ltsn generic centre
Learning and Teaching Support Network

Assessment Series No.9

A Briefing on Self, Peer and Group Assessment

Phil Race

Professor Phil Race currently works part-time at the University of Leeds, and spends the rest of his time running staff development workshops in the UK and abroad on teaching, learning and assessment methods. His recent publications include 'The Lecturer's Toolkit (2nd Edition)' (Kogan Page, London, 2001), and '2000 Tips for Lecturers' (Kogan Page, 1999). He also works with students on study-skills development, and has published 'How to Get a Good Degree' and 'How to Win as a Final Year Student' with Open University Press (1999 and 2000 respectively). He is an ILT Member, elected Council Member, and Accreditor.

Contents

Generic Centre Guides and Briefings

Welcome to the Learning and Teaching Support Network Generic Centre's series of Assessment Guides and Briefings. They aim to provide a series of overviews of important issues and practices in the field of assessment for the higher education community.

The Assessment Guides are intended for colleagues with particular roles and for students, as their titles suggest. The Briefings are primarily intended for lecturers and other staff involved in supporting learning.

The Assessment Series is a snapshot of a field in which development is likely to be rapid, and will be supplemented by specific case studies produced by the LTSN Subject Centres.

The series was developed by Brenda Smith and Richard Blackwell of the LTSN Generic Centre with the support of Professor Mantz Yorke. Experts in the field were commissioned for each title to ensure that the series would be authoritative. Authors were invited to approach the issue in their own way and no attempt was made to impose a uniform template.

The series editors are grateful to colleagues in LTSN Subject Centres and other senior colleagues who refereed the series, and of course to the authors for enabling its publication.

We hope that you will enjoy the Assessment Series and find it interesting and thought-provoking. We welcome your feedback and any suggestions you may have for future work in the area of assessment.

Professor Brenda Smith
Head, LTSN Generic Centre

Richard Blackwell, Senior Adviser,
LTSN Generic Centre

Professor Mantz Yorke,
Liverpool John Moores University

November 2001

Summary

Abstract

The purposes of this briefing are to alert you to the roles which can be played by student self-assessment, peer-assessment and group assessment to enhance students' learning, and to diversify the range of assessment approaches and formats used in higher education.

The briefing explores how to introduce and implement these different forms of assessment, and discusses the extent to which they can be made valid, reliable and transparent to students (and to those reviewing the quality of assessment in higher education). A further aim of this briefing is to increase your own belief in the value of adopting one or more of these assessment tactics into your own assessment strategy.

Self-assessment, peer-assessment and group assessment

It is useful to clarify the meanings of each of these terms. They often tend to be lumped together, as things different to 'normal' assessment – tutor assessment, but they are in fact quite different from each other as well as from tutor assessment. Moreover, discussion of self-, peer- and group assessment is all too readily taken over by what is being assessed, and it is vital to bear in mind that the real benefits of such forms of assessment are to do with how assessment is being approached. In other words, the processes by which self, peer and group assessment can be achieved are in fact much more important than the products being used for such assessments.

Self-assessment

This involves students making judgements about their own work. As with conventional assessment, the judgements can range from 'pass-fail' (or pass – not-yet-pass) decisions, to percentage marks or grades A to E, and so on. The assessment decisions can be made by students on their own essays, reports, presentations, performances, projects, dissertations, and even exam scripts. However, student self-assessment can be even more valuable when the evidence to be assessed is intrinsically personal in nature, such as reflective logs, diaries, action plans, and so on, where it can be argued that only the student really knows how well the evidence meets the purposes or criteria designed to specify it. Self-assessment cannot of course be anonymous, and where self-assessment is part of an overall assessment profile this means that assessment decisions are made in a more 'exposed' setting than where anonymous peer-assessment is used. Self-assessment can be influenced by the tendency to make judgements on what was meant rather than what was actually achieved.

Peer-assessment

Peer-assessment is quite different from self-assessment, however, in that students are making assessment decisions on other students' work. Once more, student peer-assessment can be used for almost any aspect of student performance, including essays, reports and so on, or exam scripts. It is more usual, however, to use student peer-assessment for evidence relating to presentations, performances, practical work and so on. Student peer-assessment can be anonymous, with assessors randomly chosen so that friendship factors are less likely to distort the results. Student peer-assessment can be single or multiple, and is usually regarded as working most effectively when more than one assessor assesses each element of work, so that consistency can be demonstrated (or lack of consistency can alert tutors to problems with the assessment criteria, or the commitment of students to peer-assess fairly).

Peer-assessment can be summative, and taken into account as a component of the overall assessment picture. Alternatively, peer-assessment can also pay dividends when used in a purely formative way, where the real purpose is to allow students to gain feedback from each other, and any scoring or grading is just a means towards this feedback rationale.

Group assessment

This can refer to the assessment by a tutor of the products of student group work, or to the assessment of the product by students from other groups (inter-peer assessment), or the assessment of the product of group work by students within a group (intra-peer assessment), and can include self-assessment by individuals or by the group as a whole of the product they have generated, and/or their respective contributions towards the product. Therefore it is usual for group assessment to involve at least some elements of peer-assessment and self-assessment.

Why involve students in their own assessment?

There are many reasons for involving students in self-, peer- or group assessment. The following list contains a range of such reasons. Which reasons are the most important or the most relevant depend on the circumstances in which the assessment processes are being used.

Because students are already self-assessing and peer-assessing quite naturally

Students learn a great deal from each other, both in classes and outside classes. They naturally compare what they have achieved with each other, and use this to reflect on their own learning progress. Including student self-assessment and peer-assessment in our assessment profile legitimates what students already do spontaneously, and can help them to do it much more effectively.

