

Educational Technology Expertise Centre Otec
Open University of the Netherlands

The use of self-, peer- and co- assessment in higher education

A review of literature

D. Sluijsmans
F. Dochy
G. Moerkerke

Otec report series

The Open University of the Netherlands develops higher distance education and is a central partner in a consortium for the renewal of higher education. Educational technological innovation is one of the main fields of interest. The educational and educational technological expertise of the Open University of the Netherlands are bundled in the Educational technology expertise centre (Otec). This centre is involved with tasks concerning the development, innovation, research and evaluation of the Open University and its consortium partners. These tasks are performed in close collaboration with directorates and faculties of the Open University and/or its consortium partners.

Otec publishes a report series. This report is part of that series.

Otec reports and some documents can be obtained from:

Open University of the Netherlands
secretariaat Otec
P.O. Box 2960
6401 DL Heerlen
the Netherlands
Tel. +31 45 5762406 or +31 45 5762942
Fax. +31 45 5762802

Internet: <http://www.ouh.nl/otec>

Educational Technology Expertise Centre (Otec)
Open University of the Netherlands

The use of self-, peer- and co-assessment in higher education

A review of literature

D. Sluijsmans
F. Dochy
G. Moerkerke

ISBN 9035817907

© 1998,
Educational Technology Expertise Centre,
Open University of the Netherlands

Save exceptions stated by the law no part of this publication may be reproduced in any form, by print, photoprint, microfilm or other means, included a complete or partial transcription, without the prior written permission of the publisher.

Contents

1. Introduction 7

- 1.1 Research questions 8
- 1.2 Method 8
 - 1.2.1 *Selection of studies* 8
 - 1.2.2 *Analysis procedure* 9
 - 1.2.3 *Classification and coding* 9

2. How are self- peer- en co-assessment applied in higher education 11

- 2.1 Self-assessment 11
 - 2.1.1 *Definition* 11
 - 2.1.2 *Main findings* 11
 - 2.1.3 *Self-assessment in practice* 12
 - 2.1.4 *Conclusion* 13
- 2.2 Peer-assessment 14
 - 2.2.1 *Definition* 14
 - 2.2.2 *Main findings* 14
 - 2.2.3 *Peer-assessment in practice* 15
 - 2.2.4 *Conclusion* 16
- 2.3 Self- and peer-assessment 16
 - 2.3.1 *Main findings* 16
 - 2.3.2 *Self- and peer-assessment in practice* 16
- 2.4 Self- and peer-assessment related to co-assessment 17
 - 2.4.1 *Definition* 17
 - 2.4.2 *Main findings* 18
 - 2.4.3 *Self-, peer- and co-assessment in practice* 18
 - 2.4.4 *Conclusion* 21

3. The effect of the use of self-, peer- and co-assessment on the quality of learning 23

4. Conclusion and discussion 27

References 33

Summary

To develop the skills and competencies required in professional organizations, students have to reflect on their own behavior. Many current assessment practices do not answer this need. The recent interest for new assessment forms, such as self-, peer- and co-assessment can be seen as a means to tackle this problem. In the present report, a review was conducted to find the answers to the questions: (1) How are self-, peer- and co-assessment applied in higher education; (2) What are the effects of the use of these forms of assessment on the quality of learning. Analyses of 62 studies showed that self-, peer- and co-assessment can be effective tools to develop competencies needed as a professional. These forms of assessment are often being used in combination. Implementation of these forms of assessments accelerates the developments of a curriculum based on competencies (knowledge as a tool) rather than knowledge (as a goal) and leads towards the integration of instruction and assessment in higher education. As such this development contributes to the education of responsible and reflective professionals. Significant evidence though for this statement cannot be found in this study. Further research is necessary to endorse the view that self-, peer- and co-assessment need to be trained as independent skills (by students and tutors) and that assessment systems have to be developed to accomplish further integration of learning and assessment.

1 Introduction

It is recognized that the main goal of professional higher education is to help students develop into reflective practitioners who are able to reflect critically upon their own professional practice (Schön, 1987; Falchikov & Boud, 1989; Kwan & Leung, 1996). Students in modern organizations must be able to analyze information, to improve their problem-solving skills and communication and to reflect on their own role in the learning process.

The need for lifelong learning in modern society will increase (Sambell & McDowell, 1997). It is more and more realized that the acquisition of knowledge and skills cannot be restricted to the phase of initial education; rather, it has to be a process continuing throughout one's entire working life. Traditional testing methods do not fit such goals as lifelong learning, reflective thinking, being critical, evaluate oneself, problem solving, etcetera (Dochy & Moerkerke, 1997).

Alternatives in assessment have received many attention in the last decade and several forms of more authentic assessment have been introduced in higher education (Birenbaum & Dochy, 1996). The skill of self-, peer- and co-assessment is very important in the development of lifelong learning and the development into autonomous individuals (Sambell & McDowell, 1997). Assessment procedures should not only serve as tool for crediting students with recognized certificates. Such procedures should also be used to monitor the progress and to direct students, if needed, to remedial learning activities. Research (Beckwith, 1991) showed that the nature of assessment tasks influences the approaches to learning which students adopt. The existing assessment approaches can have effects contrary to those desired. Eisner (in Boud, 1995) identified the features of the new assessment in education:

- assessment tasks need to reflect the tasks students will encounter in the world outside schools, not merely those limited to the schools themselves;
- assessment tasks should not be restricted to the solutions students formulate but also reveal how students go about solving a problem;
- assessment tasks should reflect the values of the intellectual community from which tasks are derived;
- assessment tasks need not be limited to a solo performance;
- assessment tasks should have more than one acceptable solution to a problem and more than one acceptable answer to a question;
- assessment tasks should have curricular relevance, but not be limited to the curriculum as taught;
- assessment tasks should permit the student to select a form of representation he or she chooses to display what has been learned.

The view that the assessment of students' achievements is something which happens at the end of a process of learning is no longer widespread. Assessment is now represented as a tool for learning (Arter, 1997; Dochy & McDowell, 1997). The present contribution will focus at one new dimension of assessment innovation, namely the changing place and function of the assessor. The new educational stream on alternative assessment also questions if the assessor must be the tutor or teacher himself. Self and peers are often introduced as assessors in different settings. This literature review will focus at forms of self-, peer- and co-assessment from the point of view of their applicability in higher education.

1.1 Research questions

This report aims at finding out how self-, peer- en co-assessment are used in research and in educational practice. It answers the following research questions:

- how are self- peer- en co-assessment applied in higher education;
- what is the effect of the use of these forms of assessment on the quality of learning.

1.2 Method

1.2.1 Selection of studies

In order to answer the questions a literature search was conducted. The following sources were used:

Source 1: the database of the Educational Resources Information Center (ERIC). This database was searched, using the keywords self-assessment, peer-assessment and co-assessment. The search was conducted on the publications from 1987 to 1997.

Source 2: the database of Current Contents on Disk. The years 1996 and 1997 were searched in order to find recent studies which was not yet included in ERIC. The same keywords were used as in the ERIC-search.

Source 3: through the so-called snowball-method the references in all the above materials were checked for other studies.

Source 4: the Internet was searched with the search engine Alta Vista. The result were a number of WWW-addresses on assessment with abstracts of conference papers, information of workshops, etcetera.

Next, the content of the abstracts was analyzed. The following criteria were set to determine whether literature would be included:

- the assessment form had to be predominantly self-, peer- or co-assessment. Portfolio-assessment and performance assessment for example were not central themes, although there was often a strong relationship with self-, peer- en co-assessment;
- the literature about the relationship between assessment and learning/reflection is included in this search, to find answers to the second research question of this review;
- the subjects of study had to be relevant for the research questions. Relevant subjects were decided to be students in higher education. For instance, studies dealing with peer assessment of university personnel were excluded.

