

Peer assessment in university teaching: evaluating seven course designs

Introduction

Learning to write at a proper academic level is an integral part of university education, and cannot be successful when it is isolated from the particular disciplinary content. In addition, the acquisition of academic writing skills is a long-term process. So, in many curricula, teachers search for proper methods of providing more support to students in developing their writing competence.

Research on collaborative learning, and studies on peer assessment, show that students learn from peers by collaboratively studying educational materials and assessing each other's work. They learn more when the assessment procedure includes feedback on the products and processes. Saunders (1989) distinguished between different forms of organising students' assistance during this process of learning to write, each determining the degree of student interaction and thus the results of their collaboration. In a review study, Topping (1998) reports positive effects of peer assessment, especially on writing in higher education. In this study, a definition and typology of peer assessment is being proposed. Peer assessment is defined as 'an arrangement in which individuals consider the amount, level, value, worth, quality, or success of the products or outcomes of learning of peers of similar status' (Topping, 1998, pp. 250). Flower,

Hayes, Carey, Schriver and Stratman (1986) explain that in the process of learning to write, feedback is especially profitable in the revision phase.

According to Flower et al. (1986) adequate feedback on writing consists of at least four types of utterances. To assess a writing product, it is necessary to analyse it, and to provide explanations. The receiver needs arguments to take the step to revision. So, there are four feedback functions to be distinguished: analysis, evaluation, explanation, and revision. This theory of revision seems useful to determine whether the feedback meets the requirements. It will be used in the present study to determine and evaluate the composition of the peer feedback. In addition, Lockhart and Ng (1995) found that students interact in four different ways when providing feedback: in authoritative, interpretative, probing or collaborative ways. Feedback provided by an authoritative reader points out mistakes or deficiencies. The interpretative reader wants to discuss ideas emerging when reading the text. The probing and collaborative readers on the other hand take into consideration the writer's intention. The probing reader wants to clarify the text, and tries to make it correspond more to what the writer wants to communicate. Together with the writer the collaborative reader is in search of means to enrich the text and deepen it. The authoritative and the interpretative position are both instances of the 'evaluative mode'; the probing and the collaborative position are instances of the 'discovery mode'. The discovery mode seems to be more fruitful for students' achievements,

because students communicate more about their products and the intentions they want to realize (Lockhart & Ng (1995). For the characterization of the interaction between students when orally explaining their written feedback, we will use Lockhart & Ng's typology.

The aim of our study is to find an effective way of organising peer assessment of writing in the context of university teaching. As Topping indicates, peer assessment can be organised in many different ways, so it is important to be explicit about the variables. For that reason, he proposes a typology of peer assessment, consisting of a survey of variables found in reported systems of peer assessment. To discover factors that are crucial to the organisation of peer assessment in a course, we have developed and implemented several designs of peer assessment. The designs are evaluated on the following aspects:

- Realization of peer assessment activities;
- Components of feedback;
- Type of interaction during oral feedback, and
- Learning outcomes and evaluation of peer assessment.

Toppings' typology, presented in Figure 1, will be used as a framework for the description and evaluation of the designs of peer assessment. To develop

the seven course designs, we have clustered Toppings' variables into four groups, and have varied one or more within each cluster.

====Insert Figure 1 about here====

Cluster 1 (variables 1-6) refers to the function of peer assessment as an *assessment instrument*. The focus can be formative or summative, resulting in a grade or a qualitative report (product/output). Peer assessment is meant to fulfil an additional role or as a substitute of the assessment by the teacher (relation to staff assessment). Cluster 2 concerns variables of *interaction* (7-9). The direction of the assessment can be one-way or two-way. Two-way assessment means that assessors and assessees in turns change roles in the small feedback group. By one-way assessment is meant that the assessor is to be assessed by students other than the one(s) he has assessed (directionality). The outcome of the assessment may be presented in plenary session, or in the feedback group (privacy). The assessment can be done partly or entirely out of class (place), and with or without face-to-face contact (contact). Cluster 3 refers to the *composition of the feedback group* (variables 10-13). The size of the peer groups can vary from two to more participants. All the students can have studied the same materials and have written about the same topic, or each can have a different topic. The feedback groups can be formed at random, or according to a plan in which

the differences between students are used (ability). The students can individually assess the products of their fellow students in the feedback group, or e.g. first have to reach consensus about their judgments before communicating it to the assessed (constellation assessors and assessed). Peer assessment can be carried out in and out of class (place). Cluster 4 regards the variables concerning *requirement and reward* (14-17). Students may or may not be free to decide if they want to use peer assessment (requirement); the teacher can decide to encourage participation by giving course credits (reward).

