

Individualising students' scores using blind and holistic peer assessment

Mark Russell, George Haritos and Alan Combes

Abstract

Group work has the potential to be both educationally effective and resource efficient. It provides opportunities to set more integrative assessment tasks whilst also providing opportunities for the students to learn more than they would by simply responding to the coursework briefing sheet. Issues arise, however, with the necessity to individualise the group score. Teaching teams, having constructed authentic assessments for learning, now have to turn their attention to the assessment of learning on a student-by-student basis. This article presents some of the possible approaches to undertaking this individualisation exercise and includes information on 'how' (*open versus blind*) as well as 'what' (*category-based versus holistic*) data can be used to create an individualisation framework.

The application of a 'blind and holistic' process is reported, as are examples of the students' responses. Worryingly, in many instances the groups' cohesion appeared to diverge towards the end of the project and yet the students still believed that many of them were working at a degree classification of 2.i or above.

Background

There is little doubt that assessment has a profound effect on student behaviour and ultimately their learning. This has long been recognised and is often used strategically to drive student activity (Russell, 2005) as well as influence teaching strategies. Constructive alignment (Biggs, 2003) and Just-in-Time-Teaching (Novak et al., 1999) are two strategies that actively seek to embed the assessment *within* the teaching experience. Furthermore, carefully constructed assessment can do more than just require students to respond to the coursework briefing sheet. In addition to the *expected* student response latent and tacit learning can also be enhanced. This is particularly true of group work where the

assessment task also develops skills including decision making, negotiation, communication, empathy and delegation. Such transferable skills are of value outside the immediate assessment task and support notions of life-long learning, personal development planning and employability. Group work, therefore, has the potential to be extremely educationally effective and can positively impact on the students' 'graduateness' (i.e. a combination of analytical abilities, problem solving and communication skills as well as discipline-specific knowledge).

In addition to its educational effectiveness, group work may also be increasing as a consequence of the UK government's drive to increase participation in higher education. By its very nature, group work has the potential to have less demand on staff time and, as such, may also be seen by some teaching teams as a resource-efficient form of assessment.

However, one of the major issues associated with group work arises from the recognition that the group members are individual students and hence will bring their own differing levels of knowledge and commitment to the task goal. The group score needs to be individualised so that low commitment is not over-rewarded and high commitment under-rewarded, ensuring that each student is assessed fairly as an individual.

Whatever the driver for its adoption, be it educational effectiveness or resource-efficiency, teaching teams adopting group work have to reconcile the possible benefits with the necessity to individualise the group score. Whilst group work and its associated social-constructivism will present many opportunities *for learning*, ultimately there is a need to undertake an assessment *of and contributions to learning* on a student-by-student basis.

This paper presents some of the possibilities for individualising group scores and reports on the use of one approach.

Possible approaches to individual scoring

Traditional forms of assessment rely on the teacher to grade the students' work. With group work, the grading of the work often remains with the teacher but an additional data set is required to help inform the individualisation exercise. This additional data set is not the grading of the work *per se*, but rather a view on how the grade might be distributed amongst the group.

For differing reasons it is common for the group members and not the teacher to provide the individualising data. This results in a *peer assessment* of the individuals' contribution and performance. In this sense peer assessment does not refer to the correctness of the students work (that is judged by the teaching team) but rather the students' contributions in responding to the task goal. The arguments for privileging the students over the teacher include, *inter alia*, that it is only the students that really see the inner workings of the group and hence know who genuinely makes the contributions and who doesn't. Although it could be argued that it is the observeability of real contribution which is important and not who makes that judgment, privileging the students over the teachers raises other issues, these being presumptions that the students will engage in the process with honesty and provide informed and objective assessments of their peers' contributions. Hence, having noted the importance of the students and their ability to see the inner group workings as well as some of the potential issues, some studies use individualisation data from the group *and* the teachers, for example Wilmot and Crawford (2005). Whilst this joint approach will provide more data it also raises additional questions which might include: Which data-set is the most reliable? How are both data-sets to be weighted? Will the students' individualisation data be moderated? And what happens if there is no agreement between the two data-sets? Discussion of the observeability of group workings as well as problems associated with using the teachers' individualisation scores as a benchmark can be found in Magin (2001).

