

Peer Rating for Feedback in Group Projects

J Paul Hermon

Charles D McCartan

Queen's University Belfast

ABSTRACT

The National Student Survey (NSS) in the UK has since 2005 questioned final year undergraduate students on a broad range of issues relating to their university experience. Across disciplines and universities students have expressed least satisfaction in the areas of assessment and feedback. In response to these results many educational practitioners have reviewed and revised their procedures and the UK Higher Education Academy (HEA) has produced guidelines of best practice to assist academics in improving these specific areas. The Product Design and Development (PDD) degree at Queen's University Belfast is structured with an integrated curriculum with group Design Build Test (DBT) projects as the core of each year of the undergraduate programme. Based on the CDIO syllabus and standards the overall learning outcomes for the programme are defined and developed in a staged manner, guided by Bloom's taxonomy of learning domains.

Feedback in group DBT projects, especially in relation to the development of personal and professional skills, represents a different challenge to that of individual assignment feedback. A review of best practice was carried out to establish techniques which could be applied to the particular context of the PDD degree without modification and also to identify areas where a different approach would need to be applied.

A revised procedure was then developed which utilised the structure of the PDD degree to provide a mechanism for enhanced feedback in group project work, while at the same time increasing student development of self and peer evaluation skills. Key to this improvement was the separation of peer ratings from assessment in the perception of the students and the introduction of more frequent face to face feedback interviews.

This paper details the new procedures developed and additional issues which have been raised and addressed, with reference to the published literature, during 3 years of operation.

KEYWORDS

feedback, peer rating, group projects, skills development.

BACKGROUND

The UK National Student Survey (NSS) is a voluntary, anonymous, online survey of Higher Education students, administered by Ipsos Mori, which has been conducted each year since 2005. It asks final year students to rate their educational experience (on a scale of 1 to 5) on an overall basis and in 21 more specific areas, which are grouped into 6 categories. The stated purposes of the survey are twofold: firstly to publish the statistics (unistats.direct.gov.uk) so that prospective students can be better informed about what and where they might study; and secondly to provide information for educators that could assist them in enhancing the student learning experience. The survey has not been without its

critics and there were some boycotts by students and institutions during the early years. However, almost all Higher Education Institutions (HEIs) in England, Wales and Northern Ireland now have over 50% of their graduating cohorts responding to the survey and national and institutional trends have been identified from several years of statistically significant data. Figures 1 and 2 show data for the Mechanical Engineering and PDD degrees at Queen's University Belfast, which are grouped together in the NSS. Figure 1 shows that the category of 'Assessment and Feedback' is consistently the area with which the students are least satisfied. Figure 2 shows that they find the promptness, clarity and helpfulness of feedback received the most unsatisfactory elements of their entire educational experience. These trends have been consistent over the last 4 years of continuous data. The profile across all Schools in the university is similar and across institutions similar trends have also been identified.

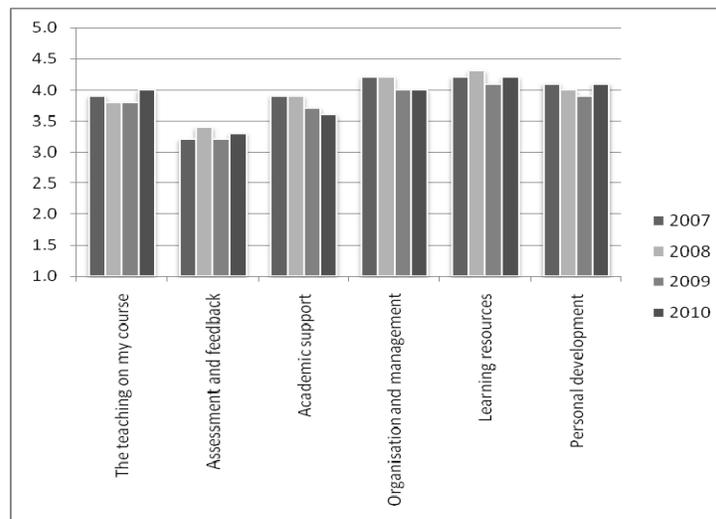


Figure 1. NSS category averages for QUB Mechanical Engineering and Product Design and Development final year students 2007 – 2010

