Higher Education Research & Development, Vol. 20, No. 1, 2001

Assessing Self- and Peer-assessment: the students' views

STEPHANIE J. HANRAHAN
The University of Queensland

GEOFF ISAACS
The University of Queensland

ABSTRACT *Self- and peer-assessment are being used increasingly in higher education, to help assign grades to students' work and to help students to learn more effectively. However, in spite of this trend there is little in the published literature on how students view these methods. In this paper we present an analysis of the views of a large number of students (*N = *233) who had just experienced self- and peer-feedback as part of one of their subjects. It is a rarely questioned commonplace in the literature that in order to gain benefit from peer and self-assessment schemes students first need training in the specific scheme being used; ideally they will play a role in devising the scheme. The intervention reported here, which involved a large (*N = *233) group of students, included no such measures. The results show that students felt, nonetheless, that they benefited from the intervention. The results also present* prima facie *evidence that training or other measures to further involve the students in the peer and self-assessment scheme might be beneficial. Our analysis of students' views revealed eight general dimensions under which are grouped twenty higher order themes. The results both support and extend previous research and give a more detailed picture than previously available. The general dimensions found were: Difficult; Gained Better Understanding of Marking; Discomfort; Productive (including learning benefits and improved work); Problems with Implementation; Read Others' Work; Develop Empathy (with assessing staff); and, Motivation (especially motivation to impress peers). The practical implications of these findings are discussed.*

Introduction

As emphasis in university education has switched from teaching to learning and from teacher management to student self-direction (Argyris & Schön, 1974; Boud, 1995; Boud & Feletti, 1997; Knowles, 1984; Rogers & Freiberg, 1994; Schön, 1987), interest has mounted in the educational advantages of students assessing their own work (self-assessment) and that of other students (peer-assessment). Nowadays most degree courses have goals such as "students will become lifelong learners" and "students will be able to function effectively in teams". Such new goals reflect changing expectations of graduates in the workplace and are consonant with

ISSN 0729-4360 print; ISSN 1469-8360 online/01/010053-18 © 2001 HERDSA
DOI: 10.1080/07924360120043658

the newer approaches to education mentioned earlier. Both self- and peer-assessment skills are needed by graduates in the course of their working (and private) lives. Self-assessment will help students to set goals and thus to learn for themselves. Peer- assessment will help them to contribute constructively in collaborative efforts. As the climate in educational institutions moves steadily in the direction of increased efficiency (i.e., reducing the cost of educating students, frequently by decreasing staffing) self- and peer-assessment also are increasingly under the microscope for reasons other than the purely academic. Arguably they are ways of assessing students' work either for feedback or for grading purposes while minimising the cost in staff time. Efficiency is of special interest to administrators; ideally effectiveness (here the capacity of peer- and self-assessment to enhance students' learning) will be of at least equal interest to them and will be the foremost concern of teaching staff.

The purpose of the study reported here was to develop understanding of tertiary students' perceptions of the benefits and problematic aspects of peer- and self-assessment. Once these perceptions are understood, it may be possible to develop assessment techniques that maintain the perceived positives and limit the perceived negatives. The process may also introduce issues that previously have not been considered in the literature.

One issue that may be further illuminated by such an analysis is the benefit to be gained by training students in the assessment techniques to be used, or even in involving them in the development of such techniques. It is generally acknowledged and asserted that such activity is necessary. To take but a few examples, Boud (1989) expressed concern that marks derived from self-assessment may be unreliable and therefore recommended training through practice. He also pointed out that often staff did not do this for their *own* marking. Mowl and Pain (1995) emphasised "The research [reported in their paper] shows that even with subjective methods of assessment such as essays, students are generally capable and conscientious self- and peer-assessors, as long as they are adequately prepared and reassured about the value of the exercise". In fact the research in their paper seems to indicate that such preparation helps, but does not prove that it is necessary. Oldfield and MacAlpine (1995) commented: "Experience had led us to believe that, in a new situation, students must have concepts introduced to them in absorbable and achievable steps, they must receive understandable feedback at each stage and their confidence must be built from experience".