The search of ERIC resulted in 191 articles; the articles reporting the application of assessment in higher education were selected for this study by analyzing titles and abstracts. This resulted in 34 articles. The search of Current Contents on Disks resulted in 53 articles. Articles already found in ERIC were excluded. The articles reporting the application of assessment in higher education were selected for this study by analyzing titles and abstracts. This resulted in 10 articles. The snowball method led to the selection of 18 articles. In total 62 articles were selected for further analysis. The search on Internet led to the selection of three WWW-addresses.

1.2.2 Analysis procedure

In our first search the 62 studies were recorded in tables according to the following variables:

- the authors and the year reported;
- the title;
- the aim of the study;
- the type and number of subjects;
- the strengths (of the assessmentform discussed);
- the weaknesses (of the assessmentform discussed);
- the treatment/instrument;
- the assessment(criteria);
- the value of statistics reported;
- the conclusions.

1.2.3 Classification and coding

The studies were classified first in empirical and non-empirical reports. Then the studies were classified for both categories in self-, peer- and co-assessment. However, since these forms of assessment are sometimes strongly interrelated, it was decided to divide the studies into five categories:

- 14 studies about self-assessment;
- 10 studies about peer-assessment;
- 7 studies about self/peer-assessment;
- 19 studies about self/peer/co-assessment;
- 12 studies about the relationship between assessment and learning/reflection.

In the present report a narrative review of the literature is presented. This form of conventional literature review implies careful reading of separate studies and integrating them. Of course, since patterns in the results are sought, this integration will be an intuitive undertaking (Knoors, Dochy & Moerkerke, 1993; Slavin, 1986). Therefore, in a next study an alternative approach will be applied to deal with the task of integrating the findings.

2 How are self-, peer- en co-assessment applied in higher education

In this section, the results related to the first research question will be described as follows. In four separate sections, the different combinations of self-, peer- and co-assessment will be treated. Within these sections, the similar structure will be maintained. First, definitions are given. Then, the main findings are presented and the way the assessment forms are used are outlined. Finally, a conclusion ends each section.

2.1 Self-assessment

2.1.1 Definition

Self-assessment refers to the involvement of learners in making judgments about their own learning, particularly about their achievements and the outcomes of their learning (Boud & Falchikov, 1989).

Self-assessment is not a new technique, but a way of increasing the role of students as active participants in their own learning (Boud, 1995), and is mostly used for formative assessment in order to foster reflection on one's own learning processes and results (Sluijsmans & Dochy, 1998).

2.1.2 Main findings

Boud and Falchikov (1989) classified the literature of self assessment under three headings: conceptual, practical qualitative, and quantitative. In the first heading, one of the most important parts in the conceptual framework is the literature about the reflective practitioner (Schön, 1987). Secondly, the practical qualitative group includes the processes involved in introducing and using self-assessment in different situations. Thirdly they classified the quantitative group, which focuses on studies of student self-ratings compared to the ratings of students by teachers. Boud and Falchikov (1989) analyzed studies from 1932 till 1988 and reported the overrating and the underrating of students. They related these findings to the different abilities of students. The finding was that good students tended to underrate themselves and that weaker students overrated themselves. Students in higher levels classes could better predict their performance than students in lower levels classes. Griffiee (1995) investigated also the question whether there is a difference in student self-assessment between first-year, second-year, and third year classes in a university department. The general answer to this question was that there was no difference. All classes tended to rate themselves lower at the beginning of the school year and higher as the semester progressed. Students gained more confidence as the semester progressed in their ability to perform. Another explanation for the fact that there is no difference between the self-assessments of the three classes, is the teacher intervention during the year. Several studies obviously show that the ability of students to rate themselves improves in the light of feedback or development over time (Birenbaum & Dochy, 1996; Griffiee, 1995; Boud and Falchikov, 1989). Moreover, students' interpretations are not just dependent on the form of the assessment process, but on how these tasks

are embedded within the total context of the subject and within their total experience of educational life.

2.1.3 Self-assessment in practice

In educational practice, we see that different instruments are used for self-assessment. Harrington (1995) used three different self-assessments. One is simply a listing of abilities with definitions and directions to indicate those areas you feel are your best or strongest. A second approach is to apply a Likert scale to a group of designated abilities (in comparison to others of the same age, my art ability is excellent, above average, average, below average, or poor). Another approach is, for each ability, to provide different examples of the ability s applications on which individuals rate their performance level from high to low, and subsequently these are summed to obtain a total score. The self-assessment forms Harrington described are cheaper and less time intrusive than traditional ways of assessing students.

An electronic interactive advice system for self-assessment is provided by Gentle (1994). The aim of this system is to see how accurate students are able to assess their own work without the involvement of their supervisor. The system is based on question and answer screens for 38 skills. These skills are arranged in four sections: (1) approach to the project - effort, time management, etcetera. (2) quality of day to day work (3) quality of the description of the work and (4) quality of presentation. The procedure is as follows: The user moves a cursor on a continuous scale of performance on that aspect of the work. The middle and end points on the scale are picked out by written statements to help the user and there is also a full advice screen available for each question. This feature makes this system much more than just an assessment program, since it includes large tranches of practical assistance, useful at any point in the project work. The output also provides much more than a mark; the five best and the five weakest points, selected by their weighted contribution to the mark, are extracted and displayed (Gentle, 1994, p. 1159). Results of the use of the system show that students can assess themselves to within five percentage points. Students become more aware of the quality of their own work. They can predict their own mark and while they are doing this, they reflect on their behavior (reflective practitioner). Because the student reflects more often than once on his work, this will lead to a higher quality of the products. According to Gentle, the system is less time spending than the conventional self-assessment: the supervisor has a minor part in the assessment.

In a research conducted by Hassmén, Sams and Hunt (1997) 128 women learned the correct answers on a specific task by either performing or observing. Participants took either a performance or a written test, with or without making self-assessments about how sure they were that their selected answer was correct. Findings of the research support the hypothesis that those participants who engage in overt self-assessment responding while learning will obtain a higher percentage of correct responses during learning trials on a test than those who learn without self-assessments.

Self-assessment is also important for successful language learning. McNamara and Deane (1995) designed a variety of activities that foster self-assessment. Three of them are: writing letters to the teacher, keeping a daily language learning log, and preparing an English portfolio. These activities can help students to identify their strengths and weaknesses in English, to document their progress and to identify effective language learning strategies and materials. They also become aware of the language learning contexts that works best for them and to establish goals for future independent learning. The idea of self-assessment for use in portfolio is described by

Keith (1996). She suggests self-assessment assignments which ask students to report on their own learning. Assignments include sharing preconceptions about teaching and learning, comparing goals, creating a community of learners, generating student explanations and improving communication, group quizzes, challenging thinking dispositions, posttest evaluations, and collaborative assessing. The roots of all the described assignments lie in collaborative learning. The biggest variable for effective learning is the amount of meaningful energy the students put in. The assignments have to encourage students to feel responsible for their own learning.

Another instrument was used by Anderson and Freiberg (1995). These authors used an audiotape self-assessment instrument for student teachers to reflect on their teaching. This instrument - called the Low Inference Self-Assessment Measure (LISAM) - has been developed to let student teachers analyze their instruction. Ten secondary student teachers completed four stages in the study. In the first stage students learned to record themselves during a lesson. In the second stage students were trained to analyze their own audiotapes. In the third stage findings and suggestions for effective use of the LISAM were discussed. The students set goals for future use of the self-assessment instrument. In the last stage there was an interview with every student teacher. Anderson and Freiberg describe three reasons why the LISAM is practical and effective: (1) the use of LISAM makes student teachers more independent, provides feedback and stimulates them to reflect on their own teaching, (2) student teachers can practice LISAM immediately and (3) the LISAM teaching behaviors are observable and alterable.