Designs of peer assessment

Seven peer assessment designs were developed, and implemented in a history curriculum. The curriculum is a four-year combined bachelor and master programme, which is the standard university curriculum in The Netherlands. The designs were implemented in seven courses, distributed over the entire programme, including various competency levels, and various writing products. The procedure is the same in all designs; the differences between the designs consist in the combination of ten variables selected from Toppings' typology. Seven variables are not used for practical reasons (curriculum area/subject, year, time, requirement), or educational reasons (objectives, focus), or because the teachers did not match any variation of this element (official weighting).

Peer assessment procedure

At the first meeting of the course, the students were verbally informed about the peer assessment objectives, and the procedure to be followed. They had the opportunity to ask questions. The teacher explained the assessment criteria. This information was presented in a written form, at a meeting held before the peer-assessment meetings. As a basic method of peer assessment, we adopted Bean's elaboration of the concept 'advice-centred feedback' (Bean, 1996) in all seven designs. Students were asked to exchange their draft versions, and assess these using the same criteria the teacher would use for the final versions. They were asked not to mark, but merely to record their findings in the form of written comments for the assessed student. At the end of their report, they should reflect on their judgements and make a selection by formulating at least three recommendations to the writer on how to improve his writing. After this peer assessment, students were given the opportunity to rewrite their draft. The teacher monitored the proper execution of these steps. To facilitate the teachers' monitoring of this process, copies of the drafts and the written feedback reports were also handed in. The teacher's strategy was to give his comment only after peer feedback had been given, in the form of a supplement.

Seven designs of peer assessment

We have developed designs for all levels of the university programme:

- A first-year course in which students take their first steps in learning how to report on historical research (case 1);
- Two second-year courses. In one of them (case 2), students plan, carry out and report on a limited historical research. In the other course (case 3), which is scheduled after case 2, students have to do a more extensive historical study.
- A third-year course, in which students have to write a biography of an important historian (case 4);
- A third/fourth-year specialisation course, in which students learn to write to a strict deadline (case 5), and
- Two third-/fourth-year specialization courses. One of them is an introduction to cultural education, for which students have to write an exhibition review (case 6); in the other course students have to process literature into an article (case 7).

In the latter two courses, parallel groups were set up following the regular educational programme without peer assessment. The differences between the designs refer to the operationalization and combination of ten variables taken from Toppings' typology. The seven designs and the operationalization of Topping's variables are summarised in Figure 2.

With respect to the variables of cluster 1, the products of our design varied in length and completeness. They also varied in the relationship with teachers'

assessment. In four designs (cases 1, 2, 3 and 5), peer assessment is supplementary. In case 3 and 5, the teacher also grades the version that is to be assessed by peers. In the other ones (cases 4, 6 and 7), this serves as a second assessment, coming before the teacher's end of course assessment. Concerning the variables of cluster 2, the course designs differ in their directionality of the assessment: one-way (case 7) or two-way. In case 1, the assessment was public, as students presented their oral feedback in a plenary session. In all courses students performed the written part of the peer assessment out of class, including reading the draft, making notes and completing the assessment form. The oral feedback was provided face-to-face. With respect to the variables of cluster 3, the size of the feedback groups varied from two (case 3) to three (case 5) or four (case 2, 4 and 6). In the cases 5 and 6, students studied the same materials and were given the same assignment. In case 2, the teacher formed feedback groups by putting together students with a related subject. In the other courses, students were grouped at random. In case 4, the assessors first had to reach consensus on their judgement, before communicating it to the assessed student. The oral part of peer assessment was carried out in class. In case 7 there was no oral feedback scheduled. As to the variables of cluster 4, participation in peer assessment was only rewarded in case 5. In this course, the teacher assessed the written feedback. When of sufficient quality, the students were graded a quarter of a point extra.

===Insert Figure 2 about here

Method

This multiple case study (see for more information on case studies, Creswell, 1998) involved 168 students following the History programme of the Faculty of Arts, and nine teachers of the History Department of Utrecht University. From this student group, 131 participated in peer assessment groups and 37 in a parallel group (without peer assessment). The parallel groups had the same teacher as the peer assessment groups. Both teachers and students had no previous experience with peer assessment.

