



Universiteit  
Leiden  
The Netherlands

## **Assessing together : Peer assessment from an interpersonal perspective**

Gennip, A.E. van

### **Citation**

Gennip, A. E. van. (2012, October 23). *Assessing together : Peer assessment from an interpersonal perspective*. Retrieved from <https://hdl.handle.net/1887/20012>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/20012>

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/20012> holds various files of this Leiden University dissertation.

**Author:** Gennip, Anna Eva (Nanine) van

**Title:** Assessing together. Peer assessment from an interpersonal perspective

**Issue Date:** 2012-10-23

## CHAPTER 3

# Arranging Peer Assessment: The Role of Interpersonal Beliefs<sup>6</sup>

The goal of the present study was to focus on the role of psychological safety and value congruency among peers and the ways in which it is affected by differentially arranged peer assessment conditions. We compare three conditions: (1) a teacher-based assessment condition; (2) a peer assessment condition, and (3) a peer assessment<sup>+</sup> condition, where the peer assessment and peer assessment<sup>+</sup> condition differ in the amount of involvement in peer assessment. Results indicate that teacher based condition differed significantly from both peer assessment conditions on psychological safety as well as value congruency.

### 1 Introduction

Peer assessment has become popular in school settings at different levels of education. It is described 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, 250). This mode of assessment is being advocated as a strong tool to enhance learning (Dochy, Segers, & Sluijsmans, 1999; Falchikov, 1995; Sluijsmans, Brand-Gruwel, & Van Merriënboer, 2002; Van Gennip, Segers, & Tillema, 2009). As Tillema (2009) argues, peer assessment arranged as an exchange of appraisal information serves to scaffold each individual learner to (1) accept provided feedback, and (2) follow recommendations (Tillema, 2009). It is this informative scaffolding of further learning especially that makes peer assessment a powerful tool for the promotion of learning (James et al., 2006).

In addition to learning gains, multiple other benefits of peer assessment are claimed (Brown & Glasner, 1999; Lui & Carless, 2006; Falchikov, 1995; Pond, Ul-Haq, & Wade, 1995). It is argued that peer assessment increases student engage-

---

<sup>6</sup> Based on: Van Gennip, N.A.E., Segers, M. S. R., & Tillema (2011, submitted). Arranging peer assessment: the role of interpersonal variables.

ment in instruction and therefore encourages student autonomy (Sluijsmans et al., 2002). Moreover, it supports students in the development of critical judgment skills (Falchikov, 1995).

The definition of peer assessment as stated above, indicates that peer assessment is a collaborative activity between peers. Peer assessment is an activity taking place within the interpersonal context formed by the peer team. In their review study on team learning Akkerman et al. (2007) indicated that interpersonal beliefs or shared beliefs of the team characteristics emerge in groups from the interaction among the team members. Subsequently, it is shown that they form a context that stimulates or inhibits learning behavior. Given the interaction among peers differs according to the peer assessment arrangement they participate in, the main question to be dealt with is: How do students participating in peer assessment perceive the interpersonal context formed by their team and to what extent differ these interpersonal beliefs in different peer assessment arrangements?

Former research on peer assessment has shown some evidence that students' interpersonal beliefs in peer assessment settings differ to the extent to which students are involved in the different stages of the peer assessment process (Tillema et al., 2010). More concretely, within a peer assessment arrangement that has a low level of student involvement (e.g., peer marking in which students only give marks), students find it difficult to evaluate their friends and therefore show antagonism towards peer assessment (e.g., Hanrahan & Isaacs, 2001); they also perceive the grading of peers as risky and unfair (Kwan & Leung, 1996). Other formats of peer assessment arrangements (i.e., peer feedback or peer evaluation), however, deal more effectively with interpersonal relationships. More precisely, in the formats presented in these studies (Strijbos, Narciss, & Dünnebier, 2010) students are actively involved in different steps of the peer assessment process and in some cases receive training (Sluijsmans, Prins, & Martens, 2006) helping them understand the goals and criteria as well as practice the skills needed to assess peers. The results of these studies indicate that such peer assessment arrangements show an increased feeling of trust among the peers (Butler & Hodge, 2001; Keaten & Richardson, 1992; Pond & Ul-Haq, 1997). These findings indicate that the involvement of students in the different stages in a peer assessment process influences the degree to which students perceive the assessment setting as safe and free from interpersonal risk-taking. This finding is referred to as 'psychological safety' (Edmondson, 2002).

