Benchmarking of final year Engineering projects

Peter Gibbings and Chris Snook
Faculty of Health, Engineering and Sciences, University of Southern Queensland
Corresponding Author Email: snook@usq.edu.au

Structured abstract

BACKGROUND
In countries that are signatories to the Washington Accord, the final year (often honours) projects in four-year engineering degree programs are considered the culmination of the students' undergraduate work and are used as one of the key sources of evidence of students' design ability. Not surprisingly there is a lot of variation in how these projects are conducted and assessed. What academia and accrediting bodies need is evidence of fundamental discipline-specific technical knowledge as well as professional skills. These competencies need to be demonstrated on an authentic realistic open-ended capstone project, providing useful feedback to the student and facilitating significant reflection. Also, this grading had to be carried out in a consistent, repeatable, and reliable manner.

PURPOSE
The aim of this research is to critically analyse the project assessment practice at an Engineering Faculty at a regional Australian university and assess this practice against the criteria above.

DESIGN/METHOD
Twelve mechanical engineering projects were selected from the 2011 cohort for analysis. Each of these projects was assessed independently by two senior academics of the discipline. A full rigorous assessment of the dissertation was made, with the assessment instrument completed and a final grade awarded. A senior member of the mechanical engineering discipline from another university assessed four of these projects. This assessment followed the USQ assessment process using the USQ assessment instrument so that the same elements were assessed and final grades could be compared. This provided cross-institutional benchmarking and feedback on the USQ process. Finally, an academic from another discipline at USQ, who was already familiar with the general assessment process, carried out full grading as well as order ranking of the 12 projects. This assessment did not focus on technical details, but was designed to provide valuable insights into the value of the assessment instrument. Each project had the original and several independently determined grades.

RESULTS
Use of the assessment instrument:
- Within the mechanical engineering discipline at USQ produced repeatable assessment grades.
- Within the mechanical engineering discipline across institutions produced acceptable results (within half a grade of the formal assessment).
- Across discipline boundaries at USQ produced assessment grades that followed the general trend of the formal assessments but with potential variation of one grade about this trend.

CONCLUSIONS
The several phases of the moderation practice make this process robust and significant outliers which may be presented from a single assessor are identified early and addressed at that time. The assessment instrument provided repeatable assessment grades across independent assessors within the same discipline. External benchmarking did not indicate any large discrepancy in assessment, although greater clarity and exemplars may be required in the assessment instrument and the expected level of achievement in individual elements in the rubric and also the overall grade. This paper highlighted the need to use project assessors who are well versed in the technical discipline area. Overall it can be concluded that the USQ holistic method of assessment of competencies and grading of the final year capstone project can be considered consistent, repeatable, and reliable.

KEYWORDS
Capstone project, assessment, benchmarking

Introduction

In countries that are signatories to the Washington Accord (International Engineering Alliance, n d), the final year (often honours) projects in four-year engineering degree programs are considered the culmination of the students' undergraduate work and are used as one of the key sources of evidence by professional accreditation bodies of students' design ability (Geske, 2009). In these capstone projects, students commonly work on one or more major projects providing the opportunity to improve professional skills and requiring them to demonstrate high level mastery and application of graduate capabilities, techniques and knowledge that are fundamental to a specific discipline (Geske, 2009). Importantly to accrediting bodies (particularly signatories to the Washington Accord) the capstone engineering projects provide evidence of the development and refinement of generic competencies required of engineering graduates (McCormack et al., 2011). These capstone projects are widely regarded as a necessary component in engineering programs, but they come with a unique set of challenges, particularly with respect to assessment of skills and capabilities.

Not surprisingly there is a lot of variation in how these projects are conducted. Howe & Wilbarger (2006) reported results from two broad studies (carried out in 1994 and again in 2005 in the United States of America) into capstone engineering projects. One consistent trend they reported was the tremendous variation between institutions, and even within the same institution. A common commitment in these projects seems to be that students must creatively analyse, synthesise, and apply knowledge gained from prior learning as well as gaining additional skills and knowledge required to complete their projects (Brackin, Knudson, Nassersharif, & O'Bannon, 2011). Projects that focus on design aspects require students to make empirically-based judgments based on a broad range of knowledge, skills, and experience (Steiner, Kanai, Cheng, Alben, & Gerhardt, 2011). Academics face a challenge to design pedagogically sound capstone project courses that are open-ended enough to provide a valuable educational experience for the students (Steiner, et al., 2011), challenging enough to support development of design skills, and rigorous enough to satisfy accreditation requirements of universities and professional accrediting bodies. Making objective judgements on the design of these projects is academically challenging.

