A CONCEPTUAL FRAMEWORK FOR INTEGRATING PEER ASSESSMENT IN TEACHER EDUCATION

Dominique Sluijsmans and Frans Prins

Open University of the Netherlands, Heerlen, the Netherlands

Abstract

Peer assessment can be a valuable learning tool in teacher education because it supports student teachers to acquire skills that are essential in their professional working life. This article presents a conceptual framework in which the training of peer assessment skills by means of peer assessment tasks is integrated in teacher education courses. Theories about constructive alignment, student involvement, instructional design, and performance assessment underlie the framework. Furthermore, two recently published empirical studies will be briefly described to provide empirical support for the value of the framework. Results of these studies show that the framework offers powerful guidelines for the design and integration of peer assessment activities in teacher training courses. In general, the peer assessment tasks that were embedded in the courses led to a general improvement in students' peer assessment skills as well as their task performance in the domain of the course. Implications for course and curriculum design are discussed.

Teacher training colleges face the complex task of educating student teachers who, in turn, have to educate pupils in elementary schools in the future. Two recent trends in education, i.e., the design of more competency-based curricula and the involvement of students in assessment (Verloop & Wubbels, 2000), urge teacher training colleges to modify their educational practices. In this article we argue that the use of peer assessment in the curriculum of student teachers fits well in a competency-based curriculum and that it fosters student involvement in assessment. To be effective, however, peer assessment training should be embedded in the existing course material that is designed according to a performance-based approach (Mehrens, Popham, & Ryan, 1998). We present a conceptual framework for integrating peer assessment in teacher education. Then two studies will be briefly described to provide empirical support for the ecological value of the framework.
Institutions of higher education in general are continuously challenged with a demand for competency-based learning. A curriculum should focus more on competencies such as learning to learn, interactive skills, communication skills, information processing, problem-solving, and reflective skills (Tillema, Kessels, & Meijers, 2000). Skill-based learning is an ongoing issue in the domain of teacher education (Darling-Hammond & Snyder, 2000; James, 2000; Kremer-Hayon & Tillema, 1999; Willems, Stakenborg, & Veugelers, 2000). A number of teacher training colleges collaboratively formulated a broad scale of skills student teachers need to develop. These skills of a primary school teacher are reported in a vocational training profile (LPC, 1995), which consists of 41 skills. These skills represent the overall accepted knowledge, proficiency and attitudes a primary school teacher needs to acquire.

The skill to assess the work of peers is a specific skill of the vocational training profile of primary school teachers. The process whereby individuals evaluate the performance of their peer is called peer assessment (Falchikov, 1995; Freeman, 1995). In our view, peer assessment is a powerful didactical method for teaching skills which is important for the teaching domain for at least four reasons. First, teachers have to work together, learn from each other and become a member of a learning organisation (Verloop & Wubbels, 2000). Besides, the importance of communication between teachers in schools has been endorsed by many researchers (Cohen, 1994; Johnson, Johnson, & Johnson-Holubec, 1992; Sharan & Sharan, 1994; Slavin, 1995). In a peer assessment task, students have to communicate and collaborate and thus they are able to acquire communication and collaboration skills. Second, discussion about reflection is an ongoing issue in teacher education (e.g., Korthagen, 1985, 2001; Newman, 1996; Reilly Freese, 1999; Richert, 1999). Encouraging students to assess each other's contributions to discussion and discourse, as in peer assessment, further exposes them to the skills of critical reflection and analysis (Birenbaum, 1996; Sambell & McDowell, 1998). Reflection skills are necessary for making reliable judgments about peers' work. Thus, peer assessment fosters reflection and the development of reflection skills. Third, student teachers will become assessors in their own classroom and, therefore, they will have to design assessments as prospective teachers of children in primary schools. It is therefore advisable to teach student teachers how to make critical judgements about the performance of their peers, and, later on, about performances of children. The last reason for the importance of peer assessment in teacher education is that after students have left higher education, they are likely to rely heavily on the judgement of their peers to estimate how effective their performances in the school are (Brown, Rust, & Gibbs, 1994). Being able to interpret the work of colleagues and peers is a necessary prerequisite for professional development and for improving one's own functioning (Verloop & Wubbels, 2000). Training in peer assessment skills stimulates this mutual influence to take place at a professional level.