Because tutor assessment is not sufficiently valid, reliable or transparent

Sometimes students are better placed to assess their own or each other's work. For example, when students have been thinking deeply about something because they have been involved in actually doing it, perhaps for the first time, they are often able to make peer-assessment judgements on other students' work, (or indeed self-assessment judgements on their own work) more objectively than would be made by someone (for example a tutor) who already knew how to do the task involved, and had not just learned how to do it. For example, someone who has just passed a driving test is usually better informed about the factors involved in the test than someone who has passed it years ago, even though the latter may be the better driver by far.

To deepen students' learning experiences

This is probably the most powerful reason for involving students in their own, and each others assessment. The act of applying assessment criteria to evidence such as essays, reports, presentations, and so on is a much deeper learning experience in itself than just reading or observing the assessment artefacts. Students can learn a great deal about their own attempt at a task by assessing two or three other students' attempts at the same task. Students can learn a lot about a task by applying assessment criteria themselves on their own evidence. Students can learn even more about a task by comparing their own judgements about it with those of fellow-students.

To let students in to the assessment culture

Getting students to participate in peer-assessment (and to a lesser extent, self-assessment) can help students to understand how tutor assessment works. The act of applying assessment criteria to other students' evidence such as essays, reports, solutions to problems, helps students to see how tutors' minds work when assessing their own evidence in more formal circumstances, including assessed coursework, and (not least) exams. Students can learn even more about the assessment culture of higher education when they are involved in designing assessment criteria to use in peer assessment.

To help students towards becoming autonomous learners

It is increasingly realised that one of the main purposes of higher education is to help students to develop themselves, and acquire a range of transferable skills (key skills) including skills relating to organising their own learning, time management, task management, problem solving, reflecting, and so on. Self-assessment, in particular, can be a vital way of causing students to reflect on their progress, and take stock of their learning. Students who know how their learning is going are much better placed to prepare to demonstrate their optimum potential in traditional exams, and in tutor-assessed coursework.

To help students develop skills relating to life-long learning

Self-assessment skills are invaluable in the context of life-long learning, and are useful to students in their continuing professional development long after they have gained their university qualifications. Similarly, peer-assessment skills are needed by lifelong learners, not only in their continuing learning, but also in contexts such as performance appraisal, team building, and so on, which need people who have become adept at assessing each other's work and contribution fairly, sensitively and appropriately.

To help students gain much more feedback than would otherwise be possible

In many disciplines, student numbers have grown, and with large class sizes tutors are hard-pressed to maintain the quality and quantity of feedback which they once used to give to students about their tutor-assessed work. While feedback from fellow-students may not be as authoritative as from an expert witness (tutor), there is much to be gained by having a lot more feedback. When a large amount of peer feedback is combined with a restricted amount of tutor feedback, students can benefit from both quality and quantity.

Why diversify assessment?

It is increasingly recognised that higher education, in the UK and in many other countries, tends to have become dominated by a relatively narrow range of assessment instruments, practices and processes. For example, as Brown and Glasner (1999) remind us, something like 90% of a typical UK degree depends on unseen time-constrained written examinations, and tutor-marked essays and/or reports. While teaching staff in higher education are generally very experienced at designing and implementing these particular forms of assessment, it can be argued that familiarity can blind us to the weaknesses that can be embodied in this small range of assessment formats.

More importantly, it is now widely understood that any assessment format or process disadvantages some candidates, and that using the same few formats disadvantages the same candidates time and time again. Therefore, students' success in higher education depends disproportionately in mastering these few assessment formats, and this gets in the way of our attempts to use assessment to measure their subject-related knowledge and skills. It is therefore important to diversify assessment, so that candidates have a greater opportunity to demonstrate their true potential on at least some of the assessment occasions they encounter.

One approach to diversifying assessment is to increase the range of assessment instruments and formats. While unseen time-constrained written examinations and tutor-marked essays and reports each have their own advantages as well as drawbacks, and can continue to play their part in a scenario of diversified assessment, modern developments in higher education now add to this basic repertory a wide range of alternative assessment instruments and formats, including:

• 'seen' exams, where students can prepare for known questions;

• 'open book exams, less dependent on memory than 'unseen' exams;

• 'open notes exams, where students can systematically prepare their own summaries and aides-memoirs for the exam;

• 'time-unconstrained examinations, reducing the penalties imposed by time-constrained exams for slower handwriting, and slower thinking;

• 'OSCEs – objective structured clinical examinations – used in medical education, where candidates under exam conditions encounter a series of short, practical, real-life situations, and measured on their reactions to these;

• ''in-tray' examinations, for example where candidates are presented with portfolios of information and data at their exam desks, and use these under exam conditions to address a series of carefully-planned short tasks and activities involving them in making decisions and judgements with the data available to them;

• 'multiple-choice questions, and other structured question types, where candidates' decision-making skills and broad knowledge levels can be measured without their speed of handwriting, or legibility, being involved in their performance; such examinations are increasingly computer-marked, taking much of the drudgery out of assessing;

• 'vivas or oral examinations, which can measure skills and knowledge which can elude written examinations altogether, and have always been a much more important part of assessment in many other parts of the world.

Increasingly, assessment for a degree involves more of these alternatives to their traditional predecessors, and this contributes to ensuring that the same candidates are less liable to being disadvantaged repetitively. However, an alternative way to diversifying assessment is to increase the range of assessment approaches, by bringing in at least some self-assessment, peer-assessment and group assessment. This can be achieved both within the traditional assessment formats such as essays, reports, and some of the types of formal exam, but is more often introduced beyond these traditional formats, including (for example) the assessment of:

• presentations;

• posters;

• exhibitions;

• performances;

• portfolios of evidence;

• reflective logs;

• diaries;

• artefacts such as paintings, sculptures, designs, and so on;

• computer programs.

One important aim of this briefing is to encourage colleagues to include at least some tactics based on self-assessment, peer-assessment, and group assessment in their strategy for diversifying assessment. This briefing is also intended to help colleagues make wise and informed decisions about which elements of the curriculum may best lend themselves to these alternative assessment approaches, and to alert colleagues both to the benefits which these approaches can deliver, as well as to at least some of the risks inherent in adopting these approaches.