The challenges faced by academics (and students for that matter) in the open-ended nature of the design projects are compounded by the challenges for academics in the final grading of the projects (Steiner, et al., 2011). Although these authors were specifically focussing on multidisciplinary team-based projects, their theory on taking a holistic approach to evaluation, assessment and feedback are relevant to all such projects. Steiner, et al. (2011) report a dichotomy in terms of the project supervisors' role: on the one hand they have to act as a coach and mentor to support the student, but on the other hand they need to monitor performance and grade the projects. To ensure a consistent approach to assessment, many authors (for example McCormack, et al., 2011; Sobek & Jain, 2004; Steiner, et al., 2011) recommend the judicious use of a marking rubric to evaluate student performance in design projects based on professional engineering expectations. In most cases the marking rubric is used for assessment of learning outcomes as well as providing feedback to the students. When these instruments were used the results indicated that they were considered useful by students and staff, and most importantly, research reports they lead to sufficient scoring consistency to be used reliably for assessment purposes, which resulted in a high degree of consistency in grading (McCormack, et al., 2011; Steiner, et al., 2011).

The capstone project addresses professional development as well as technical design competence. As a consequence, rather than just assessing technical skills, Howe & Wilbarger (2006) reported a leaning towards the development and assessment of professional skills, and it seems this is still the case with a significant number of projects in the USA being sponsored by industry (Brackin, et al., 2011). The professional real-life nature of these projects is important, but one element to consider with the assessment of these

projects is what to consider a pass and a fail grade. From industry perspective, this would be judged on the final outcome, however, as identified by Brackin, et al (2011) failure of the final design may not indicate a failure to learn. They note many instances where the fact that students tried something and failed taught them the importance of certain elements in the design process. Sobek and Jain (2004) who also question whether the quality of a design process can adequately be judged by the adequacy of its outcomes. This is supported in general educational literature, particularly when students are given the opportunity to reflect on their mistakes and rework (though this is not often possible in a capstone design project).

So, in summary, what is needed for both academia and accrediting bodies is evidence of fundamental discipline-specific technical knowledge as well as professional skills. These competencies need to be demonstrated on an authentic realistic open-ended capstone project. And the competencies need to be assessed in a holistic manner providing useful feedback to the student, facilitating significant reflection on the part of the student, and this grading has to be carried out in a consistent, repeatable, and reliable manner. How hard can that be?

It should not be a surprise that a study in 2006 revealed a lack of confidence of supervisors in assessing student performance in capstone design projects (Davis, Beyerlein, Trevisan, Thompson, & Harrison, 2006). They point to the lack of clear defensible framework for the assessment of these types of projects. Rather than simply assessing performance, whether using a rubric or not, at the completion of the project, they recommend assessment of performance repeatedly throughout the project (Davis, et al., 2006). Beyerlein, Davis, Trevisan, Thompson, and Harrison (2006) describe the development and implementation of such an assessment framework and associated assessment instruments which provide alignment of learning outcomes, assessment of performance, and student feedback on their learning in those objectives (Beyerlein, et al., 2006).

Geske (2009) also suggests assessment in phases, at key milestones, as the project progresses noting that students will concentrate on those things that are assessed and will expend highest energy on elements at are awarded the highest marks. Such assessment should incorporate clear student feedback on the strengths and weaknesses at each milestone. Three (of their four) elements of the assessment are: clear student expectations (so they are clear on what is expected), consistent evaluation (though concerned mainly with team work this is still important for individuals), and uniform assessment.

Establishing clear student expectation is important from an assessment perspective. Unfortunately, capstone projects tend to be poorly defined and the problem statement is vague or ambiguous, but this is to be expected if students are to be given a realistic project experience (Geske, 2009). As the design process evolves, requirements become clearer, but this provides a significant learning experience for the students. Rather than assess outcomes, Geske (2009) recommends grading of various components based on student behaviours (example in this publication allows from 5 to 15 % for intangibles such as professionalism and effort). It is therefore considered important to provide students with past examples of student work to illustrate what is an acceptable standard.