Performance Assessment as a Foundation for Peer Assessment Tasks

Peer assessment is regarded as a learning tool that may have positive effects on skills that are relevant for teachers. In our view, performance assessment should be the
foundation for peer assessment tasks. Performance assessments are described in terms of a certain performance that is content related and is perceived as worthwhile and relevant to the student in relation to their future profession. This performance may or may not represent an authentic situation (Wiggins, 1989). Performance assessment focuses on the ability to use combinations of acquired skills and knowledge, and therefore fits in well with the theory of constructive alignment and powerful learning environments (Birenbaum, 2003; Linn, Baker, & Dunbar, 1991).

Performance assessments require individuals to apply relevant knowledge and skills in context, not merely completing a task on cue. Students are observed while they are performing, products they create are examined, and the level of proficiency demonstrated is judged. Performance assessment can be based on multiple products or processes, for example essays, reflection papers, oral assessments, simulations, process-analyses, group-products, and work-samples. Judgments are made about the level of achievement attained by comparing student performance to predetermined standards. All students have the opportunity to attain the standards, whereby they can play a crucial role in making judgments about the performance of their peers and defining appropriate criteria for these performances. The importance of the negotiation about criteria has already been stressed in several studies (Boud, 1995; Orsmond, Merry, & Reiling, 1996; 1997; 2000). Or as Stiggins stated: "Once students internalise performance criteria and see how those criteria come into play in their own and each other's performance, students often become better performers" (1991, p. 38).

As opposed to most traditional forms of testing, performance assessments do not provide clear-cut right or wrong answers. The performance is evaluated in a way that allows for informative scoring on multiple criteria. This is accomplished by creating assessment forms. In these forms teachers determine at what level of proficiency a student is able to perform a task or display knowledge of a concept. For example, the different levels of proficiency for each criterion can be defined. Using the information yielded by the assessment form, feedback is given on a student's performance either in the form of a narrative report or a grade. A criterion-referenced qualitative approach is desirable, so that the assessment will be carried out against previously specified performance criteria. An analytic or holistic judgment then is given on the basis of the standard the student has achieved on each of the criteria. Effective application of performance assessment methodology relies on thoroughly trained raters who in their turn rely on sound performance criteria to observe and evaluate student responses to quality exercises (Stiggins, 1994).

Designing Performance Assessments

A common error in designing a course or unit of study is to leave the development of the performance assessment as a final activity (Airasian, 1991). The compatibility between learning, instruction and assessment is a basic assumption for our framework. Biggs' (1996, 1999, 2001) theory of constructive alignment and Stiggins' (1987) approach are useful to design courses and performance assessments. Four steps can be taken to design courses in which instruction and assessments are completely aligned. First, teachers must have a clearly defined purpose of a course. The concepts, skills, and knowledge to be assessed, as well as the level at which students should be performing, must be determined (Stiggins,
1987). Second, it must be decided what type of activity best suits the assessment needs. This can result in a skill decomposition in which the relevant skills are hierarchically ordered, or in which they are organized in a concept map. Third, decisions should be made concerning the assessment task. Issues that must be taken into account are time constraints, availability of resources, and how many data are necessary in order to make an informed decision about the quality of a student's performance. Finally, after the assessment task is determined, the elements of the task that determine the measure of success of the student's performance need to be defined. Sometimes, these can be found in so-called job profiles.

Most of the time, teachers have to analyse skills or products to identify performance criteria upon which to judge achievement, which is not an easy task. Criteria should be significant, specifying important performance components, represent standards that would apply naturally to determine the quality of performance when it typically occurs (Quellmalz, 1991). The criteria must be communicated clearly to, and be able to be understood by, all involved. Communicating information about performance criteria provides a basis for the improvement of that performance. When a teacher has passed through this procedure, study tasks can be designed in which students are prepared for the performance assessment. These study tasks are directly related to the performance assessment task at the end of the course.