In addition, studies also indicate (e.g., Sluismans et al., 2002) that peer assessment arrangements in which students at the start of the peer assessment practice collaboratively define learning objectives result in an enhanced shared understanding of these learning objectives. According to Jehn, Northcraft and Neale (1999), this could be referred to as the degree to which students share values or 'value congruency'.

In sum, although evidence is still scarce, the aforementioned studies indicate that peer assessment arrangements which involve students in the early stages of

the peer assessment process positively affect interpersonal beliefs, which is an important condition for peers to learn from each other in a peer assessment setting. Based on these findings, in the present study, we focus on the role of psychological safety and value congruency among peers and the ways in which it is affected by differentially arranged peer assessments.

### 1.1 The interpersonal nature of peer assessment

Several authors (Falchikov, 1995; Marshall & Drummond, 2006) warned against the possible dangers or problems in ignoring the role of interpersonal relations in peer assessment when students are inexperienced with or involved only in parts of the peer assessment process. Topping (2003, p. 67) for example claims: “Peer assessments might be partly determined by: friendship bonds, enmity or other power processes, group popularity levels of individuals, perception of criticism as socially uncomfortable or even socially rejecting and inviting reciprocation, or collusion leading to lack of differentiation.” Studies of Dochy et al. (1999), Falchikov (1995), and Sluijsmans et al. (2002) refer to problems that arise in the arrangement of peer assessments. They mention students’ hostility towards peer assessment when they first experience it, caused by a lack of trust in the self and the other as assessors, and point to friendship marking where peers mark their friends higher regardless of their performance (Dochy et al., 1999). These studies refer to the problem of psychological safety, a sense of confidence that the team will not embarrass, reject, or punish someone for speaking up (Edmondson, 1999, p. 354).

Moreover, several studies indicate (Sluijsmans, Dochy & Moerkerke, 1999; Rust Price & O’Donovan, 2003) that setting purpose and goals of assessments needs to involve students, since without explaining the rationale of an assessment it is hard to encompass its appraisal (Bloxham & West, 2007). These studies refer to the importance of value congruency for increasing learning benefits of participating in peer assessment.

#### 1.1.1 *Psychological safety*

Although psychological safety has not been explicitly referred to in peer assessment studies, the aforementioned studies (e.g., Sluijsmans et al., 2002; Dochy et al., 1999; Falchikov & Goldfinch, 2000) stress the importance of variables such as friendship and hostility and underpin the relevance of psychological safety. Peer assessment, as Edmondson states (2002, 3) “carries a risk for the individual of being seen as ignorant, incompetent, or perhaps just disruptive.” There is a personal need to manage this risk to minimize harm to one’s self-esteem especially in case one is being evaluated by another (Edmondson, 2002). Peer assessment arrangements, especially peer feedback and peer evaluation, can provide the conditions needed for low interpersonal risk-taking and therefore high psychological safety.

This is achieved, mainly, by involving students in the different steps of the peer assessment process. In this respect, Edmondson (2002) argues that seeking others' input and invite feedback and ideas from peers, which is de facto the core of a peer assessment practice, peers have a powerful positive effect on psychological safety. Moreover, she argues that organising a reflection meeting empowers those in lower-status positions to speak up and to minimize the domineering tendencies of high-power individuals. We concur that when students are involved in the peer assessment arrangement - giving input, sharing ideas and giving feedback - this can contribute to students' safety in assessing peers and in turn support their learning. In addition, according to Edmondson (2002), the introduction of reflective sessions might affect the perception of psychological safety in a positive way.

### *1.1.2 Value congruency*

Value congruency refers to the degree to which peers agree on values about group processes and group tasks. The importance of a shared understanding between peers about the nature of assessment and the criteria on what and how to assess has been highlighted in many reviews on peer assessment (Dochy et al., 1999; Falchikov & Goldfinch, 2000; Topping, 1998; Topping, 2003). The study of Sluijsmans et al. (2002) confirms that practicing peer assessment enhances a shared understanding of the task. Reviewing, clarifying, and evaluating other persons' work is a cognitively demanding task for students which requires not only a common framework on objectives and standards in the assessment task (Boud, 1995; Topping, Smith, Swanson, & Elliot, 2000), but also the consideration (and acceptance) of multiple perspectives on assessing each others' work (Searby & Ewers, 1997). We assume that peer assessment arrangements which actively involve students foster greater value congruency. Moreover, a reflective session will support the feeling of value congruency as well.