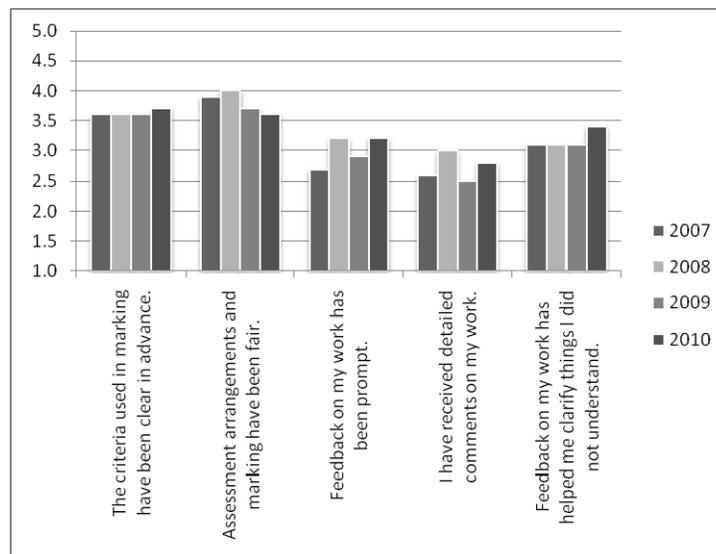


Figure 2. NSS Assessment and Feedback category - question averages for QUB Mechanical Engineering and Product Design and Development final year students 2007 – 2010

In response the Higher Education Academy (HEA) published in 2008 [1] a comprehensive study which looked at longitudinal changes to students' experience of Higher Education in the UK. This included a list of 5 key recommendations and 13 practices which were identified as effective in increasing student satisfaction in the areas of assessment and feedback. Many HEIs subsequently initiated a process of support for academics with particular focus on improving feedback and assessment within their institutions by encouraging the adoption of these best practices.

One of the HEA's key recommendations is that the NSS data is best used to identify areas that require further investigation. Each institution is encouraged to analyse and understand their own context before action to address any deficiencies is taken. An investigation into the feedback procedures in the PDD degree programmes was therefore undertaken by the authors, since the data available from the NSS suggested this was the area in which students were least satisfied.

SOME RELEVANT AND INFLUENTIAL LITERATURE

The DBT projects in the PDD degree require considerable periods of group work in addition to direct contact lectures, tutorials and design review meetings. A review of literature relating to peer assessment was carried out to assess the appropriateness of using this method to assist with the assessment of personal and professional skills such as time management, communication and collaboration in group projects. It was recognised that the students have a different perspective from the tutor on how well these skills are being developed by their peers since they are experiencing the outcomes first hand. Short of being fully embedded in a student group the tutor is restricted to taking snapshots of associated activities on which to base any evaluation of such skills. The approach of using the students involved as a resource to assist with assessment is therefore an attractive option that could be very time efficient.

Many studies have focussed on validating the accuracy of peer assessment when compared with the grades awarded by tutors. These have generally found reliable, accurate and consistent correlations and this has led many tutors to use peer assessment to award or modify marks for individual students in group projects. On the basis of these findings an approach of using peer assessment was adopted on the PDD degree. An earlier study in 1994 at Queen's Belfast by Stefani [2] had found students' assessments of laboratory reports, where the students had drawn up the marking criteria, to be as reliable as their lecturers. In a broad review carried out in 1998 Topping [3] found adequate reliability in the majority of 31 applications of peer assessment but with the caveats that unreliable findings may be less likely to be published and that peer assessment tended to be more reliable than self assessment. Topping also noted that there was considerable variety in how such studies had been carried out making direct comparisons difficult. Others such as Boud [4] have suggested that self and peer assessment is a skill that needs to be nurtured, with guidance from those already skilled in the discipline, and developed over a period of time before students can be considered competent in the practice. Kruger and Dunning [5] noted a significant relationship between a student's competence in a particular domain and the same student's ability to assess their own and their peers' competence in that same domain. The poorest performers were found to be the least accurate assessors and also the most likely to overestimate their own abilities. It was found that in peer assessment situations a clear majority of participants tend to rate their own performance as above average. This could be considered indicative of a general inability to assess accurately, particularly when self assessing, as identified by Falchikov [6] in her review of many such practices. The work of Kruger and Dunning in particular and follow up work by Ehrlinger et al [7] raised issues that caused sufficient concern to prompt a reevaluation of existing procedures.