Designing Courses in Which Peer Assessment is Integrated

According to Sluijsmans, Dochy, and Moerkerke (1999), teacher educators should be supported in the design of learning activities in which peer assessment is integrated. Stiggins's (1987) abovementioned design guidelines are helpful. Step 1 of the design process is to define the purpose of a course. It should be emphasized that a course that includes peer assessment tasks contains multiple learning goals. The performance of the student at the end of the course is content related and can be labelled as the first order goal of a course. Acquiring peer assessment skills is subsequently integrated as a higher order goal in a particular course. Students learn to evaluate the course content-related performances of peers at the end of a course. Peer assessment can thus be considered as a performance assessment that is superimposed on the content-related performance assessment. When the acquisition of peer assessment skills is one of the purposes of a course, students should be capable, at the end of the course, of making arrangements in which they negotiate with students of similar status about the design and appropriate criteria of specific study tasks and performances. Each student should also be able to take the responsibility to make critical judgements about the performances of a peer applying the appropriate criteria. It should be noted that peer assessment skills are not easily and automatically acquired. Peer assessment is considered a complex skill that needs to be developed (Birenbaum, 1996; Reilly Freese, 1999; Sluijsmans, Dochy, Moerkerke, & Van Merriënboer, 2001). Students who are novices in assessment are insecure about their ability to assess and indicate that they need more guidance on the marking criteria (Cheng & Warren, 1997; Woolhouse, 1999). Normally students need explicit training in assessment techniques during the course to make reliable and acceptable assessment reports (Boud, 1990; Hanrahan & Isaacs, 2001).
The method of skill decomposition is applied to identify constituent skills (Van Merriënboer, 1997). Here, peer assessment is broken down into separate skills and these skills are practiced one at a time, before being recombined and practiced as a complete task. Figure 1 represents the skill of peer assessment. Each constituent skill of the peer assessment is further described in Table 1. Data for this decomposition were gathered through literature review and feedback from experts in the field of peer assessment. The horizontal relationship in Figure 1 illustrates which specific skills are necessary in order to
be able to perform the skill under consideration. The vertical relationship illustrates which other skills are necessary to be able to perform the peer assessment skill.

Table 1: Description of the Constituent Peer Assessment Skills

<table>
<thead>
<tr>
<th>First level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Define assessment criteria</td>
<td>The student actively participates in a group discussion to reach a common understanding about the assessment criteria for the product to be assessed</td>
</tr>
<tr>
<td>Judge the performance of a peer</td>
<td>The student individually assesses a product of a peer in the product and then formulating the discrepancies between the product and the criteria. The formulated discrepancies are written down in a peer assessment report.</td>
</tr>
<tr>
<td>Provide (anonymous) feedback for future learning</td>
<td>The student writes a feedback report that provides feedback for future courses. This feedback: confirm the peer’s understanding of what the product required was correct; helps the student to add information to his or her own knowledge when they experience an information gap; helps the peer to replace the erroneous information with more accurate information.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Second level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Develop <em>personal</em> course objectives on the basis of given course objectives and group discussion</td>
<td>The student presents his/her personal interpretations of the course objectives and argues his/her view in a group session.</td>
</tr>
<tr>
<td>Write a personal report on course objectives</td>
<td>The student individually writes a report that reflects his/her interpretation of the course objectives.</td>
</tr>
<tr>
<td>Couple course objectives to study tasks</td>
<td>In collaboration with peers, student relates the defined course objectives to the different tasks to be carried out to reach the course objectives, and formulates which part of the task contributes to which course objective.</td>
</tr>
<tr>
<td>Develop measurable criteria for each study task</td>
<td>In collaboration with peers, the student lists the criteria that were decided for the task; these criteria are the result of the task analysis.</td>
</tr>
<tr>
<td>Analyse the performance of a peer</td>
<td>The student individually applies the assessment criteria to the product of the peer after reading the product and marks the evidence for the presence of the criteria.</td>
</tr>
<tr>
<td>Formulate discrepancies in a peer assessment report</td>
<td>The student writes an assessment report on the quality of the product which reflects evidence for reaching the desired criteria at a certain level.</td>
</tr>
<tr>
<td>Formulate points for improvement</td>
<td>The student writes individually a number of points for improvement based on the assessment criteria and the group discussions in which the assessment criteria were decided.</td>
</tr>
<tr>
<td>Reflect on points of improvement for the peer</td>
<td>Based on the assessed product, the student individually presents and argues points for improvement to the peer.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Third level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analyse given course objectives</td>
<td>The student interprets given course objectives based on prior knowledge and personal values.</td>
</tr>
<tr>
<td>Summarise results of the group discussion</td>
<td>The student takes an active role in the group discussion and writes a report which represents the outcomes of the discussions.</td>
</tr>
<tr>
<td>Analyse the study task</td>
<td>The student discusses the study task with the peers and formulates common criteria that the student must meet in order to carry out the task in a proper way.</td>
</tr>
</tbody>
</table>

The performance assessment task for determining the quality of the peer assessment skill should then be chosen (step 2: Choosing the format of peer assessment). Normally this task is to write an assessment report about the performance of a peer at the end of the
course. This assessment report can be used for summative assessment purposes, while the embedded peer assessment tasks have a more supportive function in developing the skills that are conditional for conducting a peer assessment. Both the quality of the assessment reports and the performance assessments can be examined by the teacher educator. Assessments of peer assessment skill are however still very rare in teacher education.

