports, but the maximum marks for the criteria points were adjusted to change their relative weight. For example, ‘project management’ and ‘conclusions/recommendations’ were now weighted at two marks each, while the ‘overall project technical assessment’ was weighted at eight marks. It is believed that this final version of the rubric captures most of the lessons learned and will be used in the next year.

The experience gained in reaching this final report rubric influenced the rubric for grading the poster each project team created for the design fair. However, here a three point scale was used – above average, average, and marginal – with varied maximum marks to emphasize different criteria. The instructor team tried something new by, as a group, practicing the rubric’s application on posters created by prior year project teams. This helped to further reduce subjectivity by clarifying the meaning of the criteria, which was especially helpful for the engineers assessing the message delivery (artistic) criteria. Other rubrics utilized during the year were for the design logbooks, the design binder, and oral design reviews, and varied in their incorporation of the lessons learned depending on when they were created.

5.2 Opinions on grading with the rubrics

The teaching assistants generally liked using the rubrics because they felt its guidelines made the grading process easier or faster, but also had mixed feeling about them. One TA felt it was time consuming to continually match the grade to the grade description given in the rubric, stating “I could usually tell what grade to give, but was constantly re-reading the rubric to make sure.” They also agreed with the students (see section 5.5) that the rubric did not always accurately reflect the teams’ performance, as sometimes there was no match between elements in the rubric and the report, or it only reflected whether or not they read and followed the grading criteria. One TA commented after grading the second report “I think a lot of the reports were done to show that they ‘went through the motions’ and didn't necessarily reflect the teams actual process.” However, they felt they were able to adapt to these shortcomings.

5.3 Influence of rubrics on final grades

The rubrics should have improved grading fairness by reducing the influence of subjectivity in team grades. Final course grades for each individual were weighted 75% for team assignments (group reports, design reviews, etc.) and 25% for individual assignments (exams, assessment by their peers for their team contribution, project logbooks). Course instructors were responsible for reviewing the grades assigned by the TA’s for the first three reports, and for grading the design reviews and entire final report of the teams to which they acted as project advisors. Therefore it is expected that course instructors have an unusually large influence on the team grade.

To test the hypothesis that rubrics reduce subjectivity between instructors, the final team grades were compared. However, two elements may influence team grades, the strength of the students on the team, and whether the students were mechanical engineering or manufacturing engineering majors. The
latter concern stems from a historically lower grade point average held by manufacturing majors. Therefore, the final team grade was corrected to reflect the contribution of individual grades and majors.

A multi-variant linear regression analysis was performed on the entire class, using each student’s final team grade as the dependent variable, and a classification of the student’s major (1=mechanical, 2=manufacturing) and their individual grade as the two independent variables. This gives a model of the contribution of the two independent variables to team grades:

\[ \text{Team Grade} = a_1 + a_2(\text{Major Class}) + a_3(\text{Indiv. Grade}) \]

Each individual’s team grade was then corrected for the influence of these two independent factors by subtracting \( a_2(\text{major class})+a_3(\text{their individual grade}) \) from it. Note that these two factors may not be truly independent, but were treated as such. A difference in means with unequal variance test was then conducted on the average corrected team grades assigned by each instructor.

The means test found that one instructor’s average corrected team scores were 2.8 marks higher than each of the other two instructors, who differed by only 0.1 marks. This was at a p-value of \( 2.7 \times 10^{-7} \) or better under a hypothesis that there should be zero difference between the means. The difference at a 95% level of confidence between the corrected means of the instructors is 0.32 to 0.45 marks, which is within the margin used to round-up borderline student grades. This analysis indicates that the rubrics did not entirely reduce the subjectivity between instructors, but did help, as two of the instructors corrected grades were well matched. The shortcoming may have been because the one instructor was not as diligent in using the rubrics as the other instructors.

5.4 Student survey

In the middle of the second term, an anonymous, voluntary survey was conducted in the class to assess the students’ opinion of the use of rubrics for grading.

When the survey was administered, the students had received and used the rubrics for three reports and two presentations. Twenty-four students of a class size of 123 filled out the survey, giving a response rate of 19.5%. One student’s survey was removed from the data sample, as the student used it as an opportunity to vent his general frustrations about the class, which was clear from his written responses to the open ended questions.

The survey asked for both a quantitative and an open ended assessment of the course and rubrics. The quantitative assessment used a six point scale (1 = strongly agree, 2 = agree, 3 = somewhat agree, 4 = somewhat disagree, 5 = disagree, 6 = strongly disagree) and asked to evaluate the following seven statements:

- This was an excellent course
- I learned a lot in this course
- I really wanted to take this course
- I understand what a rubric is
- The rubrics used in this course are aligned with the course objectives
- I found the rubrics useful toward meeting the course objectives
- I feel that the rubrics used in this course were applied fairly.

The open ended questions asked:

- Complete the statement: I found the rubrics used in this course useful because…
- Complete the statement: The rubrics used in this course could be improved by …
- What do you enjoy about the course?
- What do you see as an area for improvement this semester?