Another argument for using the students as 'raters' centres on the benefits that arise by bringing them closer to the assessment process (Rust et al., 2003). As with all forms of assessment, there is much to be gained by providing clear and transparent advice

on the assessment requirements as well as constructing opportunities for students to apply the requirements in the role of an assessor.

Accepting that the students are better positioned to see the inner workings of the group and that they can learn from the process too (i.e. by developing critical and evaluative skills as well as their meta-learning) suggests it is the students that need to take the lead in the individualisation process. Such a decision leaves the teacher to concentrate on the methodology to be used by the students and hence resolve:

- i) *how* the group work will be individualised; and
- ii) *what* criteria will be used to individualise the group score.

An overview of two available alternatives for both areas, together with their associated benefits and potential drawbacks, is presented in Tables 1 and 2.

For completeness, comparisons of the application of holistic and category-based peer assessments can be found in Lejk and Wyvill (2002). A useful briefing document on self, peer and group assessment is given by Race (2001).

Application

Having presented a background to the peer assessment of group work, the remainder of this article presents the findings from its application to a final year, design-based automotive engineering module. The module, Vehicle Engineering Design, is taught in one semester and has a value of 15 credit points. The teaching programme combines lectures, guest lectures, seminars and group tutorials.

The assessment of the module is entirely group and coursework based and comprises three separate, yet integrated, assessment tasks. These tasks challenge the students to develop a design idea through three evolutionary phases: 1) needs analysis (15%); 2) outline designs (15%); and 3) drawing together, detailed designs and 'where next' (70%).

A formal submission of work is required for each of these tasks. Variety of assessment is maintained by requiring the different submissions to be presented in different formats, i.e. an executive summary (phase 1), an oral presentation (phase 2) and a final integrative formal report (phase 3).

Table 1. How the group work will be individualised (assuming students provide the individualisation data).

Available alternative	Brief description	Possible benefits	Potential drawbacks
Open (negotiated)	<p>Group members are required to openly discuss their thoughts on the group members' performance and their contribution to the task goal, the output from the discussion being a group-agreed score for each group member. It is common with an open approach to require a signed statement from the participants acknowledging their agreement with the distributed scores.</p> <p><i>This is a dialogic activity.</i></p>	<p>Requires students to provide a rationale for their suggested distribution of grades and their justification to each other.</p> <p>All peer assessment is open, thus allowing a dialogue on the assessment score to take place (i.e. an aggrieved group member has an opportunity to reply and possibly justify his/her performance/contribution).</p> <p>Carefully constructed dialogue on any group 'issues' may help resolve them and not allow inappropriate commitment or contributions to go unchallenged.</p>	<p>Because of the openness of this approach, lazy-yet-popular students may be judged more on their popularity rather than on their commitment and contribution to the task goal.</p> <p>Explicit discussion of any group issues and students poor performance may disrupt the cohesion of the group.</p> <p>The easy/soft option for the group to take would be to avoid conflict. Hence some group members may feel constrained and not give their honest views of their group members' commitment and contribution.</p> <p>Less vocally confident group members may be less willing/able to articulate their justifications for others to dissect and possibly challenge.</p>
Blind (non-negotiated)	<p>Group members respond to the peer assessment task in complete isolation from each other. It is also expected that no discussions are entered into and no 'deals' are made between group members.</p> <p><i>This is a monologic activity.</i></p>	<p>Quieter/less confident students have a voice that is equal to the noisier, more confident students.</p> <p>Any group conflict is not aggravated due to issues not explicitly being raised within the group.</p> <p>Students may be more prepared to give honest views as they do not have to justify or defend them to their group members.</p>	<p>Group issues may never get resolved (i.e. lazy group members may never see their performance or contribution challenged through formal mechanisms).</p> <p>Further, because a discussion never takes place, it is possible that a student who believes their behaviour to be appropriate may be completely oblivious that it is not acceptable to the rest of the group.</p> <p>Where a student is awarded a low individualisation score it might be difficult for staff to be satisfied that the assessment was based on objective reasoning rather than being out of spite.</p> <p>No opportunity for a right-to-reply exists, thereby removing any opportunity for students to either challenge assertions or seek to justify their perceived behaviour/commitment.</p>

Table 2. What criteria will be used to individualise the group score.

