Chapter XVIII
Fostering Interactivity through Formative Peer Assessment in (Web-Based) Collaborative Learning Environments

Jan-Willem Strijbos
Leiden University, The Netherlands

Theresa A. Ochoa
Indiana University, USA

Dominique M. A. Sluijsmans
HAN University, The Netherlands, & Open University of the Netherlands, The Netherlands

Mien S. R. Segers
Leiden University, The Netherlands

Harm H. Tillema
Leiden University, The Netherlands

ABSTRACT

Extant literature on collaborative learning shows that this instructional approach is widely used. In this chapter, the authors discuss the lack of alignment between collaborative learning and assessment practices. They will argue that peer assessment is a form of collaborative learning and a mode of assessment that perfectly fits the purpose of collaborative learning. As such, the authors purposefully depart from the more traditional application of assessment as a summative tool and advocate the consideration of formative peer assessment in collaborative learning. This shift towards formative assessment they believe has the potential to enhance learning. Their goal in this chapter is to review both shortcomings of current peer assessment practice as well as its potential for collaborative learning. Interactivity is central to foster the alignment between assessment and collaborative learning and the authors present a set of guidelines derived from research for increasing interactivity through formative peer assessment among peers in collaborative learning contexts.
INTRODUCTION

Most research on collaborative learning has focused on investigating the effectiveness of instructions to support and scaffold learning in small groups, as well as the applicability of web-based technology to foster such collaboration (see Fischer, Kollar, Mandl, & Haake, 2007; Jones, Cook, Jones, & De Laat, 2007; Ochoa, Gottschall, & Stuart, 2004; Strijbos, Kirschner, & Martens, 2004). Interestingly, collaborative learning assessment practices have received less attention. Assessment practices tend to be teacher-directed (Chan & van Aalst, 2004; Slavin, 1995) and only few include self- (Barron et al., 1998) and/or peer assessment (Trahasch, 2004). Where peer assessment components are used, they are summative (to determine whether a criterion is met) rather than formative (to determine where a student can improve). Not surprisingly, a ‘group score’ is most often the unit of measure for any given group task, supplemented either with one or multiple individual tasks. Typically, the final score consists of the average with a weighting factor applied. However, as assessment strongly influences learning (Black & Wiliam, 1998; Crooks, 1988; Frederiksen, 1984), we propose in this chapter that any collaborative learning activity should apply an assessment that (a) reflects both the collaboration process and product, (b) promotes students’ collaboration skills as well as cognitive skills, and (c) promotes students’ self-regulation skills. In fact, we will argue that assessment in collaborative learning does not fully tap the potential benefits of the interactive setting unless it includes a mode of assessment that evokes a high and active responsibility from the learner, a component we hold is a critical aspect in collaborative learning.

As a team of researchers from the education and behavioural science disciplines who share an interest and expertise in collaborative learning, web-based instruction, and peer assessment, we first outline the value of peer assessment for collaborative learning settings, illustrated by recent developments in assessment. Subsequently, we present the main shortcomings of current peer assessment practice in light of collaborative learning and provide some reasons why these shortcomings hamper a formative and interactive use of assessment. From the perspective of interactivity, which we define in depth, we present various interactive forms of peer assessment (face-to-face and web-based) and also illustrate how peer assessment can be used to elicit interactivity and subsequently learning (i.e., in terms of collaboration skills, cognitive skills, and self-regulation skills). It should be noted that this chapter predominantly focuses on the pedagogical design issues of peer assessment in collaborative learning (which apply to classroom and web-based settings), and the technology design issues are beyond the scope of this chapter. We close the chapter with a set of guidelines for the application of peer assessment in collaborative learning, and directions for future research and practice of (web-based) peer assessment.

BACKGROUND

Collaborative Learning and Assessment: The Lack of Constructive Alignment

Collaborative learning refers to an instructional approach in which students work together in small groups toward a common goal. The assumption is that learning processes are more effective and productive when students solve problems in collaboration, as compared to when they work alone or only with the teacher (Webb, 1992; Slavin, 1995). In sum, advocates of collaborative learning hold that students learn more in groups than they do in traditional lecture-based instruction (Dochy, Segers, Van den Bossche, & Gijbels, 2003; Ochoa & Robinson, 2005). Research has also shown that collaboration enhances students’
motivation, social skills, and self-efficacy (Johnson & Johnson, 1994).

According to Ochoa and Robinson (2005) a cornerstone of collaborative learning is the assumption that negotiating multiple points of view enhances learning. Under this assumption, problem-solving in groups should produce discussion and cognitive dissonance that prompt students to reconsider and revise individual and collective beliefs and opinions. Furthermore, it is assumed that all members of the group will contribute to the group process. Not surprisingly, reality defies theory and assumptions do not always play out in practice – and simply asking students to interact will not automatically enhance performance.

Further examination of assessment practices used in collaborative learning, reveals that these two are rarely interlinked in research. In general, assessment is the process whereby information on a students’ performance is collected and interpreted (Brookhart, 1999). Formal as well as informal procedures, according to Mehrens and Lehman (1991) are used to collect data in order to develop a comprehensive picture of the characteristics of the learner (in terms of knowledge, skills, attitudes or competences). The effectiveness of assessment depends on the quality of assessment and how it is incorporated by students in subsequent performance. In the past two decades educational research and practice has witnessed a rapid development in the area of assessment. Specifically in terms of why, what, when, how and who should (be) assess(ed) (Segers, Dochy, & Cascallar, 2003). These developments are often characterised by the shift from a testing culture to an assessment culture (Birenbaum, 2003).