In essence the literature referring to training for peer- and self-assessment falls into two classes: those papers that argue that students ought to be taught peer- and self-assessment skills because they are useful life skills, and those who assert the seemingly obvious point that peer- and self-assessment are likely to be more reliable (and, possibly more valid) if students are trained in the use of the relevant techniques. The first argument is only tangential to the present study, as it concerns more general skills than those specialised examples involved in assessing achievements in a formal course. In the second case, as noted above, we doubt that the point has been conclusively proved. And, even if it were proven, the question remains as to whether the size of the problem is significant.

Much of the literature on peer-assessment in higher education is concerned either with the assessment of individual contributions to group work (Conway, *et al*., 1993; Earl, 1986; Freeman, 1995; Goldfinch & Raeside, 1990; Rafic & Fullerton, 1996) or with the extent to which such marks may be valid for grading purposes (Falchikov, 1986; Stefani, 1992, 1994). The literature on self-assessment, well reviewed in Boud's recent book (1995), frequently shows similar emphases. Of interest to us however are the purported benefits and difficulties of peer- and self-assessment from the students' point of view. Here the coverage is somewhat sparser.

Falchikov (1986) carried out a small study ($N = 48$) of peer-, self- and tutor-assessment of essays. Her study looked at students' views of peer- and self-evaluation schemes in which criteria were negotiated among staff and students, the marks gained counted towards the students' final grades and participation was mandatory. Semantic differential items were used to look at students' feelings both about the effects the peer- and self-assessment schemes had on them and about the schemes themselves. The most clear cut opinions were that students felt both schemes made them think, learn more, be critical, and be structured (in this last case much more so for self- than for peer-assessment). The schemes themselves were seen as time consuming, hard (especially self-assessment), challenging, helpful and beneficial. Stefani (1992, 1994) with a slightly larger group of students (between 54 and 67 respondents) essentially reproduced these results, albeit with greater agreement among her group. Case studies, largely of self-assessment, reported by Boud and his co-authors (Boud, 1995) also tend to confirm this picture.

Some qualitative data were reported by Falchikov (1986), who asked students to write down what they liked best and least about "this system" (of peer- and self-assessment), how knowledge of participation affected the student's writing process, and where they found difficulty with the application of the criteria and the like. Her report, however, covers principally students' likes and dislikes and issues concerned with criteria. Influence on the writing process is, unfortunately, reported only as a percentage and a quote—47% of students reported they were influenced; "It made you more aware of what you were writing, i.e. it wasn't just a case of getting the information down—you had to plan it".

The best-liked feature of the system that Falchikov (1986) investigated (reported by 36% of her 48 students) was the "provision of an outline [of the essay to be written] as an aid to writing". Next most popular was "increased awareness" [of what the task of essay writing actually is] (19%) and "benefits of reading a peer essay" (17%). Other features mentioned favourably by four or five of Falchikov's students were: "Less biased mark results" (11%); "Learning about mistakes and possibility of subsequent improvement" (8%); "Guidelines for marking" (8%).

There was less agreement among Falchikov's students as to the *least* liked features of the system. "Difficulty of task", divided into "Lack of knowledge of peer topic" (23%) and more general difficulty (14%) dominated, with only "Weightings were wrong" (14%), "possibility of marking down/failing a peer" (11%) and "System was too rigid/clinical" (9%) being cited by more than two students.

Thus, at present, we have the picture, based principally on structured question-

naire data, that students do see learning benefits in peer- and self-assessment. The schemes themselves they tend to see as beneficial, but difficult, challenging and time consuming. It seems appropriate, then, that we look at the unprompted views of a larger number of students on the features of self- and peer-assessment. This examination will be able to confirm the more tentative results of previous studies. However it will also yield an organised framework of categories (some of them new to this context) which are useful in generating greater insights into peer- and self-assessment and which will give a firm basis for organising and comparing results among researchers in the future.

Method

Participants

Participants were students in a third-year tertiary health psychology subject. The 244 students in the class were enrolled in the following courses:

Bachelor of Arts	139
Bachelor of Science	53
Bachelor of Applied Science (Human Movement Studies—Exercise Management)	31
Other degree programs	21

Of the 244 students enrolled in the subject, 233 participated in the study (a response rate of 95%).

Structure of the Subject

The subject under consideration here is a 10 credit point subject which would constitute roughly one quarter of a student's load during the semester in which it is studied. The subject was organised so that each week there was one two hour lecture session given to the whole class and one tutorial group meeting (roughly 25 students in each tutorial) per week. Students were also assigned to learning groups which were the unit for doing some assigned work. Learning groups were also designed to help students to gain support from each other and to avoid students becoming isolated from their peers. Degrees at this university are modular and it is not uncommon to find two students doing a particular subject who have no other subjects in common.

