British Journal of Educational Technology *Vol 41 No 3 2010* 525–536
doi:10.1111/j.1467-8535.2009.00968.x

Assessor or assessee: How student learning improves by giving and receiving peer feedback

Lan Li, Xiongyi Liu and Allen L. Steckelberg

Lan Li is an assistant professor for the College of Education and Human Development at Bowling Green State University, Bowling Green, OH, USA. Xiongyi Liu is an assistant professor for the College of Education and Human Services at Cleveland State University, Cleveland, OH, USA. Allen L. Steckelberg is an associate professor for the College of Education and Human Sciences at the University of Nebraska-Lincoln, Lincoln, NE, USA. Address for correspondence: Dr Lan Li, Classroom Technology, School of Teaching and Learning, 529 Education Building, Bowling Green State University, Bowling Green, OH, USA. Email: lli@bgsu.edu or lilan2003@hotmail.com

Abstract
This study investigated the relationship between the quality of peer assessment and the quality of student projects in a technology application course for teacher education students. Forty-three undergraduate student participants completed the assigned projects. During the peer assessment process, students first anonymously rated and commented on two randomly assigned peers' projects, and they were then asked to improve their projects based on the feedback they received. Two independent raters blindly evaluated student initial and final projects. Data analysis indicated that when controlling for the quality of the initial projects, there was a significant relationship between the quality of peer feedback students provided for others and the quality of the students' own final projects. However, no significant relationship was found between the quality of peer feedback students received and the quality of their own final projects. This finding supported a prior research claim that active engagement in reviewing peers' projects may facilitate student learning.

Theoretical framework

Peer assessment is a process in which students evaluate the performance or achievement of peers (Topping, Smith, Swanson & Elliot, 2000). This innovative assessment approach aims to empower students and foster active learning. Peer assessment can provide summative grading when students rate each other's performance. But many cases focus also on formative goals, such as promoting student learning. In most formative peer assessment models, students act as both assessors and assessees. As assessors, they review peers' work and provide constructive feedback. As assessees, students receive feedback and may make improvements accordingly. Through such processes, peer assessment becomes a strategy for formative assessment and a tool for reflection by students (Cheng & Warren, 1999).

A number of studies confirm the benefits of peer assessment for student learning. For example, Pope (2001) considers peer rating as a tool to facilitate learning. Forty master's students at an Australian university showed improved writing and reporting skills following peer rating, particularly in the area of 'spelling and grammar, referencing and logic' (p. 242). In Venables and Summit's (2003) study, computer science students, with initial reservations about peer review, credited the process with enhanced knowledge of subject matter. In Stefani's (1994) study, students commented that peer assessment made them think more; a majority of students (85%) were in favour of peer assessment regarding learning, when comparing this approach with traditional assessments. After reviewing 109 papers focusing on peer assessment, Topping (1998) summarised that peer assessment yields cognitive benefits for both assessors and assessees in multiple ways: constructive reflection, increased time on task, attention on crucial elements of quality work and greater sense of accountability and responsibility, etc.

These findings have also been supported by other technology-facilitated studies. In a study conducted by Sung, Chen-Shan Lin, Chi-Lung Lee and Chang (2003), 34 undergraduate students utilised both web-based peer and self-assessment to evaluate research proposals in a one-semester experimental psychology class. After uploading their proposals on the Web, students performed peer and self-review online, and then revised their own proposals. The comparison of the original versions and the revised versions of student proposals indicated a significant improvement of quality. Li and Steckelberg (2005) investigated the impact of an anonymous technology-mediated peer assessment on students' project quality. Forty-seven teacher education students learned and created WebQuest projects. Afterwards, they were randomly assigned into two groups—a control group and an experimental group. Students in the experimental group reviewed each other's projects in a web-based peer assessment support system, and made revisions in their own projects based on peer feedback, while the control group pursued WebQuest-related activities, such as reviewing WebQuest and the rubric that they had learned and discussed, or read WebQuest examples or articles regarding this topic. Both groups were asked to spend the same amount of time on the task. Projects from these two groups were collected and blind assessed by an independent grader. Data analysis revealed that the student group with peer-assessing experience outperformed the other.