In a testing culture the main purpose of assessment is to make evaluative decisions for summative purposes. A shortcoming of summative assessment is that it is decontextualised and atomic, isolated from the learning process and takes place only at the end of a course to judge how well a student performed. Summative assessment focuses strongly on the cognitive aspects of learning, often applies a single performance score, and it is designed and conducted by the instructor. In contrast, the features of a formative assessment culture are that assessment does not only serve summative but also (and to a large extent) formative purposes. Formative assessment is contextualised and intends to build a comprehensive picture of the learners’ characteristics, it is an integral part of the learning process and it is iterative, takes place at several moments during a course rather than only at the end. Formative assessment focuses on cognitive, social, affective, as well as meta-cognitive aspects of learning, often applies a multi-method approach and therefore leads to a profile instead of a single score. Notably, students are involved in the assessment process. Despite our strong support for formative assessment, we do not deny the important aspects of summative assessment. In fact, we like Shute (2007) promote a unified approach that uses traditional (summative; assessment “of” learning) and progressive (formative; assessment “for” learning) perspectives.

In closing our discussion on the background on assessment practices used in collaborative learning we note that despite the strong attention for the instructional and pedagogical aspects of collaborative learning (e.g., research on collaboration scripts; see Fischer et al., 2007), assessment has only recently received more interest in collaborative learning research (Birenbaum, 2005; Meier, Spada, & Rummel, 2007). Yet, it is striking that the assessment of collaborative learning has hardly evolved beyond summative practices. In most cases, assessment focuses on the final group product and is only conducted by the instructor. Criteria are typically predefined by the instructor and the assessment process is treated as a ‘black box’ without any involvement of the learners (Sluijsmans, Strijbos, & Van de Watering, 2007; Sluijsmans, 2008). Because the assessment in collaborative learning is typically disconnected from the instructional setting, a lack of ‘constructive alignment’ exists (Biggs, 1996).
Fostering Interactivity through Formative Peer Assessment

of assessment with the collaborative learning process is therefore crucial (Freeman & McKenzie, 2002). Because an active participation of learners in the assessment process corresponds with active involvement in learning in the collaboration process, this alignment may be achieved through involvement of students using peer assessment (Orsmond, Merry, & Reiling, 2002).

Peer Assessment and Collaborative Learning

According to Topping (1998) peer assessment is an arrangement where equal status students judge a peers' performance with a rating scheme or qualitative report. Under this definition, peer assessment is expected to stimulate students to share responsibility, reflect, discuss and collaborate with peers (Boud, 1990; Orsmond, Merry, & Callaghan, 2004). Peer assessment can be summative or formative, one-way or reciprocal, face-to-face or online, anonymous or non-anonymous, the assessor(s) and assessee(s) may be individuals, pairs or groups, and they may be of the same or different age and/or of the same or different knowledge level (Topping, 2003). Peer assessment is often perceived as a radical change in the assessment process. A primary criticism of peer assessment relates to reliability. Skeptics question the reliability of peer assessment compared to instructor assessment. Refuting this concern, pivotal research conducted by Falchikov and Goldfinch (2000) investigated forty-eight studies reporting on the agreement between instructor and peer marks in higher education and found an overall correlation of .69 – which is substantial and signifies that peer assessment can be applied as an alternative instructional approach.

Peer assessment incorporates varying features of collaborative learning. Regardless of the educational setting, peer assessment thrives on the interaction among group members. In the case of summative peer assessment this usually consists of peer rating using a predefined set of criteria, whereas during formative peer assessment students are involved in discussing and negotiating criteria, and providing and receiving feedback. Irrespective of a summative or formative focus, peer assessment is in essence a specific type of collaborative learning. As a consequence individual accountability and positive interdependence – two core mechanisms of any type of collaborative learning (Strijbos, Martens, & Jochems, 2004) – are important aspects for peer assessment as well; and particularly when applied in a collaborative learning context.

Individual accountability refers to the extent to which group members are held individually accountable for tasks or duties, central to group performance or efficiency. Slavin (1980) introduced this mechanism to counter the well-known free-rider effect (a student deliberately does not invest any or limited effort into group performance (see Kerr & Bruun, 1983), and social loafing (a student assumes that his/her work will be carried out by a more motivated group member (see Williams & Karau, 1991). Presence of a free-rider typically often disrupts the collaborative process and compromises the group’s dynamics or ability to complete the task (Cohen, 1994; Salomon & Globerson, 1989). The likelihood of the presence of a free-rider increases in large groups. However, the impact of the free-rider will be stronger in a small group if such a group member exists. Thus the need for accountability. Individual accountability necessitates that individual responsibilities for the group task are clearly specified. In collaborative learning, peer assessment can be used to hold students individually accountable for their contribution to the group process and product. In addition, when every group member is required to reach the same learning objective, it is in the interest of all group members to provide assistance and feedback to their peers (Slavin, 1995). Hence, interaction and interdependence are essential.

