Exploring the Development of Web-based Peer Assessment System *

Shiau-Ping Yeh
Institute of Graduate Institute of Learning & Instruction, National Central University
Taiwan
961207005@cc.ncu.edu.tw

Tzu-Chien Liu
Institute of Graduate Institute of Learning & Instruction, National Central University
Taiwan
ltc@cc.ncu.edu.tw

Sabine Graf
Institute of Graduate Institute of Learning & Instruction, National Central University
Taiwan
sabine.graf@ieee.org

Yu Wang
Institute of Graduate Institute of Learning & Instruction, National Central University
Taiwan
smallsher@mail2000.com.tw

Abstract: Peer assessment becomes a popular and effective approach to enhance students' critical thinking skills and make them reflect more about their and others' work. In recent years, several web-based peer assessment systems were developed in order to support peer assessment through technology. These systems have different functionality, characteristics, strengths, and weaknesses. This paper aims at discussing new trends in the development of web-based peer assessment systems by taking a closer look to ten systems. By providing an overview about the systems' characteristics and its application model as well as discussing the target users and the changing role of teachers and students through web-based peer assessment, this paper contributes in providing an overview of web-based peer assessment, helps to understand the today's developments, and gives recommendations for new innovations in web-based peer assessment.

Introduction

In a conventional classroom environment, students do an assigned work within a limited period of time. After receiving submissions from students, the teacher assesses their works. Students generally receive feedback only from the teacher. However, limited by time and workload, the instructor can not put much focus on the assessment, and students' improvement is confined (Cho & Schunn, 2007). Peer assessing is an innovative assessment method in today's courses, which has gradually replaced teacher assessments. Peer assessment can help instructors to spend more time on other teaching activities by reducing the instructors' workload (Cho & Schunn, 2007). Otherwise, peer assessment enhances student interpretation and reflection; students are able to observe their peers' work and learning process and come to better understanding of how their peers learn. In the process of peer assessment, students are capable of learning how to make judgment, criticize peer work, and accept peer criticism, thereby developing their critical thinking skills (Lan et al., 2007; Kwok & Ma, 1999).

However, there are several problems in administrating reciprocal peer assessments in courses. First, student peer assessors are novices in the disciplines. Thus, their feedback and evaluation can be inaccurate relative to the

* The authors would like to thank the National Science Council of the Republic of China, Taiwan, for financially supporting this research under Contract No. NSC 097-2811-S-008-001- and Contract No. NSC 96-2520-S-008 -007 -MY2.

feedback generated by a subject-matter expert or instructor (Cho & Schunn, 2007). The second constraint is that there is a relative inconvenience of displaying multimedia works such as music and paintings using pens and paper (Sung et al., 2005). Thirdly, executing peer assessment in the class poses a challenge, since there is little time for mutual discussions (Sung et al., 2005). Aiming at enhancing education by the use of technology, many web-based peer assessment systems were rapidly developed to solve the above problems in the recent years. But, there are few researches on exploring them. This paper aims at presenting the current developments of web-based peer assessment systems by exploring several systems and discussing their characteristics, application model and target users. The goal of this study is to show the current situation of web-based peer assessment system and give recommendation for future researches.

This paper is organized the following way: the next section describes an analysis of web-based peer assessment systems. Subsequently, the application model is explained. In the fourth section, we briefly introduce the target users of the investigated peer-assessment systems and subsequently the role of instructors and students in such systems is discussed. The last section concludes the paper.

An analysis of web-based peer assessment systems

As shown in Table 1, this study compares the difference and consistency between ten web-based peer assessment systems. We especially looked at web-based systems supporting peer assessments in school/university activities. In the following subsections, we describe the key features used in our investigation in more detail and discuss how the systems implement them.

Table 1: The characteristics of web-based peer assessment systems

	NetPeas (Lin et al., 2001)	SPARK (Freeman & Mckenzie, 2002)	SWoRD (Cho & Schunn, 2007)	(Yang et al., 2005)	Web-SPA (Sung et al., 2005)	Vee heuristic (Tsai et al., 2002)	CAP (Davies, 2000)	OASYS (Bhalerao & Ward, 2001)	GSS (Kwok & Ma, 1999)	(Yeh, 2001)
Anonymity	●	●	●	●	●	●	●	●	●	●
Individual works	●	●	●	●	●	●	●	●	●	●
Multiple peer feedback	●	●	●	●	●	●	●	●	●	●
Assessor support	●		●			●	●			●
Multidimensional evaluation	●	●	●	●	●	●			●	●
Verbal comments	●	●	●	●	●	●	●	●	●	●
Instructors actions	●	●	●	●	●	●	●		●	●

Anonymity

All of the investigated systems emphasize anonymity in application. Anonymity is important because writers and assessors are likely to be less critical when identities are known (Cho & Schunn, 2007；Lin et al., 2001；Yang et al., 2005；Wen & Tsai, 2006；Sung et al., 2005). Pseudonyms are usually created while using the systems, which can prevent any biased evaluations.