Any of the four possible combinations of the how and what alternatives are available i.e. - blind and holistic, blind and category-based, open and holistic and open and category-based.

Available alternative	Brief description	Possible benefits	Potential drawbacks
Holistic	Students are expected to evaluate their peers' contribution and performance by looking at the 'big-picture'. Although holistic evaluation can be supported by the provision of guidance notes on acceptable performance and contribution etc, the students are not asked explicitly to respond to them in any detail. As such they may act more as guidance notes rather than evaluation criteria.	<p>Not constrained to responding to pre-defined category lists. This may allow for a more integrative perspective of student performance and contribution.</p> <p>Each student can derive their own evaluating criteria and weight them according to their own beliefs and values.</p>	<p>Students may have a different view of what is expected of them. Ground rules may not be clearly articulated or agreed before embarking on the group task.</p> <p>Social or cultural backgrounds from previous group work activity will be varied and may bring some mis-matched ideas as to what is important/ acceptable and what is not.</p> <p>It may be difficult for the students to see the importance of both the 'process' and 'product' associated with the assessment task.</p>
Category-based	Students are expected to evaluate their peers' contribution and performance against pre-defined criteria. The criteria categories may be uniform across the cohort or could be group specific.	<p>Provides a framework for the students' evaluations.</p> <p>Provides detailed information for the students at the start of the project on what behaviours and characteristics will be prized and what will not.</p> <p>Ownership of criteria could be increased by inviting groups to set their own assessment criteria.</p> <p>It is easier to explicitly note the importance of tasks/ activities associated with the process of undertaking the assessment as well as the resulting 'product'.</p>	<p>It is possible that the category list, developed at the start of the task, may not capture all of the activities associated with the successful completion of the group task.</p> <p>It may be too 'piecemeal' in its approach and less integrative.</p> <p>It may have the potential to drive strategic student behaviours, i.e. a student may simply undertake work and exhibit behaviours that are 'prized' (i.e. noted in the category list) and thus will score well, rather than work for the good of the group on items that do not form part of the category list.</p>

Background

Drawing on Tables 1 and 2 for guidance, the work reported here applies a *blind-and-holistic* peer-assessment. The emphasis on group and in-course assessment led to the choice of the blind approach (i.e. being entirely group-assessed may place a risk on students working within a dysfunctional group; the blind peer-assessment provides those students with a chance to identify the issue and hence reduce the potential for poor performers to impact on their grade).

In this instance the students were required to self-select themselves into groups of six or

seven students. Within appropriate guidelines the groups were then required to propose, and subsequently undertake, a design study of their choice.

In addition to using randomisation or existing friendships to construct groups other approaches are possible. These include various forms of social engineering using results from psychometric tests such as Belbin or Myers-Briggs Type Indicators or simply by mixing genders, age or previous performance, the justification for social engineering being the benefits likely to arise from having a cross-section of personality trait, gender, age or ability within the groups.

Within the present study, justification for the self-selecting of groups and the study area was made in an attempt to be more student-centred and hence derive more learning potential from the students' intrinsic motivation.

Although, as mentioned previously, there are many beneficial 'spin-offs' in undertaking peer-assessment of group work, one of its fundamental goals is to develop an individual weighting factor that allows the group score to be distributed on a student-by-student basis. Rather than inviting the students to simply submit student weighting factors (a weighting factor of less than 1.0 represents an under-performing student, 1.0 represents a neutral performing student and greater than 1.0 represents a high achieving student), the work reported here asked the students to submit a percentage score for themselves as well as for each of their group members. In doing so it now becomes possible to judge their perception of the relative performance of their group members, as implied by a weighting factor, as well as the level at which they believed they were working at. Collecting a percentage score, rather than a weighting factor, limits the possible scores between 0-100%. Inviting the students to reflect on both the relative weighting of their group and also how their work relates to the wider context of degree classifications was considered another valuable exercise.