Positive interdependence refers to the extent that the performance of any group member de-
depends on the performance of all other members. Johnson (1981) implemented it to foster group cohesion and a heightened sense of ‘belonging’ to a group. Positive interdependence can be achieved through the type of task, resources, goals, rewards, roles or the environment (Brush, 1998) and it has a strong influence on group cohesion. Positive interdependence and cohesion hinge on familiarity and mutual trust. Cohesion becomes important when the contributions made by one (or more) group member(s) are limited or substandard. When positive interdependence and cohesion are limited, students are prone to regard an assessment as the exclusive realm of a teacher (Cheng & Warren, 1997; Sambell, McDowell, & Brown, 1997; Sluijsmans, Brand-Gruwel, Van Merriënboer, & Martens, 2004).

Peer assessment in collaborative learning can be product-oriented (focused on the group product or individual contributions to the product), process-oriented (focused on the group process or individual contributions the group process), or a combination of the two. It should be noted that process-oriented assessment is only feasible at the individual and intra-group level, because assessing the process of another group is strained since the students themselves did not experience that particular group. In general, peer assessment of both the collaborative process and product is recommended (Divaharan & Atputhasamy, 2002; Lopez-Real & Chan, 1999; Macdonald, 2003; Prins, Sluijsmans, Kirschner, & Strijbos, 2005).

Considering how much collaborative learning is used as an instructional arrangement and how long it has been the topic of research inquiry (see for example Rieken, 1958) it is striking that, thus far, peer assessment typologies have not yet explicitly addressed the interactive and formative potential of peer assessment when applied in a collaborative learning context (Gielen, 2007; Topping, 1998; Van den Berg, Admiraal, & Pilot, 2006). Thus, the main question that we address in this chapter is: How can alignment of collaborative learning and formative peer assessment foster interactivity and enhance learning?

**Shortcomings of Current Peer Assessment Practice in Collaborative Learning**

In an effort to gain insight in peer assessment practices in collaborative learning Sluijsmans et al. (2007) reviewed fifteen studies on peer assessment in collaborative learning. They found that nearly all studies applied peer assessment as a tool to construct individual grades from group grades. The peer assessment usually involves rating scales and appears especially appealing for the instruction of large numbers of students (when an instructor lectures one hundred plus students handing-in individual assignments to be assessed). This format of peer assessment reflects a narrow perspective on the application and benefits of peer assessment, as well as poor alignment with the interaction taking place during collaborative learning.

However, aligning peer assessment with collaborative learning and utilising its interactive and formative potential, requires that three major shortcomings in current peer assessment practice are addressed: a) the lack of attention for students’ ability, b) the lack of attention for the quality of formative feedback, and c) the lack of attention for interpersonal variables.

**Students’ Ability**

In most studies on peer assessment students’ proficiency (their level of achievement) is not clarified, which makes it difficult to determine whether high and low ability students equally benefit from peer assessment. In addition, students’ ability appears to affect peer assessment: less able students may experience more difficulty to self-assess or assess peers (Webb, 1992). Moreover, ability is not included when assessor-assessee pairs are
established (whether these are instructor assigned or self-selected).

Furthermore, ability proved to be related to the learner’s self- and peer assessment. Low achievers tend to overestimate the performance by a peer, whereas high achievers tend to underestimate the performance (Davies, 2006; Lejk & Wyvill, 2001; Patri, 2002). These high ability students are also self-critical rather than judgmental, whereas the opposite is observed for low ability students (Orsmond, Merry, & Reiling, 1997). Yet, recent work by Patri (2002) and Orsmond et al. (2002) reveals that students make more accurate judgments – comparable to those by teachers – when assessment criteria are clearly set. Especially the relevance of the criteria appears to be crucial for students’ understanding of the assessment process, as well as subsequent application in peer assessment (Freeman & McKenzie, 2002).

Finally, students’ are concerned about a potential rating bias (Divaharan & Atputhasamy, 2002). Particularly, naïve assessors may affect peer assessment validity (Langan et al., 2005; Ploegh, Segers, & Tillema, 2007). Students’ also doubt their own and theirs peers’ knowledge within a given subject area (Hanrahan & Isaacs, 2001) as well as their own and peers’ skill to peer assess (Walker, 2001).

Quality of Formative Peer Feedback

Characteristic for formative peer assessment is a strong reliance on the role and importance of feedback. However, feedback does not automatically produce a positive result. Kluger and DeNisi’s (1996) review revealed that one-third of the studies reported negative effects. These outcomes could not be solely explained by cognitive processes or the feedback sign (positive or negative). In education, feedback is systematically investigated in the context of ‘instructor feedback’ and ‘(computer-assisted) tutoring’. Interestingly, in a recent review by Hattie and Timperley (2007) a peer is mentioned as a feedback source, but peer feedback and associated implications are not addressed explicitly. Whereas an instructor, book, parent or computer is regarded as an authoritative source, a peer is not readily considered to be a reliable source for feedback by their fellow students (emphasising the role of students’ ability). Moreover, since students are not experts in a subject area, peer feedback is susceptible to assessor variation.