Assessment was on the basis of six tasks to be completed by learning groups (25% of total marks—including 1% for individuals who answered some reflective questions on how their learning groups were functioning), an assignment in the form of a research essay (25%), and a mid-semester examination and a final examination worth 50% in total (students had some control over how their possible marks were divided between mid semester and end of semester examinations).

Procedure

Given the size of the group and the limited time available for tutorial support it was judged to be impractical to develop peer- and self-assessment criteria collaboratively. Similarly, the workload involved for staff ruled out any practice marking before the "real thing". We therefore chose to instruct students clearly in writing about the process to be used and to discuss the marking criteria in tutorials.

We opted to make the process a gatekeeper task that had to be completed before the relevant assignment was counted in the final assessment for the subject. The goal was to expose students to peer- and self-assessment and to encourage them to give careful consideration to their work before submitting it, and to give feedback to their colleagues. Students were encouraged to assess and, if possible, improve their own work before submitting it for assessment by either the tutor or their peer.

The assessment task used was the writing of a 1500 word research essay worth 25% of the final grade. The subject manual included two pages describing the purpose, topics, format, and evaluation criteria for the assignment. The manual also informed the students that, "You will be assessing your own assignment as well as the assignment of another student (to be given to you the week of March 25th). Your self-assessment needs to be submitted no later than March 27th. Your peer-assessment is due April 3rd. You will not get credit for your assignment until sincere self- and peer-assessments have been submitted". Students were provided with the same marking sheet as their tutors (Appendix 1). This sheet broke down the task into the assessment of six aspects of the work, detailed the marks allocated to each aspect and the criteria to be used in marking it, and provided space for comment on each aspect as well as a mark for that aspect.

Students submitted two copies of their assignment. Tutors marked one, the other was given to a student to mark for peer-assessment after first removing, as much as possible, any identifying information on the assignment. To decrease the chances of students recognising the work of people they knew, no students assessed assignments for peers who were in their own tutorial group. Only the marks given by the tutors counted towards final grades.

Approximately two weeks after students received the marking sheets from both the tutors and their peers, they were asked to answer four questions in writing. The first three questions referred to learning groups that they were using within tutorials and are not relevant to the study at hand. The fourth question was, "What do you think were the pros and cons of doing peer and self-assessment on the essay assignment?" Students received 1% credit towards their final grades for completing the four questions.

Data Analysis

All of the responses to question 4 were collated verbatim resulting in 41 pages of single-spaced text. Raw data themes were identified via content analyses (Patton, 1987). Raw data themes, in the form of direct quotations, served as the primary unit

of analysis. A hierarchical inductive analysis was then conducted. One author grouped the original raw data themes into higher order themes that shared similar meaning. This author then gave a descriptive name to each theme. The second author then looked at the higher order themes without the assigned names and described them. These descriptions were then compared with those of the first author. There was a high degree of agreement. This provides some evidence of the validity and reliability of the grouping into themes. The higher order themes were then examined for similarities for further combination as dimensions. When higher order themes could not be meaningfully grouped into dimensions, they were carried forward independently. Eventually the original raw data themes (direct quotations from students' comments) were reexamined to ensure that the meaning associated with raw data themes had not been misconstrued in the analytic process. The assignment of higher order themes to dimensions was carried out jointly by the authors and the checking of the original data was done separately by each author and then discussed until consensus was reached.

Results

The inductive content analysis of the benefits and problematic aspects of peer and self-assessment revealed eight general dimensions: Difficult, Gained better understanding of marking, Discomfort, Productive, Problems with implementation, Read others' work, Develop empathy, Motivation. The higher order themes and selected raw data themes for each dimension are set out below.

General Dimension: difficult

Four higher order themes appeared under this dimension: "Difficult to be objective", "No experience with marking/not sure of standards", "Unfamiliar with articles in other areas", "Own assignment already meets the criteria". Examples of students' comments for each of these themes are given in Table 1.

The higher order themes "Difficult to be objective" and "No experience with marking/unsure of standards" are applicable to both peer- and self-assessment. The theme "Unfamiliar with articles in other areas" applies only to peer-assessment, while the theme "Own assignment already meets the criteria" applies only to—assessment.

