SELF AND PEER ASSESSMENT AS AN ASSESSMENT TOOL IN PROBLEM-BASED LEARNING

Kelvin Tan

Learning Academy

Leng Ho Keat

Diploma in Marketing
Temasek Business School

Student self- and peer-assessment and Problem-based Learning have become increasingly popular in higher education. The evolving needs of the global employment market have encouraged universities to train their students to take responsibility for their own learning and continuous personal and professional development. Problem-based Learning has been recommended as an effective methodology enabling students to assume that responsibility. This paper reports the findings of an investigation into the quality of students' self- and peer-assessment in Problem-based Learning. It seeks to understand the uses and limitations of self- and peer-assessment as an assessment tool in problem-based learning contexts, through an examination of the relationships between self- and peer-assessment scores on the one hand, and tutor assessment on the other.

INTRODUCTION

The evolving needs of the global employment market have encouraged universities to train their students to take responsibility for their own learning and continuous personal and professional development. Likewise, Problem-based Learning has been recommended as an effective methodology for enabling students to assume that responsibility. In the problem-based learning context, the learning and assessment processes focus strongly on student group work.

The use of student groups in marketing classes as well as in many other disciplines is now an accepted practice in many colleges and universities (Haas *et al*., 1998). At the same time, self- and peer-assessment are increasingly used as assessment tools. This move has been driven by the expectation that self- and peer-assessment are more reflective of student performance, since individual group members spend substantial time working with one another and are therefore in a good position to recognize and assess the efforts and contributions of their peers (Ghorpade & Lackritz, 2001; Haas *et al.*, 1998).

Self- and peer-assessment are also commonly used for courses based on Problem-based Learning because, in addition to the benefits identified above, they can also be used to assess students' problem solving, teamwork and self-directed learning skills (Wee & Kek, 2002). However, their effectiveness and limitations as assessment tools have yet to be established, as the literature indicates that self- and peer-assessments can be subjective. The aim of this research is thus to help establish what their limitations might be.

LITERATURE REVIEW

Savin-Baden (2004) noted that assessment is one of the most controversial concerns in Problem-based Learning. In Problem-based Learning courses, student learning occurs mostly through group interaction. However, this is often not rewarded in academic terms. Effort and time spent on the subject did not necessarily translate into better academic grades. Self- and peer-assessment have been implemented in many courses in an effort to recognise the learning process in groups.

The literature on self- and peer-assessment may be understood as falling within three broad categories – in terms of their use as an assessment tool, in terms of enhancing learning, and in terms of practices and meanings of power. The common denominator of these three areas is that there is a degree of student involvement, and this involvement takes the form of making judgements concerning their learning.

As an assessment tool, self- and peer-assessment is generally depicted as a practice in which students take on some, if not all, of the assessment responsibilities in their subject. The literature in this area frames student self-assessment in terms of grading, and emphasizes the level of agreement between the students' and the teacher's assessed score. In terms of learning, the literature on self- and peer-assessment deals with the involvement of students in making judgments to enhance their learning. Such literature describes student self-assessment practices in terms of the benefits that they bring to student learning. The emphasis is on how students' judgements of their learning are evaluated in terms of the consequences for learning. The contrasting notions of student self-assessment in terms of power describe and problematise the different forms, and consequences, of sharing power with students in assessment. Some of the literature questions whether empowerment can or should be defined as a commodity in quantifiable terms. This paper confines itself to the uses and understanding of self- and peer-assessment as an assessment tool in problem-based learning.

Self- and peer-assessment practices have been recognized as meeting a variety of extraneous needs in various contexts of higher education. Firstly, Boud (1989) observes that self-assessment practices may possess practical advantages for assessing large numbers of students, thereby allowing students to partially relieve their teachers of the time and effort of testing. The argument follows that if students were able to take on some assessment tasks, then academic staff would have more time to prepare and conduct educationally worthwhile learning and assessment activities for them.

Secondly, Taras (2001) alludes to increasing student consumerism manifesting itself in a greater demand for involvement and control of the assessment process. In the context of students being customers of an institution of higher learning, assessment grades "represent the final package that students want or expect to be delivered" (p. 612).