Along with the percentage score, the students were required to submit a free-text justification for their grades (i.e. whilst it might be possible for all of the group to be performing at the 65% level, the justification for that score was likely to be different for each of the students). Equally, allocating a 70% score to a high performer or a 40% score to a low performer also required justification. The addition of the justification statement encouraged the students to think a little harder about what scores they were awarding; they now had to describe *why* they were awarding the score. Although blind, this feature of the peer assessment sought to reap some of the benefits of an open peer assessment by forming a kind of *closed justification*. The score and the associated justification statements were seen only by the submitting student and the teaching team. At no point during the process were these made available, either in full or in summary, to the groups.

To gain an insight into any transient group behaviours the peer assessment process was repeated three times and timed to coincide with each of the group coursework deadlines.

Individualisation algorithm

From the individual percentage ratings, a group mean (\bar{G}) as well as a set of student mean scores (\bar{S}) were calculated. The group mean score formed a benchmark with which each of the group members' student mean scores were compared. This allowed a deviation between group mean and student mean to be calculated. A student recognised by the group to be under-performing/under-contributing would attract a negative score, the magnitude of which being determined by how much the group thought s/he was under-performing, whereas a student doing better than the group average would receive a positive score. This difference in rating could then be used to individualise the group score awarded by the teacher to each student. In this case a linear adjustment was adopted.

The linear individualisation algorithm can be written as:

$(GroupScore = C_{S-G} * \bar{S} - \bar{G}) * PhaseWeighting Factor$, where:

GroupScore is the grade awarded by teacher for the group work. C_{S-G} modifies the impact of the $\bar{S} - \bar{G}$ value. $C_{S-G} < 1.0$ brings the individual scores closer to each other whereas $C_{S-G} > 1.0$ separates the individual scores. *PhaseWeighting Factor* is the contribution, as a decimal, that this piece of work makes to the overall coursework grade.

Tables 3-5 show the data for group 7 across the three assessment phases. S43, S44 etc. refers to student 43, student 44 and so on. Their results are chosen to demonstrate a typical profile but also the variance of justification statements from S44 and S48 (see section on *students justification statements*).

For reference, the leading diagonal in the central part of Tables 3-6 represents the students' rating of their own performance. A self review, with justification, was also considered an important part of the task and hence the students' learning.

Table 6 summarises the difference between the students' individual means (rated scores) and

Table 3. Phase 1 - group 7 peer assessment scores.

	Student	Rated					
		S43	S44	S45	S46	S47	S48
Rater	S43	70	69	70	70	70	70
	S44	78	70	75	72	70	71
	S45	75	71	72	70	73	73
	S46	80	60	70	75	50	75
	S47	75	74	75	75	75	75
	S48	80	70	75	73	77	77
	Student mean (\bar{S})	76	69	73	73	69	74
	Group mean (\bar{G})	72	72	72	72	72	72
	$\bar{S} - \bar{G}$	4	-3	1	0	-3	1

Table 4. Phase 2 - group 7 peer assessment scores.

	Student	Rated					
		S43	S44	S45	S46	S47	S48
Rater	S43	70	69	70	68	69	70
	S44	75	75	75	75	75	75
	S45	71	73	71	72	71	74
	S46	75	65	75	75	65	75
	S47	69	71	70	72	71	72
	S48	75	68	72	74	75	80
	Student mean (\bar{S})	70	69	70	68	69	70
	Group mean (\bar{G})	72	72	72	72	72	72
	$\bar{S} - \bar{G}$	0	-2	0	1	-1	2

Table 5. Phase 3 - group 7 peer assessment scores.