All respondents answered the quantitative questions, while a large majority of the respondents answered the open ended questions.

**Quantitative analysis**

Table 2 summarizes the descriptive statistics of the quantitative assessment, while Figure 5 gives a histogram of their responses. It shows that, on average, the respondents were satisfied with or neutral about the course and rubrics, with a representative sample of respondents who felt both positive and negative about the course and rubrics. Because a correlation between the questions was expected, a factor analysis was performed on the correlation matrix of the responses [12]. Figure 6 presents a biplot of the correlation between the quantitative questions grouped into two factors, which can be described as ‘Opinion about the Course’ and ‘Opinion about the Rubrics’.

<table>
<thead>
<tr>
<th>Question</th>
<th>Mean</th>
<th>Stdev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent Course</td>
<td>4.1</td>
<td>0.98</td>
</tr>
<tr>
<td>Leaned A Lot</td>
<td>3.5</td>
<td>0.90</td>
</tr>
<tr>
<td>Wanted Course</td>
<td>3.5</td>
<td>0.85</td>
</tr>
<tr>
<td>Understand Rubric</td>
<td>2.7</td>
<td>1.06</td>
</tr>
<tr>
<td>Aligned Rubric</td>
<td>1.9</td>
<td>0.73</td>
</tr>
<tr>
<td>Useful Rubric</td>
<td>2.6</td>
<td>1.08</td>
</tr>
<tr>
<td>Fair Rubric</td>
<td>3.0</td>
<td>1.43</td>
</tr>
</tbody>
</table>
The biplot shows that two factors appear to be adequate to characterize the correlation between the quantitative questions. (A three factor analysis was also performed, but did not derive any additional insight). The fairly strong grouping of the questions along each axis indicates that the respondents’ feelings about the course and about the rubrics are relatively independent. The plot also shows, not surprisingly, that the questions about ‘want to take the course’ have a negative loading toward their ‘Opinion about the Course’ factor, and the questions on ‘excellent course’ and ‘learned a lot’ have a positive loading toward this factor. For the factor ‘Opinion about the Rubrics’, all questions have a positive loading. Interestingly, as will be made clearer from the analysis of the open ended comments, there is a minor contribution for the questions between the factors. For the rubrics questions, there is a minor contribution along the class opinion factor, with ‘understood the rubric’ loading positively, and ‘fair rubric’ loading negatively. For the class questions, there is also a minor contribution along the rubric opinion factor, with ‘excellent course’ contributing to the loading positively, and ‘learned a lot’ contributing to the loading negatively.

**Open-ended analysis**

The survey respondents, whether they agreed or disagreed in their quantitative assessment of the rubrics, generally had the same range of comments in their open ended assessment. Overwhelmingly, the respondents stated they liked the rubrics because they explicitly stated the requirements of the reports. On the other hand, they felt the rubrics could be improved by a more timely release, increased flexibility, and simplifying their language.

A common concern noted in the open response was that the rubrics could be improved by releasing them sooner before the report was due. This concern was a hold-over from the previous semester, when the course instructors released one rubric late, and was corrected for all future reports. For this late rubric, while the students did not have the specific grading weights for each section, and detailed comments about the expectations for the report, general report requirements had been provided to the students earlier. This experience highlights the importance students placed on the explicitness provided by the rubrics when composing their reports.

Another common concern was that students perceived the rubric as being a generic measure focused on the methodology of the course. They wanted a rubric that was tailored to the specifics of their project, and that better reflected their perception of whether a particular phase of the design and development process was appropriate to their project.
This resulted in the concern that the rubrics were not fairly applied because when a project team felt a section (design process step) requested in the rubric was not applicable (with students stating that they justified its exclusion in the report), they were penalized in their marks. This concern reflects a student misunderstanding of the course objective to teach a design methodology, and their perceived objective of “doing a project”. However, it does highlight a weakness in the rubric where if there truly is a justifiable reason to exclude a design phase due to the nature of the project, how does one grade for its absence when there is a line item in the rubric?

One respondent commented that the rubric did not give adequate feedback for improvement of their report. This is a valid point, in that the format of the rubric did not facilitate this adequately. This concern may have arisen from an experiment that was made concerning how the report was delivered to and graded by the instructors. One report was graded solely from an e-submission of the report. Therefore the grader did not have the opportunity to write comments directly next to the text in the report. Whereas there was limited space to put comments on the rubric assessment sheet returned to the students, comments tended to be general, summarizing an opinion of the report content, rather than specific points for report improvement.

Despite the overwhelmingly support of the rubrics for their explicit statement of requirements, two respondents commented that the requirements were not detailed enough, and should use a more direct and basic language. An example cited was the use of the word ‘succinctness’ as a heading for one rubric requirement. As fourth year college students, who presumably know how to use a dictionary, this comment is perplexing. However, these concerns can perhaps be interpreted as inconsistency in grading, with the students having trouble connecting the marks received back to the requirements, but the anonymity of the survey does not permit a test of this explanation.