Feedback can vary in functional, content-related and formal characteristics. Narciss (2006, 2008) concludes that the nature and quality of an external feedback message (by an instructor, computer or by a peer) is determined by at least three facets: (a) functional aspects related to instructional objectives (e.g. cognitive functions, such as promoting information processing; motivational functions, such as reinforcing correct responses or sustaining persistence), (b) semantic aspects related to the content of the feedback message, and (c) formal and technical aspects related to feedback presentation (e.g., frequency, timing, mode, amount, form). As an example of peer feedback variation we mention the feedback stances (authoritative, interpretative, probing and collaborative) described by Lockhart and Ng (1995). The stances reflect two primary types of feedback: evaluative and informative (Van den Berg, 2003). The widely applied use of rating in peer assessment is typical for evaluative feedback (knowledge of result, e.g. “3 out of 5 are correct”), whereas the qualitative report (hints/suggestions, e.g. “give an example”) coincides with informative feedback.

Current research on tutoring focuses on informative feedback, and investigates internal and external feedback loops and the role of persistence and motivation (Narciss & Huth, 2006). In peer assessment, however, the impact of (in) formative feedback on student learning is hardly investigated in relation to student learning (Van Zundert, Sluijsmans, & Van Merriënboer, 2007). Moreover, Van Gennip, Segers and Tillema (2009) could only find fifteen studies, since 1990, which
Empirically investigated learning effects of peer assessment. Although both quasi-experimental and non-experimental studies do indicate positive effects of peer assessment on learning, their review revealed only one of six quasi-experimental studies controlling for the presence of additional (in)formative peer feedback.

**Interpersonal Variables**

In current peer assessment practice the term ‘reciprocity effects’ refers to the bias in peer assessment caused by interpersonal processes. Well-known concepts are: friendship marking (high ratings to friends), collusive marking (high ratings to fellow group members), decibel marking (high ratings to dominant group members) and parasite marking (profiting from the efforts invested by fellow group members) (Cheng & Warren, 1997; Pond, Ul-Haq, & Wade, 1995; Williams, 1992). Reciprocity effects can also influence the social atmosphere in a group (Lejk, Wyvill, & Farrow, 1999) and decrease reliability of peer assessment (Magin, 2001).

Many students also express concerns about the fairness of peer assessment (Dochy, Segers & Sluijsmans, 1999; Sambell et al., 1997; Sluijsmans, Dochy, & Moerkerke, 1999). An increased concern for fairness invokes a central role for the teacher. Zhang (1995) and Sengupta (1998) both found that students’ voiced the perspective that evaluation was the role of the teacher. Furthermore, the perceived or actual ability of both assessor and assessee (Kali & Ronen, 2008; Lin, Liu, & Yuan, 2001; Strijbos, Narciss, & Dünnebier, 2007) appears to affect peer ratings and may have an important effect on both acceptance and application of formative peer feedback in relation to performance (e.g., revision induced by comments) and learning benefits.

In essence, reciprocity effects are grounded in interpersonal variables, such as familiarity, trust, psychological safety, dependence on self and other, and the ability of self and other (see Van Gennip, Segers, & Tillema, 2009). Trust in the self and in peers appears to be especially important for peer assessment and a prerequisite for peer assessment in collaborative learning contexts as most students tend to be hesitant to assess peers, especially when a group member provides a limited or substandard contribution.

Finally, an increasing number of studies indicate that students’ emotional state can mediate the impact of feedback on their performance (Shute, 2008). Overall, students express positive attitudes towards peer assessment and state that they benefit from peer assessment (Divaharan & Atputhasamy, 2002; Hanrahan & Isaacs, 2001; Parikh, McReelis, & Hodges, 2001). Those students who gained a better understanding of the assessment process were more productive in terms of reading peers’ work, developed empathy, they were more motivated—especially to ‘impress peers’ (Hanrahan & Isaacs, 2001) and were encouraged to collaborate, as well as to improve their interpersonal skills (Divaharan & Atputhasamy, 2002).

**INTERACTIVITY IN FORMATIVE PEER ASSESSMENT**

The three shortcomings of current peer assessment practice discussed in the previous section can hamper the alignment between peer assessment and collaborative learning. The extent, to which these shortcomings occur and affect both the alignment of the peer assessment process and collaboration process, appears to depend heavily on the level and intensity of interactivity in a specific peer assessment setting.

Thus far, typologies for peer assessment have expressed the relation between assessor and assessee only in terms of the directionality, i.e. who assesses whom: uni-directional (from an assessor to an assessee; not vice versa), reciprocal (two students or groups assess each other) and mutual (more than two students or groups assess each other) (see Gielen, 2007; Topping, 1998; Van den
Fostering Interactivity through Formative Peer Assessment

Berg et al., 2006). However, directionality is but one aspect of interactivity.

To foster interactivity in peer assessment and thus better alignment with collaborative learning, we extend the definition of interactivity by adding two extra dimensions besides directionality: directionality (uni-directional versus bi-directional), frequency (one-off versus iterative) and constellation (one to one versus one to many; student or group level). This leads to four types of interactivity – each of which can vary in terms of directionality, frequency and constellation – one-way, reactive, reciprocal and negotiated (Figure 1).