Based on the skills presented in the model, criteria have to be defined for a good assessment report (step 3: Definition of criteria). Assessments written by expert assessors can be used to determine these criteria. Criteria are determined regarding the use of adequate criteria, giving feedback and the style of a written assessment report. In practice, students write a qualitative assessment report about a performance of one or more peers on a blank peer assessment form. A rating form has to be developed to analyse the quality of students’ peer assessments (step 4). Naturally, this rating form is based on the criteria for a good assessment report. Teacher educators use the rating form to determine the quality of the assessment skill.

An Integrated Framework For Training Assessment Skills

Figure 2 illustrates how the concepts presented in the previous sections are integrated in a framework that underlies our two empirical studies.

Surveying the concepts discussed in the previous sections, it can be concluded that there are two parallel paths, illustrated by the dotted arrows. In the first-order course design path, students are guided in the acquisition of content-related skills through study tasks with the aim to meet the criteria for the content-based performance assessment. The second path is the higher-order course design path, in which students are supported in the acquisition of peer assessment skills, by means of peer assessment tasks (PA-tasks). These peer assessment tasks, which are superimposed on the regular study tasks, are characterised by collaborative learning, more specifically by social interaction, individual accountability and positive interdependence (Slavin, 1989). Students work towards two assessments: a content-related assessment (the first order course goal) and a peer assessment (the higher order course goal).

The two paths are integrated (see the two dotted arrows), in other words, the peer assessment tasks are completely embedded in the study tasks of the course, because the content of the study tasks provide input for the peer assessment tasks. The first-order and higher-order course design are the basic components of the framework and are defined by reference to the theory of student involvement, the constructive alignment theory and the design principles of Stiggins (1987). At the end of a course, students have to carry out a performance assessment, which is subsequently object of the peer assessment.
Researchers Sluijsmans and colleagues conducted two studies within the context of teacher education to examine the effects of an embedded training in peer assessment skills on students' performance in their peer assessment skills and content-based skills. Below, these two studies are briefly described to illustrate how peer assessment can be integrated in a course according to the conceptual framework presented in Figure 2. The following research questions were explored:

1. Does training in peer assessment lead to the development of the skill to assess the work of peers (the higher order goal)?
2. Does following training in peer assessment lead to an improved task performance in the domain of a course (the first order goal)?
3. What are of students' and teachers' perceptions regarding the implementation of the framework?

Figure 2: Student Involvement and Course Design for Powerful Learning Environments – An Integrated Framework

Empirical Support for Implementing the Framework

Recently, Sluijsmans and colleagues conducted two studies within the context of teacher education to examine the effects of an embedded training in peer assessment skills on students' performance in their peer assessment skills and content-based skills. Below, these two studies are briefly described to illustrate how peer assessment can be integrated in a course according to the conceptual framework presented in Figure 2. The following research questions were explored:

1. Does training in peer assessment lead to the development of the skill to assess the work of peers (the higher order goal)?
2. Does following training in peer assessment lead to an improved task performance in the domain of a course (the first order goal)?
3. What are of students' and teachers' perceptions regarding the implementation of the framework?
The rationale underlying the conceptual framework led us to expect that the training in peer assessment would have positive effects on the development of peer assessment skills as well as on task performance in the domain of the course.

For a complete overview of these studies we refer to Sluijsmans, Brand-Gruwel, and Van Merriënoer (2002) for Study I, and Sluijsmans, Brand-Gruwel, Van Merriënoer, and Bastiaens (2003) for Study II.