Although general learning gains have been frequently reported, how the different role of being assessors or assessees affects learning is not clear. Would students benefit more from performing as assessors (reviewing peers' work and providing feedback) or assessees (receiving feedback from peers)? Or would both roles contribute to students' learning? Existing literature on this issue is often limited to students' perceptions, rather than empirical data. Regarding the assessor role, some researchers (eg, Althauser & Darnall, 2001) asserted that the higher the quality of the feedback students gave their peers, the better they themselves performed. This claim is congruent with student perceptions. For example, Davies (2000) found that, after a computerised peer assessment, students acknowledged a significant benefit from marking peers' work. Over 60% of the students felt that they 'had worked at a deeper level of understanding' (p. 350). In another study

by Li and Steckelberg (2006), students acknowledged that it was helpful to 'look at what others are doing', and some of them felt 'inspired' by peers' work (p. 268).

From the perspective of assessees, however, findings are mixed. While students acknowledged the value of peer feedback, some students complained about the poor quality of peer feedback that they received. In a third study, 38 undergraduate students (Li, Steckelberg & Srinivasan, in press), after experiencing an anonymous technology-mediated peer assessment, responded to a post-assessment survey and marked peer assessment as a worthwhile activity. However, when asked to report their least liked features, students called for more constructive and more detailed feedback.

This response occurs in a number of other studies. Orsmond and Merry (1996) found that students were doubtful about the value of their peers' marks. Brindley and Scoffield (1998) reported that students criticised other students' capacity to perform assessments. In summary, literature on roles of assessor and assessee regarding student learning is inadequate and limited to student perceptions and feelings. The present study addressed this issue by empirically examining two relationships: (1) the relationship between the quality of students' final projects and the quality of peer feedback which students provide, and (2) the relationship between the quality of students' final projects and the quality of peer feedback which students receive. Specifically, we ask the following questions: (1) When the quality of students' initial projects (prior to peer assessment) is controlled for, is there a relationship between the quality of students' final projects (post-peer assessment) and the quality of peer feedback students provide to others? (2). When the quality of student initial project is controlled for, is there a relationship between the quality of students' final projects and the quality of peer feedback these students receive?

Facilitating website

To facilitate this study, we built a web-based peer assessment support system (Figure 1), which provided two interfaces. In the student interface, students first uploaded their projects to this system. Students' projects were then randomly assigned to each other. Each student reviewed two peers' projects. Likewise, each student received feedback for his/her own project from two peers. Students then revised their own projects based on peer feedback they had received. Both project authors and reviewers remained anonymous during the review process. In the instructor interface, instructors were able to manage students' accounts and track the peer assessment process.

Method

Participants

Forty-three undergraduate teacher education students enrolled in a required, entry-level educational technology application course at a Midwestern university participated in the study. Most of these students were traditional, female (75% among reported participants) and ranged from freshman to senior standing, and represented varied educational backgrounds.

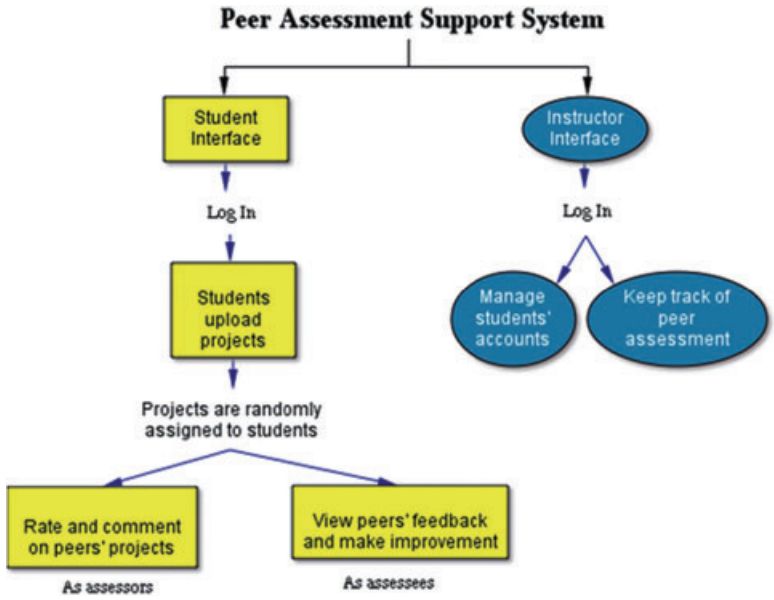

Figure 1: Web-based peer assessment support system

Project and rubric

Participants in this study created a WebQuest project in a word processing document. WebQuest, as defined by Bernie Dodge, is 'an inquiry-oriented activity in which most or all of the information used by learners is drawn from the web' (Some thoughts about WebQuests, n.d.). A WebQuest provides a series of scaffolding activities to aid student learning. Because a WebQuest is frequently used in educational settings, it is important for preservice teachers to understand what a WebQuest is, and to possess the ability to construct a good quality WebQuest.