Data on peer assessment activities and learning outcomes were gathered from two student questionnaires, a semi-structured interview with each teacher at the end of the course, and class observations. The class observations concentrated on the introduction of peer assessment by the teacher, students' responses to this, how the feedback groups were formed, the exchange of writing products and written feedback, the participation in oral feedback, the plenary discussion after oral feedback, and the interaction during classes between students, and between students and teacher. In addition, writing products with and without revisions were collected, together with all the accompanying feedback. Moreover, additional information on peer feedback was gathered using the completed feedback

forms. Learning outcomes are defined in terms of the revisions students made, the grades given to their writing products, and students' evaluations of their own progress. To determine the type of revision, and the elaboration of the feedback, we studied the writing products before and after revision. The differences between both versions were compared to the received feedback.

Instruments

Student evaluation

Student evaluations of peer assessment were measured by means of a questionnaire, administered directly after the oral feedback and before students were credited. This 'questionnaire about peer assessment' consisted of 16 open questions that were identical for all courses, and one or two questions on a specific element of the course design. The items of this questionnaire referred to the amount of time students had spent on reading and assessing their fellow students' work, their approximation and appreciation of time invested, the usefulness of the assessment form, their ability to assess the work of other students, their evaluation of their fellow students' feedback, and the progress in their work after revision. In addition, students were asked about their thoughts on working with peer assessment in the future, and their suggestions for improving the method. The answers to the open questions were scaled on a three-point scale (with 1= mainly

negative and 3= mainly positive). The average inter rater-reliability, in terms of Cohen's kappa was $\kappa \geq .70$, varying from .64 to .90.

At the end of the course, students in peer assessment groups, as well as those in control groups, completed a second questionnaire. In this questionnaire, they were asked about the total amount of time they had spent on writing and revising, the measure of satisfaction with the interaction in class to evaluate their own writing competence, especially in terms of having a better understanding of the criteria, and the progress they had made in writing. The answers were given in the form of scores on a five-point scale (with 1=mainly negative and 5= mainly positive). This was done directly after the exam, before students knew the grade of their final version. The questions of both questionnaires were checked on relevancy and formulation by the research team.

Teacher evaluation

At the end of the course, immediately after they had graded the final versions of the students' writings, a semi-structured interview was conducted with each teacher. The interviewer was not involved in the research. The interview included topics similar to those in the 'questionnaire about peer assessment', including teachers' time spent on assessment, the usefulness of the assessment form, perceived effects of peer assessment, and the idea of making peer assessment a structural part of the course. The research team

checked the questions on relevancy, and the answers were checked by means of ‘member check’. At times this resulted in corrections and additions.

Peer feedback

The peer feedback has been coded from two major perspectives: function and aspect. According to Flower et al. (1986), adequate feedback on writing fulfils at least four functions: analysis, evaluation, explanation, and revision. Other utterances are not product-oriented, for example utterances aimed at discussing the writing process (coded as ‘method’), and those meant to structure the discussion (coded as ‘orientation’).

The term ‘feedback aspect’ refers to the subject of feedback: content, structure, or style (Steehouder, Jansen, Maat, Staak, van der & Woudstra, 1992). ‘Content’ refers to the relevancy of the information, the clarity of the problem, argumentation, and explanation of concepts. ‘Structure’ refers to the inner consistency, for example the relation of the main problem to the subdivided research questions, the explanation, and the conclusion. ‘Style’ refers to the outer form of the text, for example the language, grammar, spelling and layout. These categories refer to the product-oriented feedback. Feedback which is not product-oriented includes remarks structuring the discourse, or remarks on the subject of the writing products. The inter-rater reliability, in terms of Cohen’s kappa, was $\kappa = .85$ for feedback function and $\kappa = .93$ for feedback aspect.

Type of interaction

In order to determine the interaction mode of students in our study we adapted Lockhart and Ng's (1995) typology. We coded the interaction data into an evaluative mode, embracing Lockhart's and Ng's authoritative, respectively interpretative reader stance, and a discovery mode, which includes the probing and collaborative stance. The following criteria were used: if evaluation was the main feedback function carried out in 50% or more of the discussions per product and the functions analysis, explanation and revision in 15% or less, the interaction was characterized as evaluative; if evaluation was the main feedback function in 10% or less of the discussions and the functions analysis, explanation and revision each in 30% or more, the interaction was characterized as discovery.

Quality of revisions

We collected two versions of each writing product: the one assessed by peers, and the revised one handed in to the teacher. Both were compared, and the differences were categorized in terms of content, structure, and style (see above). The inter-rater reliability, in terms of Cohen's kappa, was $\kappa = .89$.