Generally, next to addressing the instruments used for self-assessment, the content could be addressed. At the content level, it is striking that self-assessments are mostly used to foster skills and abilities, next to knowledge and that assessments are used in a formative or diagnostic way (Birenbaum & Dochy, 1996). For example, students of the Alverno College have to develop problem-solving as one of the eight abilities in order to graduate (Loacker and Jensen, 1988). At the heart of the educational process at Alverno stands assessment. The faculties see it as a natural part of encouraging, directing, and providing for development of abilities. Since self-assessment is required to integrate into the students' problem solving process, faculties have found that students show increasing understanding of inter-relationships of ability, content, and context. Students take responsibility for their learning as a dynamic, continuing process. They gradually internalize their practice of both problem solving and self-assessment abilities.

2.1.4 Conclusion

Overall, it can be concluded that research reports positive findings concerning the use of self-assessment in educational practice. Students in higher education are well able to self-assess accurately (see Gentle), and this ability improves with feedback and development over time. Moreover, students who engage in self-assessment tend to score higher on tests (see Hassmén). Self-assessment, used in most cases to promote the learning of skills and abilities, leads to more reflection on one's own work, higher quality of products, responsibility of one's own learning and increased understanding of problem-solving. Instruments for self-assessment vary from Likert scales, ability listings, and written tests to portfolios, audiotape assessments or electronic interactive systems.

2.2 Peer-assessment

2.2.1 Definition

Falchikov (1995) defines peer assessment as the process whereby groups of individuals rate their peers. This exercise may or may not entail previous discussion or agreement over criteria. It may involve the use of rating instruments or checklists, which have been designed, by others, before the peer assessment exercise, or be designed by the user group to meet their particular needs.

2.2.2 Main findings

More theoretical studies found in our search do come with goals and forms of peer assessment, Somervell (1993) found that peer assessment engages students in making judgments about the work or the performance of other students. At one end of the spectrum it may involve them giving feedback of a qualitative nature or, at the other, may involve them in marking. The assessment may be formative or summative and could form part of a larger scheme whereby peer feedback is given prior to self assessment by the recipient¹. Peer assessment is not only a grading procedure, it is part of a learning process where skills are developed. Peer assessment can be seen as a part of the self-assessment process and serves to inform self assessment. The contribution of other students can be a very useful input into the self-assessment process. They have an opportunity to observe their peers throughout the learning process and often have a more detailed knowledge of the work of others than do their teachers (Somervell, 1993).

Keaten et al. (1993) reported that peer assessment is a practice that can foster high levels of responsibility among students; the students must be fair and accurate with the judgments they make regarding their peers.

Peer-evaluation is also an alternative term to peer-assessment (Weaver and Cotrell, 1986). Peer evaluation emphasizes skills, encourages involvement, focuses on learning, establishes a reference, promotes excellence, provides increased feedback, fosters attendance, and teaches responsibility (Weaver and Cotrell, p. 25). Dancer and Dancer (1992) indicate that research studies have not shown the validity of peer rating. Peers are prone to produce ratings based on uniformity, race and friendship if there is no extensive training in peer rating. Based on this assumption it is sometimes important to determine a individual s contribution to a group project.

Different forms of assessment are distinguished by Kane and Lawler III (1978):

- *peer ranking*: consists of having each group member rank all others from best to worst on one or more factors;
- *peer nomination*: consists of having each member of group members as being the highest in the group on a particular characteristic or dimension of performance;
- *peer rating*: consists of having each group member rate each other group member on a given set of performance or personal characteristics, using any one of several kinds of rating scales.

¹ <http://141.163.121.36/TeachingLearning/Ethos/Assessment/PeerAss1.html>

2.2.3 Peer-assessment in practice

The more practical oriented studies focus on one of the main advantages of peer-assessment, being fairness (construction of assessment criteria, more objective assessment by more subjects, assessment of output and process such as assessing student contribution).

Conway et al. (1993) indicate that students found group projects (more) interesting than traditional methods of teaching. Since the fairness of the assessment was found to be the only negative aspect of this type of working, peer assessment was introduced. First, each group's presentation was assessed by the other members of the group. Secondly, the students assessed the contribution of their fellow group members to the work of the project. The aim of the study was to examine ways in which students may be awarded individual marks, reflecting personal effort, for group projects. Conway et al. found good elements between the scheme of Goldfinch and Raeside (1990) and simplified this scheme, combining the elements which could be very effective. The method Conway et al. used will not be outlined here, but the results showed that students felt that the peer assessment is a good method and fair enough. Students felt that they should play a part in the assessment in order to make assessment more objective.

Cutler and Price (1995) describe an investigation where presentations and seminars, built into each of the three years of the geography program, are peer-assessed against a set of criteria. Self-appraisal forms are also a part of the assessment procedure. The results of the peer-assessment showed that the majority of the students were happy and confident being assessed by their peers. Half of the students felt that their assessment of their peers was accurate. A third of the students thought that they had improved in confidence, organization of materials and use in voice.

Boud (1995) investigated students' ability to make assessments of themselves and their peers in a one session first year class entitled 'The Legal System' taught at the University of New South Wales. A series of instructions was provided for each student. These gave guidelines for formulating assessment criteria. Once the students had constructed the criteria they would use, they made an assessment of both their own performance in class and that of the other students in the class. There were two methods of scaling used. Method A: the scale 1, 2, 3..10 with 5 as pass mark and method B: the scale -2, -1, 0, 1, 2 with the rule that there should be equal numbers of students above and below the mean of 0. At the end of the exercise the self-mark, the peer-mark and the teacher-mark were available for each student on each criterion using both methods. Results show that the student found constructing the assessment criteria very useful. Other results were that students rated themselves higher than they are rated by their peers and that students rated themselves lower than the teacher using method A but higher using method B. In general, there was a very high level of agreement between the marks given by peers and those given by the teacher.

Experience from peer-assessment revealed that peer assessment as a formative assessment method and as a part of the learning process can be seen as a valuable instrument. Students are more involved, both in the learning and in the assessment process. They find peer assessment fair enough and accurate. Though the following can be though see that the following often can be observed during peer-assessment (Pond et al., 1992):

- friendship marking: resulting in overmarking;
- collusive marking: resulting in a lack of differentiation within groups;
- decibel marking: where individuals dominate groups and get the highest marks;
- parasite marking: where students fail to contribute but benefit from group marks.

These problems can be prevented by combining peer-assessment with self-assessment or co-assessment. It can be observed that this is exactly the reason why the majority of studies investigate these combinations of assessment forms. The following sections will illustrate this.

2.2.4 Conclusion

Experiences from peer assessment revealed that peer assessment as a formative assessment method and as a part of the learning process can be seen as a valuable instrument, since students are more involved both in learning and in the assessment process and do find it a fair and accurate form of assessment. Disadvantages of peer assessment such as friendship marking and decibel marking are mostly solved by using combinations with self- and co-assessment as will be explained further.

2.3 Self- and peer-assessment

Self- and peer-assessment are combined when students are assessing peers but the self is also included as a member of the group and must be assessed. This combination fosters reflection on one's own learning process and one's learning activities compared to the other members in the group or class.

2.3.1 Main findings

Because of the above stated disadvantages of peer assessment, almost all studies found on combination of assessment forms were practically oriented and were seeking for more proof on validity, interrater (or inter-peer) reliability and positive involvement of students in the process.

2.3.2 Self- and peer-assessment in practice

In a study described by Burnett and Cavaye (1978) fifth year medical students had to assess their peers as part of the examination. They also were asked to assess their own performance. Results show that the peer assessment highly correlated with the final grade ($r=.99$) and the staff-assessment ($r=.93$) and that the self-assessments highly correlated with the results of the peer assessments ($r=.99$).