The conclusion of this reflection was that using peer assessment as a method of generating or moderating grades was potentially less reliable than had first been assumed, particularly when students first use it, due to their inexperience and lack of skill in assessment. Instead it was considered that with only minor modification there existed an opportunity to enhance the educational environment and help students develop the required skills. The procedural changes implemented as a result included discussion of self and peer assessment with students as part of the feedback process. This aimed to provide a structured and supported mechanism for students to develop their own assessment skills in order that they would as Boud [4] contended become more effective continuing learners and practitioners.

PEER RATING FOR FEEDBACK

The PDD degree has been designed using the CDIO integrated curriculum model and has group Design Build and Test (DBT) projects as the core activity of each year of the programme [8]. The majority of modules are continually assessed and the average cohort size is around 25 students. Only 25% of the modules are co-taught with other degree programmes in the School. As shown in Table 1 the stage 1 PDD students undertake 3 short group projects as part of the Introduction to Product Design module and a further 12 week group design project in semester 2. In stage 2 there are 3 x 8 week DBT projects running in series and in stage 3 a 24 week major group project that runs across both semesters of the academic year. This includes the development of a functioning proof of concept prototype and an associated business plan for introduction of the product into the market.

Table 1. Group projects, Supervisors, Peer Rating and Feedback in PDD stages 1, 2 & 3

	Stage1 x 6 week (1) McCartan		Stage1 x 6 week (2) McCartan	Stage1 x 6 week (3) McCartan
	Stage1 x 12 week Hermon (PR1)			
Stage 2 x 8 week (1) (FB1) Hermon (PR2)	Stage 2 x 8 week (2) (FB2) McCartan (PR3)	Stage 2 x 8 week (3) (FB3) Hermon (PR4)		
(FB4)	Stage 3 x 24 week Hermon & McCartan (PR5) (FB5)			(PR6)
week 1	week 12		week 24	

The curriculum structure with a core of DBT group projects coordinated by the same 2 members of academic staff through the first 3 years of the programme provides several opportunities for enhancing feedback. Project groups are constructed at all stages so that by the end of the second year all students within a cohort will have experienced working with all other students in their year group. Table 1 also shows the 6 instances where the students complete a peer rating for feedback spreadsheet (PR1-6) and the 5 instances where face to face meetings take place to discuss the processed results from this process (FB 1-5).

The phrase 'Peer Rating for Feedback' is used instead of peer assessment in order to remove any confusion in the students' minds that their marking is influencing their own grades and the grades of others in their group. Previously it had been presumed that a minority of students had apparently been attempting to distort the process by inflating their own grades. It had not been considered that this effect might be due to their lack of ability to assess. Others who failed to participate fully in group activities tended to score all members very evenly and were reluctant to use the full range of the available scale, in marked contrast to their peers. The current procedures state explicitly that the peer ratings are not used to