## **1.2 Peer assessment arrangements**

The many studies on peer assessment (Van Gennip et al., 2009) show high variety in the arrangement of peer assessment intervention and the degree of engagement or participation of students in the different parts of the assessment process (Birenbaum, 2007; Tillema et al., 2010). While in some cases students are only involved in the scoring or marking (Sivan, 2000; Topping et al., 2000) other arrangements include students as well in the criteria setting and feedback giving (Orsmond & Merry, 1996). Student involvement in and along the different stages of peer assessment adds to how they perceive the interpersonal context of peer assessment.

Based on an earlier review of studies on peer assessment (e.g., Van Gennip et al., 2009) at least five different arrangements of peer assessment can be identified in relevant literature. They vary in the extent to which students are involved in the

different stages of the process from goal formulation to decision making, including their role as assessor, from grader to feedback giver. Firstly, in *peer marking* students score each others' work against a set of criteria, without further commenting on the work. Students are only involved in the scoring. *Peer feedback* takes this a step further, and allows students to comment on each others' work as well, often supplementing the scoring itself. *Peer grading*, thirdly, grants students the responsibility to make decisions about the outcomes of the assessment. However, feedback is not included in peer grading, while *peer appraisal* does include feedback. Finally, in *peer evaluation* students are not only involved in formulation of peer assessment criteria, scoring, giving feedback and the decision-making, but usually get to give input for the task formulation as well, at the start of the peer assessment process. Figure 1 gives an overview of the different peer assessment arrangements.

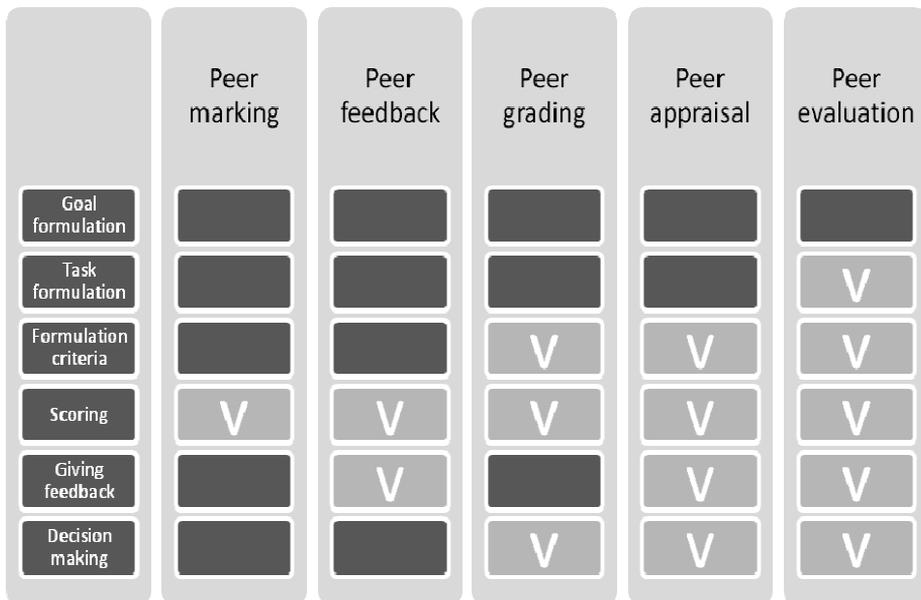


Figure 1  
Overview of peer assessment arrangements

It was this variety in peer assessment conditions which challenged us to compare interpersonal beliefs not only between teacher assessment and peer assessment, but between different peer assessment arrangements as well. We therefore implemented a peer assessment arrangement with a reflective session, trying to create a more shared mindset between students using peer assessment, examining whether this leads to differences in perceptions of interpersonal beliefs (peer as-

assessment<sup>+</sup> condition). Consequently, in this study, we focus on the relation between assessment arrangement and how students experience psychological safety and value congruency. In order to answer this question, we formulated the following research questions

- 1a Do students in a peer assessment condition hold more positive beliefs of psychological safety and value congruency than students in a teacher-based assessment condition?
- 1b Do students who have been participating in a reflective session (peer assessment<sup>+</sup> condition) hold more positive interpersonal beliefs than students who did not participate (peer assessment condition)?
- 2 In which stage of the peer assessment process do students in the peer assessment and peer assessment<sup>+</sup> conditions differ in interpersonal beliefs?