Strachan and Wilcox (1996) describe a peer- and self-assessment strategy in a third year in microclimatology. Thirty students were asked to form groups of three. Each group had to do a seminar presentation. The students were informed that they would undergo a peer assessment. Students had the opportunity to develop their own criteria. After this brainstorming a Group Project Peer- and Self-assessment sheet was established. It was also a self-assessment sheet since every student had to include himself in the assessment. Each member of each group completed the sheet after turning in their papers. The students also were asked to give written feedback on this way of assessment. Some preferred written evaluative comments to number ratings and some students felt that this way of assessment was not necessary because in a group you always have a certain responsibility. Strachan and Wilcox recommend however that it is important to give the student an active role in the development of assessment criteria. The process is thereby equally important as the product for the quality of learning.

Warkentin et al. (1995) investigated self- and peer-assessment in a study with 83 undergraduate educational psychology students. Warkentin et al. hypothesized that students taking tests using individual and group assessments would perform significantly better on a posttest of educational psychology course concepts than students who took the traditional tests (individual exams). The effects on student knowledge structure representations were examined. The results indicated that there were no significant differences between the two groups on achievement and knowledge structure. Warkentin et al. (1995) though found that the reactions to a cooperative assessment procedure they used were overwhelmingly positive. The students did like the group assessment and thought it contributed to their learning through this process as they discussed and debated test items.

Sambell and McDowell (1997) studied six cases which included peer- and/or self-assessment. Findings show that students were generally positive towards an involvement in the assessment process. Students awareness that self- and peer-assessment helped them to develop important skills, e.g. problem solving, was high.

Under- and overmarking in a study whereby self- and peer-assessment was investigated by Falchikov (1991). The process of working together in a small group project is assessed by the group members, seven developmental psychology students. In the study the development of a self/peer group process assessment checklist was carried out and was designed to compare the assessments of task and maintenance functions to be made (Falchikov, 1991). Students were informed to work on a piece of coursework. The checklist contained sixteen task functions and eight maintenance functions. This list was developed with cooperation of the students; this enabled the students to become familiar with the assessment checklist. The students had - after finishing the coursework - to rate their peers and themselves on the checklist. They rated the level of activity (high, medium, low) to which each group member including self had carried the sixteen functions (group activities). The results showed that there was no consistency of over- or undermarking when comparing self-ratings with peer-ratings. There was also a high level of agreement between peers. Falchikov states that the process of working in a group is a simple and effective way of moving towards a solution to the problem of evaluation of group work (p. 15).

2.4 Self- and peer-assessment related to co-assessment

In the prior sections the use of self-assessment, peer-assessment, and the use of a combination of these two forms is described. One step closer to the current mainstream in traditional educational practice are assessment procedures whereby the tutor plays a significant role in the process. The term co-assessment is here introduced.

2.4.1 Definition

Co-assessment, the participation of the students with the staff in the assessment process, is a way of providing an opportunity for students to assess themselves while allowing the staff to maintain the necessary control over the final assessments (Hall, 1995). Synonyms for co-assessment are collaborative assessment and cooperative assessment.

2.4.2 Main findings

Co-assessment can be used for summative purposes while self- and peer-assessment are rather used in a formative way. Somervell (1993) sees collaborative assessment as a teaching and learning process in which the student and instructor meet to clarify objectives and standards. In this case the student is not necessarily responsible for the assessment, but rather that the student has collaborated in the process of determining what will be assessed and perhaps, by whom. Pain, Bull and Brna² argue that the term 'collaborative assessment' can be applied to an assessor and an assessee working together to form a mutual understanding of the student's knowledge. It is a true collaboration in so far as both parties work on the shared goal of providing a mutually agreed assessment of the student's knowledge. This entails both parties negotiating over the details of the assessment, and discussing any misunderstandings that exist. This is consistent with the less confrontational approach to assessment that we seek to adopt while stressing the need to develop an ongoing relationship between the assessor and assessee.

2.4.3 Self-, peer- and co-assessment in practice

Co-assessment is often related to forms of self- and peer-assessment. In the co-assessment Hall (1995) described, the students and staff set the criteria. The process involved a double side face-sheet. On the back of this sheet the students had the opportunity to give their own self-assessment of the piece of work and then hand it to the staff member. The staff member used the outside of the sheet to record his assessment of the student's work. Then the staff member turned it over to see whether or not the student had chosen to offer their own assessment on the other side. The findings showed that generally the staff member's grade being higher than the student's grade. The experiment Hall described identifies three purposes of co-assessment. One is to assist the student-teachers in making the role-changing from being student to being a teacher, a second is to provide insights into the assessment process which may be of use to them in assessing their own students, and a third is to provide a skill-development step towards self-assessment.

Many other studies do combine self-, peer- and co-assessment in some way and mostly striking is, apart from the assessment forms, the content of the assessments. In the studies we found, there is a majority of positive experiences for assessing essays and assignments and to a smaller extent assessing presentations. Falchikov (1986) and Stefani (1992;1994) also described studies with collaborative, and self- and peer-assessment. The study of Falchikov aimed to implement and evaluate a method of collaborative self and peer assessment. First the tutors set criteria and ranked these criteria in terms of their relative importance. Then students set criteria and tutor-student criteria comparisons were made. An essay marking schedule was drawn up. Students marked their own essays and then each group member and the tutor marked the essays. Self-, peer and tutor marks were compared. Results show that collaborative, and self-assessment does appear to be comparable to traditional tutor methods of assessment, while collaborative, and peer assessment corresponds less well with either tutor or self grading. Stefani (1992) carried out an experiment in collaborative, self- and peer-assessment of a first year undergraduate biochemistry laboratory practical experiment. The students themselves defined the marking schedule for a scientific report. The results show that students have a realistic perception of their own abilities and can make rational judgments on the achievements

² <http://cbl.leeds.ac.uk/~paul/papers/euroaiedpapers96/smpaper/smpaper.html>

of their peers. Many tutors expressed their fears in handing over the assessment to the student. Concerning the evaluation of the learning benefits, almost every student said that the scheme made them think more, learn more and was challenging.

Freeman (1995) conducted a peer assessment experiment with 210 final-year undergraduate business students. Students were divided into 41 teams, and each team had to complete two of the four assessable tasks. The presentation, one of the two tasks, was chosen by staff to experiment with a peer assessment worth 25% of the overall grade of the students. In the first week of semester each student was given the presentation marking and feedback sheet with 22 items, eight items related to the content and fourteen related to the presentation, weighted 60% and 40% respectively. In the results of the team presentations rated by staff and peers Freeman found that the quality of the presentations was very high. There was no significant difference between the average staff ratings and average peer ratings. Students though tended to undermark the good presentations and overmark the poor presentations.

Longhurst and Norton (1997) designed a study to investigate how accurately 67 second year psychology students would be able to assess their own essays and thereby ascertain whether or not they understand what taking a deep approach in their essays actually means. Student grades with tutor grades are compared. The students were asked to rate themselves on tutor-specified criteria which were designed to measure a deep approach, essay grade and level of motivation in completing one specified essay. The tutor didn't see these self-assessments, since the self-assessment sheet was removed from each essay. The tutors also marked the essays on the deep processing criteria. Results show that the tutor grade for the essay highly correlated with each of the five criteria. There also was a positive correlation between students and tutor grades ($r=.43$). The results further indicated that overall, students were very accurate in grading their own essays but less accurate in assessing their own deep processing. Less motivated and weaker students appeared to be less clear on understanding the individual criteria.

Oldfield and Macalpine (1995) investigated the competence of students in making assessments. The peer assessment was designed in steps from individual tasks to group assignments. Each task was assessed by the peer group and compared with the assessment of the lecturer. Results show high correlations between student marks and lecturers marks of individual essays and presentations.