Analyses

The qualitative data were analysed by means of content analysis. The quantitative data were analysed with the help of descriptive statistics. To

determine the significance of differences in frequencies of feedback function and feedback aspect (within each design), we used the χ^2 test. Cross case analyses (between the seven designs) included t-tests and non-parametric analyses of variance. The significance of the relationship between the manner of providing feedback and the components of the feedback is expressed in terms of Cramér's C. In order to relate the results to design features, we used the matrix-method of Miles and Huberman (1994), with the columns containing the seven designs, and the rows containing the topics and issues that were studied.

Results

Assessment activities

The realization of the peer assessment procedures is summarized in Table 1.

=== Insert Table 1 about here ===

In general, the procedures were carried out as scheduled. About 80% of the students handed in a draft version on time, received feedback from at least one other student and assessed the work of at least one fellow student. About two thirds received the number of comments prescribed in the peer assessment design. Only a few did not receive any feedback from their peers;

all of these were participants in case 2. Not all students completed the planned number of assessments. In the cases 2 and 4 student' discipline did not match our requirements.

Regarding the time students had to invest minimally in reading and assessing (see Van den Berg, 2003, p. 217 for an explanation of our criteria), we found that in two designs (case 3 and 7) students spent too little time on reading and assessing.

Written feedback

The frequencies of function of students' written feedback are presented in Table 2.

=== Insert Table 2 about here ===

In general, the written feedback is product oriented and students concentrated on evaluating the product. They were not much engaged in asking questions about the text, explaining their judgments, and proposing revisions.

The designs differ significantly in the frequencies of the different feedback functions ($\chi^2= 143$; $df =48$; $p \leq .001$ for all feedback functions, $\chi^2= 77$; $df =24$; $p \leq .001$ for product oriented feedback).

The results on aspect of students' written feedback are presented in Table 3.

==== Insert Table 3 ====

Again, the written feedback is mostly product oriented. Most feedback is concentrated on the aspects 'content ' and 'style', and hardly any on the 'structure' of the text.

In respect to the feedback aspect there are also significant differences between the designs ($\chi^2= 116$; $df =40$; $p \leq .001$ for all feedback; $\chi^2= 86$; $df =24$; $p \leq .001$ for product oriented feedback).

Oral feedback

The frequencies of function of students' oral feedback are presented in Table 4.

==== Insert Table 4 about here====

In contrast to the written feedback, the oral feedback contains more non-product oriented utterances, in which students, for example, structure the discourse or discuss subjects that are not related to the text. In their product-oriented oral feedback, students were more engaged in offering arguments

for the evaluation than in their written feedback, asking questions about the text and proposing revisions.

Also in the oral feedback the designs differ significantly in frequencies of different feedback functions ($\chi^2= 351$; $df =42$; $p \leq .001$ for all feedback functions, $\chi^2= 93$; $df =21$; $p \leq .001$ for product oriented feedback).

The results on the aspect of students' oral peer feedback are presented in Table 5.

==== Insert Table 5 ====

As in the written feedback, the product-oriented feedback is mainly concentrated on the content and the style of the writing. Feedback on the structure is hardly ever provided, even less than in the written feedback.

Regarding the feedback aspect there are also differences between the designs in the oral feedback ($\chi^2= 330$; $df =35$; $p \leq .01$ for all feedback; $\chi^2= 251$; $df =21$; $p \leq .001$ for product oriented feedback).

As these outcomes suggest that written and oral feedback differ in focus, we analysed the relationship between the type of feedback, written or oral, and the kind of feedback functions and feedback aspects students concentrated on. The analysis is done for the product and non-product oriented feedback,

and only repeated for the product-oriented feedback. There is a significant relationship, expressed in Cramér's C , ($C = .30$, $p \leq .001$), between how feedback is given (written or orally), and the kind of feedback functions that are being fulfilled. This applies to the feedback aspects as well. This leads to the conclusion that it seems to be important to have a combination of both written and oral feedback in order to achieve a spread of all product oriented feedback functions, along with all product oriented feedback aspects.

Interaction

To determine the type of interaction between students during oral feedback, this was analyzed for each writing product. There were two, three, or four writings to be discussed in each feedback group. Per discussion, utterances were counted of the different types of product-oriented feedback. Afterwards, the proportions of frequencies were related to the previously mentioned typology of Lockhart & Ng (1995). The results are summarized in Table 6.