## 2 Method

### 2.1 Setting

Our peer assessment study took place in a large institute of vocational education in the Netherlands. 106 second-year male students participated in the study. Their ages ranged from 15 to 18. Students took a project-based course, teaching detailed technical skills in metal work and electronics. The assessment task was to design and construct a mechanical robot artefact: a moving device containing pneumatic and hydraulic elements.

We compare three conditions: (1) a teacher-based assessment condition; (2) a peer assessment condition, and (3) a peer assessment<sup>+</sup> condition, where the peer assessment and peer assessment<sup>+</sup> condition differ in the support students receive to develop beliefs of psychological safety and value congruency.

#### 2.1.1 *Peer assessment condition*

This group consisted of 25 second-year students. During the project, groups received instruction in plenary sessions and worked on their artefacts as a group. At the start of the project students received a two-hour in-class instruction on the nature of peer assessment. The instruction included an overall explanation of the concept of peer assessment by the researcher. Additionally, fourteen appraisal criteria were formulated and students were instructed to use these in the appraisal of each other's work. At the end of the six-week project all groups presented their (robot) artefacts. After this presentation each group was assessed by their peers (i.e., not belonging to that group). Assessment was done on a special form that listed all eleven criteria, and students could rate their peers on a scale of 0 (poor) to 1 (good) for each of the criteria. The completed forms were collected after-

wards. The researcher calculated the average ratings and returned these to the project groups one week later.

### 2.1.2 *Peer assessment<sup>+</sup> condition*

Forty-five students participated in this condition. Peer assessment procedures in this condition were identical to those in the peer assessment condition. However, to enhance psychological safety and value congruency, a reflection meeting was organized where students reflected on and discussed among themselves how they approached the role of being an assessor and how they dealt with grading each other's work and hence, each other. This reflective session was implemented half-way down the course, in week 3 (out of a total of six weeks) and before the actual appraisal at the end of the course. At the start of the reflection meeting, the students completed a questionnaire with statements belonging to the scales 'psychological safety' and 'value congruency'. They received three cards with different colours: green, yellow and red. After the students had completed the questionnaires, the researcher read aloud the questionnaire statements one by one, after which every student showed one of the cards: red when the student disagreed, green when he or she agreed, and yellow in case of doubt. After this exercise, the researcher provoked a discussion among the students which he then moderated by posing questions like: 'Why do you agree/disagree?', 'Why does someone else think differently?', and 'How can you reach consensus in the group about this statement?' Meetings lasted approximately 1.5 hours, and all statements were discussed.

### 2.1.3 *Teacher-based (or baseline) condition*

This is the regular appraisal condition for the course, lacking any kind of peer assessment. This group consisted of 36 second-year students. As a control group, these students received no training in peer assessment, but were assessed by the teacher only, on the same criteria the students in both peer assessment groups were instructed to use.

## 2.2 **Research design**

An experimental comparison of conditions was adopted for the peer assessment condition, the peer assessment<sup>+</sup> condition and the teacher-based condition. For the teacher-based condition, end-of-course measurements are available. In both peer assessment conditions, prior measures (T0) as well as end-of-course measures (T2) of both interpersonal beliefs (psychological safety and value congruency) were included. Additionally, for both peer assessment conditions, psychological safety and value congruency were measured before the intervention (T1). Table 1 shows an overview of the research design of this study.

## Chapter 3

Table 1  
*Overview of research design*

	T0	T1	T2
Teacher assessment baseline (n = 36)			Psychological safety Value congruency
Peer assessment condition (n = 25)	Psychological safety Value congruency	Psychological safety Value congruency	Psychological safety Value congruency
Peer assessment+ condition (n = 45)	Psychological safety Value congruency	Psychological safety Value congruency	Psychological safety Value congruency

### 2.3 Measurement instruments

All items were measured using 5-point Likert scales, and anchored by 1 (totally true) and 5 (totally untrue).