What's in it for tutors?

The benefits to students can also be thought of as benefits to tutors. It is easier for tutors to work with students who are aware of their own progress, and who understand how assessment works, and are therefore better able to prepare for it. There are, however, some further benefits which can be realised by tutors themselves. These include:

Some of the drudgery can be saved by peer-assessment

Few people who have been involved in implementing peer-assessment would claim that it saved them time. It takes time to set up, time to train students to do it well, and time to oversee it and debrief students after it. The benefits start to accrue when large numbers of students need to do assessed tasks which are relatively straightforward in nature, and where the students can quite easily make assessment judgements on each others' attempts at such tasks.

Self-assessment gets students to reflect on their own work, and can open up productive student-tutor dialogues.

This is considered further later in this briefing. The main point is that giving students feedback on their self-assessment (rather than just on their work) causes deep learning for students, and can be very useful for tutors, allowing them to see in much clearer perspective where the strengths and weaknesses of a cohort of students lie.

Students performance in traditional assessments is enhanced

Most tutors who have involved students in peer-assessment and/or self-assessment report that when students face similar tasks in formal elements of assessment, such as exams, students perform better than expected. This reflects a deepening of their learning of the elements addressed by self- or peer-assessment. This is a benefit for tutors too, as tutors' performance is often judged by the performance of their students.

What are the risks?

Most of this briefing is advocating involving students in their own assessment, using self-assessment, peer-assessment and group assessment. However, it is important that they are used in carefully selected curriculum elements, where it is possible for students to be sufficiently knowledgeable to be able to make informed judgements. When feedback from an expert witness is really needed, there is no substitute for tutor assessment.

With assessment under quite intense scrutiny (for example, in the UK by the Quality Assurance Agency), anything which could be seen to compromise the reliability of assessment is hazardous. Therefore, when using self-, peer- or group assessment, care needs to be taken to ensure that moderation processes are sufficient to guarantee a reasonable level of reliability, but without the moderation being so intrusive that the benefits of involving students in their own assessment are undermined.

Further problems can arise when external examiners are not themselves experienced in using self-assessment or peer-assessment. External examiners are usually expected to assert themselves as guarantors not only of the standards of courses, but also as scrutineers of assessment processes. Even when peer-assessment (for example) is facilitated as ideally as possible, there are still likely to be instances where external scrutiny can 'pick holes' in the quality of such assessment. That is not to say that at least as many 'holes' exist in conventional tutor-assessment, but these can be more difficult to pinpoint. In other words, peer-assessment can become an easy target for criticism, especially by external examiners who are not themselves convinced of the benefits. In all cases of changing assessment approaches, it is wise to involve external examiners in discussions prior to the design and implementation of new approaches.

There are further issues which need to be confronted in developing self-, peer- and group assessment. Not least of these is the time it takes to set up these alternative forms of assessment. They should not be regarded as a 'quick fix' addressing assessment overload. Although in the long run, when experience has been gained at implementing these assessment processes, staff time can indeed be saved. It can be argued that for students (already overloaded with assessment) it takes more time to self-assess and/or peer-assess. The counter argument has to be made in terms of increased learning payoff. When students can appreciate that their learning is being deepened by being involved in their own, and each other's assessment, they can be persuaded that they are indeed benefiting from the processes.

Principles of assessment

It is increasingly required of us that our assessment instruments and processes are valid, reliable and transparent. For example, in the UK, the Quality Assurance Agency (QAA) Code of Practice for Assessment gives considerable detail and guidance on how assessment should be designed to achieve these qualities. Though most attention tends to be focused at making sure that tutor assessment is valid, fair and reliable, it is important to strive to extend these qualities to embrace self-assessment, peer-assessment and group assessment as far as is reasonably practicable.

• Validity: assessment should demonstrably measure that which it sets out to measure. In other words students' achievement of the published intended learning outcomes of their courses, modules or programmes. Including student self-assessment can increase validity in some situations, for example students themselves are best able to judge to what extent they have evidenced their achievement of some kinds of learning outcomes, particularly those relating to reflection and personal development. Student peer-assessment can allow greater validity to be delivered for some kinds of assessment, for example that relating to presentations, performances, practical competences, as well as to some of the more traditional formats such as essays, reports and dissertations. Group-based assessment can extend the range of assessment to include cooperative and collaborative skills, teamwork, and so on, and increase assessment validity by bringing into the assessment framework skills and competences which are more closely connected to real-life vocational situations than are traditional exams and essays.

• Reliability: assessment should be independent of which assessor is involved (inter-assessor reliability) and independent of where and when a particular assessor marks students' work (intra-assessor reliability). Reliability is synonymous with consistency, and fairness, and (as far as can be achieved) lack of subjectivity. Both student self-assessment, and (particularly) student peer-assessment can increase the reliability and objectivity of assessment, by making assessment decisions and gradings less dependent on the judgments of a single assessor, and exposing the extent of possible inter-assessor unreliability. Clearly, for self-assessment and peer-assessment to work well in such respects, they need to be demonstrably objective and fair, and appropriate moderation needs to be in place to assure this. Group-based assessment, while contributing strongly to extending the validity of some kinds of assessment, is less suitable for assuring the reliability of assessment.

• Transparency: this is about the goalposts being clearly defined, so that students are well aware of the standards expected of them to gain particular awards, and the nature of the evidence that they will need to furnish to demonstrate their achievement of the published intended learning outcomes. Both self-assessment and peer-assessment can play a strong part in enhancing the overall transparency of assessment, not least as both require much greater care in formulating and expressing assessment criteria than is needed for tutor-assessment.

Implementing student self-assessment

What lends itself particularly to student self-assessment? While students can self-assess just about any aspect of their work (including essays, reports, presentations, posters, portfolios, artefacts such as drawings, paintings, and so on) there are some things that are arguably best self-assessed, including some which in many respects can only be self-assessed. Some of these are mentioned below.