The self-assessment Oldfield and Macalpine describe is also an approach in achievable steps, first comparisons of contributions to group activities excluding self, then including self and finally a self-assessment of individual work. The students first make a peer assessment of all the groups' achievements. To train the self-assessment, students also had to do this for their own group. The same procedure takes place within the group: assessing the group members and then assessing one's own contribution. Orpen and Macalpine found that this assessment procedure strengthens the confidence of students to assess the work of others and that of themselves.

Orpen (1982) describes a study with 21 students in an Organizational Behavior-course and 21 students in a Political Philosophy-course. These students had to write an essay. The students were informed that their papers would be marked by five lecturers later in the year, and that their final grade would be the average of the marks they received from their fellow-students and from the lecturers (p. 568). The marks were given according to the criteria (1) coverage of the relevant material, (2) coherence and strength of the underlying argument, and (3) fluency and clarity of expression. Results show that there was no difference between the lecturers and students in their average

marks, in the variation of their marks, in the extent to which their marks agreed with each other and in the relationship between their marks and the writer's performance in course-end examinations.

Rushton, Ramsey and Rada (1993) developed a computer-based peer assessment tool. 32 Computer Science undergraduates wrote an essay on the viability of peer assessment. They typed their essays on the subject of peer assessment into the system. The class was split into groups of 3 or 4 students. Each group member used the peer assessment window to mark each other's work. Contrary to expectations, the marks awarded by the peers were remarkably similar to those awarded by the tutors, so peer and teacher assessment seem to be equally reliable forms of assessment.

Zoller and Ben-Chaim (1997) investigated the self-assessment ability of 71 biology majors enrolled in a four-year college program, with respect to HOCS (Higher Order Cognitive Skills) and their confidence in self-assessing. A specially designed self-assessment questionnaire is described. This questionnaire consisted HOCS-questions, interdisciplinary science-technology-environment-society (STES)-oriented questions and Likert-type questions with respect to students' confidence. Students assessed their knowledge and understanding on this questionnaire. Results indicated that the students evaluate themselves as quite knowledgeable. The results further showed that 75% of the students thought they were capable in self-assessing and peer-assessing. Zoller and Ben-Chaim found a discrepancy between the student assessment and the teacher assessment, which they accounted for by the lack of integration between assessment and learning in contemporary science teaching.

Kwan and Leung (1996) report the results of tutor and peer group assessment of student performance of 96 students in a simulation exercise on hotel personnel training. The group was divided in five tutorial groups. Then students were paired and each student had to conduct a training session with the partner to an audience. The performance of each student was assessed by the tutor and the peers according a checklist. Results show that there was some agreement between tutor and peer group markings, but somewhat less reported by Falchikov (1986) and Stefani (1994). Arguments for this finding are that student aren't capable enough to assess since this was the first time they did it. Secondly, the students had no contribution in identifying the criteria, there was no negotiation between tutor and students in understanding the criteria.

Orsmond, Merry and Reiling (1996) describe an experiment in peer assessment for a first-year undergraduate animal physiology poster assignment. The 39 pair groups of students completed a poster assignment. Students were informed about the poster requirements. At the end of the 12-week lecture course, the students were divided between two laboratories. Later the students of each laboratory were asked to mark all posters in the other laboratory on five criteria. Each criterion had a grade point 0-4. Each poster was marked and a peer assessment grade point for each criterion was calculated as that with the maximum number of ticks, the modal value. After the students marked the posters the tutor also marked the posters without seeing the marks the students had given. Orsmond et al. (1996) found that there was 18% agreement between students and tutor, with 56% students-overmarking and 26% students undermarking.

The correlation was .54. The students had also to fill in a questionnaire which showed that 76% of the students thought that the peer assessment had made them think more, and work in a more structured way (p. 243).

Fry (1990) describes a study whereby the tutor introduced peer marking. The tutor first marked the script of the students and then handed the scripts over to the students. The tutor asked the students to mark each others work (peer assessment) according a marking scheme. The agreement between the tutor marks and the students marks were generally very good. Fry further found five strong aspects of peer marking. These aspects will be discussed in a later section.

The relationships among self-, peer- and co-assessments are examined by Horgan, Bol and Hacker (1997). They used the predictions of grades, actual grades, peer reviews, and reflective essays on self-assessment of undergraduate teacher education students to analyze these relationships. The students were trained in self-assessment abilities. The students had to complete three multiple choice exams, from which the third was a cumulative final. Students had to predict their grade and after the exam they had to reflect on their performance. The students also had to do a written analysis of a case study, self-assessed and reviewed by three peers and the instructor based on five criteria. A third part in the assessment procedure Horgan et al. (1997) described was an oral case analysis as part of a group. These presentations were also reviewed by peers. The final part was an essay about the reflection on the self-assessment activities. Results of the assessments described above showed (1) agreement across assessors; (2) little consistency of self-assessment across tasks; (3) improvement in accuracy over the semester; (4) increased accuracy with increased performance; and (5) better students used self-assessments to guide work, while weaker students used feedback to find the errors.

2.4.4 Conclusion

The results described in this paragraph illustrate how effective self-, peer-, and co-assessments are used in combination. Self- and peer-assessment can be used for summative purposes as apart of the co-assessment, by giving the tutor the power to express the final decision about a process or a product. In this way the traditional assessment, where the tutor made an autonomous decision, is not comparable with co-assessment. The combination of self-, peer- and co-assessment makes tutors and students work together in a constructive way and as a result they come to higher levels of understanding by negotiation. When the student becomes teacher, this role-changing provides him insights into the assessment process.

Several studies indicate that the marks given by the tutors and those given by the students highly correlate (Fry, 1990; Freeman, 1995; Longhurst and Norton, 1997; Oldfield and Macalpine, 1995; Orpen, 1982). Only a few researchers found lower correlations between student and tutor marks (Rushton et al., 1993; Kwan and Leung, 1996). Arguments for these findings were that the assessment and the learning weren't sufficiently integrated or the fact that students are not capable enough to assess themselves the first time they have to do it.

It can be concluded that the use of self-, peer- and collaborative assessment are important to remove the student/tutor barrier, to develop enterprising competencies in students and can lead to greater motivation and deeper learning (Somervell, 1993). Where application of self-assessment and peer-assessment were mostly used for formative purposes, combinations of these forms with co-assessment do work out well for summative purposes. Various applications show various possibilities, ranging from using the peer-assessments as a contribution of e.g. 25% to the overall score to using peer-assessments as a correction for tutor-assessments.

Developments in this area do clearly open possibilities to assess skills and abilities where higher education had traditionally problems in assessing or (perhaps mainly) in the costs for assessing. If peer- and co-assessment indeed is a valid, fair and useful method for assessing essays and assignments, it may become a very widespread assessment method in the near future.

3 The effect of the use of self-, peer- and co-assessment on the quality of learning

In order to find answers to this second research question, we can refer to several previously named studies, showing the quantitative effects. Studies, such as the Hasmén et al. study (1997) show that students who use self-assessment procedures do get better scores on a final test. However, such studies mostly see assessments as support devices embedded in the learning materials and use a traditional final test (Martens, 1997).

Most studies given in earlier sections stress the agreement between self-/peermarks and teacher-marks (e.g. Falchikov & Boud, 1989; Stefani, 1992, 1994). These findings though give us less information about the impact the form of assessment has on the learning process. Above all, this is the main reason why these assessment forms need to be integrated in curricula in higher education. Falchikov and Boud (1989) stated in their discussion the following: although we have focused on student-teacher agreement over rating, we must not be distracted by the search to maximize congruence at all costs. Self-assessment can be a valuable learning activity, even in the absence of significant agreement between student and teacher, and can provide potent feedback to the student about both learning and educational and professional standards (p. 427). According to this view, self- and peer assessment can be regarded as independent competencies, rather than pure assessments. The studies in this literature review are more concerned with integrating learning and instruction, which means not only using assessment instruments to improve the learning process, but also adapting the assessment instruments (including the final examination) to the main goals of higher education. In this case we look at effects of new assessment forms on the quality of the learning process. Therefore, the studies in primarily the fifth category in the classification were analyzed: studies about the relationship between assessment and learning/ reflection.