#### 2.3.1 *Psychological safety*

This scale measures the degree to which students perceive their group as safe for interpersonal risk-taking and was derived from Edmondson (1999). It consists of seven items. All items were measured using 5-point Likert scales, and anchored by 1 (totally true) and 5 (totally untrue). A sample item is: 'It is easy to ask my peers for help' (Cronbach's  $\alpha = .86$ ).

#### 2.3.2 *Value congruency*

The scale, adopted from a study by Jehn et al. (1999), measures value congruency as the differences that different group members perceive on group task and goal or mission. It consists of six items which were measured using 5-point Likert scales, and anchored by 1 (totally true) and 5 (totally untrue). Sample items here are: 'The group as a whole has one single goal' and 'Group members agree on what is important for the group' (Cronbach's  $\alpha = .83$ ).

### 2.4 Analysis

In order to answer research questions 1a and 1b, the three conditions (teacher-based assessment, peer assessment, peer assessment<sup>+</sup>) were compared on the interpersonal beliefs (psychological safety and value congruency by means of ANOVA'S. For research question 2, in order to test at which stage in the peer assessment differences in interpersonal beliefs occur, multivariate analyses of variances were conducted (MANOVA).

### 3 Results

Means and standard deviations of variables measured in all groups are presented in Table 2.

Table 2  
*Means and standard deviations*

	Means (SD)		
	Teacher assessment (n = 36)	Peer assessment (n = 25)	Peer assessment+ (n = 45)
Psychological safety	2.40 (.68)	2.79 (.42)	2.95 (.54)
Value congruency	2.51 (.78)	2.88 (.61)	3.14 (.62)

Table 2 shows low mean scores (i.e., below 3 on the 5-point Likert scale) for both beliefs safety and congruency in all conditions. The variance in the teacher-based condition is found to be higher than in the peer assessment conditions.

#### 3.1 Differences in interpersonal beliefs between conditions

To answer research questions 1a and 1b, we firstly investigated the differences between teacher assessment condition and peer assessment conditions: to what extent does teacher assessment differ from peer assessment, and more specific the various arrangements of peer assessment with regard to psychological safety and value congruency. Results show that conditions differed significantly on psychological safety  $F(2, 99) = 9.11, p < .01$  as well as value congruency  $F(2, 98) = 8.08, p < .01$ . Further, results of Bonferroni corrected posthoc analyses show that, in the case of psychological safety, these differences appear between control condition and peer assessment condition ( $M = .40; p < .05$ ), as well as between control condition and peer assessment<sup>+</sup> condition ( $M = .55; p < .01$ ). In the case of value congruency, these differences only appear between control condition and peer assessment<sup>+</sup> condition ( $M = .62; p < .01$ ).

#### 3.2 In which stage of the peer assessment process do students in the peer assessment and peer assessment<sup>+</sup> conditions differ in interpersonal beliefs?

In answer to research question 2, a MANOVA of the two dependent measures was performed, which showed that there is an overall effect of time on psychological safety ( $F(2,44) = 3.52; p = .034; \text{partial } \eta^2 = .074$ ) and value congruency ( $F(2,44) = 5.43; p = .006; \text{partial } \eta^2 = .110$ ). Bonferroni corrected posthoc analyses show that there is a significant difference in value congruency between the start of the project and the intervention ( $M = .23; p < .05$ ), but there is barely a difference between the intervention and the end of the project ( $M = .03; p = \text{ns}$ ). However, these post hoc analyses show no significant changes for psychological safety between the start

of the project and the intervention ( $M = .14$ ;  $p = ns$ ), as well as between the intervention and the end of the project ( $M = .04$ ;  $p = ns$ ).

## 4 Discussion

The focus of this study is on whether interpersonal beliefs (psychological safety and value congruency) are differentially affected by different formats of peer assessment arrangement. Previous peer assessment research (Dochy et al., 1999; Tillema et al., 2010) points out the importance of a climate or assessment culture which is perceived as safe and in which participants agree on the goals and values of the assessment practice. In this study we focused on two interpersonal beliefs in particular: psychological safety and value congruency. First, we hypothesized that a difference between a peer assessment condition and a teacher-based assessment condition would occur in these two interpersonal beliefs. Second, we expected students in the peer assessment<sup>+</sup> condition would show higher levels of psychological safety and value congruency than students in the regular peer appraisal condition, exactly because they got a reflective intervention that raised awareness for the interpersonal beliefs that play a role in the assessment. Finally, we wanted to know where in the process of peer assessment differences in the interpersonal beliefs would occur.