Reflective logs and diaries

While tutors can attempt to assess how well students have expressed their reflections, or how well they have structured their diaries, in truth only the students themselves know how deeply they have reflected while putting together such evidence of their reflection.

Learning payoff

While tutors can assess how well students have evidenced their learning, only students themselves know how much they have really learned. It is useful to get students to self-assess the state of their learning, not least because this alerts them sharply to that which they have not yet learned, but may need to learn in time for exams. Students can, of course, delude themselves about how much they have learned, but for those who strive to self-assess honestly and objectively, the rewards are worth the effort.

Presentations

It can be particularly useful to get students to self-assess their own presentations, using the criteria that are concurrently being used to peer-assess them. This can cause students to reflect quite deeply on what they think are the strengths and weaknesses of their approach and performance. Usually, students tend to be more critical of their own presentations than their peers, and the comparison of self- and peer-assessment data can help them to feel more encouraged about their performances.

Links to intended learning outcomes

Where the curriculum is expressed by means of published intended learning outcomes, it is important that these are seen by students to be useful in helping them to set their sights and work towards achieving clear targets. This applies to assessment in general, and in the UK the Quality Assurance Agency (QAA) in subject review, and the newer methodology of academic review places considerable emphasis on causing institutions to make clear explicit links between assessment criteria and learning outcomes. When using student self-assessment, it is every bit as important to make the same clear explicit links, and to allow students themselves to see how the assessment criteria they are applying to their own work relate directly to the curriculum as defined by learning outcomes. This helps to convince students that the self-assessment they are doing is just as 'real' as tutor assessment.

To moderate, or to let go?

This is an issue which has been discussed thoroughly by those who have worked extensively in facilitating student self-assessment (for example Boud 1995). In short, if tutors moderate student self-assessments with anything other than a light touch, students do not put their hearts into being objective in their self-assessment, and depend on being put down by a tutor if they are too hard on themselves, or try giving themselves more credit than they deserve. If tutors let go entirely, placing the full responsibility on students, the danger is that there will always be some students whose self-assessment is not justified, but the payoff is that the majority of students undertake their self-assessment much more seriously, and therefore learn a great deal more in the process of doing it.

Feedback on self-assessment

This can be very powerful. Students who have engaged conscientiously with self-assessment, and then receive feedback from a tutor on how objectively they have self-assessed, take the feedback very seriously. This, after all, is feedback not just on their work, but on their thinking about their work. In particular, when tutor feedback is to encourage students that they have underestimated the worth of their own work, it can give students a significant confidence boost. On the other hand, when students have over-valued their own work, feedback from tutors needs to be relatively gentle, and explanatory rather than accusatory, otherwise students can lose faith in the value of engaging in self-assessment.

Some prompts to consider for self-assessment dialogues with tutors

It is useful to give students a questionnaire to fill in and submit alongside tutor-marked work (essays, reports, almost anything), causing students to self-assess their own work at the point of submission, and opening up dialogue agendas with the tutors who assess their work. It is, however, vitally important not to give students the same questionnaire time after time, or the effect (like any other repeated questionnaire) pales into insignificance as surface responses from students take over from considered ones. Ideally, the prompts on a self-assessment proforma should include elements which are specific to the particular task that the students have just undertaken, but it is useful to illustrate some of the relatively general prompts which can be included from time to time. Some such prompts are listed below, with a few words about why they can be useful dialogue starters.

• *What do you think is a fair score or grade for the work you have handed in?*
Surprisingly often, this turns out to be in close agreement with the actual scores tutors award the work, and feedback to students can confirm this. When there is a significant difference between the two scores, it can be useful for tutors to discuss this on a one-to-one basis with students. Usually in such cases either students have underestimated the quality of their own work, and need a little encouragement along the lines that 'this is a perfectly good way of addressing Question 3' or (more rarely) when students have overestimated their work, they need help in tracking down exactly where the discrepancy arises. In the latter situation, it is usually a blind spot in students' learning – something they thought they had 'cracked' that they have not quite yet 'cracked'.

• *What was the thing you think you did best in this assignment?*
Most often, this turns out to be what they actually did best, and it is quick and easy for tutors to give feedback to this effect. From time to time, however, students did something else even better, and giving feedback on this can be most encouraging.

• *What was the thing that you think you did least well in this assignment?*
Once again, most students reply with the aspect that they actually did least well, and feedback is straightforward and confirmatory. However, from time to time, when students have actually done the stated part perfectly well enough, feedback to this effect is very important.

• *What did you find the hardest part of this assignment?*
Most students will identify the part they did least well, and feedback is straightforward. However, quite often, students will identify something they actually did perfectly well, and then it is really important to be able to give them appropriate feedback along the lines 'you say you found this hard – but it doesn't show at all. Well done, you have mastered this!'.

• *What was the most important thing you learned in doing this assignment?*
Here, the real learning points are often for the tutors who set the assignments in the first place. Feedback to students needs to reassure them that their learning will turn out to be useful to them.

A range of further 'prompt' possibilities is proposed in Race, 2001.

Implementing student peer-assessment

What lends itself to student peer-assessment? Just about anything! However, most work on student peer-assessment seems so far to have been focused on presentations, performances, portfolios, essays, reports, and designs.

Even though in essence peer-assessment is quite different than self-assessment, some of the same considerations continue to apply. For example, the issue of whether tutors should moderate or 'let go' are just as important – peer-assessment is unlikely to be taken seriously by students if they know that their tutors will have the final say. Similarly, the criteria used for peer-assessment need to be linked closely and demonstrably to the intended learning outcomes of the elements of curriculum being peer-assessed.