Study I

In the first study, second-year students of a Primary Teacher Training College in the Netherlands were trained in peer assessment skills. For the purpose of this study, a six-week course on creative learning was chosen. The teachers who were jointly responsible for this course first redefined the course objective since it had not been revised for several years and teachers had developed multiple perspectives on what the content should be. It was decided that students would be guided in the content skill designing a creative lesson (first order goal). At the end of the course, students had to make a videotape of a creative lesson, designed and conducted by them. The four teachers collaboratively decomposed the skill of designing a creative lesson. This resulted in a concept map with a number of constituent skills. Four one-hour study tasks were defined, based on the constituent skills. In these tasks, students learned how the content of different domains such as art, Dutch language, and music, were related to creative learning and the design of creative lessons. The whole course included an introductory class, sixteen study tasks, and a concluding class in which the peer assessment was organized.

Ninety-three student teachers were randomly assigned to control groups and experimental groups. The experimental groups were trained in peer assessment skills. Before the start of the course, all students filled out a questionnaire about their perceptions on instruction and assessment. Items were divided among several variables related to instruction, vision on instruction and assessment and the role of the student in assessment. The pre-test was carried out to investigate the students’ perceptions on prior courses that were comparable to the courses selected in the present studies. These prior courses were not designed in a skill-based way. The post-test was focused on students’ perceptions on the redesigned, skill-based course. The students had to score the items on a five-point Likert scale, varying from I totally disagree to I totally agree.

During the course, all students worked in subgroups of five or six students on their design of a creative lesson and the group report. They prepared their lesson, which was taped on video and subject to the peer assessment. In between classes, each student worked individually on the individual report and the study task-related assignments. During the course the students in the experimental groups performed the four peer assessment tasks of one hour each. These tasks were embedded in the study tasks concerning the design of creative lessons. The training focused on the three main constituent skills of the peer assessment model (see Figure 1).

At the end of the course, a peer assessment session was organized for each group (approximately 25 students), in which the video lessons of each subgroup were shown (four video lessons in each group). The peers were instructed to write a qualitative peer
assessment report with regard to the content of the video lesson of each group on a blank peer assessment form. The experimental groups were free to use the output of the peer assessment tasks. The students from the control group used the regular course materials from the study tasks. Each student wrote three peer assessment reports, because in each group there were three other subgroups to assess. After the course, all students filled out the same questionnaire as in the pre-test.

The teacher educators who were involved in this course evaluated the four peer assessment tasks by means of a short questionnaire. The questions concerned issues related to the framework on the basis of which the courses were redesigned and the peer assessment tasks were integrated. The questions related to two phases: the design phase of the course and the implementation phase. Regarding the design phase, questions were asked about participants' experiences with the redesign of the course and their joint work with the other colleagues. Questions concerning the implementation phase referred to the experiences with the instruction of the peer assessment tasks and participants' view on assessment and instruction, and their own role as well as that of the students.

Study II

The findings of the first study set the design and goals of the second experiment. In this longitudinal study, 110 first-year student teachers were trained in peer assessment skills within three courses on mathematics. The study was designed according to a within-subject repeated-measures design. Students participated in the experiment for a period of seven months. In a two-hour intake session that took place a day before the start of the first mathematics course, the students carried out three activities: filling out the student questionnaire (identical to Study I); writing a reflection paper about prior experiences in mathematics, and assessing an anonymous reflection paper. This anonymous reflection paper was previously marked "unsatisfactory" by the mathematics teacher. After the intake, all students attended three successive courses on mathematics. Within the three courses students were confronted with basic skills that are required for teaching mathematics. In addition, the students had to write a reflection paper after the first course, which could be improved after the second and third course, to submit a final version two weeks after the last feedback session. All students received training in the assessment skills during the courses. The assessment training was directed at three topics: what are important criteria for a reflection paper (four tasks in the first course), how to give feedback (two tasks in the second course), and how to write an assessment report (two tasks in the third course). In this third course, for example, students developed a peer assessment form based on an expert assessment report that was written by the mathematics teacher.

The output of the first part of the training was a list of 19 criteria for a reflection paper. Students agreed in negotiation with the mathematics teacher that a good reflection paper contains, for example, self-criticism, work field experiences, personal expectations, and strengths/weaknesses.

In the second part of the training, integrated in the second course, students developed guidelines for giving feedback. One guideline that students agreed on was that it
would be positive for a peer to mention their own learning experiences in the assessment report.

In the third and last part of the training, which was embedded in the third course, students worked on a peer assessment form and decided what is important in the writing of an assessment report. An expert assessment report served as an example.