All illustrations in Figure 1 represent a one-to-one scenario, that is, one student or group assesses a fellow student or group (a one-to-many representation would be too complicated and provides redundant information). Directionality is depicted by a straight arrow, and the broken arrow represents a reply to the received peer assessment. The frequency is depicted by the number of tasks – with negotiated being an exception as here the peer assessment is an ongoing process. The boxed-letters (A, B, C) illustrate the constellation, and may represent either the individual level (a student assesses another student) or the

---

**Figure 1. Four types of interactivity in peer assessment**

<table>
<thead>
<tr>
<th>Interactivity</th>
<th>One-way</th>
<th>Reactive</th>
<th>Reciprocal</th>
<th>Negotiated</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st task</td>
<td><img src="image1.png" alt="Diagram" /></td>
<td><img src="image2.png" alt="Diagram" /></td>
<td><img src="image3.png" alt="Diagram" /></td>
<td><img src="image4.png" alt="Diagram" /></td>
</tr>
<tr>
<td>2nd task</td>
<td><img src="image5.png" alt="Diagram" /></td>
<td><img src="image6.png" alt="Diagram" /></td>
<td><img src="image7.png" alt="Diagram" /></td>
<td><img src="image8.png" alt="Diagram" /></td>
</tr>
<tr>
<td>3rd task</td>
<td><img src="image9.png" alt="Diagram" /></td>
<td><img src="image10.png" alt="Diagram" /></td>
<td><img src="image11.png" alt="Diagram" /></td>
<td><img src="image12.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>

P = peer assessment (subscript indicates who is assessed ), R = response, E= elaborate, A/D = agree/disagree
Fostering Interactivity through Formative Peer Assessment

group level. The latter may consist of the intra-group level (each student assess the contribution to a shared product by all other group members) and inter-group level (one or more members of a group assess the performance or the product by another group; Sivan, 2000).

One-way peer assessment is the most commonly applied type of interactivity. One student assesses a fellow student. One-way peer assessment is usually different for consecutive tasks. In Figure 1, student A assesses student B during the first task, and student B assesses student C during the second task. In this case each peer assessment task is a one-off event. That is to say, if student A assesses student B during consecutive tasks, interaction between students A and B would be iterative.

Reactive peer assessment is gaining more attention and allows the assessee to reply to the received assessment. In Figure 1 the reply is depicted by the broken arrow. When people feel threatened, for example when they are judged or an assessment is not in line with a view of self-performance, the fairness of assessment is often questioned. Perceived unfairness induces stronger negative reactions as compared to a non-threatening context (Miedema, 2004), and makes acceptance of peer assessment (irrespective of its accuracy) less likely. In contrast, when offered a reply after an assessment, higher levels of perceived fairness and acceptance are observed (Lind, Kanfer, & Earley, 1990).

Reciprocal peer assessment is also regularly applied. Students’ or groups assess each other with respect to their performance on a given task. In essence, this type of interactivity in peer assessment is related to reciprocal teaching (Palincsar & Brown, 1984), scripted cooperation (O’Donnell & Dansereau, 1992) and reciprocal peer tutoring (Fantuzzo, Riggio, Connelly, & Dimeff, 1989). Each student assesses a fellow student and simultaneously s/he is assessed by that fellow student. In Figure 1, student A assesses student B and simultaneously student A is assessed by student B. In this representation students’ also have an opportunity to reply to the assessment. This constitutes the most interactive type of reciprocal peer assessment, although a reply is not prerequisite for an assessment to be reciprocal. The reciprocal nature is further enhanced when students’ assess each other during consecutive tasks, as is shown in Figure 1.

Negotiated peer assessment is the least applied type of interactivity. It is also an ongoing process, which is reflected in Figure 1 by the absence of consecutive tasks. Negotiated peer assessment emphasises the role of an assessee (Prins, Sluijsmans, & Kirschner, 2006) in both understanding (i.e., asking for additional clarification) as well as applying the assessment to improve performance. Figure 1 also illustrates the increased interactivity—in contrast to one-way, reactive and reciprocal peer assessment—in terms of elaboration following the reply and subsequent decision by the assessee to agree or disagree with the peer assessment.

Impact of Interactivity in Peer Assessment

Systematically conceptualising interactivity in peer assessment both highlights the interactive nature of peer assessment and provides a framework to foster constructive alignment between peer assessment and collaborative practices.

A reply to an assessment—or ongoing negotiation of an assessment—fosters the alignment because providing assessee with an opportunity to reply to the assessment and/or feedback received, enhances positive interdependence. Furthermore, a reply and negotiation are typical for formative peer assessment. With respect to assumed learning benefits of a peer assessment it is crucial that a peer understands the assessment or feedback, considers it to be fair, correct, and accepts the assessment. Similar to collaborative learning research on giving and receiving help from a peer (Webb & Farivar, 1999), students have to detect inconsistencies between his/her
own view and the peers’ judgement, understands his/her error(s), and change his/her view accordingly so s/he can meet the criteria in the future (Van de Ridder, Stokking, McGaghie, & Ten Cate, 2008). In sum, a reply may lead recipients to perceive peer assessment as more fair and/or useful, enhance the acceptance of an assessment and subsequently lead to performance improvement (Gielen, 2007; Prins et al., 2005).

An interactivity perspective on peer assessment requires a further specification of what is colluded in the concept of reciprocity effects, namely familiarity and genuine reciprocity. In all types, familiarity – i.e. knowing who you will be assessing and/or who is assessing you – can occur (although it is unlikely to occur in a one-way format). In essence – and as the word ‘reciprocity’ suggests – genuine reciprocity effects can only occur in the case of reciprocal or negotiated peer assessment. Genuine reciprocity occurs when a favourable peer assessment is provided because of the expectation of receiving a favourable assessment in return. Moreover, interpersonal variables inevitably affect intra-group peer assessment in collaborative learning as students’ will need to know the identity of fellow group members in order to judge his/her contribution to the group process and product.