This emphasis on satisfying the student customer in the form of assessment grades creates a dilemma for self- and peer-assessment practices. On the one hand, teachers or academics may desire to provide opportunities for self-assessment to meet a variety of educational objectives. Such objectives may include enhancing the students' learning and developing the students' self-assessment ability. However, if consumer satisfaction was the ultimate determiner of students' grades, then their integrity and role in the assessment process may be questionable.

A third context for self- and peer-assessment practice is its potential for providing greater transparency in the assessment process for students. Self-assessment may allow students to familiarize themselves with the context(s) in which they are assessed. It allows students to experience and interpret the often arbitrary requirements which their work needs to satisfy. In this sense, self- and peer-assessment assist students to understand how they will be assessed by their teachers, by having their assessments compared with (and discussed in terms of) the teacher's assessment.

Hence, self- and peer-assessment practices may be viewed in terms of responding to extraneous factors such as increased and diverse class populations, in terms of increased student consumerist demands for accountability for their grades, and in terms of increased transparency and understanding in the assessment process leading to their academic results. The common pre-requisite for these contexts of self-assessment practice is that students must be able to prove themselves capable of grading their work accurately, in order to be relied upon by their teachers.

It is within these contexts of assessment that the reliability of self-assessed and peer-assessed grades assumes importance. It is generally approached in the literature in terms of comparing students' ratings of their work with the teacher's, and/or with each other's. And the "accuracy" of students' self-generated grades is the single most discussed issue in the literature of student self-assessment (Boud, 1995; Boud & Falchikov, 1989).

However, the main criticism of self- and peer-assessment as assessment tools stems from students' inability to assess themselves and each other objectively. As a result, students tend to under-mark or over-mark themselves when compared with tutor assessments. While findings are inconclusive as to whether students tend to *overrate* or *underrate* themselves when compared to

the tutor assessment, there is empirical evidence pointing to the fact that the magnitude of under-marking by students is greater than the magnitude of over-marking (Falchikov 1986; Segers & Dochy, 2001).

Haas *et al.,* (1998) noted that the literature seems to suggest that self-ratings may be more accurate than peer-ratings. When comparing peer- and self-assessment scores, over-marking by peers tends to be more prevalent than under-marking. In addition, the magnitude of over-marking is slightly higher than the magnitude of under-marking (Das, Mpofu, Dunn & Lanphear, 1998; Falchikov, 1986).

Self-ratings were higher for students with high esteem, because they attributed good performance to their own abilities. However, it was also true that students who did not contribute to the group (e.g. free riders or social loafers) tended to rate themselves more favourably in anticipation of lower ratings from their peers. Consequently, the ratings given were often inflated and unobjective. This suggests that students who are weaker academically may have greater motivation to inflate their self-ratings. However, research has shown that there is no significant relationship between student achievement scores and self- or tutor-assessment scores (Das *et al.*, 1998).

The above suggests that there is a difference between self-, peer- and tutor-assessment scores. This research aims to add to the literature on the relationship among the three variables. In addition, it attempts to explore the relationship between self- and peer-assessment scores and academic results.

METHODOLOGY

Much of the literature on student/teacher mark agreement in self- and peer-assessment practices relates to the context of Australian and British universities. This study adds to that body of empirical research by focusing on investigating student-teacher mark agreement in self- and peer-assessment practices in a Problem-based Learning course taught in a polytechnic in Singapore. The study involved 131 second-year marketing students from Temasek Polytechnic. As part of their curriculum, students attended a course on consumer behaviour delivered using Problem-based Learning. There were several assessment components, including a traditional content-based test and a group project based on a problem as well as a self- and peer-assessment component.

Students form groups of between four to six members and are given approximately three months to offer a solution to a problem posed by a real client. Students are briefed on the self- and peer-assessment process at the beginning of the term, and are also required to complete an evaluation midway through the project. The purpose of the interim assessment is to identify possible free riders and dysfunctional group dynamics. This interim assessment does not count towards the final assessment. At the end of the project, the group members are reminded that the self- and peer-assessment processes will count towards the final assessment. The results in this final assessment are also used for the analysis.