In answer to the first research question, our results indicate that psychological safety is higher at the end of the project in the peer assessment conditions than in the teacher assessment condition. Value congruency is higher at the end of the project only in the peer assessment<sup>+</sup> condition compared to the teacher assessment condition: in the peer assessment<sup>+</sup> condition there is more unanimity of goals at the end of the project. This is in line with earlier research (Van Gennip, Segers, & Tillema, 2010) which showed that students in a peer assessment setting feel significantly safer and perceive more agreement in goals than students in a traditional teacher assessment setting.

In answer to the second research question, the time effect for psychological safety disappears after Bonferroni correction. Apparently, given the differences between the peer assessment conditions and the teacher-based condition, implementing peer assessment in itself is a powerful intervention in terms of increasing students' beliefs of psychological safety. For value congruency, differences seem to appear in the first half of the project; there is more unanimity in goals at the end of the project, which seems to take place in the first half of the project, before the intervention took place. This might be explained by the fact that discussion of goals and purposes with the students has been taken place before the reflection session. It seems that the reflection session has no value added to the stage of goal and purposes discussion in terms of increasing value congruency.

For future research, we suggest to measure the differential effects of all five peer assessment arrangements mentioned at the start of this study. They differ in terms of student involvement as well as the amount of feedback given which can be expected to influence the interpersonal context in which peer assessment takes place. Such research might advance our insights in how to organize peer assessment interventions in such a way that one can make optimal use of the advantages of students' interpersonal beliefs, and in addition to investigate the role of change in interpersonal beliefs on students' learning outcomes.

This study was conducted in a setting of secondary vocational education, where, given the educational setting and age of the students, peer pressure is a significant factor. It might be interesting to repeat our study in different educational settings to reveal the context-specificity of the role of interpersonal beliefs in peer assessment settings.

Moreover, due to the choice of studying interpersonal beliefs in real classroom peer assessment setting, the sample size in our study is rather small. Repeated studies are necessary to confirm our results.

Finally, although surveys are a valid method to detect relations between variables, quantitative research is necessary in order to have a deeper understanding of the meaning of the results. Observations of classroom behaviour of students, combined with the thinking aloud method to reveal students' motives for their behaviour, are interesting methodologies to pursue in future research.

Peer assessment has been introduced in different educational levels, although the practices differ in many aspects. Our study indicates that peer appraisal as a specific peer assessment arrangement is a powerful tool to enhance psychological safety and value congruency. Involving students in the different steps of the peer assessment process seems to be worthwhile to invest in.

### References

- Akkerman, S., Van den Bossche, P., Admiraal, W., Gijsselaers, W., Segers, M., Simons, R. J., & Kirschner, P. (2007). Reconsidering group cognition: From conceptual confusion to a boundary area between cognitive and socio-cultural perspectives? *Educational Research Review*, 2, 39-63.
- Birenbaum, M. (2007). Evaluating the assessment: Sources of evidence for quality assurance. *Studies in Educational Evaluation*, 33, 29-49.
- Bloxham, S. & West, A. (2007). Learning to write in higher education: students' perceptions of an intervention in developing understanding of assessment criteria. *Teaching in Higher Education*, 12, 77-89.
- Boud, D. (1995). *Enhancing learning through self-assessment*. London: Kogan Page.
- Brown, S., & Glasner, A. (Eds.) (1999). *Assessment matters in higher education: Choosing and using diverse approaches*. Buckingham: Society for Research into Higher Education & Open University Press.
- Butler, S. A., & Hodge, S. R. (2001). Enhancing student trust through peer assessment in physical education. *Physical Educator*, 58, 30-42.
- Dochy, F., Segers, M. & Sluijsmans, D. (1999). The use of self-, peer-, and co-assessment in higher education: A review. *Higher Education*, 24, 331-350.
- Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44, 350-383.
- Edmondson, A. (2002). Managing the risk of learning: Psychological safety in work teams. In West, M. (Ed.), *International Handbook of Organizational Teamwork*. London: Blackwell.
- Falchikov, N. (1995). Peer feedback marking: Developing peer assessment. *Innovations in Education and Training International*, 32, 175-187.
- Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research*, 70, 287-322.
- Hanrahan, S., & Isaacs, G. (2001). Assessing self- and peer-assessment: The students' views. *Higher Education Research and Development*, 20, 53-70.
- James, M., Black, P., McCormick, R., Pedder, D., & William D. (2006). Learning how to learn, in classrooms, schools and networks: Aims, design and analysis. *Research Papers in Education*, 21, 101-118.
- Jehn, K. A., Northcraft, G. B. & Neale, M. A. (1999). Why differences make a difference: A field study of diversity, conflict, and performance in work groups. *Administrative Science Quarterly*, 44, 741-763.
- Keaten, J. A., & Richardson, M. E. (1992). *A field investigation of peer assessment as part of the student group grading process*. Paper presented at the Western Speech Communication Association Convention, Albuquerque, NM.
- 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, 205-214.
- Liu, N-F., & Carless, D. (2006). Peer feedback: the learning element of peer assessment. *Teaching in Higher Education*, 11, 279-290.
- Marshall, B., & Drummond, M. J. (2006). How teachers engage with assessment for learning: Lessons from the classroom. *Research Papers in Education*, 21, 133-149.
- Orsmond, P., & Merry, S. (1996). The importance of marking criteria in the use of peer assessment. *Assessment and Evaluation in Higher Education*, 21, 239-251.
- Pond, K., Ul-Haq, R. & Wade, W. (1995). Peer review: A precursor to peer assessment. *Innovations in Education and Training International*, 32, 314-323.
- Pond, K., & Ul-Haq, R. (1997). Learning to assess students using peer review. *Studies in Educational Evaluation*, 23, 331-348.