	Student	Rated					
		S43	S44	S45	S46	S47	S48
Rater	S43	80	80	80	80	80	80
	S44	75	75	80	70	70	70
	S45	75	72	76	73	76	77
	S46	75	60	75	75	65	75
	S47	72	70	73	70	72	72
	S48	74	68	72	70	64	72
	Student mean (\bar{S})	75	71	76	73	71	74
	Group mean (\bar{G})	73	73	73	73	73	73
	$\bar{S} - \bar{G}$	2	-3	3	0	-2	1

Table 6. Summary of student scores across the three phases.

	S43	S44	S45	S46	S47	S48
Phase 1 ($\bar{S} - \bar{G}$)	0	-2	0	1	-1	2
Phase 2 ($\bar{S} - \bar{G}$)	0	-2	0	1	-1	2
Phase 3 ($\bar{S} - \bar{G}$)	2	-3	3	0	-2	1

the group mean across the three phases. Note how S44 has been identified as a constant under-achiever/contributor to the three tasks, whereas S48 has been identified as an over-achiever/contributor.

Application of the individualisation algorithm to phase 1 for group 7 is given below (i.e. drawing on the $\bar{S} - \bar{G}$ data given in Table 6). For phase 1, group 7 received 60%, the PhaseWeighting Factor was 0.15 and the C_{S-G} was set at 1.0.

$$\begin{aligned} S43, (\bar{S} - \bar{G} = 0), & \quad [60 + 0] * 0.15 = 9.00 \% \\ S44, (\bar{S} - \bar{G} = -2), & \quad [60 - 2] * 0.15, = 8.70\% \\ S45, (\bar{S} - \bar{G} = 0), & \quad [60 + 0] * 0.15, = 9.00\% \\ S46, (\bar{S} - \bar{G} = 1), & \quad [60 + 1] * 0.15, = 9.15\% \\ S47, (\bar{S} - \bar{G} = -1), & \quad [60 - 1] * 0.15, = 8.85 \% \\ S48, (\bar{S} - \bar{G} = 2), & \quad [60 + 2] * 0.15, = 9.30 \% \end{aligned}$$

Although the range in the individual grades for phase 1 appears small (i.e. 0.6% (9.3–8.7)), it should be noted that the range of $\bar{S} - \bar{G}$ was small at phase 1. A larger range of individual scores arose in phase 3 (i.e. to coincide with their wider range of $\bar{S} - \bar{G}$ scores).

Students justification statements

In addition to the numeric data presented in Tables 3-6, it is also informative to review examples of the students' justification statements. Again, these justification statements were used to encourage the students to articulate *why* they were awarding the percentage grades. The following statements were written, at phase 3, by S44 and S48 (i.e. the aforementioned weaker and stronger contributors to the group activity respectively). [Note: Spelling mistakes have been corrected for ease of reading.]

Student 44 wrote:

about student 43

"Very good information for phase 3 and 4 on his chosen subject."

about student 44

"Helped put together the final project, and wrote up the control systems section."

about student 45

"Put a high amount of work into final project, and wrote up his section with James."

about student 46

"Composed final project."

about student 47

"Did a good section of FEA."

about student 48

"Generally good work overall, helped when help was needed, wrote up his section with Lee."

Student 48 wrote:

about student 43

"Andy did three-quarters of the section on the way the system works, he did loads of research and wrote a great section on the pneumatic system. He delivered his work just after the groups pre-determined times and did a fair amount of the final put-together and read-through of the project. Great work, shame it was late, spent a lot of time on it."

about student 44

"Mark did some great work for phase 3. He put the whole of the control systems part together. He delivered his work to the groups pre-determined times and did a reasonable amount of the final put together and read through of the project. Overall didn't put enough time in."

about student 45

"Lee did part of the paddle system part with myself, Lee did a lot of writing up of conversations we had, to be honest, he isn't great at CAD and didn't want to do that part and didn't have many ideas for the ergonomics study, but we talked through my ideas and he came up with some good ideas. He worked really hard and spent a lot of time writing up our section and getting the presentation correct. He also spent the most time putting the final project together and really kept the team together in that part. For me he may not have the ideas yet but he makes up for it in his dedication."