Other data analysed in this paper include tutor assessment scores, results from a traditional content-based test, and the group project score. The tutor evaluates the student based on several criteria including the student's attitude, individual learning and team work. As all the students share the same tutor, possible bias arising from different tutors is reduced. The content-based test is carried out about six weeks into the semester. It seeks to test student's understanding and comprehension of theories relevant to the subject. The group project is assessed at the end of the semester and a mark given, based on the feasibility of the solution for the problem posed. This is the major component for the subject.

RESEARCH FINDINGS

The self-, peer- and tutor-assessment scores were compared to establish the relationship among these terms.

Table 1
Comparisons Between Self- And Tutor-Ratings

	% of student population	**Average difference in marks awarded**	**Standard deviation of difference in marks awarded**
Over-marking by self	61.1	13.2	8.29
Under-marking by self	23.7	9.5	6.52
N.B. 15.2% of the students had the same ratings as the tutor gave them			

Well over half the students rated themselves higher than the tutor did. For this group of students, the difference between self- and tutor-assessment was also larger, by an average magnitude of 13.2 marks. In addition, the degree of over-marking of themselves was greater than the degree of under-marking by students who underrated themselves.

Table 2
Comparisons Between Self- And Peer- Ratings

	% of student population	**Average difference in marks awarded**	**Standard deviation of difference in marks awarded**
Over-marking by self	45.8	9.4	6.85
Under-marking by self	47.3	5.4	4.66
N.B. 6.9% of the students had the same ratings as their peers gave them			

From Table 2 above, the proportion of students under-marking and over-marking is almost the same. However, when we compare students who over-marked themselves with their corresponding peer reviews, the difference is more substantial than in the comparisons between students who underrated themselves and *their* corresponding peer reviews.

Table 3
Comparisons Between Peer- And Tutor- Ratings

	% of student population	**Average difference in marks awarded**	**Standard deviation of difference in marks awarded**
Over-marking by peer on student's performance	73.3	9.7	6.97
Under-marking by peer on student's performance	26.0	10.2	8.41
N.B. 0.7% had exactly the same score from peers as from the tutor			

While a much larger proportion of peers rated an individual's performance higher than the tutor ratings, the difference in magnitude between over-marking and under-marking is comparable.

Table 4
Comparisons Between Test Scores And Self-Ratings

	% of student population	**Average difference in marks awarded**	**Standard deviation of difference in marks awarded**
Ratings above test scores	93.1	25.3	13.68
Ratings below test scores	3.8	7.4	8.60
N.B. 3.1% of the students had the same self-ratings as test scores			

From Table 4, we see that students tend to give themselves higher scores when compared to traditional content-based tests. The magnitude of the difference in scores between self-assessment and the test is also large, especially for those self-rating above their test scores.

Table 5
Comparisons Between Test Scores And Tutor Ratings

	% of student population	Average difference in marks awarded	Standard deviation of difference in marks awarded
Ratings above test scores	84.0	21.0	11.71
Ratings below test scores	9.2	6.3	4.77
N.B. 6.8% of the students had the same tutor ratings as test scores			

Interestingly, Table 5 shows that tutors also frequently gave a higher score on tutor assessment when compared to the traditional content-based test. However, the magnitude of difference between the two scores is slightly lower than the magnitude of difference between self-assessment and the test. While it is acknowledged that the different assessment tools are assessing different skills and knowledge, it is possible that a student may use one component (e.g. the self-assessment component) to compensate for the low scores in another (e.g. the test). To test for this hypothesis, the students are ranked according to their test results. The list is then split into three groups. The group with the highest test scores is compared to the one with the lowest. Results are presented below in Table 6.

Table 6
Differences In Self-Ratings Between High And Low Scorers

	Average difference in marks awarded	Standard deviation of difference in marks awarded
Low Test Scores	33.6	14.71
High Test Scores	12.1	12.29

It was found that students in the low test score group tended to give themselves a much higher self-assessment score when compared to their test results. The magnitude of the difference between the test score and the self-assessment score is 33.6. Students with high test scores also give themselves higher self-assessment scores than their test scores. However, the magnitude of difference is much lower at 12.1.

Students based their self- and peer-assessment scores on the effort shown by the individual members for their group project. Groups with higher levels of teamwork will rate each other higher on the self- and peer-assessment scores and obtain higher scores for the group project. Hence, it was hypothesized that self and peer assessment scores are likely to be positively correlated to group project scores.