General Dimension: gained better understanding of marking

Two higher order themes appeared under this dimension: "Productive self-critique", and "Reinforced marking procedure". Examples of students' comments for each of these themes are given in Table 2. Both higher order themes are applicable to both peer- and self-assessment.

General Dimension: discomfort

Three higher order themes appeared under this dimension: "Uncomfortable having peer read own paper", "Peers can be too critical", "Uncomfortable critiquing others'

TABLE 1. General dimension: Difficult—higher order themes and illustrative raw data themes

Selected raw data themes	Higher order theme
With both peer- and self-assessment, it was hard to mark because firstly you don't want to mark yourself too low because you want to think that you have done a good assignment and then also you don't want to mark a fellow student too harshly.	Difficult to be objective
I think it's very difficult to be objective on both accounts—obviously you are going to be relatively easy on yourself, even if you think you are being objective.	
A student, with individual exceptions I am sure, is not a credible grader. It may be good practice for a student to evaluate or critique another's essay, but this student would not know how they compared to the tutor. This evaluating student would not know if his or her grading was proficient.	No experience with marking/not sure of standards
A con would be that the other marker in peer-assessment has no real marking experience so may mark somewhat inaccurately.	
On peer-assessment, I thought it was hard to interpret the analyses of others and their accuracy, because I had not read the texts that she had.	Unfamiliar with articles in other areas
It would have been much easier for the peer-assessors if the assignment they had to mark was of a similar topic to their own. This would make it easier to critically evaluate the writing, especially whether or not source articles have been scrutinised for their strengths and weaknesses.	
Self-assessment may not be very useful, because the assignments I have handed in are usually the best that I can produce, so I would find it hard to mark my own assignment.	Own assignment already meets the criteria
The disadvantages are that as the piece is in fact your own automatically you believe that you meet the criteria and may in fact fall short in others eyes.	

work". Examples of students' comments for each of these themes are given in Table 3. All three of the higher order themes under the Discomfort dimension are applicable only to peer-assessment. Students mentioned no discomfort at assessing their own papers.

TABLE 2. General dimension: Gained better understanding of marking—higher order themes and illustrative raw data themes

Selected raw data themes	Higher order theme
You realise what markers are looking for (a new experience for me and very valuable) and are forced to acknowledge whether or not the factors which must be in your essay are present. It helped me see more clearly some of the skills I need to focus on in my essay writing. There were no cons, it just helped me identify necessary structures and approaches in my own work.	Productive self-critique
It is a good idea to include the peer-assessment on the essay assignment because we can understand the strengths and weaknesses of our own assignment better when viewing the others assignment on similar topic. We can also understand our essay better by the different point of view from peers.	
More aware of the marking criteria prior to writing the assignment, so this was helpful in formulating headings and content. This also assisted with knowing weightings for certain sections.	
Peer-assessment reinforced assignment marking procedure.	Reinforced marking procedure

General Dimension: productive

Three higher order themes appeared under this dimension: "Get more feedback", "Improved own assignment prior to submission", "Helps develop critical thinking". Examples of students' comments for each of these themes are given in Table 4.

The higher order theme "Get more feedback" is applicable only to peer- assessment, while the theme "Improved own assignment prior to submission" is applicable only to self-assessment. However the other theme "Helps develop critical thinking" applies to both peer- and self-assessment.

General Dimension: read others' work

Two higher order themes appeared under this dimension: "Learn what others are doing", "Need to see good and bad work". Examples of students' comments for each of these themes are given in Table 5. It is implicit in the nature of this dimension that all of the higher order themes, concerned as they are with reading other students' work, will be relevant only to peer-assessment.

General Dimension: develop empathy

Only one higher order theme appeared under this dimension: "Develop empathy with lecturers/tutors". Examples of students' comments for this theme are given in Table 6. This higher order theme, as expressed in the examples, arises from

TABLE 3. General dimension: Discomfort—higher order themes and illustrative raw data themes

Selected raw data themes	Higher order theme
I felt uncomfortable about another peer reading my work and even though it was anonymously marked I still felt pressured and awkward while writing my assignment.	Uncomfortable having peer read own paper
Many students, like myself, feel uncomfortable having another student evaluate his or her paper (It was a great and simple idea to have it done anonymously).	
Cons: they were very critical (more so than the tutor).	Peers can be too critical
My own experience is that assessing others work gave me a good feel of my own work. However, I felt I was assessed savagely by my peer.	
The undesirable task of picking another student's work to pieces, and the thought of bringing their marks down.	Uncomfortable critiquing others' work
As the person who wrote the essay had a much more thorough knowledge than myself, I feel that it is almost inappropriate for me to be fairly critical in my own evaluation of the essay.	