A study with responses from 119 engineering faculties in the USA reported that at that time the project criteria required by the relevant accreditation agency were not well assessed in capstone design courses (McKenzie, Trevisan, Davis, & Beyerlein, 2004). Of particular concern was that the assessment practices used may lead to the mis-measurement of student achievement (McKenzie, et al., 2004) and false and inaccurate conclusions about the quality of the end design (Sobek & Jain, 2004). In 2004, the assessment of capstone projects was considered as perhaps among the most under-researched topic in engineering education at that time (Sobek & Jain, 2004).

Capstone project assessment was considered so important in the USA that a conference was held in 2007 to share experiences and improve capstone design and assessment with a follow up conference held in 2010 (Howe, Goldberg, Palo, & Terpenny, 2011). Clearly

academics need a user-friendly, consistent, reliable, repeatable, and valid method of assessing these widely disparate and unique design projects (Geske, 2009).

The aim of this research is to critically analyse the current project assessment practice at an Engineering Faculty at a small regional university in Australia to see how it measured up against the criteria established above. Cross-institutional benchmarking is included in this evaluation to provide feedback on both the assessment process and the standard of student work. Analysis of the usefulness of feedback to the students is beyond the scope of this paper.

Background and current process

The newly formed Faculty of Health, Engineering and Sciences (HES) has responsibility for delivery of engineering programs at the University of Southern Queensland (USQ). At USQ students may elect to study in the on-campus (internal) or off-campus (distance or external) modes. Approximately 75% of the 2,600 engineering and surveying students in HES study by distance education from various geographic locations around the world. Engineering and surveying students in HES may be studying any of nine majors: agricultural, civil, computing/software, environmental, electrical/electronic, mechanical, mechatronic, surveying, and GIS; and can complete an Associate Degree (2-year), Bachelor of Engineering Technology (3-year), or Bachelor of Engineering Degree with Honours (4-year). Although the two and three year degrees have capstone projects, in this study we restrict our attention to the assessment of the capstone project in the Bachelor of Engineering Degree with Honours. To reduce the scale of the project we only concentrate on the Mechanical and Mechatronic (MEC) disciplines. Since the assessment process is the same for all majors, we consider that the findings are representative of the wider cohort.

Within the project courses, students are provided clear instructions and assessment criteria for their projects and they are provided past examples of what is an acceptable standard. Each student carries out an individual project, which is assessed in phases. Assessment during the project is formative and designed to keep the students on schedule. Final assessment is summative and is carried out with the assistance of a mature instrument (including a detailed rubric) that provides considerable feedback to the student.

After completion, project supervisors carry out a detailed assessment on the submitted project and supporting documentation. Part of this assessment is the awarding of a grade. Strictly, this grade only needs to be on a scale F (fail), C (pass), B (credit), A (distinction), and HD (high distinction), but from a practical perspective and to assist moderation, supervisors add a plus and minus grade to all but the fails to differentiate at a finer level. For the purposes of this research, each grade is represented on a scale to 15 as indicated in the following table.

Table 1: Numeric and Alpha Grades

F	C-	C	C+	B-	B	B+	A-	A	A+	HD-	HD	HD+
3	4	5	6	7	8	9	10	11	12	13	14	15

In addition students can receive an IM grade (Incomplete requiring Make-up work), which means they need to complete significant remedial work on their project. If this is completed in the time given they receive either an F or a C grade depending on the quality of the final product. For this reason the IM grade does not appear in the above table.

Following this initial assessment, the projects are blind marked by a second staff member from the same discipline. The two grades are then moderated by the course moderator and any large anomalies are discussed and resolved with the two assessors. All projects for all disciplines are then presented in an open forum for further moderation. This involves members of each discipline commenting on anomalies and making recommendations on grades with justifications. Large discrepancies between the first two assessors are picked up

and dealt with in a diplomatic manner using this process. It also involves cross-disciplinary moderation to ensure consistency of grading (a B grade in one discipline is equivalent to a B grade in another discipline).