DISCUSSION

The results of this study clearly support the view that student-teacher mark agreement is difficult to achieve. Self-ratings can be inflated compared to peer-ratings, tutor assessments and test scores. This suggests that the use of self- and peer-assessment must be considered carefully. As discussed above, if students are not able to give themselves and each other a fair assessment, then it puts into doubt both self- and peer-assessment as assessment tools.

In the past two decades, there have been numerous quantitative studies involving complex statistical data on the comparison between student-derived and tutor-derived marks for a wide range of collaborative assessment practices (Stefani, 1998). It is an important area of research to appraise, since many commentators insist that a reasonable level of student-teacher mark agreement is a pre-condition for student self-grading to play any role in the formal assessment process (Boud, 1992, 1995; Boud & Falchikov, 1989; Dochy & Segers, 1999; Stefani, 1994; Ward *et al.*, 2002). The extreme proposition is posed by Sluijmans, Dochy & Moerkerke (1999) who argued that "if it is not possible to demonstrate that students can produce marks that are acceptable for teachers, then self-assessment should be restricted to a purely learning role and as a skill to be developed" (p. 314).

Boud and Falchikov's (1989) critical review of 48 quantitative studies on student self-assessment concluded that no clear tendency for students to underrate or overrate themselves could be determined. The review found that 17 studies reported overrating and 12 studies reported underrating but because of methodological flaws discovered in the studies, no overall pattern could be identified. Instead, Boud and Falchikov suggested that "there is no consistent tendency to over- or underestimate performance. Some students in some circumstances tend towards one direction, others in the same or different situations towards the other" (p. 543). Thus, it was concluded that student-teacher mark agreement could not be generalised but depended on the specific circumstances of each self-assessment practice.

Student-teacher mark agreement in PBL

The sample in this study consisted of second-year students on the Diploma in Marketing course. As such students are relatively new to Problem-based Learning, it might be expected that they would have difficulty in adjusting to the philosophy of self-directed learning. More importantly, the idea of self- and peer-assessment may be novel to the students, increasing the difficulty that they may have in evaluating themselves and each other accurately. Most of the students in the sample were more familiar with performing self- and peer-ratings which did not permit their involvement in the formulation of assessment criteria. Hence, the opportunity to formulate assessment criteria may not increase the level of student-tutor mark agreement if students do not have the necessary level of experience with self- and peer-assessment and Problem-based Learning pedagogy to formulate relevant and realistic criteria. Direct training and guidance for such students on the construction of assessment criteria is recommended for future practice.

It must be noted, however, that students *were* briefed on the requirements for, and the operation of, the self- and peer-assessment exercise during the course. An interim self- and peer-assessment

exercise was also carried out during the course. It was assumed that students were therefore aware of the requirements for the assessment exercise. However, student-tutor mark incongruency still persisted. This suggests that there may be other reasons for the mark difference between student and tutor besides a lack of understanding of the assessment criteria. It suggests that as this group of students have limited experience with self- and peer-assessments, they may have adopted other covert criteria in awarding marks.

Inflation of self-assessed scores by academically weak students

In addition, the study has also found that students who are poorer academically have a higher tendency to inflate their scores when compared to students who are stronger academically. This finding is consistent with the trend noted by Boud & Falchichikov's meta-analysis of studies on student/teacher mark agreement in self-assessment, whereby students with low overall academic scores tended to have lower student/teacher mark agreement. One of the reasons suggested by the authors were that such students lack understanding of the assessment criteria used for measuring their performance. Their lack of academic ability was therefore concomitant with their poor self-assessment ability.

Such an observation is reinforced by current emphases in higher education on emphasising self-assessment *ability* as opposed to seeing it as merely self-assessment *activity.* The development of self-assessment ability is recognized as a specific outcome of higher education (Boud, 1989; Dearing, 1997; Stefani, 1998) and a critical educational tool for learning beyond university education (Tamir, 1999; Taras, 2001).