TABLE 4. General dimension: Productive—higher order themes and illustrative raw data themes

Selected raw data themes	Higher order theme
More feedback is better for students.	Get more feedback
Get feedback from others.	
Helps you to examine your own work more thoroughly and to follow the criteria of the assignment more closely.	Improved own assignment prior to submission
Self-assessment made me more aware of what I needed to do to improve my assignment.	
Peer-assessment allowed you to view another persons assignment which helps to develop skills for critical thinking as well as comparing own work.	Helps develop critical thinking
Doing a self-assessment made me look at my assignment more critically then I normally would have.	

TABLE 5. General dimension: Read others' work—higher order themes and illustrative raw data themes

Selected raw data themes	Higher order theme
A good point was that the peer-assessment allowed you to read someone else's assignment where you previously would have no idea about fellow students writing styles and ideas.	Learn what others are doing
Gives a good idea of the standards of others work.	
Assessing a poor assignment is difficult because 'things' don't flow and you have to read certain parts over and over to understand and it gets rather confusing. It would be good to read a well-written, well-thought out assignment and a poor one, then compare and contrast. That would be far more educational.	Need to see good and bad work
I had no idea of what I was doing—having nothing to compare it to apart from mine.	

TABLE 6. General dimension: Develop empathy—higher order theme and illustrative raw data themes

Selected raw data themes	Higher order theme
Peer-assessment was a very good idea - > it gives you insight into the difficulty tutors experience in assessing students. It is easier to relate to them as well.	Develop empathy with lecturers/tutors
I was astounded that despite the fact that all students were required to complete a self-assessment and must at least, have been at 3rd year level, that such poor quality work could be produced. I felt sympathy for the instructors who had to mark large numbers of assignments—it must be quite disheartening.	

TABLE 7. General dimension: Motivation—higher order theme and illustrative raw data themes

Selected raw data themes	Higher order theme
The fact that my peers were marking my assignment encouraged me to put more effort in.	Motivated to impress peers
Peer-assessment made me work harder on my own assignment in order to make it more understandable and interesting for the reader.	

students' experiences of peer-assessment. It is interesting that no students said that empathy with their peers was an outcome of the experience!

General Dimension: motivation

Again only one higher order theme appeared under this dimension: "Motivated to impress peers". Examples of students' comments for this theme are given in Table 7. This theme is (and is intrinsically) concerned only with peer-assessment.

General Dimension: problems with implementation

Four higher order themes appeared under this dimension: "Time-consuming", "Process not taken seriously/doesn't count for marks", "Should get feedback for peer/self-assessment", "Peer assessment too late to be useful". Examples of stu-

TABLE 8. General dimension: Problems with implementation—higher order themes and illustrative raw data themes

Selected raw data themes	Higher order theme
Can be slightly time consuming, especially when there are a lot of other assessments due from other subjects.	Time-consuming
Reading and trying to give comments are time-consuming. (Note: A responsible person would not just read it ONCE & give comments without trying to understand the rationale of the structure and contents.)	
Some just pick out average scores so there is little effort put into the assessment (get it over and done with).	Process not taken seriously/doesn't count for marks
I don't believe that the peer- & self-assessments were taken seriously enough as they were not worth marks.	
No feedback as to comments or marks provided by self. Co-incidentally, my marks provided by student who marked my assignment were nearly the same as tutors, but no comments were made about this assessment either.	Should get feedback for peer-/self-assessment
No feedback was provided on marking process or comments made on the peer-assessment.	
I didn't find the peer-assessment valuable because as it was done after our own assignments had been handed in (necessarily, I know), we couldn't apply anything we may have learnt from marking somebody else's work to any assessment during the course.	Peer-assessment too late to be useful
It would have been nice to have the feedback of a peer before the essay had to be handed in.	

dents' comments for each of these themes are given in Table 8. The first three themes are relevant to both peer and self-assessment, while the last is obviously relevant only to the way in which peer assessment was implemented.