An important concept that links assessment with the quality of learning is that of consequential validity (Boud, 1995, p. 41). This refers to the effects of assessment on learning and other educational matters. Boud describes it as follows:

Consequential validity is high when there is a positive backwash effect on learning and low when it encourages ways of learning which re counter to what are desired. Assessment procedures of high consequential validity should be developed.

Encouraging deep approaches to learning is one aspect which can be explored in considering consequences. Another is the impact which assessment has on the competencies and skills students have in being able to assess themselves.

The importance of developing students into reflective practitioners is already pointed out. Boud (1992, 1995) developed a self-assessment schedule to provide a comprehensive and analytical record of learning in situations where students have substantial responsibility for what they do. The main guidance is a handout which suggest the headings students might use. The headings are goals, criteria, evidence, judgments and further action. Self-assessment schedules are effective tools to use in enabling students to bring together a wide range of their learning, to reflect on their achievements and to examine the implications for further learning (Boud & Knights, 1994; Boud, 1992). Boud (1990) further recognized the gap between what is required of students in higher education and what happens in real life. He stresses the need for examination of assessment practices to see if they are compatible with

the goals in higher education as described in the introduction of this report. The two main goals of student assessment in his view are improvement of the quality of learning (by e.g. reflection) and secondly the (still existing) need to certify. Boud sees self- and peer assessment as fundamental to learning. Existing assessment practices might be more defensible if they could bear some relationship to the ways in which academic and other professional work is assessed in actual working environments and the situation in which knowledge is used.

Adams and King (1995) investigated the perception of self-assessment by different student groups and the skills required for self-assessing. They also recognize that employment at a professional level usually requires specialist knowledge. An important part of this knowledge is the ability to have a continual knowledge of one's own capabilities and to be able to update weaknesses as appropriate. Adults for example like to learn and are capable of considerable amounts of quality learning of their own. Adams and King identified activities that may develop self-assessment skills. A framework helps students to develop self-assessment skills to be competent at self-assessment. Adams and King identified three levels. At the first level students work on having an idea about the assessment process. Students perform activities such as: discussing good and bad characteristics of sample work, discussing what is required in an assessment, critical reviews on literature, etc. At the second level students work on identifying important criteria for assessment. At the third level students work towards playing an active part in identifying and agreeing assessment criteria and being able to competently assess peers and themselves.

The relation between reflection and self-assessment is also pointed out by Sobral (1997). Self-assessment of self-directed learning supported reflection and learning partnerships and facilitated by discussions and exercises. It is thereby stated by Longhurst and Norton (1997) that self-assessment is clearly an important part of helping students to improve their own learning, as it focuses students' attention on the metacognitive aspects of their learning and teaches them to be more effective at monitoring their own performance.

In some studies the perceptions from students towards innovative assessment and the impact on learning are investigated. Sambell, McDowell and Brown for example (1997) investigated the perceptions of students towards different aspects of innovative assessment. When discussing innovative assessment many students believed that success more fairly depended on consistent application and hard work, not a last minute burst of effort or sheer luck. Many students felt that openness and clarity were fundamental requirements of a fair and valid assessment system. Students were very positive about the effects of alternative assessment on their learning.

A small-scale study of the views of a group of newly enrolled Open University students in London resulted in a mixed response to alternative methods of assessment (Peters, 1996). The majority of the students disagreed with self- and peer-assessment. This finding did not mean that the students were committed to traditional forms of assessment. The possibility of being able to re-draft assignments after tutor feedback were viewed more favorably. Williams (1992) found that the vast majority of students see benefits in peer-assessment (90%). Benefits are seen in three main categories: in comparison of approaches, in comparison of standards and in exchange of information. However students found criticizing their friends to be difficult (see also Strachan & Wilcox, 1996). Students found also peer-assessment to be difficult or undesirable when guidelines for evaluation are not established first. The two major findings in the study of Williams (1992) were that (1) students like to have more say in how they approach their learning and its assessment and (2) that students need guidance and training in new role behaviors before this can actually happen. Orsmond et al. (1996) found that students did enjoy carrying out the peer assessment and beneficial to their learning. Keaten and Richardson (1993) also affirmed that peer assessment fosters an appreciation for internal awards and interpersonal relationships in the classroom.

Cheng and Warren (1997) conducted a research in the English Department of the Hong Kong Polytechnic University to gauge the students' attitudes prior to a peer assessment and after the peer assessment. The students and the teacher assessed each group seminar and oral presentation. Before and after the peer assessment the students filled out a questionnaire with four items. The results of the questionnaire show that the students were mostly positive towards the peer assessment, but that just a few students thought that beginning students were able to conduct the assessment in a fair and responsible manner. We already found the same in Falchikov and Boud (1989). Further, the students were not entirely confident in their ability to assess their peers. There was however a positive shift overall in both attitudes and confidence. Finally, Cheng and Warren concluded that there exists a need for giving students systematic and comprehensive training in how they can assess their peers on how to establish criteria (see also Williams, 1992).

4 Conclusion and discussion

Self-, peer- and co-assessment in higher education can be used in different ways and can be reliable and valid to a large extent. Assessment as a tool for learning has a great impact on the students' learning and development into reflective practitioners. In this final part of the report the main strengths and weaknesses of self-, peer- and co-assessment that arise from the studies described earlier will be discussed.

The strengths and weaknesses of the innovative assessment forms mentioned earlier do generally fit with the ones that are given by McDowell (1995). The main strengths are that (1) there is a development of evaluative and critical abilities; (2) there are opportunities for skills development; (3) knowledge is more integrated and (4) students collaborate, are motivated and satisfied. The weaker points of innovative assessment lie sometimes in organizational issues: cheating occurs and there can be stress and time constraints. At last there could be a mismatch between learning and marks whereby feedback is not always provided. Self-assessment improves the independence in students' learning, responsibility for decision making related to assignments, proactivity, and creativity in taking charge of their own work (Klenowski, 1995). The constraints on student self-assessment were lack of time, the paucity of professional development and support for student self-evaluation, and the change process itself. Adams and King (1995) remind us of a real problem of perception. The idea that teachers do the teaching and the marking is hard to change. Then there is the possibility that students take advantage of their role and that they become strategic in their approach to their studies. Students may also assess themselves too high. This though is not always the case, especially with better students (see for example Boud & Falchikov, 1989; Stefani, 1992). At last students may not always understand what is asked and that the tutors do not know how to mark the self-assessment. The idea of the self-assessment schedule described by Boud (1992, 1995) shows that students, once they experienced with it, see its value. The majority of students have been initially supportive and become enthusiastic having been through the process of constructing a schedule. A range of student responses affirm this fact (Boud, 1992, p. 191). There is much more concern about the self-grading aspect than there is about the qualitative assessment. If it is not possible to demonstrate that students can produce marks which are acceptable teachers, the self-assessment should be restricted to a purely learning role and as a skill to be developed (Boud, 1989). Until now, we still look thus at self-assessment as a formative tool. Moreover, it should be clear that students have to know the criteria clearly and that peer-assessment can be time consuming.