This study utilised a WebQuest rubric published by San Diego State University to evaluate the quality of students' WebQuest projects (A rubric for evaluating WebQuest, n.d.). This rubric, which is available online (http://webquest.sdsu.edu/webquestrubric.html), includes 13 items (maximum 50 points) addressing five critical components of WebQuest activities: introduction, task, process, evaluation and conclusion. Three levels of performance indicators (beginning, developing and accomplished) and corresponding points are provided for each item in this rubric. This rubric was used for both peers' assessment (students rated and commented upon peers' projects according to this rubric) and raters' assessments (raters assessed the quality of students' initial and final versions of WebQuest projects based on this rubric).

Procedures
All participants followed the same peer assessment process.

Step 1: discussing marking criteria
The instructor provided an introduction to WebQuests that included WebQuest content and the rubric describing basic elements for good performance. Because both peer assessment and instructor assessment would utilise this rubric, it was vital for them to understand what each item in the rubric meant and how to utilise it when providing peer feedback. Three steps were followed in class in order to strengthen students' understanding of rubric and foster their assessment skills. First, abstract terms were explained in class in order to clarify possible misunderstandings. Second, students in small groups discussed each item of the rubric, with the instructor available to answer questions. In addition, because most students had never had any prior experience judging other people's work, students practised assessment of two example projects to develop their assessment skills.

Step 2: composing and submitting projects
Students were introduced to the web-based peer assessment website, and registered for their own accounts. They began composing their initial WebQuest projects and uploaded their projects to the site. Typical WebQuest projects may, for example, first introduce water quality issues, then provide detailed steps, guiding target audience to assess water quality in their community, and ask target audience to compose a letter to the mayor regarding suggestions on how to resolve these issues. Completed WebQuest projects ranged from two to three pages in length in Microsoft Word.

Step 3: assessing peers' projects and providing feedback
Upon submission to the website, students' projects were randomly assigned for peer review by the website software. Each student was assigned to two peers' projects. The website provided access to projects to be reviewed and prompted students to rate and provide comments/suggestions on these peers' work. The website provided the structure for presenting the rubric and gathering comments/suggestions. During the review process, students were instructed to provide constructive forward-looking feedback, as 'good feedback comprises not just commentary about what has been done, but suggestions for what can be done next' (Brown, 2007, p. 1). For any issue identified in peers' work, students should not just give an opinion, but also back it up with clearly articulated reasons why and how a change should be made. Their feedback should also be expressed in an honest, respectful and supportive way to their peers.

Step 4: viewing peers' feedback and improving the project
The web-based system returned peer feedback to respective project authors. Students considered their peers' suggestions and revised their projects. They were reminded that the quality of peer assessment would vary. Two peers might offer contradictory comments. It was important for authors to evaluate the quality of suggestions and decide which to apply. Because of this circumstance, the proportion of suggestions taken by students varied greatly. While some students adopted most of the comments, a few

students decided that their peers did not understand WebQuest, the rubric or their projects, therefore, took no or only a small amount of suggestions. Students submitted their final revised projects for instructor assessment.

Grading

To ensure the reliability and validity of grading, one researcher and an independent rater evaluated the initial WebQuest projects (prior to peer assessment), final WebQuest projects (post-peer assessment) and the quality of the peer feedback. The independent rater was an instructional designer who had at least 4 years of experience of building WebQuests. He participated in a 10-hour training to learn how to assess the quality of WebQuests according to the specific rubric utilised in this study. Students' projects and peer assessments were mixed together and identified by ID numbers. The researcher and the independent rater were not able to identify whose assessments or projects they evaluated.

Students' initial and final projects were assessed according to the WebQuest rubric, which students used in peer assessment. The quality of peer feedback was judged on the identification of major issues and the quality of constructive suggestions for improvement (Appendix A). When assessing the quality of peer feedback, the researchers calculated their own scores and ignored the scores which peer assessors had assigned to projects, as rating can be affected by other issues, such as friendship or fear of offending people (Li & Steckelberg, 2006), and may not represent students' real judgement. This present study was concerned with students' ability to identify problems and provide constructive suggestions. Because each student reviewed two peer projects and received feedback from two different peers, the outgoing and incoming peer review pairs were both likely to differ.