Discussion

It is important to note that our concern here has been to identify *themes* and *dimensions* underlying students' comments, rather than with how widespread comments along these lines might be. The dimensions found then lead to issues for further consideration and, in some cases for possible action or further investigation. The results from this study support and extend previous research by Falchikov (1986), Stefani (1992, 1994) and Boud (1995). However the qualitative data in the present study, which are derived from responses freely written by students (rather than selections from Likert or similar types of question), give a more detailed picture of the benefits and drawbacks students saw in peer- and self-assessment.

The dimensions themselves have been detailed in the Results section. Their implications are addressed here.

Training and Practice in Assessment

The dimension Difficult reveals three themes concerned with students' uncertainty about their ability to mark ("Difficult to be objective", "No experience with marking/not sure of standards", "Unfamiliar with articles in other areas"). At the very least, the concerns expressed in the first two might be addressed by some training in assessment and by practice in the use of the marking scheme. The Discomfort some students feel about peer-assessment seems addressable by practice and by training. Two of the higher order themes in this dimension seem to be about students' sensitivities—to exposing their own work and to being critical of others'. Such sensitivities might be addressed by desensitisation; that is, by practice in peer-assessment. The third higher order theme, that "Peers can be too critical", might also be addressed by practice and by training. Students' responses on this dimension, like their responses on the dimension Difficult, provide a prima facie case for training in peer-assessment.

Practical and Learning Issues

The third theme under Difficult ("Unfamiliar with articles in other areas") is a different kind of concern and shows that novices will have problems marking in some cases. Staff are expected to know or to get to know the area; are we to expect the student to get to know the ground covered by a peer's essay as well as (in both senses) the ground covered by their own? Or might the problem be better addressed (at some cost in administration) by making sure students only assess work on the same topic as their own? The first solution may have learning benefits for students (indeed, one might choose to make a virtue of necessity by *requiring* students to mark an essay on a topic other than their own), while the second may help to keep the assessment workload in bounds.

Learning Benefits

The other theme under Difficult, "Own assignment already meets the criteria", may for some students represent reality—given that they had the assessment criteria available while writing their essay they may well have engaged in a process of self-assessment as they wrote the essay. On the other hand, as one student said, "… the assignments I have handed in are usually the best I can produce …". That is, self-assessment may not lead to an improved essay; however it may lead to an enhanced insight into the strengths and weaknesses of the essay the student will submit. Perhaps some students need to be made more aware of the possible learning benefits.

Gained better understanding of marking has as its main them "Productive self-critique". The comments here illustrate students' appreciation of the learning benefit to be gained in analysing work according to clearly stated criteria and also the added benefit gained when one's own essay was seen in the context of another. Given the isolation of some students in diverse modular degree courses, peer-assessment may be the only opportunity they have to see the work of other students. That the "seeing" in such a case is enhanced by the structure of assessment criteria is a bonus.

The nature of the other benefits students claimed to have gained is quite extensive and is reflected in the dimensions Productive, Read others' work and Develop empathy. Productive benefits include the instrumental benefit of getting feedback, the practical one of improving one's own assignment (because of self-assessment), and the more abstract one of enhanced critical thinking. Under Read others' work the benefits amount to gaining an appreciation of the standards required and of what might be achieved. The benefit under Develop empathy is that students may come to appreciate the difficulties their teachers have in carrying out assessments. Such a benefit is relevant to students' lifelong learning; especially to their learning to be responsible individuals. Eventually students leave their dependent role in the university and must become more independent agents in the community. For most this role change happens almost overnight at the point of graduation. If the assessment process helps them to gain some appreciation of the role of a teacher this is likely to help them achieve the change.

Implementation Problems and Opportunities

Of course, as the dimension Problems with implementation shows, students may see problems with the way these processes are put into practice. That some students find these processes time consuming might perhaps be interpreted that some students feel they do not get enough learning benefit from the time cost of (usually) peer- and (sometimes) self-assessment. The feeling of some that there should be feedback on students' attempts at assessment may be taken as a further justification of the need for training; alternatively it could be seen as an opportunity to further integrate these assessment activities into students' learning. The idea that peer-assessment is not taken seriously because it does not count for marks has two (in some cases related) aspects: markers not putting significant effort into the process and recipients of the

marks not taking note of them. Unfortunately, at present students in the Australian university system generally seem so pressured that their learning activities are almost completely driven by assessment. Whether teachers should collaborate (collude?) in this process by allocating marks to all tasks they seriously want completed, or whether they should resist it is an issue that might profitably be discussed elsewhere.