Establishing student ownership of the peer-assessment agenda

This can make all the difference between peer-assessment working well and working badly. When students feel a sense of ownership of the criteria they are applying to each others' work, they apply the criteria much more objectively than when they are merely working to a checklist of someone else's (for example, their tutor's) criteria.

However, it is not straightforward to get students to think of a set of criteria to use in peer-assessment. It is well worth spending a little time doing this with students, first of all leading them to defining the evidence which will constitute a 'good' attempt at the task concerned, then teasing out from them what makes it a 'good' attempt.

A step-by-step account of no less than eighteen stages by which student ownership of peer-assessment criteria can be achieved is proposed in Race, 2001, and is reproduced in Appendix 1 at the end of this guide. In practice, using such a process as outlined in Appendix 1 can tease out, refine, and weight criteria from a relatively large group of students in less than an hour of contact time. The increased quality of the peer-assessment justifies the time spent getting the criteria right in the first place.

Assessing student group work

Anyone who has attempted student group work assessment will have confronted the problems of 'free-riding', and the difficulties involved in sorting out fairly the different contributions students may have made to the products of the group work. There are no easy solutions to these problems, but a range of approaches, as discussed below, can help to address the problems.

There are many reasons why student group work needs to be part of the curriculum. In their careers beyond university, most students are going to be required to be able to work in groups or teams, and student group work is a valuable training ground. It is in group contexts that students can best develop and flex their interpersonal skills, leadership skills, and indeed 'followership' skills which can be even more vital.

However, students nowadays are (everyone seems to be saying) more strategic than they used to be. In the context of group work, this boils down to the fact that if it isn't assessed, many students will not take it very seriously, and therefore will not learn much from doing it. So the onus is on us to try to assess group work as validly, reliably and transparently as we can (not least because these qualities in our assessment design are as ever under external scrutiny in quality assurance procedures).

It is relatively straightforward to assess the products of student group work. Such products can include presentations, project reports, exhibitions, and all manner of artefacts and supporting evidence. We can assess the products in very similar ways to those used to assess the products of individual student work. The hard part is to assess the group work itself, and the relative contribution of individual members of the group to the development of the assessed product.

In the final analysis, when it comes to measuring individuals relative contribution to group work, the only people who really know what the respective contributions actually are, are the group members themselves. We can try to get students to tell us about such contribution levels, but human nature being what it is, if students decide to 'carry' passengers, there is not much we can do about it. Sometimes, even when students are very reluctant to tolerate such 'bystanders', they are still reluctant to 'shop' the offenders, and often unite to present a picture of agreement about the contribution levels being more-or-less equal.

We can investigate further, and find out who contributed strongly, and who did not. For example, we can set exam questions which will be easily answered by those who engaged fully with the set group work, and which will reveal those candidates who did not participate. Similarly, we can have individual face-to-face oral exams (or vivas) either with the whole group, or with individual members of the group, and we can often tease out the real levels of contribution to the final group product.

Alternatively, we can pass the buck back to the students themselves. For example, we can say to a group of four students "your groupwork product is assessed as scoring a total of 340 marks out of a possible 400. How do you want to distribute these 340 marks between the four of you?". Some groups will attempt to establish differentials to reflect the levels of contribution which they know were involved; others will hide the discrepancies and agree to have exactly 340/4 marks each. Then, of course, it becomes unfair, in that some groups are recognising differentials and others are not; in other words some students are assessed more generously than they should be, and others do not get as much credit as they deserve.

Problems such as those touched on above often dissuade tutors from building assessed group work into the overall formal assessment picture at all. This is unfortunate, if we remind ourselves that traditional assessed coursework is subject to parallel problems. For example, is the excellent essay we mark all the own work of the candidate whose name it bears, or is it a product of discussion, research of other people's work, and so on? Has it been adapted from something downloaded from the Internet – or even commissioned and purchased over the Internet or elsewhere? We can not always tell! The most skilful plagiarists will always escape our notice (unless we triangulate our assessment with convergent exam questions, or face-to-face oral tests or vivas). It is the relatively clumsy plagiarists whose efforts we see through. And furthermore, should we face up to the fact that in the world outside education, people continue to benefit from other people's efforts – whether it be a manager who takes the credit for the work of a team, or the head of department who takes the credit from the efforts of its members. Thus can be the nature of group work!

Seven approaches to assessing group work were explored by Race, 2000. Each of these has both advantages and disadvantages. The seven approaches can be summarised as follows.

Take the simplest path – just use the same group mark for all involved.
This approach is easy to manage and is worth considering if it is primarily the product of the group learning which is to be assessed, and not the processes leading up to this product. However, giving the same mark for all can be perceived as unfair, encouraging passengers, giving no bonus for excellence.

'Divide and concur' – divide up the assessed group task, and assess each component separately.
This can help groups to avoid disagreement, as everyone knows that their assessment will depend primarily on their own work. This approach enables individuals to shine, and to know that the success of their work will be attributed directly to them. However, it can be difficult to find equivalent tasks for all, and disputes may break out if some members of the group feel that they have been burdened with more demanding tasks than others. Moreover, the overall assessment load is increased, and it can be difficult to balance assessment decisions across the group, particularly if some members were set more demanding tasks than others.

Add differentials – give a mark for the overall group product, but negotiate differentials between group members. (For example, in a group of four members, award the group product 65%, then ask the group to divide up (4 x 65%) according to the way they feel the work was shared. You may need to decide whether to leave the differentials entirely to the group, or to make a ruling (for example) that a maximum differential should be 20%).

This approach is perceived to be fair, and places value on individual contribution to the work of the group. It also gives ownership to the group of the method of differentiation of the assessment of their overall work.

A further advantage is that the onus of awarding credit for group processes is taken away from the assessor, who may not in any case be in a good position to estimate the equivalence of contribution of members of the group. However, such an approach needs a mature group to achieve consensus, and can be found very intimidating to groups whose members do not know each other very well. It can also result in everyone just agreeing to have the same mark, while causing internal resentments to build up inside the groups where contributions have not been equivalent, destabilizing the group in future collaborative work together.