In terms of using self- and peer-assessment as an assessment activity, some writers argue that the primary problem of finding a role for student self-grading practices is the relative inaccuracy of the students' assessment (Rainsbury & Hodges, 1998) whilst others see a reasonable level of student-teacher mark agreement as a pre-condition for any value of self-assessment as a learning tool (Stefani, 1994; Sullivan & Hall, 1997). Part of the problem arises from differing interpretations of studies on student-teacher mark agreement. For example, there are varying interpretations of Boud & Falchikov's (1989) critical review of 48 studies from a variety of disciplines which concluded that student-teacher mark agreement could not be generalized but depended on the specific circumstances of each self-assessment practice. This conclusion was interpreted by Lejk & Wyvill (2001) to mean that "agreements between student self-assessment and tutor assessment vary considerably" (p. 558) and by Ward *et al.,* (2002) to mean "that the measurement of self-assessment often yields less than promising results" (p. 64).

Hence, the findings on self- and peer-assessments should not be accepted at face value. They must be adjusted and standardized so as to be more appropriate in a more qualitative sense to obtain insight into group behaviour, social loafing and other such factors (Ghorpade & Lackritz, 2001; Haas, *et al.*, 1998). Clearly, there are multiple and diverse influences on the way students perform their self- and peer-assessment.

Conversely, the different forms of assessment influence the way students learn. Savin-Baden (2004) notes that forms of assessment can encourage students to adopt methods of learning which

ensure they pass the course with high grades rather than those that will be in their best interests as an individual and group member. While many tutors see assessment as providing students with the results of their performance and an award for intellectual or vocational competence, this is too narrow for Problem-based Learning approaches, where the process of learning is also seen as important. There is a need for forms of assessment that are valued by both staff and students, are meaningful to students, do not de-contextualise the subject, and are aligned to Problem-based Learning.

Towards more student autonomy in self- and peer-assessment in Problem-based Learning

Self- and peer-assessment has been often associated with the notion of enhancing students' autonomy in assessment and learning. Heron's (1981; 1988) notion of the teacher's unilateral assessment practice has been a popular premise for advocating student self-assessment as a means for redistributing the power of teachers to students. Boud (1995), Butcher & Stefani (1995), McMahon (1999), Rainsbury & Hodges (1998), Somervell (1993) and Stefani (1992; 1998) have all cited redistributing or sharing the teacher's unilateral power over students as a basis for the practice of self-assessment and peer-assessment.

In this paper, students' self- and peer-assessment outcomes were compared to the tutor's assessment. Problem-based Learning approaches permit students to exercise self-direction in their learning practices. These practices are congruent with the notion of an autonomous and self-directed student. Such autonomy should not be limited to the awarding of marks for their own work and for their peer's work, but should extend to the formulation of criteria for such self- and peer judgments.

We have noted that academically weaker students tend to rate themselves higher compared to academically stronger students. A useful way for the academically weak students to achieve self-assessment scores that are closer to their tutors' score would be to involve these students in formulating and modifying the self-assessment criteria. Stefani (1994) investigated the self-assessment processes and outcomes of 87 students in their laboratory reports, based on student-generated marking schemes, and concluded that a reasonable level of student/teacher and student/peer mark agreement existed when students were involved in the formulation of the assessment criteria.

In addition, a higher level of student responsibility and therefore performance was attributed to the practice of involving students in the generation of assessment criteria in addition to the measurement of their performance. The involvement of students in generating and developing criteria for their self-assessment is recognized as good practice by many writers (Hanrahan & Isaacs, 2001; Mowl & Pain, 1995; Orsmond, Merry, & Reiling, 1997, 2000; Reynolds & Trehan, 2000; Sambell, McDowell & Brown, 1997; Sluijmans, Dochy, & Moerkerke, 1998; Stefani, 1998; Strachan & Wilcox, 1996; Woods, Marshall, & Hrymark, 1988). This is endorsed by Brew, McCulloch, & Barrie (1999) and Longhurst (1997) who cite the involvement of students in the formulation and explanation of assessment criteria as being critical to achieving general agreement between students' and teachers' assessment results. Student involvement in formulating and negotiating self-assessment criteria led to greater clarity in the resulting criteria.

CONCLUSION

Clearly, the research shows that there is a need to re-evaluate the use of self- and peer-assessment in the Problem-based Learning context. If such assessment leads to self/tutor mark inconsistency and opportunities for academically weaker students to award themselves higher marks, then its value may be compromised. Educators need to consider these negative factors before implementing such forms of assessment in their Problem-based Learning courses.