Practical Implications

In the process of carrying out the present study, we have demonstrated that peer- and self-assessment are feasible in classes with over 200 students. However, depending on how they are implemented, the administrative load can be fairly heavy. In our case the load was over 40 person hours. Most of this load was incurred in running a system that preserved the anonymity of both the assignment writer and the assignment marker, while still allowing teaching staff to track the process. This enabled us to know who had completed assessing a peer, and to return the peer-assessed assignment to the original author, while at all times keeping the identity of the peer-assessor and the author confidential. Relaxing any of these requirements would lessen the administrative load. Nonetheless, the larger the class, the greater the load.

Some of the problems with the process noted are remediable, whereas others may not be. Providing clearer and more detailed standards may lessen students' uncertainty about the standards required. Similarly, providing students with the opportunity to practice assessing assignments may enhance their self-efficacy and make them feel they are experienced. Unfortunately, this process is also likely to result in significantly increased workload for the students and a major increase in administrative workload for the staff.

Making other changes to the implementation process may help overcome some of the negatives reported by students in this study. For example, providing exemplars of good work or supplying completed assignments of varying quality may aid the assessment process as well as aid learning in its own right. Changing the timing of the peer-assessment may help to reduce the administrative load while maintaining the learning benefits of the process. Encouraging students to establish writing partnerships where they provide feedback to each other prior to the submission of the assignment would allow them to make changes before marks are allocated.

For practical purposes, getting students to take peer-assessment seriously implies adopting, at the very least, a criterion-referenced marking scheme (as used in this case) and preferably a fully-fledged criterion-referenced assessment system. Without a criterion-referenced marking scheme students would have no rational basis on which to judge their own or others' work. Moreover, under a norm-referenced system (not used in this case), students are in a sense competing with each other, so there is a disincentive for them to give helpful feedback to their peers.

The Future

The data used in this paper are statements made by students responding freely to a straightforward question about the pros and cons of self- and peer-assessment. Their

responses were in no way prompted. This freedom, especially given that there were a large number of respondents, enhances the value of the results. The results, when summarised as dimensions and higher order themes, are likely to be of considerable use to help future investigators construct structured items to look at this area.

This research has served to delineate clearly the dimensions of students' perceptions of peer- and self-assessment. Further studies should verify these findings and refine them further. A major benefit of peer- and self-assessment processes demonstrated in this study is the positive effect they have on students' learning. This effect in particular might profitably be investigated further, especially its generality across subject areas and types of assessment. Both peer- and self-assessment obviously have effects on students, but we do not know either the details or the extent of these effects. For example, do students' general expectations of assessment change as a result of the experience? Having experienced the benefits of self-assessment in one subject, do they then proceed to apply the method themselves to improve their work in other subjects? Do they expect or request that these methods be used in other subjects?

The research issues above are fairly general and, in addition, have clear implications for both theory and practice. However, peer- and self-assessment still will need to be implemented on a case-by-case basis in varying subjects and contexts. The "case based" literature in this area is still alarmingly sparse. In our opinion it would be worthwhile to carry out research, probably using the action research model, looking at ways, within specific subjects, of maintaining the learning benefits mentioned while minimising the problem areas. Such research would have clear practical benefits for university students and teachers and might also serve to illuminate the more general research issues.

Address for correspondence: Stephanie J. Hanrahan, School of Human Movement Studies and School of Psychology; The University of Queensland, St Lucia, QLD 4072, Australia. E-mail: steph@hms.uq.edu.au

References

ARGYRIS, C., & SCHÖN, D. A. (1974). *Theory in practice: Increasing professional effectiveness.* (1st ed.). San Francisco: Jossey-Bass.

BOUD, D. (1989). The Role of Self-Assessment in Student Grading. *Assessment and Evaluation in Higher Education*, *14*, 20–30.

BOUD, D. (1995). *Enhancing learning through self-assessment.* (1st edn). London: Kogan Page.