about student 46

"Anthony decided to re-work all of our previous sections for this section, he isn't strong in design and doesn't really understand how the system will work. He did a good re-write for this section, but if it was me I would have put more new information into this section because we had talked about it in previous sections and made notes. He delivered on time and was one

of the main few who really got on with putting the project together. Put in good amount of time, work was not as good as I would expect."

about student 47

"David did the section on FEA, for me it was late in and only the study was completed, none of the write up for the section had been completed, we all agreed to do the lot for the week back. I personally if I was in charge or it was just my project would have done more than 3 iterations as well. This did cause us to lose some time at the end because we were a man down on finishing up the project and had to wait for his section. We could have had it done much earlier that week. Also David didn't take a role in putting together of the project, for me not really a team player at the end when it counts, but did do a good section in the end. Good work, didn't spend enough time and late."

about student 48

"in this section I feel I put in just over average time into the project, I did all the design work and CAD for the paddles, wrote that section up and did 70% of the ergonomics study including creating the questions. For me I wish everyone had put in my effort, we would get a better grade because I feel our project isn't our best work, it is good but with a designated team leader I feel we would have organised our time better and spent the last week and Christmas improving on a basic project. I have said that I was more of a leader before but for this section we did our own sections so there wasn't any need to take control. For me I feel I did my best work and the team would not have done as well without me."

What is immediately apparent from the above is the difference in the length of the justification statements. This in itself might also suggest their differing commitment to the task. These levels of contribution are not unusual with the less committed students generally providing less information than the more committed students. What is interesting also from S48 is the language of his submission; it appears thoughtful and considered. He notes his disappointment in some of his group members but also looks for positive statements about their work too. Note also in his description of his own performance how he mentions the task goal as well as the issue of group leadership (i.e. a useful, non-task related, learning experience that is consistent with the previously noted spin-off benefits of group

work). His level of response also evidences his engagement with the peer-assessment process and the seriousness with which he participated. The seriousness of engagement in the peer assessment process is in agreement with experiences elsewhere, such as Johnson and Miles (2004). Their claim for the seriousness of student engagement comes from the differences in the submitted data amongst the group, and they suggest that *'there appeared to be no instances of group collaboration in the ratings given.'*

Additional analyses

In addition to individualising the students within the groups, the wealth of data collected allows additional analyses to be undertaken. These additional analyses use the individual and group data and provide useful insights into:

- i) the transient nature of the peer-assessment through the phases;
- ii) the level at which the students believed they were working.

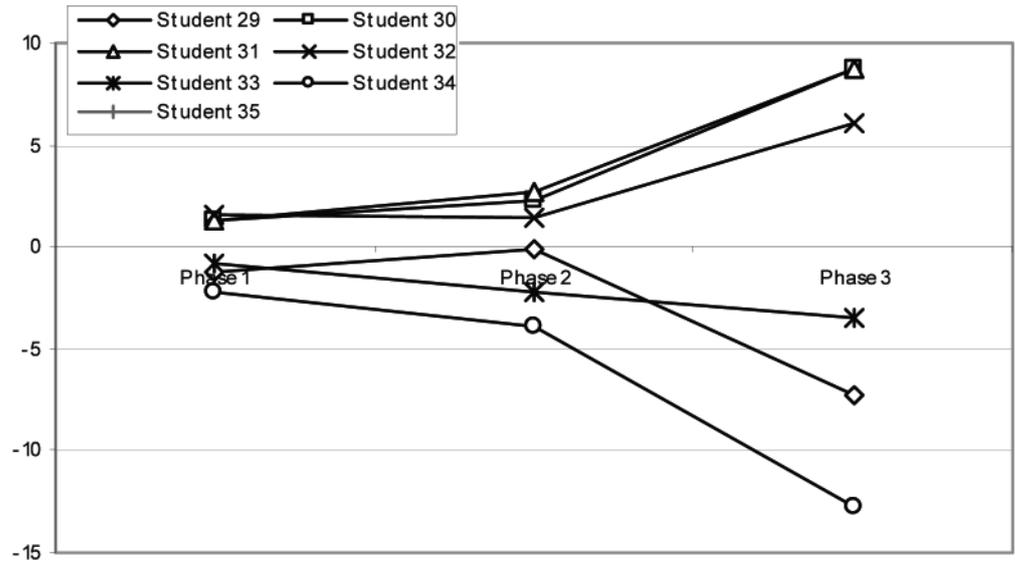
The transient nature of the peer assessment through the phases