The following strengths in using self and peer assessment could be stated (see also Brown and Dove, 1991): (1) it can foster students' feelings of ownership for their own learning; (2) it can motivate them and encourage their active involvement in learning; (3) it makes assessment a shared activity rather than a lone one; (4) it promotes a genuine interchange of ideas; (5) it leads to more directed and effective learning; (6) it encourages students to become more autonomous in learning; (7) it signals to students that their experiences are valued and their judgment is respected; (8) it develops transferable personal skills; (9) it produces a community of learning in which students feel they have influence and involvement; (10) it reduces the teacher's workload (Rushton et al., 1993) and (11) it makes students think more deeply and see how others tackle problems, pick up points and learn to criticize constructively. Reading the strengths and weaknesses it may be concluded that the probably most difficult part in self-, peer- or co-assessment is to determine the criteria. Criteria are the basis of evaluating student progress; they identify the critical aspects of a

performance or a product that describe in specific terms what is involved in meeting the learning outcomes³. It is necessary for the concept that the criteria are presented in operational terms with which all participants are familiar. Criteria should include information about the area to be assessed, the aims to be pursued and the standards to be reached (Boud, 1995). Boud and Falchikov (1989) identified two elements in any assessment decision: the identification of criteria or standards to be applied to one's work, and the making of judgements about the extent to which work meets this criteria. In self-assessment the student judges his own performance and products against his own assessment criteria (Falchikov, 1986). Students also have to be trained in self- and peer-assessment. They have to learn and understand their role in the assessment process. Assessment should only be used in a summative assessment system whose outcome is not a grade or label but a *profile* of the student to which all who are able to speak about him can contribute what they know - and in which conflicting assessments are highlighted rather than ironed out (p. 297). Involvement in learning, including assessment, is vital to effective learning and that the teacher is the key person to help students develop this learner autonomy (Dickinson, 1992).

An important shift is that the future labor market will play a distinguished part in the way the curriculum and the goals are revised (Pilot, 1997). The form of assessment determines whether the student achieved the skills required for the working field. This working field will thus also influence the content of the assessments (Moerkerke & Terlouw, 1998). The curricula nowadays are more and more competency-based. This redesign of the curriculum will require a redesign of the assessment. Self-, peer and co-assessment can discourage passive, reproductive forms of learning. By integrating these forms of assessment in the curriculum students will develop into competent persons and lifelong learners who reflect continuously on their behavior and learning process. These alternative forms of assessment should be a part of a process of change towards a student-centered approach. This change requires a shift in emphasis from the norm-referenced to the criterion-referenced, from the purely summative to the formative and summative, from the external to internal and from the assessment of product to the assessment of process as well. In order for student self-assessment to be successful, the following supporting factors to be necessary: pedagogical change, a shared value system between students and teachers and a organization wide evaluation ethic.

Overall, one could state that self-, peer- and co-assessment do improve different aspects in the learning of students. However, the skill to self-assess or to peer-assess has to be trained for an optimal impact on the learning process. The prior section stresses the use of self-, peer- and co-assessment as tools for learning, i.e. formative ways to develop reflective practitioners in higher education.

³ <http://www.est.gov.bc.ca/.curriculum/www/irps/dance810/abdintro.htm>

References

- Adams, C., & King, K. (1995). Towards a framework for student self-assessment. *Innovations in Education and Training International*, 32, 336-343.
- Anderson, J.B., & Freiberg, H.J. (1995). Using Self-Assessment as a Reflective Tool to Enhance the Student Teaching Experience. *Teacher Education Quarterly*, 22(1), 77-91.
- Arter, J. (1996). Using Assessment as a Tool for Learning. In Blum, R. and Arter, J. (Eds.), *Student Performance Assessment in an Era of Restructuring*, Alexandria, VA: ASCD, 1-6.
- Beckwith, J.B. (1991). Approaches to learning, their context and relationship to assessment performance. *Higher Education*, 22(1), 17-30.
- Birenbaum, M. & Dochy, F. (Eds.) (1996). *Alternatives in Assessment of Achievement, Learning Processes and prior Knowledge*. Boston: Kluwer Academic.
- Boud, D. (1989). The role of self-assessment in student grading. *Assessment and evaluation in Higher Education*, 14(1), 20-30.
- Boud, D. (1990). Assessment and the Promotion of Academic Values. *Studies in Higher Education*, 15(1), 101-111.
- 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*. Kogan Page. London: Philadelphia.
- Boud, D. & Falchikov, N. (1989). Quantitative Studies of Self-Assessment in Higher Education: a Critical Analysis of Findings. *Higher Education*, 18(5), 529-549.
- Boud, D. & Knights, S. (1994). Designing courses to promote reflective practice. *Research and Development in Higher Education*, 16, 229-234.
- Brown, S. & Dove, P. (1991). Opening mouths to change feet: some views on self and peer assessment. In Brown, S., & Dove, P. (Eds). *Self and Peer Assessment*, Birmingham, Standing Conference on Educational Development, 59-65.
- Burnett, W. & Cavaye, G. (1980). Peer assessment by fifth year students of surgery. *Assessment in Higher Education*, 5(3), 273-278.
- Cheng, W., & Warren, M. (1997). Having Second Thoughts: student perceptions before and after a peer assessment exercise.
- 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(1), 45-56.

- Cutler, H., & Price, J. (1995). The Development of Skills through Peer-Assessment. In Edwards, A., & Knight, P. (Eds). *Assessing Competence in Higher Education*, Birmingham, Staff and Educational Development Series, 150-159.
- Dancer, W.T., & Dancer, J. (1992). Peer Rating in Higher Education. *Journal of Education for Business*, 67(5), 306-309.
- Dickinson, L. (1992). *Collaboration in Assessment: Empowering the Individual Course Member*. U.K.: England.
- Dochy, F.J.R.C., & McDowell, L. (1997). Assessment as a tool for learning. *Studies in Educational Evaluation*, 23(4), 279-298.
- Dochy, F., & Moerkerke, G. (1997). The present, the past and the future of achievement testing and performance assessment. *International Journal of Educational Research*, 27(5), 415 - 432.
- 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.
- Falchikov, N. (1991). Group Process Analysis. In Brown, S. and Dove P. (Eds.), *Self and Peer Assessment*. Birmingham, SCED Paper 63.
- Falchikov, N. (1995). Peer Feedback Marking: Developing Peer Assessment. *Innovations in Education and Training International*, 32(2), 175-187.
- Falchikov, N. & Boud, D. (1989). Student Self-Assessment in Higher Education: A Meta-Analysis. *Review of Educational Research*, 59(4), 395-430.
- Freeman, M. (1995). Peer assessment by Groups of Group Work. *Assessment and Evaluation in Higher Education*, 20(3), 289-300.
- Fry, S.A. (1990). Implementation and Evaluation of Peer Marking in Higher Education. *Assessment and Evaluation in Higher Education*, 15(3), 177-189.
- Gentle, C.R. (1994). Thesys: an expert system for assessing undergraduate projects. In M. Thomas, T. Sechrest & N. Estes (Eds.), *Deciding our future: Technological imperatives for education*, (pp. 1158-1160). Austin, Texas: The University of Texas.
- 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(3), 210-231.
- Griffee, D.T. (1995). *A Longitudinal Study of Student Feedback: Self-Assessment, Course Evaluation and Teacher Evaluation*. U.S.: Alabama.
- Hall, K. (1995). *Co-assessment: participation of students with staff in the assessment process. A report of Work in Progress*. 2nd European Electronic Conference On Assessment And Evaluation.
- Harrington, T. F. (1995). *Assessment of Abilities*. U.S.: North Carolina.