Method

From the projects completed and assessed in 2011, 12 MEC projects were selected for further analysis that covered a range of grades from IM to HD. For this research, the IM grades were considered equivalent to a fail grade (even though there was a chance to upgrade) and equated to a 3 in Table 1. Each of the 12 projects was assessed independently by two markers who are senior members of the discipline (C and B in Table 2). Some of the projects were assessed by a further two members of the discipline (S and A in Table 2). A full rigorous assessment of the dissertation was made, with the assessment instrument completed and a final grade awarded (but no student feedback or notes were required for obvious reasons). The assessments were completely independent and no correspondence was entered into between the markers, nor between markers and the initial assessor or co-assessor of the project. If the marker was the original assessor of the project, then they were not required to reassess that project. Consequently, each project has two to four independent grades including the original assessment grade (the original assessment mark is noted as USQ in Table 2).

In addition, a senior member of the MEC discipline from another university was asked to assess four of the projects covering the full range of grades. This assessment followed the USQ assessment process using the USQ assessment instrument so that the same elements were assessed and final grades could be compared. This provided cross-institutional benchmarking and feedback on the USQ process.

Finally, an academic from another discipline at USQ carried out full grading as well as order ranking of the 12 projects (this staff member was familiar with the general assessment process at USQ). Logically this assessment could not have anything to do with technical details, was designed to provide valuable insights into the value of the assessment instrument as well as providing guidance for any possible changes to the assessment process.

Results

In Table 2, the ‘Project’ column represents the different student project unique identifiers; USQ represents the official final grade issued for the project (note these official grades are ‘rounded’ to not include plus and minus granularity); and the last four columns represent the first initial of the four USQ assessors as described in the Method section.

Each of these grades from the four independent assessors were assigned a representative number based on Table 1. The grades for each student were then averaged and used as the standard of comparison. The grade circled with a full line is clearly an outlier. We argue that because this would have been identified and amended during the rigorous moderation process explained earlier, we are justified in leaving this grade out of the calculation of average grades. Nevertheless, we still leave this data point in for the graphing and analysis. The second potential outlier shown as a dashed circle has been left in the calculation of average scores since it would be been left to moderation to decide how to deal with this.

1. Comparison within Discipline and within Faculty

Figure 1 depicts the USQ and independent assessor grades for the 12 dissertations under consideration in the present investigation. The horizontal axis represents the twelve projects sorted in ascending order by average grade point. The vertical axis indicates the grade allocated by the assessors as noted.

Table 2: Independent grading of projects

Project	USQ	C	B	S	A
11_M06	C	C-	C-	IM	
11_M11	C	C-	C-		
11_M02	C	C	B-	C+	A+
11_M08	B	C+	B-		
11_M04	B	B	B+	B+	C
11_M09	B	A	B+		
11_M07	A	A-	B+		
11_M10	A	A-	A		
11_M01	A	HD-	A	A+	HD-
11_M05	HD	A-	A+	A+	
11_M03	HD	A+	HD-	HD	HD-
11_M12	HD	HD	HD		