adjust grades but rather will primarily be used in the feedback interviews. Carried out on a 1 to 1 basis these interviews concentrate on comparing their rating of the group with the combined totals of all group members, and that of the tutor. The students are told upfront that the interviews will address the accuracy by which they complete the self and peer assessment. Along with the removal of any incentive to inflate their own grades, by separating the process from the assessment, the completion of the peer rating spreadsheet now provides a better mechanism for evaluating student aptitudes in this area. The correct completion of a peer rating spreadsheet is a mandatory requirement for each group project. The interviews are also used to discuss the tutor's assessment of the individual with them, as well as the assessment of them in the context of their group. This provides an opportunity for the student to self evaluate their own ratings when compared to an experienced practitioner. To facilitate this the tutors score each individual on a weekly basis in the same key areas as appear on the peer rating spreadsheets, namely 'technical contributions', 'contribution to deliverables' and 'collaboration', as shown in Table 2. The categories stay the same but the 15 questions differ depending on the content and context of each project.

Table 2. Typical peer rating criteria – 15 questions split into 3 categories

Technical Contributions	Contributions to Deliverables	Collaboration
Ability to apply technical knowledge from other modules (including stages 1 & 2) to project	market research	Effectively takes charge of tasks assigned
Contribute alternative design concepts	Preparation for interim group presentation	Is fair and even in the treatment of ideas/solutions put forward by other group members
Sourcing of relevant technical information	Writing of interim group report	Produces work on time
Demonstrate an ability to apply critical thinking	Construction of concept prototype	Willing to take on tasks
Effectively troubleshoot problems and find answers	Design (sketches, CAD etc.)	Communicates clearly with other members of the team

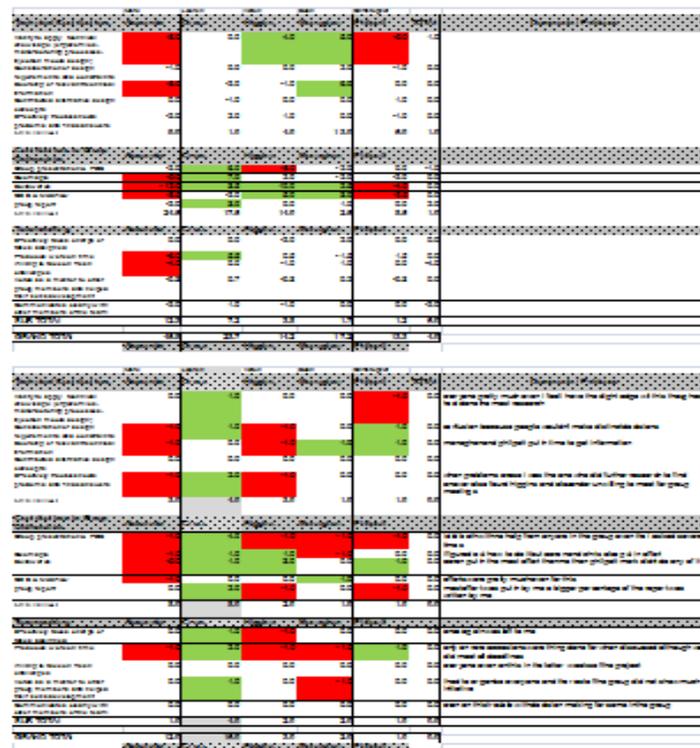


Figure 3. Group total (top) and individual (bottom) peer rating spreadsheets

Figure 3 shows a typical summary sheet produced for an individual interview and tabled for discussion with the student. The top half shows the sum of the peer ratings for the whole group while the bottom half shows the individual's ratings. The columns represent the individuals in a group and there are 15 rows relating to skills, attributes and activities which the student rates on a zero mean basis; the total for each row adding to zero. Individual cells can be scored as a real number between -2 and +2. To assist the discussion, cells which are significantly positive (≥ 0.5 on an individual spreadsheet) are filled green and significantly negative cells are filled red. In this way differences between the individual and group ratings are more easily identified and form the basis for discussion with the student at interview.

Students are required to supply justifying comments for any row with non zero cells. These are printed on the individual bottom half but comments from other members of the group are not disclosed in the top half. The students are made aware before completing the spreadsheet that their comments relating to other group members will remain confidential. This has been done to encourage full disclosure of potentially sensitive interpersonal issues that might otherwise not come to the attention of the supervisors.