- Rust, C., Price, M., & O'Donovan, B. (2003). Improving students' learning by developing their understanding of assessment criteria and processes. *Assessment & Evaluation in Higher Education*, 28, 147-164.
- Searby, M., & Ewers, T. (1997). An evaluation of the use of peer assessment in higher education: A case study in the school of Music, Kingston University. *Assessment and Evaluation in Higher Education*, 22, 371-383.
- Sivan, A. (2000). The implementation of peer assessment: An action research approach. *Assessment in Education: Principles, Policy & Practice*, 7, 193-213.
- Sluijsmans, D. M. A., Brand-Gruwel, S., & van Merriënboer, J. J. G. (2002). Peer assessment training in teacher education: Effects on performance and perceptions. *Assessment and Evaluation in Higher Education*, 27, 443-454.
- Sluijsmans, D., Dochy, F., & Moerkerke, G. (1999). Creating a learning environment by using self-, peer-, and co-assessment. *Learning Environments Research*, 1, 293-319.
- Sluijsmans, D. M. A., Prins, F. J., & Martens, R. L. (2006). The design of competency-based performance assessment in e-learning. *Learning Environments Research*, 9, 45-66.
- Strijbos, J. W., Narciss, S., & Dünnebier, K. (2010). Peer feedback content and sender's competence level in academic writing revision tasks: are they critical for feedback perceptions and efficiency? *Learning and Instruction*, 20, 291-303.
- Tillema, H. H. (2009). Assessment for learning to teach: appraisal of practice teaching lessons by mentors, supervisors, and student teachers. *Journal of Teacher Education*, 60, 155-167.
- Tillema, H., Leenknecht, M., & Segers, M. (2010). Assessing assessment quality: criteria for quality assurance in design of (peer) assessment for learning - a review of research studies. *Studies in Educational Evaluation*, 37, 25-34.
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68, 249-276.
- Topping, K. J. (2003). Self and peer assessment in school and university: Reliability, validity and utility. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Optimizing new modes of assessment: In search of qualities and standards* (pp. 55-87). Dordrecht, the Netherlands, Kluwer.
- Topping, K. J., Smith, E. F., Swanson, I., & Elliot, A. (2000). Formative peer assessment of academic writing between postgraduate students. *Assessment and Evaluation in Higher Education*, 25, 146-169.
- Van Gennip, N. A. E., Segers, M. S. R., & Tillema, H. H. (2009). Peer assessment for learning from a social perspective: The influence of interpersonal variables and structural features. *Educational Research Review*, 4, 41-54.
- Van Gennip, N. A. E., Segers, M. S. R., & Tillema, H. H. (2010). Peer assessment as a collaborative learning activity: The role of interpersonal variables and conceptions. *Learning and Instruction*, 20, 280-290.