=== Insert Table 6 about here===

In general, students' interaction mode was more evaluative than discovering. This seems reasonable, as students' task was to evaluate their fellow students' work. In addition, most of the writing products to be assessed were

almost finished, which means that students might not be very willing to revise their text very thoroughly. This attitude might cause the evaluative interaction of the assessors and assessees. However, it is surprising that in the second course design, there is more interaction of the discovery type. In this design, the writing product was assessed in two phases: one during the planning of the writing of a key chapter, and one during the writing process itself. The former means that students provided feedback in an earlier phase of the writing process than in the other designs.

Revision

The results concerning the use of peer feedback in students' revision of the writing products are presented in Table 7.

==== Insert Table 7 about here====

On average, students used about one third of all suggestions for revision from their peers; the other ones were neglected. Their reasons for accepting or neglecting suggestions have not been investigated. For the revised parts of their writing products, students only revised those parts their fellow students had commented on. This means that most of the time, students did not revise text that was not commented on. When students revised the parts that were commented on, their revisions were related to the peer feedback. Almost all students used at least some of the feedback for revision. In general, 44% of

the revisions were related to the content, and 51% to the style of writing. Revisions on the structure of the writing products were hardly carried out. There are differences between the designs in the amount of feedback being processed. The stage of the draft version, and the openness of the writing task might explain these differences. In case 4, there were more unfinished writings than in the other designs, so students were more inspired to complete their work based on their fellow students' remarks. Another cause may have been the openness of the writing task. In case 6, the writing format was given in much more detail than in the other designs. This may have caused the absence of peer comments on structure, and, at any rate, of students' lack of interest to revise much.

In three designs (cases 4, 6 and 7), the grades for the final version of the writing product in the peer assessment groups could be compared to those in the control groups. We did not find any differences.

Student- and teacher evaluation

Based on the analysis of the questionnaire data, we conclude that 80% of all students considered that the progress in their writing product was the result of their processing of peer feedback. A Kruskal-Wallis analysis of variance shows that students in different designs differ in their opinions (KW=15,0; df=7; p=.04). In case 1 and in one of the two peer assessment groups of case 4 students estimated much progress in their writing, in case 7 students estimated little progress. In case 1 the students of peer assessment groups

experienced more progress in their writing skills than in the non-experimental groups.

The answers to the questionnaire about peer assessment also reveal that on average 75% of the students considered the activity of reading and assessing fellow students' work useful and not consuming too much time. In general, 83% and 78% of the students positively valued the received written feedback and oral feedback, respectively. Many students added extra remarks to emphasize that they see getting new ideas about writing as an added value of peer feedback. Some students suggested that the time necessary for assessment should be compensated in the study load. Only few students did not like to be dependent on the activities of fellow students. Only in the courses 2 and 4, students sometimes had some negative remarks on the dependability of other students' work.

On average, 61% of all students would like to work more often with peer assessment, and 40% of these students wanted to adjust the system on one or more aspects. As students' suggestions for adjustment are very different, it is hard to use them in a re-design.

The teachers also see some pros and cons in peer assessment. A first advantage mentioned is that students learn to work with colleagues in a way that they will do during their professional career. A second outcome mentioned by all teachers is that peer assessment offers a structure for students to read, think and discuss the work of their fellow students, which

makes it possible to discuss it together. One of the problems mentioned by teachers is that the role of the teacher was not clear enough. They wanted to provide more guidance and support, but felt that there was no time for it, as students needed the time to provide and discuss their feedback. Despite this problem, all teachers were positive about the idea of implementing peer assessment in their courses.

Conclusion and discussion

In this study, seven designs are evaluated in order to determine the important factors in the effective organization of peer assessment within coursework. We have investigated the peer-assessment procedures, the functions and aspects of feedback, the interaction between students during oral feedback, students' revisions of their writing products and students' and teachers' evaluations of peer assessment.

Overall, the students as well as their teachers carried out the scheduled assessment procedures. In their written feedback, the students focused strongly on the assessment of the text. In the oral feedback, students not only assessed, but also explained their judgments and proposed suggestions for revision. It appeared that chances of fulfilling all the feedback functions, and discussing all the feedback aspects, increase when both written and oral feedback are being provided. Students' interaction more often had the character of the evaluative mode than of the discovery mode.

Almost all the students used some peer feedback when revising, and almost all revisions were related to comments peers had made. On average, one third of all the suggestions were processed. About half of the revisions referred to content, and the other half to style. Of all the students, 80% were of the opinion that their revised version was better, because they had used peer feedback, but there were differences between the designs.