Add contribution marks – award a mark for the product of the group, and ask group members to peer-assess an additional mark for their contribution. In other words, for example, award each member of the group 65 marks for the product, and ask them to award each other member of the group between 0 and 10 for the extent to which they contributed to the work. This approach enables group members to feel justice is being sought in the assessment of their work, and encourages them to value process as well as outcome, giving them the message that process is regarded as being important. However, learners may turn round and say "It is your job, not mine, to assess my work". Also, training and practice is needed before group members enter into peer assessment, as they may be reluctant to mark down peers, and may agree to award each other equal (or maximum) marks for the process component of their work.

Add further tasks – award an equal mark to each member for the product of the group task, then add individual assessed tasks for each member of the group. This can be a way to accommodate the diversity of group members, and can allow them to take responsibility for allocating the additional tasks between group members. It also minimizes the amount passengers can benefit. However, this approach can make considerably more work for assessors, and deciding equivalent additional tasks can be difficult.

Test groups orally – award all group members the same mark for their product, but add an individual viva (oral exam). This enables assessors to test individual participation. Whether the viva is done with the group as a whole, or with individual members separately, it is usually fairly easy to establish a reasonably accurate impression of whether the group members contributed equally to the work of the group. However, vivas can be stressful, and some group members may not give an accurate impression of their contribution, either by (through shyness or modesty) underplaying their contribution, or by some members being able to 'fake good' in the viva when in fact their contribution was not good. This approach also makes more work for assessors.

Test them in writing – allow the group mark for the product to stand, but add a separate related assessment component to an exam. This approach makes it more difficult for most passengers to evade justice, and is perceived to be fair.

It can allow the most deserving individuals the opportunity to shine. Furthermore, knowing that group work remains on the exam agenda causes learners to include such work in their revision for exams, causing them to deepen their own learning by reflecting further on the group work. However, the exam may not be testing the same kinds of skills as the group work itself, and may unduly reward candidates who happen to be skilled at written exams. Furthermore, additional marking will be involved, and assessors are already overloaded with exam marking!.

As can be seen from the summary above, all of these methods have their pros and cons, and working out how best to assess group work is not an easy decision. The fact that it is difficult to assess group work fairly should not discourage us to include such work in the assessed curriculum. The fact remains that group work is a vital training ground for students, and can enrich their educational experience, not least from the peer-group learning which occurs when students work together towards a common goal.

Enhancing learning using self, peer and group assessment

In several parts of this briefing, it has already been suggested that the real reasons for involving students in their own assessment should be to deepen students' learning. Five factors have been identified below which underpin successful learning:

• Wanting to learn (intrinsic motivation)

• Needing to learn (extrinsic motivation)

• Learning by doing (practice, trial and error)

• Learning through feedback (praise, constructive criticism)

• Making sense or 'digesting' what has been learned.

Involving students in their own assessment can be seen to relate to all five of these factors, as summarised below.

Wanting to learn

Students can be considerably enthused by being involved in self-assessment, peer-assessment, or the assessment of group learning. Once they become excited by trying to match their achievements to the assessment criteria (particularly when they have been involved in generating the criteria in the first place) their motivation is enhanced, and their learning is deeper. In particular, the act of applying assessment criteria to their own work, and each others' work, can help students to want to achieve fully, and demonstrably, the associated learning outcomes.

Needing to learn

This is enhanced by the close encounters students have with assessment criteria, when engaging in self- and/or peer-assessment. The fact that students themselves are applying such criteria helps them to relate much more directly to the targets they need to achieve, and to be more self-aware of their progress towards achieving them.

Learning by doing

All the forms of assessment discussed in this briefing involve students in activity beyond simply undertaking the assessed tasks. Applying criteria and making judgements are high-level learning activities, and the learning payoff of self-assessing and/or peer-assessing is often much higher than that of undertaking the assessed task on its own. Even more, when students are involved in formulating and prioritising assessment criteria to use in peer-assessment, this activity in its own right has considerable learning payoff relating to the learning outcomes being addressed by the assessment.

Learning through feedback

This is arguably among the greatest of the benefits associated with peer-assessment. There is learning payoff associated with receiving feedback from one or more fellow-students, but perhaps even greater learning payoff in formulating and giving feedback to other students. It is the person who explains who really deepens their learning, rather than the person being explained to.

When student self-assessment is coupled to tutor-assessment and opening up dialogues with tutors, the feedback agenda is particularly rich. Students are not only receiving feedback about the quality of their work, but also about the quality of their thinking about their work, which deepens their thinking about everything involved, and strengthens their understanding of how assessment in higher education works in general.

Group work is also a rich learning environment regarding feedback. Students giving feedback to each other while undertaking group work helps to compensate for the limited possibility of feedback to students from hard-pressed tutors. Moreover, when someone who has just seen the light dawn, on a tricky concept for example, explains it to a fellow student, the explanation tends to be more effective than when someone who has known it for years tries to explain it. The act of explaining has high learning payoff too, as already noted above.

Making sense – 'digesting'

Assessment is the 'sharp end' of learning. Most students' learning is driven by assessment. Students continue to be strategic about learning and assessment, and after all this is an intelligent response to the situation of over-assessment most students find themselves facing. Involving students in their own assessment, in any of the ways discussed in this briefing, necessarily helps students to make sense of what they are learning. In particular, students benefit from seeing how assessment works, and how learning outcomes are linked to assessment criteria, and working out the evidence which will demonstrate that they have achieved the learning outcomes, and the criteria to judge the quality of this evidence. Many commentators have observed that curriculum elements where students have been involved in self- and/or peer-assessment are found, through formal assessments such as exams, to have been learned much more deeply (in other words, better digesting).