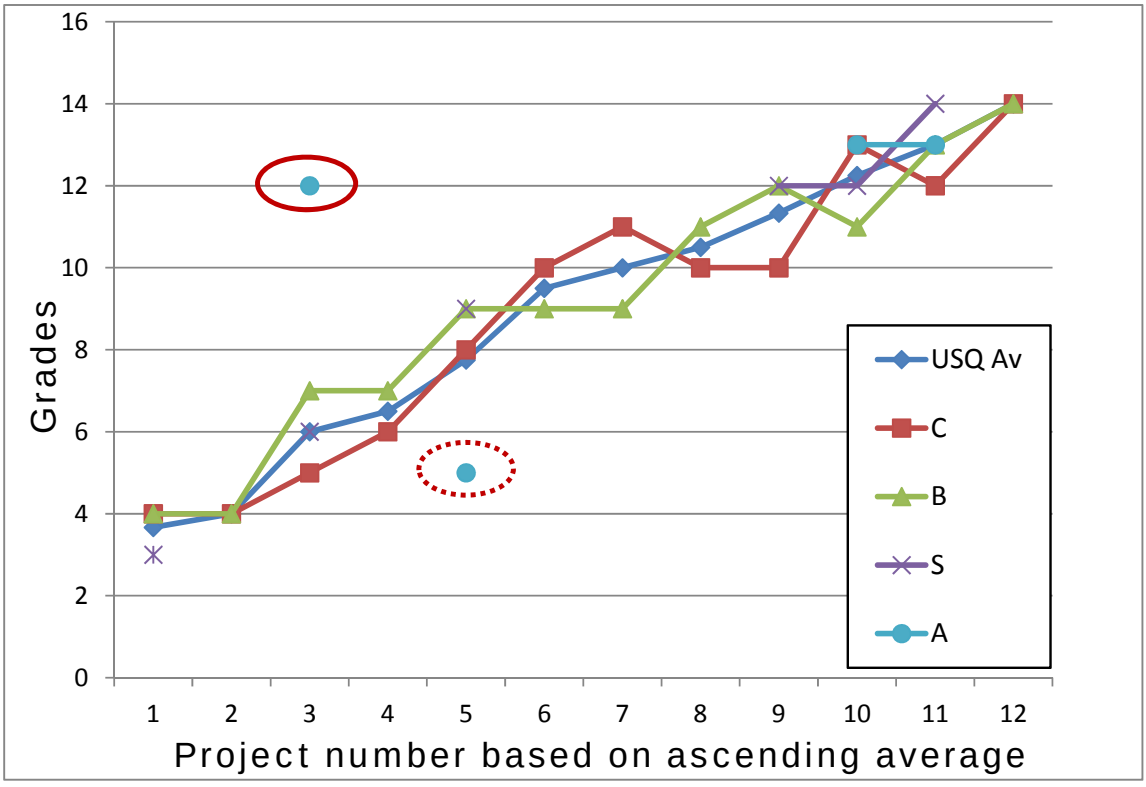

Figure 1: Comparison within discipline and within faculty

The two anomalies noted earlier and circled in Table 2, are identified in the same manner in Figure 1. Note that a spread of 3 points on the vertical axis represents a full grade change. With this in mind, and given that the moderation process would have identified and dealt with the outliers we can say in general the assessment instrument provided repeatable assessment grades across the independent assessors.

2. Comparison within discipline and outside faculty

It should be noted that the academic staff member from outside of USQ was not familiar with the assessment instrument and assessment process used at USQ, but were from the same discipline area (Mechanical). Further, they did not spend the same amount of time on the assessment task. Combined with the fact there are only four data points in Figure 2 it is hard to be definitive, but we can conclude that using our assessment instrument, staff from outside the Faculty in general marked around half a grade lower. From an external benchmarking perspective this is quite acceptable. Most likely this small difference can be attributed to some lack of clarity in the assessment instrument and the level of achievement that is generally accepted. This expected level of achievement may be discussed informally in HES and become a cultural norm, which has not been captured fully in the assessment process. In addition, there is some staff "calibration" of expectations as they are encouraged to be involved in the cross-moderation process each year. Exemplars to demonstrate expected levels of achievement and more clarity in the rubric on expected outcomes for specific grades may be a helpful addition, particularly for guidance of new staff members.