EVALUATION

In order to evaluate student opinions of the peer rating for feedback procedures an anonymous questionnaire was carried out during the second semester of the 1011 academic year on the stage 2 and stage 3 PDD cohorts. Table 3 shows the total number of responses for each of 12 questions relating to the procedures adopted, and other related issues which were noted during the literature review; such as student involvement in deciding the rating criteria.

Table 3. Peer Rating for Feedback Questionnaire - Combined Responses from QUB PDD 2010/11 Stages 2 and 3

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Peer rating for feedback is a valuable practice which has helped me reflect on my own performance	14	20	3	2	0
Since the peer rating process does not directly influence the assessment I feel it is a pointless exercise	1	6	8	18	6
I consider the overall group ratings provided an accurate assessment of the relative contributions of the group	1	16	11	9	2
I felt uncomfortable criticising the efforts of my peers	3	10	10	11	5
I would prefer that all comments were made known to the group	2	9	8	13	7
Keeping my comments concealed from the rest of the group allowed me to say what I really felt	4	22	6	5	2
The feedback received on what my peers thought of my performance was useful	7	23	6	2	1
I would like more involvement in deciding the criteria to be included in the peer rating spreadsheets	2	14	20	3	0
I was motivated to work harder knowing that my peers would be rating my contribution	6	23	7	2	1
I was honest in my marking of my peers	22	16	1	0	0
I think other members of my team may have been unfair in their rating of my contribution to the group	3	9	17	10	0
I think that the peer rating marks should be used to adjust group marks for individuals	7	12	11	7	2

Figures 4 and 5 show the results of two of the questions plotted as histograms. These indicate a strong recognition of the benefit of the procedures and also show that the peer feedback regime had a significant motivating effect, despite the fact that there were no marks linked to the process. There was less enthusiasm however for wanting the peer ratings to be used to adjust grades, with the results for the last question in Table 3 showing only moderate agreement. Despite the value placed on the process by 95% of the respondents around 30% thought that the overall ratings were inaccurate and that others may have been unfair in their ratings. It is possible that since the questionnaire was carried out after a number of cycles of feedback interviews had taken place that the students had become conscious of their own failings as assessors and possibly suspected the same shortcomings in others.