- Hassmén, P., Sams, M.R., & Hunt, D.P. (1996). Self-assessment responding and testing methods: effects on performers and observers. *Perceptual and Motor Skills*, 83, 1091-1104.
- Horgan, D.D., Bol, L. & Hacker, D. (1997). *An examination of the Relationships Among Self, Peer, and Instructor Assessments*. Paper presented at the European Association for Research on Learning and Instruction, Athens, August 1997.
- <http://141.163.121.36/TeachingLearning/Ethos/Assessment/PeerAss1.html>
- <http://cbl.leeds.ac.uk/~paul/papers/euroaiedpapers96/smpaper/smpaper.html>
- <http://www.est.gov.bc.ca/.curriculum/www/irps/dance810/abdintro.htm>
- Kane, J.S., & Lawler III, E.E. (1978). Methods of peer assessment. *Psychological Bulletin*, 85(3), 555-586.
- Keaten, J.A., & Richardson, M. E. (1993). *A Field Investigation of Peer Assessment as Part of the Student Group Grading Process*. U.S.: Colorado.
- Keith, S.Z. (1996). Self-Assessment Materials for Use in Portfolios. *Primus*, 6(2), 178-192.
- Klenowski, V. (1995). *Student Self-Evaluation Processes: Empowering Students in Learner-Centered Contexts*. Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 17-22, 1995). United Kingdom.
- Knors, E., Dochy, F., & Moerkerke, G. (1995). *Factors influencing the use of prior knowledge and progress assessment: a review of literature and a drop out study*. Heerlen: Open Universiteit Nederland.
- Kwan, K., & Leung, R. (1996). Tutor versus Peer Group Assessment of Student Performance in a Simulation Training Exercise. *Assessment and Evaluation in Higher Education*, 21(3), 205-214.
- Loacker, G., & Jensen, P. (1988). The power of performance in developing problem solving and self-assessment abilities. *Assessment and Evaluation in Higher Education*, 13(2), 128-150.
- Longhurst, N., & Norton, L.S. (1997). Self-assessment in Coursework Essays. *Studies in Educational Evaluation*, 23(4), 319-330.
- Martens, R.L. (1997). *The use and effects of embedded support devices in independent learning*. Utrecht: Lemma.
- McDowell, L. (1995). The Impact of Innovative Assessment on Student Learning. *Innovations in Education and Training International*, 32(4), 302-313.
- McNamara, M.J., & Deane, D. (1995). Self-Assessment Activities: Toward Language Autonomy in Language Learning. *TESOL Journal*, 5(1), 17-21.
- Moerkerke, G. (1996). *Assessment for flexible learning*. Utrecht: Lemma.

- Moerkerke, G. & Terlouw, C. (1998). Herontwerp van toetsing. *THEMA*, 1, 19-25.
- Oldfield, K.A., & Macalpine, J.M.K. (1995). Peer and Self-Assessment at the Tertiary Level--An Experiential Report. *Assessment and Evaluation in Higher Education*, 20(1), 125-132.
- Orpen, C. (1982). Student versus lecturer assessment of learning: a research note. *Higher education*, 11, 567-572.
- Orsmond, P., Merry, S., & Reiling, K. (1996). The importance of Marking Criteria in the Use of Peer Assessment. *Assessment and Evaluation in Higher Education*, 21(3), 239-249.
- Peters, M. (1996). Student attitudes to alternative forms of assessment and to openness. *Open Learning*, 11(3), 48-50.
- Pond, K., Ul-Haq, R., & Wade, W. (1995). Peer Review: A Precursor to Peer Assessment. *Innovations in Education and Training International*, 32(4), 314-323.
- Pilot, A. (1997). *Curricula voor een op informatie gerichte maatschappij*. Paper presented at the Consortium COO HBO and SUNCOO conference, 11th and 12th december 1997 in Utrecht.
- Rushton, C., Ramsey, P. & Rada, R. (1993). Peer Assessment in a Collaborative Hypermedia Environment: A Case Study. *Journal of Computer-Based Instruction*, 20(3), 75-80.
- Sambell, K. & McDowell, L. (1997). *The Value of Self and Peer Assessment to the Developing Lifelong Learner*. ISL Symposium 1997. To be published.
- 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.
- Schön, D.A. (1987). *Educating the Reflective Practitioner: Towards a New Design for Teaching and Learning in the Professions*. San Francisco: Jossey-Bass.
- Slavin, R. (1986). Best-evidence synthesis: An alternative to meta-analysis and traditional reviewers. *Educational Researcher*, 15(9), 5-11.
- Sluismans, D., & Dochy, F. (1998). Het gebruik van self-, peer- en co-assessment: een literatuurstudie. *Tijdschrift voor Hoger Onderwijs* (accepted).
- Sobral, D.T. (1997). Improving learning skills: a self-help group approach. *Higher Education*, 33(1), 39-50.
- 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), 221-33.
- Stefani, A.J. (1992). Comparison of collaborative, self, peer and tutor assessment in a biochemistry practical. *Biochemical Education*, 20(3), 148-151.

- Stefani, L.A.J. (1994). Peer, self and tutor assessment: relative reliabilities. *Studies in Higher Education, 19*(1), 69-75.
- Strachan, I.B., & Wilcox, S. (1996). Peer and self assessment of group work: Developing an effective response to increased enrollment in a third-year course in microclimatology. *Journal of Geography in Higher Education, 20*(3), 343-353.
- Warkentin, R.W. et al. (1995). *An Exploration of the Effects of Cooperative Assessment on Student Knowledge Structure*. Paper presented at the annual meeting of the American Educational Research Association (San Francisco, CA, April 18-22).
- Weaver II, R., & Cotrell, H.W. (1986). Peer evaluation: a case study. *Innovative Higher Education, 11*(1), 25-39.
- Williams, E. (1992). Student Attitudes towards Approaches to Learning and Assessment. *Assessment and Evaluation in Higher Education, 17*(1), 45-58.
- Zoller, Z., & Ben-Chaim, D. (1997). *Student self-assessment in HOCS Science Examinations; is it compatible with that of teachers?* Paper presented at the European Association for Research on Learning and Instruction, Athens, August 1997.

Published before

- 96/R01 Brok, E.C.T. & R.L.. Martens (1996). *Auteurservaringen met een Interactieve Leer- en CursusOntwikkelomgeving (ILCO)*.
- 96/R02 Weges, H.G. & S. J. Portier (1996). *De experimentele cursus Milieurecht: ervaringen van studenten*.
- 96/R03 Weges, H.G. (1996). *De experimentele cursus Milieurecht: het didactisch ontwerp en het interactief ondersteuningssysteem WindMill*.
- 96/R04 Portier, S.J., R.L. Martens & M.M.A. Valcke. (1996). *Ervaringen van studenten met drie flexibele, interactieve Oucursussen*.
- 96/R05 Leeuwe, M.F.J. de. (1996). *Het gebruik van animaties in een interactieve leeromgeving*.
- 96/R06 Martens, R.L. & S.J. Portier. (1996). *Dynamische begeleiding in een interactieve leeromgeving: een onderzoek bij studiebegeleiders van de Open universiteit*.
- 96/R07 Kirschner, P.A., P.P.M. van Vilsteren, H.G.K. Hummel & M.C.S. Wigman. (1996). *The design of a study environment for acquiring academic and professional competence*.
- 97/R01 Boon, J., J. Janssen & F. Cox.(1997). *ICT in de onderwijsorganisatie. Drie cases in het hoger onderwijs*
- 97/R02 Martens, R. , P. Poelmans & M. Valcke. (1997) *Students working with and reflecting about flexible embedded support devices*
- 97/R03 De Haan, D., P. Fisser, & H. de Wolf (1997) *ICT-gebruik in Hoger Onderwijs Inventariserend onderzoek*
- 98/R01 Bos, Evelijn S. (1998) *Competentie. Verheldering van een begrip*
- 98/R02 Martens, R. (1998) *Gebruikerservaringen Mercator. CD-rom uitlevering bij de cursus Natuurbeheer en natuurontwikkeling*
- 98/R03 Boon, J. & G. Joosten (1998). *Is de huidige generatie studenten klaar voor de derde generatie afstandsonderwijs? Bezit van computerapparatuur en houding ten aanzien van ICT-ontwikkelingen in het onderwijs*