Conclusion

There are considerable dividends which can result from well-planned implementation of involving students in their own, and each others, assessment. In this guide, the nature of the processes of self-, peer- and group assessment has been explored, and the ways that these processes can be implemented in particular contexts has been illustrated. The role of self-, peer- and group assessment in the context of the need to diversify assessment processes has also been discussed. The benefits and pitfalls associated with these forms of assessment have also been highlighted.

It could be concluded that the principal benefit is that students' learning payoff can be enhanced by involving them in their own, and each other's assessment. Moreover, it can be argued that these forms of assessment help students to develop skills invaluable in later lifelong learning contexts, and their own ongoing continuing professional development as graduates. However, it is vitally important that when making changes to the overall assessment profile in a course or module that students can see that their tutors believe in the value of these changes. If students sense thattheir tutors distrust new approaches to assessment, or do not believe students will handle them responsibly, we should not be surprised if the dividends turn out to be losses.

Appendix 1:
<u>illustration of processes whereby students can develop ownership of criteria to be used in peer-assessment</u>

Brainstorming: Ask all students to jot down individually a few key words in response to: "What makes a really good 10-minute presentation? Jot down some of the things you would look for in an excellent example of one".

Sharing: Get students to work in groups. Even in a large lecture theatre, they can work in groups of 4 or 5 with their near neighbours. Alternatively, if students are free to move around the room where the exercise is happening, they can be put into random groups (alphabetical, or by birthday month, or allowed to form self-selecting groups). Ask the groups to share and discuss for a few minutes all of their ideas for a good presentation.

Prioritising: Ask the groups to make a shortlist of (say) "the most important five features of a good 10-minute presentation". Ask each group to appoint a scribe to note down the shortlist.

Editing: Get the groups to look carefully at the wording of each item on their shortlists. For example, tell them that when they report back an item from their list, if you can not tell exactly what it means, you will ask them to tell you "what it really means is……". Maybe mention that some of the more-academic words such as 'coherence', 'structure' and 'delivery' may need some translation into everyday words (maybe along the lines of 'hangs well together, one point following on logically to the next…' , 'good interest-catching opening, logical order for the middle, and firm solid conclusion', and 'clearly-spoken, well-illustrated, backed-up by facts or figures….'). However, do not put too many words of any kind into students' minds, let them think of their own words.

Re-prioritising: Remind the groups about the shortlisting process, and to get their five features into order of priority. This may have changed during the editing process, and meanings became clearer.

Turning features into checklist questions: Suggest that the groups now edit each of their features into a question-format. For example, "was there a good finish?", "How well was the material researched?" and so on. The point of this is to pave the way for a checklist of criteria that will be more-straightforward as a basis for making judgements.

Collecting the most important questions in the room: Now start collecting 'top' feature-questions. Ask each group in turn for the thing that came top of its list. Write these up, one at a time, on a flipchart or overhead transparency, so that the whole class can see the emerging list of criteria. Where one group's highest-rating point is very similar to one that has already been given, either put a tick beside the original one (to acknowledge that the same point has been rated as important by more than one group), or (better) adjust the wording slightly so that the flipcharted criterion reflects both of the sources equally. Continue this process until each of the groups has reported its top criterion.

Fleshing out the agenda: Now go back round the groups (in reverse order) asking for "the second-most- important thing on your list". At this stage, the overlaps begin to occur thick and fast, but there will still emerge new and different checklist-questions based on further features identified by the groups. Use ticks (maybe in a different colour from the overlaps of top-rated questions) to make the degree of concurrence visible to the whole group as the picture continues to unfold. With a large class, you may need to use more than one flipchart-sheet (or overhead transparency), but it is important to try to keep all of the agenda that is unfolding visible to the whole class. This means posting up filled flipcharts where everyone can see them, or alternating the transparencies so that students remember what has already come up.

Any other business? If the degree of overlap has increased significantly, and after gaining all the second-round contributions, the flow of new ideas has slowed down, it is worth asking the whole group for "any fairly-important things that still aren't represented on your list?". Usually, there will be a further two or three significant contributions at this stage.

Numbering the agenda: When all of the criteria-questions have been noted down, number them. Simply write numbers beside each criterion, in the order that they were given. During this stage, if you notice that two criteria are more-or-less the same, it can be worth asking the class whether you can clump them together.

Weighting individually: Ask students to work individually again next. Ask them to weight each criterion, using an agreed total number of marks. Choosing the total number needs care! If there are ten criteria, 100 marks would be too tempting regarding the possibility of some students just giving each criterion ten marks, and avoiding the real business of making prioritising decisions again. Thirteen criteria and sixty marks works better, for example. Ask every student to ensure that the total marks number adds up to the agreed figure. Legitimise students regarding ignoring any criteria that they individually do not think are important: "If you think it is irrelevant, just score it zero".

Recording everyone's weighting publicly: The next stage is to record everyone's marks on the flipcharts or transparencies. This means starting with criterion number 1, and writing beneath it *everyone's* marks-rating. It is worth establishing a reporting-back order round the room first, so that every student knows who to follow (and encouraging students to nudge anyone who has lost concentration and is failing to give you a score!). "Can you shout them out as fast as I can write them up?" usually keeps everyone (including you) working at speed.

Optional separating: It can be worth starting with two flipcharts from the outset. For example, you may wish to record separately the criteria relating to content and those relating to structure. This may pave the way for peer-assessment grids which help to separate such dimensions.

Discussing divergent views: Then go through all of the remaining criteria in the same way. Do not worry that sometimes consecutive scores for the same criterion will be quite divergent. When this happens, it will be a rich agenda for discussion later, and if you're writing the scores up in the same order each time, it is not too hard to pinpoint the particular individual who gave an unusually high or low rating to any criterion. You can, for example, ask the student who rated criterion 8 highest to argue briefly with the student who rated it lowest, and see what the causes of the divergence may be.