There have been calls to move away from the notion of reliability in student self-assessment altogether, and to judge its value in terms of its learning benefits for the student. Hence, writers such as Stefani (1994) and Cowan (1988) argue that the benefits of self-assessment are so great that we should trust students to act appropriately even when there is a risk that there could be differences between the student's mark and the tutor's mark. Likewise, Orsmond *et al*., (1997) argue that "it is far better to take the risk over the marks than to deprive students of the opportunity of developing the important skill of making objective judgments about the quality of their work and of generally enhancing their skills" (p. 357).

Increasingly, student self-assessment is being viewed in terms of a learning perspective (Boud, 1995; Dochy & Segers, 1999; Stefani, 1998). The role of the numerous studies and reviews on student/teacher mark agreement may be said to have paved the way for more widespread interest in self-assessment, by acting as an appeasement to academics who were hostile to the concept of student involvement in the assessment process (Stefani, 1998). The collective benefit of the studies was thus to assist academic staff to grapple with the new paradigm of student self-assessment, which shifts the balance of power between students and themselves, by demonstrating credibility in student-generated assessment. In this regard, quantitative studies of the reliability of student-generated marks can be viewed as the formulation of a preparatory body of statistical assurances in the literature to bulwark academics' concerns about summative assessment priorities. This results of the investigation described in this paper demonstrate that the reliability of self- and peer-assessment in problem-based contexts cannot be assumed. The involvement of students in the formulation of assessment criteria is perhaps the best assurance that self- and peer-assessment may be relied upon as a valid assessment tool in Problem-based Learning.

REFERENCES

Boud, D. (1989). The role of self-assessment in student grading. *Assessment & Evaluation in Higher Education, 14*(1)*,* 20-30.

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.

Boud, D. & Falchikov, N. (1989). Quantitative Studies of Student Self-assessment in Higher Education: A Critical Analysis of Findings., *Higher Education, 18*(5), 529-549.

Brew, A., McCulloch, M., & Barrie, S. (1999). Qualitatively Different Conceptions of Criteria Used to Assess Student Learning. *Australian Association for Research in Education Annual Conference, November 29 - December 2*. Melbourne, Victoria.

Butcher, A. C., & Stefani, L. J. (1995). Analysis of peer, self- and staff-assessment in group project work. *Assessment in Education: Principles, Policy & Practice*, *2*(2), 165-186.

Cowan, J. (1988). Struggling with student self-assessment. In D. Boud (ed.), *Developing student autonomy in learning*. London: Kogan Page.

Das, M., Mpofu, D., Dunn, E. & Lanphear, J.H. (1998). Self and Tutor Evaluations in Problem-based Learning Tutorials: Is There A Relationship? *Medical Education*, *32*(4**)**, 411-418.

Dearing, R. (1997). *Higher Education in the Learning Society* (Summary Report). HMSO: The National Committee of Inquiry into Higher Education.

Dochy, F. & Segers, M. (1999). The use of self-, peer and co-assessment in higher education : A review. *Studies in Higher Education, 24*(3), 331-350.

Falchikov, N. (1986). Product Comparisons and Process Benefits of Collaborative Peer Group and Self Assessments. *Assessment and Evaluation in Higher Education*, *11*(2), 146-166.

Ghorpade, J. & Lackritz, J.R. (2001). Peer Evaluation in the Classroom: a Check for Sex and Race/Ethnicity Effects. *Journal Of Education For Business*, *75*(5), 274-281.

Haas, A.L., Haas, R.W. & Wotruba, T.R. (1998). The Use of Self Ratings and Peer Ratings to Evaluate Performances of Student Group Members. *Journal of Marketing Education*, *20*(3), 200-209.

Hanrahan, S., & Isaacs, G. (2001). Assessing self- and peer-assessment: The students' views. *Higher Education Research & Development, 20*(1), 53-70.

Heron, J. (1981). Self and peer assessment. In T. Boydell & M. Pedler (Eds.), *Management Self-development: Concepts and Practices,* pp. 111-128. Aldershot: Gower.

Heron, J. (1988). Assessment revisited. In D. Boud (Ed.), *Developing student autonomy in learning*. London: Kogan Page.