Interrater reliability for the independent rater and researcher were satisfactory for the WebQuest projects, 0.83 for the initial project and 0.80 for the final project respectively. Interrater reliability for the quality of peer review was even higher (0.85). Data analysis was based on grading by the researcher.

Data analysis

Descriptive statistics and correlational coefficients

Table 1 reports means, standard deviations and Pearson's bivariate correlations for all variables. In general, student performance on the project improved after the peer review. The average score on the project increased from 31.88 to 38.82, with a positive change of about seven points. The standard deviation of project scores was somewhat smaller for the final project (post-peer assessment) than it was for the initial project prior to peer assessment (5.82 vs. 7.85). The quality of all students' reviews on their peers' projects was moderate, with an average of 7.01 on a 10-point scale. The standard deviation of the reviews was somewhat larger for the sets of reviews students provided for peers' projects than it was for the sets of reviews students received for their own projects (2.07 vs. 1.78).

Table 1: Means, standard deviations and bivariate correlations among final project score and its predictors

			Correlations			
	Mean	*SD*	*P2*	*P1*	*PA*	*PB*
P2	38.82	5.82	—			
P1	31.88	7.85	0.45*	—		
PA	7.01	2.07	0.56**	0.15	—	
PB	7.01	1.78	0.17	–0.03	0.12	—

Note. $n = 42$.
P2, final project score; P1, initial project score; PA, quality of reviews students provided for peers' projects; PB, quality of reviews students received for their own projects.
*$p < 0.01$. **$p < 0.001$.

We examined Pearson's bivariate correlations among the dependent variable (ie, final project score) and its three predictors. Both the initial project scores and the quality of reviews students provided for peers' projects were moderately related to the final project scores ($r = 0.45$ and 0.56 respectively), while the quality of reviews students received for their own projects showed a much weaker relationship ($r = 0.17$). Initial project scores had a low correlation with the quality of reviews students provided for others ($r = 0.15$), and almost no relationship with the quality of reviews students received. Finally, the quality of reviews students provided for others had a low correlation with the quality of reviews students received ($r = 0.12$).

Multiple regression analysis

Hierarchical multiple regressions were used to analyse the data, which is appropriate when the predicting variables are not highly correlated. Hierarchical multiple regression is a major analytic strategy in multiple regression (Tabachnick & Fidell, 2007). The analysis proceeds in steps, and at each step, a new independent variable or a set of new independent variables are entered and evaluated in terms of what it adds to prediction of the dependent variable, beyond the predictability afforded by the independent variables entered at previous steps. The order by which the independent variables enter the equation or the regression model is determined by the researcher, according to logical or theoretical considerations. For example, independent variables, which are considered to be causally prior to other independent variables, are assigned with higher priority.

In the present study, our analysis consisted of three steps, with new independent variables or predictors added at each step. We then examined how much the models created at later steps add to the prediction of the dependent variable (ie, student final project grade). Table 2 shows the predictors entered at each step, and the parameter estimation for each predictor in the model created at each step. As shown in Table 2, the first model with the quality of student initial project as the only predictor fit the data well, $F(1, 40) = 9.858$, $p = 0.003$, $R^2 = 0.198$. The second model, with the quality of reviews

Table 2: Results of hierarchical multiple regression for prediction of final project score

	Model 1		*Model 2*		*Model 2*	
Predictors	*β*	*SE*	*β*	*SE*	*β*	*SE*
P1	0.330**	0.096	0.276*	0.089	0.280*	0.089
PA			1.414**	0.339	1.369**	0.341
PB					0.406	0.392

Note. β, standardised regression coefficient; P1, initial project score; PA, quality of reviews students provided for peers' projects; PB, quality of reviews students received for their own projects. *$p < 0.01$. **$p < 0.001$.

students provided for peers' projects added to the predictors, fit the data significantly better than the first model, $\Delta F(1, 39) = 17.392$, $p < 0.001$, $\Delta R^2 = 0.247$. Both the quality of student initial project and the quality of reviews students provided for peers' projects were significant predictors when controlling for each other, $t(1, 40) = 3.085$, $p = 0.004$ and $t(1, 40) = 4.170$, $p < 0.001$ respectively. When the quality of their project before peer assessment was controlled for, students who gave better feedback to their peers produced significantly better final projects than those who gave poor feedback. The third model, with the quality of reviews students received for their own projects added to the predictors, did *not* fit the data better than the second model at a significant level, $\Delta F(1, 38) = 1.070$, $p = 0.308$, $\Delta R^2 = 0.015$. While the quality of student initial project and the quality of reviews students provided for others remained significant predictors when controlling for the other two predictors, the quality of reviews students received was not significantly related to the quality of the final project, $t(1, 40) = 1.034$, $p = 0.307$.