BOUD, D., & FELETTI, G. (1997). *The challenge of problem based learning.* (2nd edn). London: Kogan Page.

CONWAY, R., KEMBER, D., SIVAN, A., & WU, M. (1993). Peer assessment of an individual's contribution to a group project. *Assessment and Evaluation in Higher Education, 18*, 45–56.

EARL, S. E. (1986). Staff and peer assessment—measuring an individual's contribution to group performance. *Assessment and Evaluation in Higher Education, 11*, 60–69.

FALCHIKOV, N. (1986). Product comparisons and process benefits of collaborative peer and self-assessments. *Assessment and Evaluation in Higher Education, 11*, 146–166.

FREEMAN, M. (1995). Peer assessment by groups of group work. *Assessment and Evaluation in Higher Education, 20*, 289–300.

GOLDFINCH, J., & RAESIDE, R. (1990). Development of a peer assessment technique for obtaining individual marks on a group project. *Assessment and Evaluation in Higher Education, 15*, 210–231.

EARL, S. E. (1986). Staff and peer assessment—measuring an individual's contribution to group performance. *Assessment and Evaluation in Higher Education, 11*, 60–69.

FALCHIKOV, N. (1986). Product comparisons and process benefits of collaborative peer and self-assessments. *Assessment and Evaluation in Higher Education, 11*, 146–166.

FREEMAN, M. (1995). Peer assessment by groups of group work. *Assessment and Evaluation in Higher Education, 20*, 289–300.

GOLDFINCH, J., & RAESIDE, R. (1990). Development of a peer assessment technique for obtaining individual marks on a group project. *Assessment and Evaluation in Higher Education, 15*, 210–231.

KNOWLES, M. (1984). *The adult learner: A neglected species.* (3rd ed.). Houston: Gulf.

MOWL, G., & PAIN, R. (1995). Using self and peer assessment to improve students' essay writing: a case study from Geography. *Innovations in Education and Training International*, *32*, 324–335.

OLDFIELD, K. A., & MACALPINE, J. M. K. (1995). Peer and self-assessment at the tertiary level: An experiential report. *Assessment & Evaluation in Higher Education*, *20*, 125–132.

PATTON, M. Q. (1987). *How to use qualitative methods in evaluation*. Newbury Park, CA: Sage.

RAFIC, Y., & FULLERTON, H. (1996). Peer assessment of group projects in civil engineering. *Assessment and Evaluation in Higher Education, 21*, 69–81.

ROGERS, C. R., & FREIBERG, H. J. (1994). *Freedom to learn*. (3rd edn). Columbus: Merrill.

SCHÖN, D. A. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions*. (1st edn). San Francisco: Jossey-Bass.

STEFANI, L. A. J. (1992). Comparison of collaborative self, peer and tutor assessment in a biochemistry practical. *Biochemical Education, 20*, 148–151.

STEFANI, L. A. J. (1994). Peer, self and tutor assessment: relative reliabilities. *Studies in Higher Education, 19*, 69–75.

PY352 ASSIGNMENT: RESEARCH ESSAY
MARKING SHEET

Marker (circle one of the following): Self Peer Tutor
Marker's student number: _____
Student number on assignment: _____

1. Organisation • Logically organised throughout, sub-headings to further organise (where appropriate) • Clear introduction, stating purpose and content of paper • Clear conclusion; brings in no new ideas Comments:_____	Sub-total: /5
2. Depth of Analysis • a number of sources utilised • includes up-to-date sources • references cited have been thoroughly explored • text backed up with relevant references Comments:_____	Sub-total: /5
3. Evidence of Critical Thinking • analysis of information (i.e., not simply reported) • sources scrutinised for strengths and weaknesses • personal opinions backed up with relevant references (where appropriate) Comments:_____	Sub-total: /10
4. Coherence • sources are not presented in a book report fashion • each paragraph is logically related to the one preceding it; the essay flows well • material; from different sources are related and integrated logically Comments:_____	Sub-total: /10

5. Quality of Writing • succinct • technically correct: spelling, grammar, punctuation Comments:_____	Sub-total: /5
6. Presentation Format • conforms to APA guidelines • neat, double-spaced, stapled, numbered pages, title page Comments:_____	Sub-total: /5
Exceeds word limit:	
TOTAL MARK:	

Additional comments:_____