Students were instructed that the criteria, feedback rules and structure guidelines derived from the peer assessment training could be helpful in writing the reflection papers and the peer assessment. After each course, the students had to send their reflection paper to the other students. This was done using the facilities of Blackboard®, a virtual learning environment. Each student was asked to assess the reflection paper of another student, in such a way that every student had to assess and was assessed by different peers. After each course, a feedback session was organized, chaired by the mathematics teacher. In these sessions, in which a group of ten to twelve students participated, each student orally presented his or her assessment report. The written report was given to the assessed student after the feedback session. The students used the peer feedback to rewrite and improve their reflection paper. The student feedback can be regarded as the formative assessment of the papers. To reduce test anxiety and to lengthen the period in which the peer assessment skills were trained, students received no grades for their reflection paper after each course. The role of the teacher was limited to coaching and chairing in the feedback sessions. The reflection paper that was written based on the given peer feedback after each course was used for the final grade given by the mathematics teacher. After the third feedback session, an outtake session took place, similar to the intake. In this session, all students filled out the student questionnaire again. They also wrote an assessment report of the same reflection paper that was presented in the intake session. During the experiment, the mathematics teacher evaluated the four peer assessment tasks by means of the questionnaire of Study I.

**Overall Results**

Two studies were described to illustrate the implementation and the effects of the framework depicted in Figure 2. To analyse the quality of the peer assessments that were written by the students, a rating form with underlying variables derived from the peer assessment model was developed. Variables were related to the skills of defining criteria, giving feedback, and writing an assessment report. In each study independent research assistants scored the peer assessment forms with the rating form. To measure an effect of the peer assessment training on students' content-related skill, the marks on the performance assessments given by the teacher were analysed. These performance assessments were a video on creative learning (Study I) and a reflection paper (Study II).

We shall summarise the effects for each of the three research questions separately. Table 2 shows the effect sizes for the variables that concern our research questions 1 (effect on the assessment skill) and 2 (effect on the content related skill). Cohen (1988) defined an effect size of approximately 0.2 as small, of 0.5 as medium, and of 0.8 as large.

Results showed positive effects of peer assessment training on the students' skill to assess the work of peers (research question 1). In the first study, the analyses of the qualitative peer assessment reports revealed that the experimental groups were more likely
to use the criteria and to give more constructive comments than the student teachers from the control groups. The students who received training also scored higher on structure and wrote in a less naive manner. In spite of the positive results of this study, we concluded that student teachers could not be regarded as expert assessors after peer assessment training in one course. The training in the second, longitudinal study was integrated in three successive mathematics courses. Analysis of the peer assessments from the intake and outtake data revealed significant progress for most variables. All students used the criteria more adequately, gave more constructive feedback, and wrote more structured assessment reports after the training period of ten months. Students also adopted a more critical attitude in the outtake than in the intake.

Table 2: Effect Sizes of Main Effects for Variables Concerning the Three Research Questions

<table>
<thead>
<tr>
<th>Variables</th>
<th>Study I</th>
<th>Study II</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Question 1: Development of peer assessment skill</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Using criteria</td>
<td>1.31</td>
<td>1.43</td>
</tr>
<tr>
<td>Constructive comments</td>
<td>1.02</td>
<td>1.22</td>
</tr>
<tr>
<td>Structure</td>
<td>0.31</td>
<td>2.61</td>
</tr>
<tr>
<td>Naive formulation</td>
<td>-0.61</td>
<td>ns</td>
</tr>
<tr>
<td><strong>Question 2: Improved task performance</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning results</td>
<td>0.72</td>
<td>2.24</td>
</tr>
</tbody>
</table>

Note: ns = not significant

The second research question focused on the effect of training peer assessment skills on students' content-related performance. In Study I, a positive effect of the peer assessment training on the actual learning results was found. The student teachers from the experimental groups outperformed the students from the control groups. This same result was found in the second study, where the total group of students wrote better reflection reports after the training than before the training.