Discussion

This study examined how the role of assessor and assessee in peer assessment impacts student learning. Two research questions were: (1) When the quality of students' initial projects (prior to peer assessment) is controlled for, can the quality of students' final projects (post-peer assessment) be predicted by the quality of peer feedback students provide? (2) When the quality of student initial project is controlled for, can the quality of students' final projects be predicted by the quality of peer feedback students receive? For Question (1), data analysis suggested, when controlling for the quality of the initial project, that there was a significant relationship between the quality of the peer feedback the students provided for others and the quality of the students' own final projects. This finding is congruent both with assertions in the literature that active involvement in the peer assessment process improves learning, and studies reporting student perceptions that reviewing peers' work facilitated their learning. This finding also supports the idea that the more constructive feedback students are able to give their peers, the better they perform on the task. However, for Question (2)—relationship between quality of peer feedback received and student learning—there was no direct link between the peer feedback students received regarding their own projects and the

quality of their final projects. This finding contradicts our common belief that high-quality feedback leads to better performance.

There might be two possible explanations for this result. One, students were instructed to assess the value of peer feedback before they accepted suggestions to revise their projects. Students were advised that peer feedback might vary in quality, and if they obtained conflicting reviews, or if they received comments contradictory to their understanding, they should go back to study the content area and rubric till they figured out what was right and what was wrong. We think this process may have engaged students in an active learning mode and balanced any negative effects of poor-quality peer feedback. Two, it might also be possible that, because all participants both gave and received feedback, what they learned about the project might have allowed them to better improve their project even if they had poor-quality feedback.

If the finding of this study regarding lack of significant relationship between quality of peer feedback students receive and students' performance in peer assessment is warranted and can be replicated, students may benefit from peer assessment even though the quality of the feedback they receive is varied.

These findings, however, do not suggest that improving the ability of students to make high-quality feedback is not important. In fact, it suggests that the ability to give high-quality feedback is a critical issue, and that improving the quality of feedback is worthwhile for the assessor. Both training and monitoring of the quality of the feedback may have played a role in improving assessors' ability to provide feedback and are important components in understanding the peer assessment process. We suggest future in-depth studies to investigate these issues. For example, future studies may examine the different impact of providing feedback and receiving feedback on their learning improvement by randomly assigning students to play the roles of either an assessor or an assessee in an experimental setting.

In most peer assessment studies, students play the roles of assessor and assessee. Although learning gains have been frequently reported in literature, it is not clear how these two roles contribute to student learning process. It was not clear if students benefited from being assessors or assessees, or both. This study helped answer this question. Students' effectiveness in assessing their peers is derived from an understanding of the content area and marking criteria, both of which contribute to performance. On the other hand, receiving poor-quality peer review may not adversely impact students' performance as long as students do not rely solely on peer comments as resources for project improvement. This study suggests that students' active involvement in assessing peers' work and evaluating peer feedback is related to learning outcomes, and how to engage students warrants attention in the implementation of a peer assessment system.

A limitation of this study is that the unit of analysis is confined to students who experienced this specific technology-assisted peer assessment system. Literature review

reveals that peer assessment models may vary dramatically in main parameters, such as objectives (eg, formative or summative), assessment delivery (eg, paper based or web based), student academic levels (eg, undergraduate or graduate), previous peer assessment experience, etc (Topping, 1998). This study will not account for all peer assessment practices. We suggest that other studies be conducted using new technology to assess other different peer assessment models. Future studies should also consider recruiting different groups of participants. In this study, participants were teacher education students enrolled in a face-to-face technology application course. Prior to this peer assessment activity, students had already acquired some basic computer skills, instructional design principles and technology integration strategies. It would be interesting to see how other groups of students' experiences as assessors would influence their performance.

This study makes an important contribution in providing empirical data on how the roles of assessor and assessee impact student learning. Although a number of researchers claim that the process of giving and receiving feedback in peer assessment process fosters students' cognitive development, statistical evidence demonstrating the value of this process on student learning is scarce. This study suggested that the quality of feedback students provided in reviewing the work of peers correlated positively with the quality of their own work. However, there was no evidence of a direct link between the quality of feedback students received and the quality of their projects. These findings provide evidence for a theoretical explanation of the value of active engagement in peer assessment. At the same time, this study shows the importance of students' assessment capability in formative peer assessment activities, that is, students' ability to judge the quality of their own work may have more influence than feedback from peers, particularly if the feedback may be of varied quality.