Repeating the peer assessment process provides an insight into the students' perceptions of their peers' performance and contribution throughout the group work, not just at the end of it. By plotting the difference between the student mean and the group mean ($\bar{S} - \bar{G}$) at the three phases, it becomes straightforward to spot those groups that are tending to collectively converge as well as those groups that are diverging from a common goal. Interestingly, only two of the seven groups appeared to be converging and working more cohesively at the end of the project than they did at the start. The remaining five appeared to be diverging. A typical plot of this data is given in Figure 1. The individual lines join up the students' individual mean scores across the three phases. In this instance, the students' scores were clustered around each other at phase 1 but became more spread out from each other through phases 2 and 3, thus indicating group dysfunction/divergence at phase 3. The general shape of this plot was repeated in five of the seven groups.

Students' perception of the level they were working at

In all, a total of 800 individual (blind) submissions were made. Table 7 summarises these submissions and shows the students'

Figure 1. Example plot of the individualised peer-assessment data through the assessment phases



perceptions of the level at which they believe their group members were working. In all three phases around 70% of the students believed their group members were working at the first class level, whereas less than 10% of the students thought their group were working at 2.ii or below.

There are many reasons why the students may have had such a high perception of their group. First, the groups were self-selecting which suggests they were likely to be working alongside people with whom they had previously had a good experience. It is unlikely that a student will knowingly choose to work with another student whom they think would be awkward. Second, within the context of the module, the students were free to select their own design project. The nature of the project would therefore be more authentic and more exciting to the students. Finally, the students were in the final year of their degree programme and hence may have been driven by the necessity to score well and increase their chance of gaining a better degree classification.

Although the above may justify some of the students' ratings and may have tapped into their intrinsic motivation, the scores may also have arisen because the students were simply not aware of the characteristics and requirements associated with, for instance, a first-class honours degree. This lack of knowledge becomes apparent when the justification statements are read alongside the awarded grades. Looking at the statements presented by S44, for instance, it becomes

difficult to see how these can conceivably justify giving all of his team members a first class grade (see Table 5).

Given that there seems to be an increase in the number of first-class degree awards within HE it will be no surprise that so many of the students believe such an award is well within their grasp. Naturally, it is likely that a combination of factors have resulted in the high ratings; the self motivation, described above, as well as the 'grade inflation' issue.

Conclusions

This article has reaffirmed the importance of assessment and, moreover, the additional learning opportunities that group-based assessment can bring. The challenge for teaching teams wishing to adopt group work, however, is how they will consider the individual students' contributions within the group. A framework of *how* and *what* could be

Table 7. Categorising the students' responses to their peers contribution/performance by degree classification.

		Degree class				Total
		1st	2.1	2.2	3rd	
Phase 1	n	220	48	13	3	284
	%	77	17	5	1	100
Phase 2	n	181	67	17	2	267
	%	68	25	6	1	100
Phase 3	n	169	57	16	7	249
	%	68	23	6	3	100

considered to individualise the work has been presented, as has the application of a blind and holistic individualisation process.

Analysis of the students' contributions suggests that the process was seen as authentic by the students and also vindicated the addition of the justification statement. Even though the students' comments and submissions would not be seen by their group members, many of their comments were considerate, thoughtful and articulate.

In terms of group cohesion, there are concerns as to why so many of the groups appeared to collectively diverge rather than converge towards the end of the project. This in itself is worth further exploration. Even without additional analysis, the data provides teaching teams with a need to pay close attention to the group cohesiveness at *all* stages of the project.