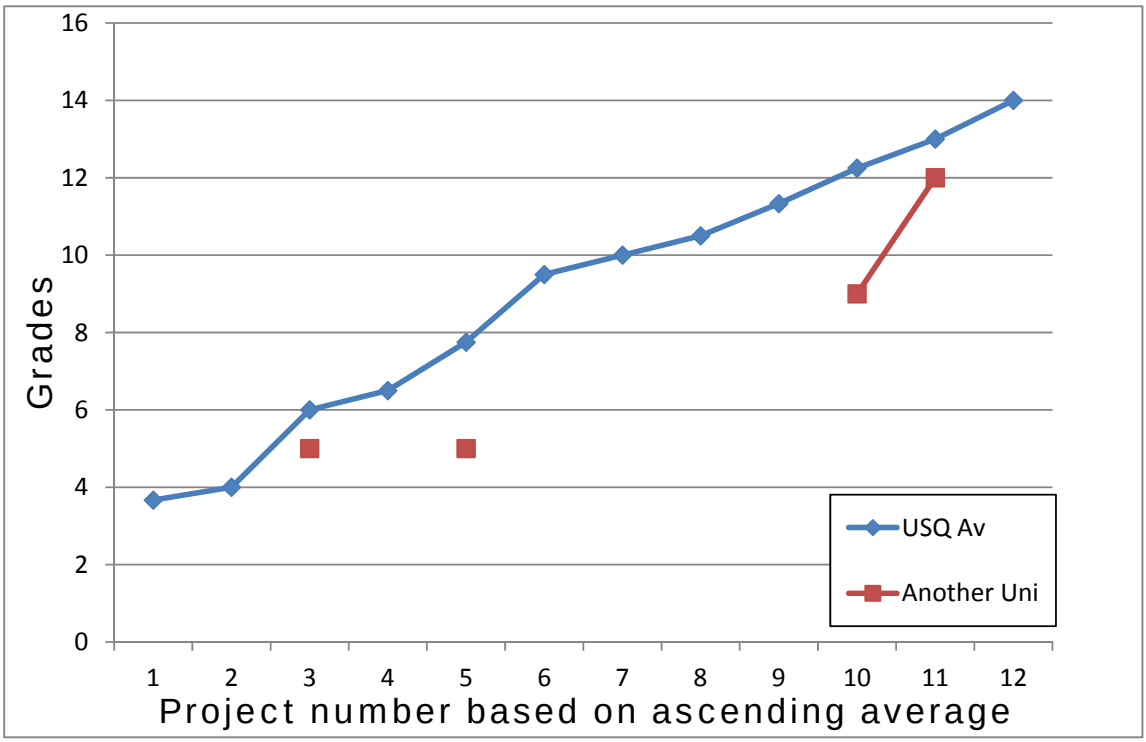

Figure 2: Comparison within discipline and outside faculty

Although we have no data to verify this, the slight difference between the USQ and external grades may indicate that the USQ assessment instrument is not weighting the types of competencies this academic staff member is used to focussing on this their own institution. It may therefore be prudent to benchmark the key elements being assessed in the USQ instrument against professional competencies to ensure the assessment is aligned to expectations of professional bodies at a national and international level.

3. Comparison outside discipline and within faculty

Figure 3 clearly demonstrates the need to have staff members who are well versed in the discipline area to assess projects. Whilst the general trend is reasonable, there is a substantial difference from the average grade in most projects. This may again tell us that the general criteria on the assessment instrument need some exemplars, but that may only benefit staff who have sufficient knowledge in the discipline.

Conclusion

It is clear that the several phases of the moderation practice make this assessment process robust and those significant outliers which may be presented from a single assessor are identified early and not allowed to progress through the remainder of the assessment process. The assessment instrument provided repeatable assessment grades across independent assessors within the same discipline. External benchmarking did not indicate any large discrepancy in assessment, although greater clarity and exemplars may be required in the assessment instrument and the expected level of achievement in individual elements in the rubric and also the overall grade. A further area of improvement is to explicitly map the key elements being assessed in the USQ instrument against professional competencies. Finally, this paper highlighted the need to only use project assessors who are well versed in the technical discipline area. Overall it can be concluded that the USQ holistic method of assessment of competencies and grading of the final year capstone project can be considered consistent, repeatable, and reliable.