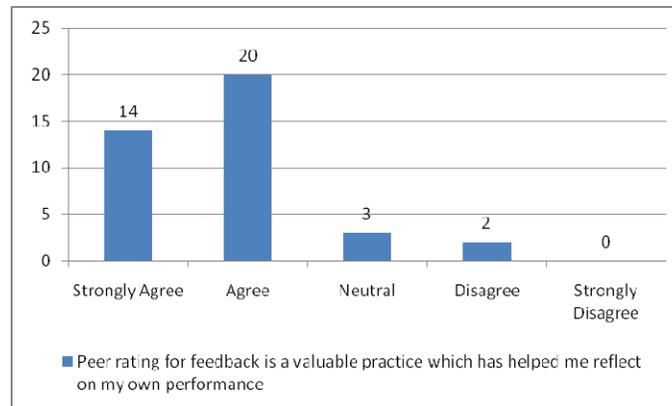


Figure 4. Student opinion of the value of peer rating for feedback

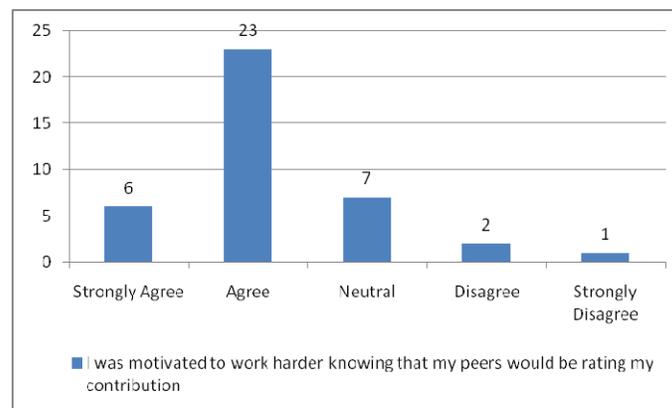


Figure 5. The influence of peer rating for feedback on student motivation

As part of its own internal Quality Assurance procedures QUB operates a system of module review which includes anonymous student questionnaires. The 16 questions cover many of the same topics as the NSS. In particular questions 12 and 15 relate to the usefulness and timely nature of feedback. The questionnaires use the same 5 point scale allowing broad comparison of scores in similar category areas to the NSS. Figure 6 shows the results from the last 3 years for module MEE2026, a stage 2 'Design and Prototyping Projects' module with 3 x 8 week DBT group projects running across 2 semesters, and which operates the peer rating for feedback procedures. It can be seen that Q12 (There was good interaction and feedback between students and lecturer) had an average score of 4.1 and Q15 (The lecturer provided me with helpful and timely feedback on my work) an average of 3.8. These figures are significantly higher than the NSS average of 3.0 calculated for the last 4 years in the same area of feedback (Figure 2).

It must be remembered, however, that the averages have been calculated from data gathered in very different ways, even though the questions are quite similar. The NSS average is for final year students asked to grade their experience of their entire degree and in this case includes data from 2 different degree programmes (Mechanical Engineering & PDD). The online NSS response rate is also typically much lower (reported as 59%) than the module questionnaires which are conducted in class (>90%) and relate to a 1.0 or 2.0 weight module out of a degree programme of 18 (BEng) or 24 (MEng) modules. This demonstrates that while the NSS may indicate general areas with which students are dissatisfied a more detailed analysis down to the module level is required to identify specific reasons for this dissatisfaction. Since the NSS only provides an average for the whole degree it is unclear if single bad experiences or the most recent experiences unduly influence how students rate their degrees.