References

A rubric for evaluating WebQuest (n.d.). A rubric for evaluating WebQuest. Retrieved September 1, 2008, from http://webquest.sdsu.edu/webquestrubric.html

Althauser, R. & Darnall, K. (2001). Enhancing critical reading and writing through peer reviews: an exploration of assisted performance. *Teaching Sociology, 29*, 1, 23–35.

Brindley, C. & Scoffield, S. (1998). Peer assessment in undergraduate programmes. *Teaching in Higher Education, 3*, 1, 79–80.

Brown, S. (2007). Feedback & feedforward. *Center for Bioscience Bulletin, 22*, 1. Retrieved December 1, 2008, from http://www.bioscience.heacademy.ac.uk/ftp/newsletters/bulletin22.pdf

Cheng, W. & Warren, M. (1999). Peer and teacher assessment of the oral and written tasks of a group project. *Assessment & Evaluation in Higher Education, 24*, 3, 301–314.

Davies, P. (2000). Computerized peer assessment. *Innovations in Education and Training International, 37*, 4, 346–354.

Li, L. & Steckelberg, A. (2006). Perceptions of web-mediated peer assessment. *Academic Exchange Quarterly, 10*, 2, 265–269.

Li, L. & Steckelberg, A. L. (2005). Impact of technology-mediated peer assessment on student project quality. In M. Simonson & M. Crawford (Eds.), *Proceedings of Association for Educational Communications and Technology International Conference 2005* (pp. 307–313). Bloomington, IN: AECT.

Li, L., Steckelberg, A. L. & Srinivasan, S. (in press). Utilizing peer interactions to promote learning through a computer-assisted peer assessment system. *Canadian Journal of Learning and Technology*.

Orsmond, P. & Merry, S. (1996). The importance of marking criteria in the use of peer assessment. *Assessment & Evaluation in Higher Education, 21*, 3, 239–250.

Pope, N. (2001). An examination of the use of peer rating for formative assessment in the context of the theory of consumption values. *Assessment & Evaluation in Higher Education, 26*, 3, 235–246.

Some thoughts about WebQuests (n.d.). Retrieved September 1, 2008, from http://webquest.sdsu.edu/about_webquests.html

Stefani, L. A. J. (1994). Peer, self and tutor assessment: relative reliabilities. *Studies in Higher Education, 19*, 1, 69–75.

Sung, Y.-T., Chen-Shan Lin, J., Chi-Lung Lee, J. & Chang, K.-E. (2003). Evaluating proposals for experiments: an application of web-based self-assessment and peer assessment. *Teaching of Psychology, 30*, 4, 331–333.

Tabachnick, B. G. & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon.

Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research, 68*, 3, 249–276.

Topping, K. J., Smith, E. F., Swanson, I. & Elliot, A. (2000). Formative peer assessment of academic writing between postgraduate students. *Assessment & Evaluation in Higher Education, 25*, 2, 149–169.

Venables, A. & Summit, R. (2003). Enhancing scientific essay writing using peer assessment. *Innovations in Education & Teaching International, 40*, 3, 281–290.

Appendix A

Peer feedback rubric

ID of assessor:

ID of the assessed project:

Introduction

(1). Was the assessor able to identify critical issue(s) in this section of the assessed project? (1 point)

(2). Was the assessor able to provide constructive suggestions for issue(s) identified in this section of the assessed project? (1 point)

Task

(3). Was the assessor able to identify critical issue(s) in this section of the assessed project? (1 point)

(4). Was the assessor able to provide constructive suggestions for issue(s) identified in this section of the assessed project? (1 point)

Process

(5). Was the assessor able to identify critical issue(s) in this section of the assessed project? (1 point)

(6). Was the assessor able to provide constructive suggestions for issue(s) identified in this section of the assessed project? (1 point)

Evaluation

(7). Was the assessor able to identify critical issue(s) in this section of the assessed project? (1 point)

(8). Was the assessor able to provide constructive suggestions for issue(s) identified in this section of the assessed project? (1 point)

Conclusion

(9). Was the assessor able to identify critical issue(s) in this section of the assessed project? (1 point)

(10). Was the assessor able to provide constructive suggestions for issue(s) identified in this section of the assessed project? (1 point)

Total points for this peer feedback: _____ out of 10