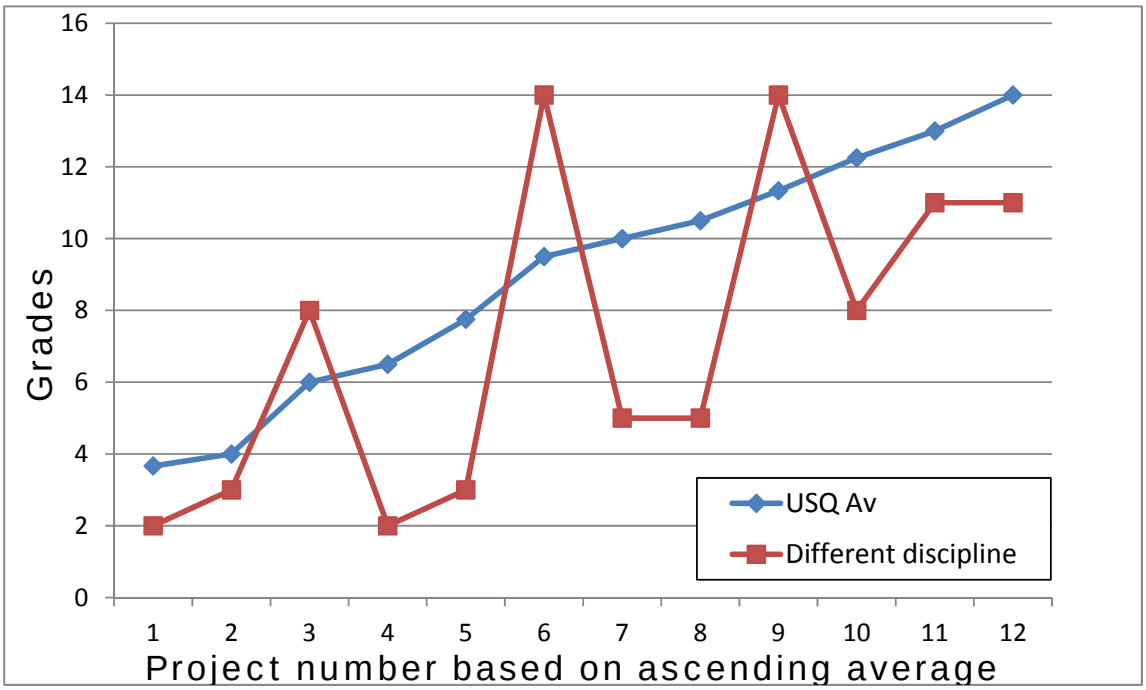

Figure 3: Comparison outside discipline and within faculty

References

Beyerlein, S., Davis, D., Trevisan, M., Thompson, P., & Harrison, K. (2006). *Assessment Framework for Capstone Design Courses*. Paper presented at the 2006 ASEE Annual Conference and Exposition, ASEE, 2006, Chicago, IL.

Brackin, P., Knudson, D., Nassersharif, B., & O'Bannon, D. (2011). Pedagogical Implications of Project Selection in Capstone Design Courses. [Article]. *International Journal of Engineering Education, 27*(6), 1164-1173.

Davis, D., Beyerlein, S., Trevisan, M., Thompson, P., & Harrison, K. (2006). *A Conceptual Model for Capstone Engineering Design Performance and Assessment*. Paper presented at the 2006 ASEE Annual Conference and Exposition, ASEE, 2006, Chicago, IL.

Geske, J. (2009). Order from Chaos: Assessment and Evaluation of Design in a Team Environment. [Article]. *International Journal of Learning, 15*(11), 23-31.

Howe, S., Goldberg, J., Palo, S., & Terpenny, J. (2011). The 2010 Capstone Design Conference. [Article]. *International Journal of Engineering Education, 27*(6), 1160-1163.

Howe, S., & Wilbarger, J. (2006). *2005* . Paper presented at the 2006 ASEE Annual Conference and Exposition, ASEE, 2006.

International Engineering Alliance. (n d). Washington Accord, from http://www.washingtonaccord.org/

McCormack, J., Beyerlein, S., Brackin, P., Davis, D., Trevisan, M., Davis, H., . . . Howe, S. (2011). Assessing Professional Skill Development in Capstone Design Courses. [Article]. *International Journal of Engineering Education, 27*(6), 1308-1323.

McKenzie, L. J., Trevisan, M. S., Davis, D. C., & Beyerlein, S. W. (2004). *Capstone Design Courses and Assessment: A National Study*. Paper presented at the 2004 ASEE Annual Conference and Exposition, ASEE, 2004.

Sobek, D. K., & Jain, V. K. (2004). *Two Instruments for Assessing Design Outcomes of Capstone Projects*. Paper presented at the 2004 ASEE Annual Conference and Exposition, ASEE, 2004.

Steiner, M., Kanai, J., Cheng, H., Alben, R., & Gerhardt, L. (2011). Holistic Assessment of Student Performance in Multidisciplinary Engineering Capstone Design Projects. [Article]. *International Journal of Engineering Education, 27*(6), 1259-1272.

Copyright statement