Students positively appreciated the method of peer assessment. In their opinion, reading and assessing fellow students' work was a useful activity, and they valued peer feedback as being helpful for revising. About two thirds of the students preferred to work more often with peer assessment.

Most teachers thought peer assessment was an effective means of stimulating subject matter discussions between students, instead of having only student-teacher interaction. Some teachers said they had some problems with their role: they wanted to provide more support and guidance, but there was no time for it, as the students had to have the opportunity to give feedback.

Optimal model

Based on a cross-case analysis of the seven designs of peer assessment and their results, we draw conclusions about the most important design features supporting effective peer assessment in university teaching. These are summarized in Figure 3.

==== Insert Figure 3 about here=====

Important design features are:

1. Product: The size of the writing is five to eight pages. The reason is that students will not be willing to invest enough time in assessing larger products.
2. Relation to staff assessment: There must be sufficient time between the peer assessment and teacher assessment, so that students can first revise their paper on the basis of peer feedback, and then hand it in to the teacher.
3. Directionality: Two-way feedback is easier to organise for teacher and students, as it is clear that the assessor will in turn be the assessee, which makes it easier to exchange products. Oral feedback during class will not take much time, because the feedback groups can discuss simultaneously.
4. Contact: Verbal explanation, analysis and suggestions for revision are necessary elements of the feedback process; these require face-to-face contact.
5. Constellation assessors/assesseees: The size of feedback groups has to be three or four. In that situation, students have an opportunity to compare their fellow students' remarks, and to determine their relevance. A group of two students is too small, because of the risk that the partner might not perform properly.

6. Place: Oral feedback must be organised during contact hours, because it is difficult to ascertain if students will organise this themselves when out of class. ICT-tools can be used to enable students to read the peer feedback before discussing it.

On our study, some critical remarks have to be made. Firstly, we could not fully control each design variable from the peer assessment table of Topping (1998), as the educational practice required some necessary conditions. Secondly, the study could have been more complete, gathering more qualitative data on, for example, the dynamics of the discussion, or performing more qualitative analyses of the available data. For example, in the longer conversations, the interaction position of students varies. We would have obtained a better insight into the process of peer feedback, and the quality of the interaction, using qualitative (sequential) analyses.

An interesting issue in further study would be to develop forms of peer assessment that evoke patterns of interaction according to the discovery mode. Referring to Saunders (1989), our conclusion that students interact in an evaluative way is not surprising. The assessor was given the task to assess someone else's writing product, a product that was almost completed. The question is how a less evaluative and more discovery-directed task could be formulated. Will such a task result in students interacting more often in a discovery mode? Will this type of interaction result in better writings? Peer

assessment is increasingly used in university teaching, in many variations, based on substantial arguments or research, or not. We think it is time to start more extensive investigations in order to arrive at methods of peer assessment evoking discovery mode patterns of interaction.

References

Bean, J. C. (1996). *Engaging ideas: the professors' guide to integrating writing, critical thinking and active learning in the classroom*. San Francisco: Jossey-Bass Publishers.

Berg, B.A.M. van den (2003). Peer assessment in universitair onderwijs. Een onderzoek naar bruikbare ontwerpen. [*Peer assessment in university teaching. An exploration of useful designs*]. Dissertation, Utrecht University, Utrecht.

Creswell, J. W. (1998). *Qualitative Inquiry and Research Design: choosing among five traditions*. Thousand Oaks: SAGE Publications.

Flower, L., Hayes, J. R., Carey, L., Schriver, K., & Stratman, J. (1986). Detection, Diagnosis, and the Strategies of Revision. *College Composition and Communication*, 37, 16-55.

Lockhart, C., & Ng, P. (1995). Analyzing Talk in ESL Peer Response Groups: Stances, Functions and Content. *Language Learning*, 45, 605-655.

- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: an expanded sourcebook* (2nd ed.). Thousand Oaks, California: SAGE.
- Saunders, W. M. (1989). Collaborative writing tasks and peer interaction. *International Journal of Educational Research*, 13, 101-112.
- Steehouder, M., Jansen, C., Maat, K., Staak, J. v.d., & Woudstra, E. (1992). *Leren communiceren* (3e herziene druk ed.). Groningen: Wolters-Noordhoff.
- Topping, K. (1998). Peer Assessment Between Students in Colleges and Universities. *Review of Educational Research*, 68, 249-276.