What was particularly concerning was the level at which the students believed they and their group were working at. In the data presented less than 10% of the students thought their peers were working at or below the 2.ii degree classification level. There are many reasons for this but it does raise the issue of grade inflation and students not being familiar with the broad characteristics of the various degree classifications. It is difficult to see how so many of the groups demonstrated divergence in cohesion and yet at the same time were apparently working at or above a 2.i level.

Recommendations for the future

The use of group work and the issues surrounding the individualisation process are not constrained to engineering education. Additional uptake and application of this methodology to other subject areas will allow cross-disciplinary comparisons to be made. Such studies will help develop the robustness of the approach and establish the transportability of the work.

Further, many questions have arisen as a consequence of this work. These include comparing the students' grades and justifications when exposed to blind and negotiated peer assessment and evaluating the method used in creating the groups. To help investigate these areas, the authors are currently undertaking a parallel (blind and negotiated) peer assessment process within a first year engineering design module.

The students will be invited to share their experiences of both processes. The intention is to construct a hybrid approach that gains more of the benefits from each approach. Presently, for instance, no mechanism exists in the blind approach to inform students how the group rates their behaviours and performance. This could be a major drawback with blind peer assessment. Furthermore, in contrast to the work reported here, the groups in the on-going first year study were formed randomly and not by self-selection. Although it was argued in this study that allowing self-selection might help the groups' cohesiveness this hypothesis needs further testing.

Follow-up work is also planned to establish the criteria used by the students to grade their peers in an holistic assessment. This work will seek to establish a criteria resource bank that could be used either to i) initiate criteria-based peer assessment or ii) share with students on a holistic peer-assessment what behaviours were previously prized and what were not. ■

References

- Biggs, J. (2003) *Teaching for Quality Learning at University*. Buckingham: Society for Research in Higher Education and Open University Press.
- Johnston, L. and Miles, L. (2004) Assessing contributions to group assignments. *Assessment and Evaluation in Higher Education*, **29** (6): 751–768.
- Lejk, M. and Wyvill, M. (2002) Peer assessment of contributions to a group project: student attitudes to holistic and category-based approaches. *Assessment and Evaluation in Higher Education*, **27** (6): 569–577.
- Magin, D. J. (2001) A novel technique for comparing the reliability of multiple peer assessments with that of single teacher assessments of group process work. *Assessment and Evaluation in Higher Education*, **26** (2): 139–152.
- Novak, G., et al. (1999), *Just-in-time teaching: blending active learning with web technology*. Upper Saddle River, New Jersey: Prentice Hall.
- Race, P. (2001) *A briefing on self, peer and group assessment*. York: LTSN Generic Centre.
- Russell, M.B. (2005) Evaluating the Weekly Assessed Tutorials Sheet approach to assessment: Background, pedagogy and impact. *Journal for the Enhancement of Learning and Teaching*, **2** (1): 26-35
- Rust, C. et al. (2003) Improving students' learning by developing their understanding of assessment criteria and processes. *Assessment and Evaluation in Higher Education*, **28** (2): 174–164.
- Wilmot, P. and Crawford, A. (2005) Validating the assessment of individuals within undergraduate teams. *Proceedings of the International Conference on Engineering Education ICEE 2005*. Poland, 501-506.

About the authors

Mark Russell BEng (Hons), PGCE, CEng, ILTM, MCIBSE,
Principal Lecturer, School of Aerospace, Automotive and Design Engineering & The Blended Learning Unit, University of Hertfordshire, AL10 9AB
Telephone: 01707 285 083 Email: m.b.russell@herts.ac.uk Fax: 01707 285 086

George Haritos BSc (Hons), MSc, PhD, CEng, ILTM, MIMechE,
Senior Lecturer,
School of Aerospace, Automotive and Design Engineering,
University of Hertfordshire, AL10 9AB

Alan Combes BSc (Hons),
Head of Automotive Centre,
School of Aerospace, Automotive and Design Engineering, University of Hertfordshire,
AL10 9AB