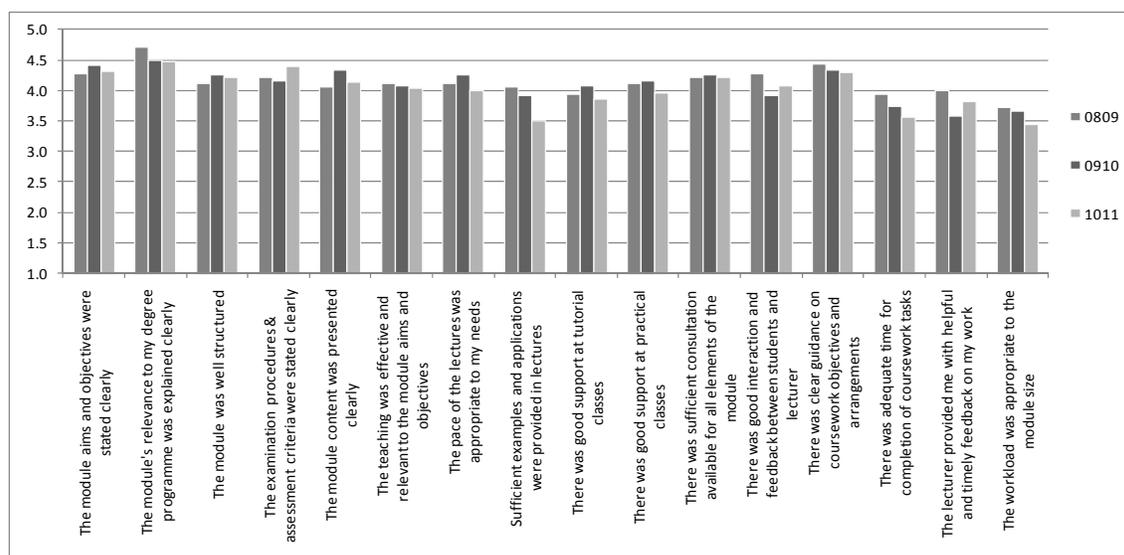


Figure 6. QUB module MEE2026 questionnaire results for 2008/9 – 2010/11

Overall the evaluations carried out so far have shown very favourable responses from the student cohorts and the results from relevant module questionnaires are encouraging. The cohort sizes for the PDD degree are currently relatively small, with an average of 25, and the 2 tutors get to know the individual students very well through supervision of projects in each of the first 3 years of their degree. This clearly improves the quality of feedback that can be provided but raises the issue of scalability of this approach. The authors suggest that by splitting larger cohorts into divisions of up to 30 students it should be possible to construct groups over a similar number of projects in the first 2 years so that each student gains experience of working with most if not all of the other students in their division. If supervisors similarly can be assigned in the same ratio of 2 per division of 30 throughout their first 2 years then the same level of intimacy in the feedback process could be provided.

CONCLUSIONS

- A peer rating for feedback process was developed with minor adjustment to an existing peer assessment regime, primarily by removing the link between peer rating and the assignment of grades.

- By adding a series of formative face to face interviews to discuss the outcomes of the peer rating process an environment was created to help students develop their self and peer assessment skills.
- Student responses showed that they valued the process of face to face feedback interviews which focused on development of self and peer assessment skills through comparison with the assessments of both their peers and tutors.
- Students reported increased motivation from knowing that their peers would be rating their performance, even though it was explicitly stated that these ratings would not be used to alter grades.

REFERENCES

- [1] Williams, K., Kane, D., 2008, "*Exploring the National Student Survey: Assessment and feedback issues*", The Higher Education Academy, York, UK, ISBN 978-1-905788-69-9
- [2] Stefani, L.A.J., 1994, *Peer, Self and Tutor Assessment: relative reliabilities*, Studies in Higher Education, Vol 19, No 1, p69-75
- [3] Topping, K., 1998, *Peer Assessment between Students in Colleges and Universities*, Review of Educational Research, Vol 68, No 3, p249-276
- [4] Boud, D. and Associates (2010), *Assessment 2020: Seven propositions for assessment reform in higher education*, Sydney; Australian Learning and Teaching Council
- [5] Kruger, J., Dunning, D., 1999, *Unskilled and Unaware of It: Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments*, Journal of Personality and Social Psychology, Vol 77, No 6, p1121-1134
- [6] Falchikov, N., 1995, *Peer Feedback Marking: Developing Peer Assessment*, Innovations in Education and Teaching International, 32: 2, p175-187
- [7] Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., Kruger, J., 2008, *Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent*, Organizational Behavior and Human Decision Processes, 105 (2008) p98-121
- [8] Hermon, P., McCartan, C., Cunningham, G., 2010, "*Group design-build-test projects as the core of an integrated curriculum in product design and development*", Engineering Education: Journal of the Higher Education Academy Engineering Subject Centre, Vol. 5, Issue 2, p50-58

Biographical Information

J. Paul Hermon is a Senior Teaching Fellow in the School of Mechanical and Aerospace Engineering at Queen's University Belfast. He holds a MEng Degree in Mechanical and Manufacturing Engineering (QUB 1987) and is Program Director for the BEng and MEng Product Design and Development degrees. He is a Fellow of the Higher Education Academy and a Fellow of the Royal Society for the encouragement of Arts, Manufactures and Commerce.

Charles D. McCartan is a Teaching Fellow in the School of Mechanical and Aerospace Engineering at Queen's University Belfast. His current scholarly interests include developing applying and evaluating active and interactive learning methods, teaching mathematics to

engineers, first year introductory courses and the assessment of group projects. In addition, he is a professional engineer with experience in industry, research and consultancy. He is a member of the Society of Automotive Engineers (SAE) and a Fellow of the Higher Education Academy.

Corresponding author

J Paul Hermon
School of Mechanical and Aerospace Engineering
Queen's University Belfast
Ashby Building
Stranmillis Road
Belfast BT9 5AH
+44 (0)28 9097 4499
p.hermon@qub.ac.uk