Towards Interactivity During Peer Assessment in (Web-Based) Collaborative Learning

Actual applications of more interactive types of peer assessment in (web-based) collaborative learning contexts are limited. We will briefly discuss three studies to further illustrate the need for and role of interactivity during peer assessment in collaborative learning.

A study by Prins et al. (2005) on formative peer assessment in multidisciplinary groups during a case-based virtual seminar in higher education revealed that students did not apply the predefined criteria for inter-group assessment. The quality of assessment reports was low, they contained more negative than positive comments, and as a consequence it was difficult for students to accept critical peer feedback. Nevertheless, those students that were actively involved in the peer assessment task expressed positive attitudes towards peer assessment and perceived peer feedback as helpful for revising their report. The results clearly reveal that the alignment of the peer assessment with the collaborative learning context was suboptimal; involving students in defining criteria and providing an opportunity to reply could have increased the quality (and acceptance) of assessment reports.

Sung, Chang, Chiou and Hou (2005) studied the use of web-based peer assessment of multimedia web pages designed by high school students. They investigated the effect of Progressively Focused Self and Peer Assessment (PFSPA) which consists of six steps: self-assessment, peer assessment, discussion in small groups, review the best and poorest work (system selected using criterion set by the instructor), second self-assessment, and discussion among groups. The PFSPA format reduced the number of peer assessments that students had to perform. Results indicated that the quality of web pages improved significantly. Students’ selection of the best and poorest work was moderately similar to that of two experts. Overall, their self-rating was closer to expert ratings after PFSPA. Although students worked individually on their web page, the PFSPA format is an expanded, iterative and interactive type of peer assessment. It provides students an opportunity to compare and negotiate the outcome of self- and peer assessment and what characterises high quality work. Unfortunately, this study presents no information on the within-group discussions, which could shed light on students’ reviewing approaches.

Finally, Lee, Chan and van Aalst (2006) describe the application of interactive portfolio assessment in secondary education. In contrast to ‘traditional’ portfolio assessment in which
students collect their best work supported by artefacts and evidence, Lee et al. applied the knowledge-building portfolio “for which students are asked to identify collective knowledge advances documenting the community’s best work and progress” (p. 283). A portfolio (meta comment) consists of four clusters of messages that provide evidence for knowledge-building principles, e.g. ‘working at the cutting edge’. Lee et al. conclude that the exposure to peers’ work and diverse conceptual models facilitates learning, application of knowledge-building principles and surpasses merely identifying good answers. The portfolio approach integrates peer assessment of the collectively acquired knowledge and community processes, although the alignment and interactivity could be further enhanced through small group discussion of students’ knowledge building portfolios.

These studies clearly illustrate the feasibility and potential of more interactive types of peer assessment in collaborative learning contexts. In the following section we will present a set of guidelines to foster interactivity, as well as deal with the implications of interactivity.

GUIDELINES FOR PEER ASSESSMENT IN (WEB-BASED) COLLABORATIVE LEARNING CONTEXTS

We previously outlined three shortcomings in peer assessment practice with special relevance for web-based collaborative learning. We also described the pivotal role of interactivity with respect to the three shortcomings. In this section we present a set of guidelines compiled from literature on peer assessment, collaborative learning and the lessons learned in studies conducted by the authors. These guidelines will address both general issues, as well issues related to interactivity and how a more intentional focus on interactivity can help address the shortcomings inherent in peer assessment.

General Guidelines for (Web-Based) Peer Assessment

The nine principles compiled by Kali and Ronen (2005) serve as a sound starting point for peer assessment. They propose: (1) involve students in the development of criteria, (2) make assessment as anonymous as possible, (3) use an overall global score rather than individual dimensions, (4) use scores generated from peer assessment only after validation, (5) minimise instructor workload, (6) enable students to state their personal, non-objective viewpoints about their peers’ work, (7) foster discussion about non-objective evaluation criteria, (8) do not grade students according to the results of peer assessment, and (9) evaluate students as evaluators using results from peer assessment.

Assuming that instructors are able to manage time effectively, web-based peer assessment in particular provides a viable way of tracking multiple groups simultaneously. In this sense it becomes an attractive approach to ameliorate the group grade with peer ratings of individuals’ contributions. However, a major pitfall for peer assessment during web-based collaborative learning is the apparent ‘ease’ of application. Instructors should be wary of overly simplistic implementations, i.e. rating is appealing in the large scale web-based courses but the minimal information of rating scores should be carefully balanced against the learning objectives.