Lejk, M., & Wyvill, M. (2001). The effect of the inclusion of self-assessment with peer assessment of contributions to a group project: A quantitative study of secret and agreed assessments. *Assessment and Evaluation in Higher Education, 26*(6), 551-561.

Longhurst, N. (1997). Self-assessment in Coursework Essays. *Studies in Educational Evaluation, 23*(4), 319-330.

McMahon, T. (1999). Using negotiation in summative assessment to encourage critical thinking. *Teaching in Higher Education, 4*(4), 549-554.

Mowl, G., & Pain, R. (1995). Using self and peer assessment to improve students' essay writing: A case study from Geography. *Innovations in Education & Training International, 32*(4), 324-335.

Orsmond, P., Merry, S., & Reiling, K. (1997). A study in self-assessment: Tutor and students' perceptions of performance criteria. *Assessment and Evaluation in Higher Education, 22*(4), 357-369.

Orsmond, P., Merry, S., & Reiling, K. (2000). The use of student derived marking criteria in peer and self-assessment. *Assessment and Evaluation in Higher Education, 25*(1), 23-38.

Rainsbury, E., & Hodges, D. (1998). Academic, employer and student collaborative assessment in a work-based cooperative education course. *Assessment and Evaluation in Higher Education, 23*(3), 313-325.

Reynolds, M., & Trehan, K. (2000). Assessment: A critical perspective. *Studies in Higher Education, 25*(3), 267-278.

Sambell, K., Mcdowell, L., & Brown, S. (1997). But is it fair?: An exploratory study of student perceptions of the consequential validity of assessment. *Studies in Educational Evaluation, 23*(4), 349-371.

Savin-Baden, M. (2004). Understanding the Impact of Assessment on Students in Problem-based Learning. *Innovations in Education and Teaching International*, *41*(2), 223-233.

Segers, M. & Dochy, F. (2001). New Assessment Forms in Problem-based Learning. *Studies In Higher Education*, *26*(3), 327-343.

Sluijmans, D., Dochy, F., & Moerkerke, G. (1998). *The use of self-, peer- and co-assessment in higher education: A review of the literature*. Educational Technology Expertise Centre, Open University of the Netherlands.

Sluijmans, D., Dochy, F., & Moerkerke, G. (1999). Creating a learning environment by using self-, peer- and co-assessment. *Learning Environments Research, 1*, 293-319.

Somervell, H. (1993). Issues in assessment, enterprise and higher education: The case for self-, peer and collaborative assessment. *Assessment and Evaluation in Higher Education, 18*(3), 221-233.

Stefani, L. (1992). Comparison of collaborative self, peer and tutor assessment in a biochemistry practical. *Biochemical Education, 20*(3), 148-151.

Stefani, L. (1994). Peer, Self and Tutor Assessment: Relative Reliabilities. *Studies in Higher Education, 19*(1), 69-75.

Stefani, L. (1998). Assessment in partnership with learners. *Assessment and Evaluation in Higher Education, 23*(4), 339-350.

Strachan, I. B., & Wilcox, S. (1996). Peer and self assessment of group work: Developing an effective response to increased enrolment in a third year course in Microclimatology. *Journal of Geography in Higher Education, 20*(3), 343-353.

Sullivan, K., & Hall, C. (1997). Introducing students to self-assessment. *Assessment and Evaluation in Higher Education, 22*(3), 289-305.

Tamir, P. (1999). Self-assessment: The use of self-report knowledge and opportunity to learn inventories. *International Journal of Science Education, 21*(4), 401-411.

Taras, M. (2001). The use of tutor feedback and student self-assessment in summative assessment tasks: Towards transparency for students and for tutors. *Assessment and Evaluation in Higher Education, 26*(6), 605-614.

Ward, M., Gruppen, L., & Regehr, G. (2002). Measuring self-assessment: Current state of the art. *Advances in Health Sciences Education, 7,* 63-80.

Wee, K.N.L. & Kek, Y.C. (2002). *Authentic Problem-based Learning.* Singapore: Prentice-Hall

Woods, D., Marshall, R. R. & Hrymark, A. (1988). Self-assessment in the context of the McMaster problem solving programme. *Assessment & Evaluation in Higher Education, 13*(2), 107-127.