Averaging: Next, average out all the scores. If there are students with calculators in the group, the average rating may be forthcoming from the group without any prompting. Otherwise, it is usually possible to do some averaging and rounding up or down to the nearest whole number just intuitively by looking at the numbers. Ask the whole group "Does criterion 7 get a 5 or a 6 please? Hands up those who make it a 5?" and so on.

Shedding weak criteria: Look back at the whole range of criteria and ratings. At this point, there will usually be one or more criteria that can safely be dropped from the agenda. They may have seemed like a good idea at the time to some of the students, but the visible ratings tell their own story.

Confirming ownership: "Are you all happy to proceed with the averaged-out version of the ratings, and with these criteria?" is the question to ask next. Mostly, there will be no dissent. Just occasionally, a student with a different view of the ratings may wish to speak out against the consensus. It is worth then offering that any individuals who feel strongly about the ratings can choose to be peer-assessed by their own idiosyncratic rating scales, but that these must now be shared with the whole group for approval. Students rarely wish to do this, particularly if the feeling of ownership of the set of weighted criteria is strong in the group as a whole.

Administrating: Turn the criteria-questions into a grid, with the criteria down the left-hand-side, and the weighting numbers in a column alongside them, with spaces for students to write in their peer-assessment ratings. If students are going to be asked to peer-assess several instances of the task involved (for example maybe 10 short presentations) the grids could be marked up so that students used the same grid for the successive presentations. Alternatively, if the peer-assessment grids are going to be used for a small number of assessments (for example, where all students mark three essays or reports, and each of theirs is to be marked by three students), it is worth having separate sheets, with a column for individual feedback comments relating to the score awarded for each of the criteria.

**(From 'The Lecturer's Toolkit', 2nd Edition,
Phil Race, Kogan Page, 2001).**

References and further reading

Andressen, L., Nightingale, P., Boud, D. and Magin, D. (1989) *Strategies for Assessing Students - Teaching with Reduced Resources.* SEDA Paper 78. Birmingham: SEDA.

Biggs, J. (1999) *Teaching for Quality Learning at University.* Buckingham: Open University Press/SRHE.

Boud, D. and Lublin, J. (1983) *Self-Assessment in Professional Education,* Australia: University of New South Wales.

Boud, D. (1986) Implementing Student Self- Assessment. *HERDSA Green Guides.* **5**. Australia: University of New South Wales.

Boud, D. (1992) The Use of Self-Assessment Schedules in Negotiated Learning. *Studies in Higher Education.* **17** (2), 185-200.

Boud, D. (1995) *Enhancing Learning through Self- Assessment.* London: Kogan Page.

Brown, G., Bull, J. and Pendlebury, M. (1997) *Assessing Student Learning in Higher Education.* London: Routledge.

Brown, S. and Glasner, A. (eds.) (1999) *Assessment Matters in Higher Education - Choosing and Using Diverse Approaches.* Buckingham: Open University Press.

Brown, S. and Knight, P. (1994) *Assessing Learners in Higher Education.* London: Kogan Page.

Brown, S., Rust, C. and Gibbs, G. (1994) *Strategies for Diversifying Assessment in Higher Education.* Oxford Brookes University: Oxford Centre for Staff Development.

Brown, S., Race, P. and Smith, B. (1995) *500 Tips on Assessment.* London: Kogan Page.

Brown, S. and Race, P. (1997) Staff Development in Action. *SEDA Paper 100.* Birmingham: SEDA Publications.

Fry, H., Ketteridge, S. and Marshall, S. (1999) *A Handbook for Teaching and Learning in Higher Education - Enhancing Academic Practice.* London: Kogan Page.

Gibbs, G. (1992) *Assessing more students.* Oxford Brookes University: Oxford Centre for Staff Development.

Gibbs, G. (1999) Using assessment strategically to change the way students learn. In: S. Brown and A. Glasner (eds.) *Assessment matters in higher education-choosing and using diverse approaches.* Buckingham: Open University Press.

Knight, P. (ed) (1995) *Assessment for Learning in Higher Education.* SEDA Series. London: Kogan Page.

Race, P. (ed.) (1999) *2000 Tips for Lecturers.* London: Kogan Page.

Race, P. (2000) 500 *Tips on Group Learning.* London: Kogan Page.

Race, P. (2001) *The Lecturer's Toolkit.* (2nd ed) London: Kogan Page.

Ramsden, P. (1992) *Learning to teach in higher education.* London: Routledge.

Rowntree, D. (1989) *Assessing students: how shall we know them.* (2nd ed) London: Kogan Page.

The Learning and Teaching Support Network Generic Centre

The Learning and Teaching Support Network (LTSN) is a network of 24 Subject Centres, based in higher education institutions throughout the UK, and a Generic Centre, based in York, offering generic information and expertise on learning and teaching issues that cross subject boundaries. It aims to promote high quality learning and teaching through the development and transfer of good practice in all subject disciplines, and to provide a 'one-stop shop' of learning and teaching resources for the HE community.

The Generic Centre, in partnership with other organisations, will broker information and knowledge to facilitate a more co-ordinated approach to enhancing learning and teaching. It will:
- Work with the Subject Centres to maximize the potential of the network;
- Work in partnership to identify and respond to key priorities within the HE community;
- Facilitate access to the development of information, expertise and resources to develop new understandings about learning and teaching.

The LTSN Generic Centre Assessment Series Guides for:
Senior Managers
Heads of Department
Lecturers
Students

Briefings:
Assessment issues arising from key skills
Assessment of portfolios
Key concepts: formative and summative, criterion and norm-referenced assessment
Assessing disabled students
Self, peer and group assessment
Plagiarism
Work-based learning
Assessment of large groups

Published by
Learning and teaching support Network (LTSN), Genesis 3, York Science Park, York, YO10 5DQ

For more information, contact the Generic Centre at the above address or
Tel: 01904 434794   Fax: 01904 434247
Email: gcenquiries@ltsn.ac.uk
www.ltsn.ac.uk/genericcentre