We also examined student teachers' perceptions regarding the implementation of the framework for integrating peer assessment in teacher education. Results of the two studies showed that the students were more positive about the instruction and the integration of assessment and instruction after they took the redesigned course. The renewed course, which was designed from a skill-based perspective and consisted of tasks that fostered collaborative learning and interaction, led to the active participation of student teachers and teachers. It can be concluded that the student teachers positively changed their view on aspects of learning and assessment. They were more satisfied about the classes and the criteria, and goals were clearer. The role of the teacher was also evaluated in a more positive way. The student teachers indicated that they felt more capable in assessment than before the course was redesigned.
Teacher experiences were investigated in the second study. It appeared that the redesign process and the implementation phase demanded a lot of effort on the part of the teachers who were involved in the courses. The revision of the courses did lead to some resistance. Some teachers doubted the value of the peer assessment and were sometimes reluctant to give up some part of their content expertise on behalf of the "higher order" skills. In both studies, however, the teachers had no major problems in instructing the peer assessment tasks. The teachers indicated that implementing the peer assessment training led to a rethinking of the existing course and stimulated them to view the content from a different perspective.

Discussion

We presented a conceptual framework for integrating peer assessment in teacher education and supported it with empirical results from two studies. Each study was conducted within a teacher training context, in which the skill to assess peers' work is considered to be important. Our hypothesis was that if student teachers were trained to assess the performance of peers, this should lead to a general improvement in their peer assessment skills as well as their task performance in the domain of the course. Results of the studies corroborated this hypothesis. It is conceivable that peer assessment training and more critical reflection about assessment might have a long-term effect for students.

A relevant question for future research is how the design of courses and the design of assessment training may be most conducive to skill acquisition. A reconsideration of the peer assessment model and the collaborative activities that were used in the framework appear to be desirable. It is also interesting to elaborate further on the relationship between peer assessment skill acquisition and content skill acquisition.

By involving students in the design of instruction and assessment, they become aware of how and on what knowledge and skills they are assessed. Peer assessment can be conceived as an evaluative device, but in our approach it is also a powerful learning activity. The student acts as an important collaborator with the teacher in the creation of tasks as well as in developing guidelines for scoring and interpretation. Until today, many tests are kept under lock and key so students do not have knowledge about them ahead of time. This causes students to study in a particular way, in hopes that they will improve their test performance, but there is virtually no way that students can "learn by doing" as happens through engaging in a performance-based assessment in which they are involved as one of the assessors (Frederiksen, 1984).

The framework has implications for course design. Within the framework of skill-based curriculum design, the educational material is no longer defined from the perspective of the content domain, but from the perspective of the relevant skills (Tillema, Kessels, & Meijers, 2000). This means that skills are trained in the context of different content domains. Working with the framework encouraged teacher educators to think about the performance assessment at the beginning of a course design process. Assessment drives the learning process and overrides practically every other aspect of curriculum design (Longhurst & Norton, 1997). Changing assessment practices towards more performance-based approaches, will inevitably lead to a revision of instruction. Instruction, assessment, and learning and teaching strategies have to be completely aligned. Educators must develop
appropriate assessments that have no single right answer and in which students' argumentation is key in defending their solution. The involvement of students in these processes implies an extra investment. Although the present studies focused mainly on the training of student teachers, it became increasingly apparent that much effort has to be put into the professional development of teacher educators. Meanwhile, initiatives are underway to define a vocational profile for teacher educators (Koster & Korthagen, 2001). The competencies of teacher educators are operationalised (Plake, Impara, Fager, 1993). Designing rich, authentic performance assessments is one of these competencies that deserves special attention. After all, assessment is the tail that wags the dog.

Changing assessment practices and views on learning, as well as the role of students in this, is a considerable challenge in teacher education and higher education in general. The success of sound assessment practices lies on the one hand in a close relationship between learning, instruction, and assessment, on the other hand in qualified (student) assessors. The framework and studies presented in this article constitute an attempt to make a contribution to both aspects. Important guidelines for practice are that students need to be guided in their skill-development; that a clear definition of performance criteria is crucial for effective assessments; that collaborative activities need to be stimulated, and that teacher educators receive training in instructional design and alternative assessment approaches. According to the practice as you preach philosophy, an important condition for successful initiatives on the student level is that teachers are receptive for self-reflection and change.

References


**The Authors**

DOMINIQUE SLUIJSMANS is assistant professor at the Educational Technology Expertise Center (OTEC) at the Open University of the Netherlands in Heerlen. Her main interests are student involvement in assessment, teacher education, and curriculum design.

FRANS PRINS is assistant professor at the Educational Technology Expertise Center (OTEC) at the Open University of the Netherlands in Heerlen. He has been involved in research in higher education with main interests in learning styles, metacognition, intellectual ability, inductive learning, and peer assessment.

Correspondence: <dominique.sluijsmans@ou.nl>