Often web-based peer assessment is applied to save instructional time and ease instructor administration. It should be noted that a sole focus on saving instructional time compromises the learning aspects of collaborative learning in general and peer assessment in particular – in practice collaborative learning and peer assessment may not really save time, as it requires considerable amount of up-front time commitment. The use of
Fostering Interactivity through Formative Peer Assessment

web-based peer assessment should be determined carefully. We argue in line with Salomon (2000) to “(...) let technology show us what can be done, and let educational considerations determine what will be done” (¶ If it ain’t technology, what is it then?) and that the pedagogical design decisions made by the instructor are central to web-based peer assessment. There are several tools available, but first and foremost the instructor should determine whether a web-based format has an added value (for example, students gain access to peer comments outside of classroom time). However, if web-based peer assessment is chosen some examples of available tools include:

- Networked Peer assessment system (NetPeas; Lin, Liu, & Yuan, 2001);
- Web-based Self and Peer Assessment system (Web-SPA; Sung, Lin, Lee, & Chang, 2003; Sung et al., 2005);
- Self and Peer Assessment Resource Kit (SPARk; Freeman & McKenzie, 2002);
- Online Peer Assessment System (OPAS; Trahasch, 2004);
- Questions Sharing and Interactive Assignments (QSIA; Barak & Rafaeli, 2004);
- Question Posing and Peer Assessment system (QPPA; Yu, Liu, & Chan, 2005);
- Collaborative e-learning Structures (CeLS; Kali & Ronen, 2005; 2008);
- Computerised Assessment by Peers (CAP; Davies, 2006);
- Scaffolded Writing and Rewriting in the Discipline (SWoRD; Cho & Schunn, 2007).

As in the general peer assessment literature, peer rating (marking) is the most common approach in these tools (Web-SPA, QSIA) and usually this is supplemented with open ended comments (NetPeas, QSIA, QPPA, CAP, SWoRD, CeLS). These systems all typically focus on peer assessment of individual student papers or assignments. To our knowledge only CeLS allows the instructor flexibility in the design of peer assessment (a choice for rating, with or without comments, individuals, groups – in isolation or combination) for alignment with the course and the specific instructional format, for example collaborative learning.

Guidelines to Foster Interactive Peer Assessment During (Web-Based) Collaborative Learning

The application of peer assessment in classroom and web-based collaborative learning is especially suited to implement the more interactive formats, and interactivity is crucial for formative peer assessment. Hence, peer assessment should take interactivity in account as well. We will discuss some guidelines to foster interactivity ordered along directionality, frequency and constellation.

Directionality

- Aim for a mix in directionality types; either through several consecutive tasks during a single course or across multiple courses. Considering peer assessment as an interactive process implies that both research and practice pay more attention to the support of the process of conducting, receiving and replying to peer assessment, as well as processes that are grounded in interaction and may mediate the impact of peer assessment, such as familiarity and genuine reciprocity;
- Of the four interactivity types, reciprocal and negotiated are the most ‘constructively aligned’ with collaborative learning as students perform the role of both assessor and assessee (and the peer assessment is subject to reflection – either through a reply or a continued dialogue);
- Students often voice that they do not have the ability and/or level of experience to perform an adequate and fair peer assessment. It is
imperative that students are trained in the ability to assess their peers (Sluijsmans et al., 2004) and that their skill as an assessor is included in the overall course assessment (Davies, 2006; Kali & Ronen, 2008). The impact of student ability on peer assessment exists at both the knowledge level as well as their collaboration skills (process) – both of which may affect peer assessment and should be considered when designing or implementing peer assessment.

**Frequency**

- If possible use multiple tasks: feedback from different peers on different task increases the variation of peer assessment and feedback that students are confronted with, which in turn may lead to performance improvement and enhance learning benefits;
- Peer assessment is most often performed only once – usually on the group product or process – and a combined focus on both the collaborative process and group product is desirable. In addition, iterative or multi-cycle peer assessment (Kali & Ronen, 2008) provides students the opportunity for gradual improvement, exposes them to different assessors, and enables the instructor to determine the intra-rater reliability and assessor consistency.
- Students become comfortable in using peer assessment and value it as contributing to learning as their experiences with it increase (Van den Berg et al., 2006). Make peer assessment an integral aspect of collaborative learning activities in courses and do not present it as an isolated activity.

**Constellation**

- Aim for a mix in assessor and assessee constellations (individuals, groups, one-to-one, one-to-many). The design of a collaborative learning setting in itself dictates the type of interactivity best suited for that setting. Negotiated peer assessment is not feasible when there are three hundred students that are all required to assess three peers during one week. Negotiated peer assessment thrives on intensive interaction and a high level of interdependence, and as such it is better suited for small scale courses. It is all about recognising the context and utilising the restrictions placed on peer assessment design by the collaborative setting;
- Variation in constellation may help to better match students based on their ability and as such create pairs or groups whose zone-of-proximal development are in each others reach – ensuring that differences in ability are not too large;
- When peer assessment is applied summatively, it should always include assessment by the instructor. In addition, multiple peer assessments of a students’ performance or the contribution to the group process is most effective (Cho & Schunn, 2007).

Purposeful design in peer assessment activities that require students to notice how much each group member contributes, as well as the quality of their own contribution to the groups’ process and product, as suggested by Ochoa and Robinson (2005), may yield better peer assessment results. Applying a peer assessment in which each members’ contribution is explicitly noted, can help group member to identify a potential free-rider and ingrains in all group members that an important component of collaborative learning is paying attention to both quality and quantity of peers’ contributions.

In line with the plea by Gielen (2007) for more systematic description of peer assessment settings (to allow for comparison and aggregation of outcomes), we call for a specification of the type of interactivity (one-way, reactive, reciprocal or
negotiated) in terms of directionality, frequency and constellation.