One respondent felt that marks were taken off his report for style issues that were perceived as insignificant, and that the rubric did not explicitly state they were a requirement. The student may have overweighed this concern, feeling unduly penalized by the lost points. Students were not given the relative weight for ‘style’ and ‘content’ rubrics, where ‘style’ had significantly less weight, and decreased with each report. Withholding this information was a deliberate choice by the instructors, to emphasize that good communication also requires good technical writing skills, but may have created a perceived problem where none existed.

The last two concerns highlight the tension between explicitness and ambiguity required for subjective grading. Whereas the intent of the rubric is to reduce the level of subjectivity in grading an assignment that requires a subjective assessment, some error will always occur in the assignment of marks. Thus the rubric must have leeway around a grade classification which can account for this error, just as one sets upper and lower limits on what is an ‘A’ or a ‘B’. To facilitate this in practice, one suggestion is that each sub-criteria of the rubric should have enough points to distribute to allow this leeway within the criteria.

6. Conclusions and next Steps

We began this paper with the challenge of creating an assessment tool to reduce the subjectivity in grading a fourth year design course which would have clear requirements, was clearly linked to the course objectives, was flexible, and was fair. In using rubrics, we were only partially successful in this goal. Through the quantitative student survey, we know that on average the students agreed that the assessment was linked to the course goals, and somewhat agreed that the assessment was fair. Through an analysis of the evolution of the design report rubric, we believe that, in the end, a rubric was developed that was clearly linked to the objectives. However, feedback from the students and the teaching assistants indicate that we failed to provide a rubric that was flexible in accommodating a wide range of design projects.

How successful was the rubric in supporting the course objectives? A major course requirement was the development or exercise of effective team interaction and communication skills. The vehicle for delivering these objectives was team presentation and report writing, to which the rubric directly contributed. Through the reports, the rubric also supported the objectives of applying the knowledge of engineering to open-ended projects and creating a design to satisfy stated requirements. The rubric was less effective in supporting the evaluation of the students’ ability to understand, apply, and justify the steps of the design and development process, as well as evaluate and optimize the quality of a design solution.

The greatest shortcoming of the rubric, as it was implemented, was in supporting a flexible method to the assessment of design projects that vary in their content and approach. This was commented on by the students. It is a great challenge to devise a generic assessment tool that can account for project variation and be supplied prior to the submission of the reports. However, one can question if this assessment objective is wholly necessary. The course is intended to deliver to the students a design and development
methodology that can be applied to any design project, and not just concentrate on the needs of their particular course project. By not assessing each project on all methodology steps, we are not serving our students properly. To make up for this detachment between the course requirements and the details of an individual course project, the project advisor must help the students to interpret each step of the methodology in the context of their course project. For example, a project advisor may suggest to the project team that is designing a consumer good that their ‘cost analysis’ focus on the product cost, while a team developing a one-of-a-kind product for a university laboratory focus their ‘cost analysis’ on the project management budget. In practice, project advisors must be aware of this need, and interpret and discuss the differences between the rubric requirements and the projects well before reports are due.

One teaching assistant also felt that there may still be a shortcoming in aligning the rubrics to the course objectives. He posed a valid question in asking “what [is] the goal of writing the [design] reports; is it to demonstrate understanding of course concepts, to demonstrate the application of course concepts, or to communicate their design project status? The rubric should reflect the goal of the report writing.” This may be because all three of these goals are intertwined, with the ‘understanding of the course concepts’ assessed both in the class exams and as an element of the group reports. Would it help to categorize these points in the rubrics to make the course objectives more clear? Probably not, as it may cause the students to structure their reports in a less effective (and less interesting) manner.

Overall, it was judged by the instructor team that the rubrics were worthwhile. However, several suggestions to improve their format and implementation are being considered. First, appropriate marks must be assigned to each criterion to emphasize where students should spend the most effort. This will change from report to report as the course progresses through each phase of its design methodology. Second, rubric criteria and sub-criteria must be better defined to support the course objective of understanding, application, and justification of the steps of the design process. Third, one teaching assistant suggested that the rubric design should include a checklist of elements expected to be found in each section or criterion, rather than just a qualitative description for each quality level. This could improve the diligence in its usage, as now it is not just a subjective assessment of the quality of the content, but an enumeration of whether the criteria were met. This would also address the teaching assistant’s difficulty with “constantly re-reading the rubric to make sure” they had the right quality level. Finally, the issue of flexibility within the rubric must be addressed, through one or more changes: change the mechanics of marks assignment (ability to discount a section when it is not appropriate), change the criteria language to be even more generic, tailor the rubric to different design projects, and/or improve the interpretation of the rubrics through the project advisors. If these concerns can be effectively addressed, we believe that the approach that was taken in assessing our fourth year design course will provide a level of success that others may wish to adopt.

7. Acknowledgements

The authors wish to thank the Natural Sciences and Engineering Research Council of Canada and the University of Calgary, Faculty of Engineering for their generous support of this research. We also wish to thank our TA’s, Valerio Giuliani, Rohit Newaskar, Chris Regier, and Jason Scarlett for their input on using the Rubrics.

8. References