FUTURE TRENDS

In addition to the guidelines to foster peer assessment in general and interactivity in particular, there are several issues we believe are important topics of future research on peer assessment in classroom and web-based settings. The issues revolve around a need to further enhance constructive alignment and interactivity of peer assessment and collaborative learning, and better understand formative peer assessment – both in terms of design and application.

Designing Flexible and Interactive Peer Assessment Formats

At present, peer rating is the dominant peer assessment approach – often in combination with open comments. More variability in the application and design of peer assessment formats is needed. Thus, for research it is important to conceptualise peer assessment also as a design problem, and not merely one of assessment.

Sluijsmans et al. (2007) argue for increased flexibility in peer assessment formats, where peer assessment is used to transform a group mark into individual marks. Current approaches to transformations suffer from five limitations: the type of peer assessment score, using the group mean for calculation, familiarity and genuine reciprocit, scaling down issues, and lack of flexibility. The results reveal that a peer assessment format strongly influences the transformation of a group mark into individual marks. Moreover, peer assessment reliability depends on the weight of criteria, the type of scoring scale, the inclusion of self-assessment, and the maximum deviation of an individual mark from the group mark. Finally, the type of peer assessment source, i.e. the peer assessments received by fellow group members and/or the quality of peer assessments provided to fellow group members (Kali & Ronen, 2008), is an aspect that demands further investigation.

We advocate a use of these formats to foster more interactivity and learning, instead of merely using them to score and calculate. The formats can also be transformed into ‘peer assessment scripts’ (similar to ‘collaboration scripts’), and flexibly applied in web-based collaborative learning contexts (Miao & Koper, 2007) to guide students interactivity, as well as control for some of the shortcomings of current peer assessment practice, for example matching students according to their zone-of-proximal development.

Interpersonal Variables and Anonymity

In response to the familiarity and genuine reciprocity effects, as well as recent indications that students’ emotional state (Shute, 2008) and characteristics of the peer assessor (Kali & Ronen, 2008; Strijbos et al., 2007) mediate the impact of peer feedback on performance, anonymous peer assessment is strongly advocated (Cheng & Warsen, 1997; Davies, 2006; Freeman & McKenzie, 2002; Kali & Ronen, 2008). Unquestioningly web-based peer assessment systems are very suited to retain anonymity.

We agree that anonymous peer assessment is preferred when the peer assessment outcomes are used summatively or have high-stake implications. However, a formative peer assessment aims to foster reflection and changes in performance and collaborative skills – including skills to cope with feedback received from peers. Furthermore, interpersonal variables inevitably affect intragroup peer assessment in collaborative learning as students’ will need to know the identity of fellow group members in order to judge his/her contribution to the group process and product.

Anonymous peer assessment may be less threatening but it simultaneously appears needlessly protective and may even hinder the development of a healthy set of behaviours to provide
feedback to fellow students that they are familiar with (or even close friends). We are aware that this position questions the second principle by Kali and Ronen (2005), but familiarity is subject to an intriguing paradox: we are inclined to rate familiar persons more positively and we are more prone to accept feedback from a familiar person than someone unfamiliar.

The call for anonymous peer assessment is predominantly motivated from the perspective of rating bias. When peer assessment is used formatively, the argument for using anonymous assessment should be reconsidered and its benefits weighted against the caveats. The choice for anonymous or non-anonymous assessment is also affected by the perspective of learning how to deal with assessment from persons that students know and learning how to cope with critical feedback, and distil how students can improve and cope with emotions. One approach could be the gradual introduction of personal and performance indicators, or specific instruction as to how students’ should deal with peer assessment scores and feedback in particular (e.g., focus on improvements that can be made from the feedback). In case of few comments students can be assigned to contrast their work again in light of the criteria, because there can always be elements that peers failed to notice. Another approach could be a combined anonymous web-based and non-anonymous classroom peer assessment approach, ameliorated with group discussions of the peer assessment outcomes—including students’ perceptions of the received peer feedback and their emotions.

CONCLUSION

Towards a Good Marriage between Collaborative Learning and Peer Assessment

In this chapter, our main focus was concerned with the alignment of collaborative learning and interactivity in formative peer assessment. We asked how instructors could enhance learning. Our conclusion, one we elaborated upon, is that interactivity in formative peer assessment (classroom or web-based) is crucial to support the development of students’ cognitive skills, collaboration skills and their self-regulation skills. Moreover, we hold that interactivity (e.g., through formative peer assessment) can foster the constructive alignment between assessment and collaborative learning, as such the assessment can be integrated in the collaborative environment instead of a stand-alone feature at the end. This may enhance the perceived relevance, usefulness and fairness of group work assessment.

The overview of shortcomings illustrates that a formative and interactive application of peer assessment to achieve these learning goals is still in its infancy, but that is it also time to reconsider our philosophy on assessment in collaborative learning. We made a first attempt in presenting a perspective on assessment in collaborative learning that warrants for alignment between collaborative learning and assessment, but also addresses the responsibility of the learner. After all, it becomes increasingly important that the learner develops him- or herself as a lifelong learner, who is able to act professionally in each (collaborative learning) context (Boud, 2000).

REFERENCES


Fostering Interactivity through Formative Peer Assessment


Fostering Interactivity through Formative Peer Assessment


Fostering Interactivity through Formative Peer Assessment


Sivan, A. (2000). The implementation of peer assessment: An action research approach. *As-
Fostering Interactivity through Formative Peer Assessment