Individual works

Some studies showed that students benefit most when they work individually for solving a specific task and afterwards cooperate, but not when they continuously cooperate (Cho & Schunn, 2007). In all investigated systems, students are asked to perform assignments individually rather than in collaboration with others. After submitting their own works to the systems, each student is assigned with peers' works to grade and comment. After providing grades and comments, the feedback is sent back to the authors, who can reflect on their own works, and revise their works based on the peers' suggestions.

Multiple peer feedback

Some studies questioned that there are situations in which the quality of assessing is impaired owing to the assessors' lack of knowledge or incorrect judgments (Lin et al., 2001；Cho & Schunn, 2007；Tsai et al., 2001). In order to reduce the impact of peer assessments weakness, multiple peer feedback is stressed in all systems. Multiple peer feedback means that every piece of work is assessed by multiple peers. Besides the aspect of fairness and possible weakness of assessments, students can better understand the readers' point of view through getting more than only one feedback.

Assessor support

Another approach for improving the fairness and accuracy of the assessments' feedback is implemented in SWoRD (Cho & Schunn, 2007) through back-evaluations about the helpfulness of assessors' feedback. In the process of assessing, assessors also have to learn how to give constructive comments on peers' works and therefore should receive feedback on the quality of their assessments. SWoRD uses assessment accuracy indices which measure each assessor's overall accuracy of their ratings and generate feedback to assessors on their assessment accuracies, compute assessors' grades, and finally weight the contribution of each assessment to the performance scores. Vee Heuristic (Tsai et al., 2001), NetPeas (Lin et al., 2001), CAP (Davies, 2000) and Yeh's system (Yeh, 2001) also establish a mechanism to explore students' or the instructor's views about the feedback from the assessors, and the assessors are provided with an evaluation on their assessing works by the instructor, which aims at encouraging students to provide useful comments to their peers and help them to improve their assessing skills.

Multidimensional evaluation

Students think that a holistic peer feedback without concrete evaluation criteria is often too vague or useless for doing modification and self-improvement (Lin et al., 2001). So, to make the peers' assessments more valid, multiple-specific feedback has been introduced in several systems (Lin et al., 2001；Cho & Schunn, 2007；Yang et al., 2005；Sung et al., 2005；Tsai et al., 2001). Multidimensional evaluation means that there are some criteria for evaluating. Take Yang's system for example, which is used in a writing course, the assessment criteria for narrative mode are "Elegant words," "Clear paragraph," "Coherence," "Title consistent," "New and original" (Yang et al., 2005). By using multidimensional evaluation, the assessors can give grades and comments based on the criteria for evaluation, which provides authors with more detailed feedback and makes it easier for them to understand how to improve their works.

Verbal comments

Moreover, all investigated systems stress that students do not only see the grades but also get some more specific comments and suggestions on their works. They propose that assessors can not only give a grade to evaluate peers'

works but also need to give some verbal suggestions to justify their assessment. Taking Vee Heuristic (Tsai et al., 2001), which is used in a teacher preparation course, as an example, in addition to give a score to every designed classroom activity, the students have to provide some qualitative comments responding to their quantitative evaluation to his or her peers.

Instructors actions

Although peers take the important roles in the assessments, some instructions of the teacher are necessary while executing the system. Most of the researchers recommend that before using the system, the instructor should clearly inform students about the educational goals, give them some guidelines, and specify the procedures and policies such as the due dates. In this way, students can fully understand the mechanism, which is a requirement for successfully using peer assessment (Cho & Schunn, 2007；Liu et al., 2002).

The Application Model

Every system has its appropriate application model, according to its characteristics and target users. But roughly, web-base peer assessment procedures comprise of the following steps:

1. After registering the course, the instructor clearly manifests the educational goals of peer assessment and the whole procedures, and set the topics for works and some policies, such as the penalty for late submissions.
2. The instructor provides course materials which is necessary to complete the assignments.
3. The instructor discusses the assignments with students and gives them some guidelines.
4. The assignments completed by students are uploaded to the system.
5. The system randomly assigns assessors for each work (mostly each assignment is assigned to multiple assessors).
6. Assessors evaluate the assigned works and give some comment on it to justify the assessment.
7. The system returns the assessment back to the original authors.
8. Based on the comments, students make corrections and modifications.

However, some of the investigated systems use an extended procedure which is described in the following paragraphs in more detail.

Because SWoRD (Cho & Schunn, 2007) uses assessment accuracy indices, before sending back the evaluation to the authors, SWoRD automatically determines the accuracy of each assessor's numerical ratings using three scales and provides the assessors with the feedback. Furthermore, an additional step is included with respect to providing authors the possibility to provide feedback to the usefulness of the assessors' comments. While seeing the evaluation and comments from the assessors, the authors can provide back-assessment with respect to their assessors' feedback in terms of how helpful the verbal feedback (not the numerical rating) is. So, the system allows not only that the assessors generate feedback to the authors, but also that the authors give some evaluation on the feedback they got.
Web-SPA seeks a balance between the thoroughness of evaluating works and the economy of time allocation (Sung et al., 2005) and uses an eclectic approach for providing students with good and bad examples of assignments. First, the teacher randomly groups the students, and let the group members evaluate the works of each other, as described in the general application model. Additionally to the general application model, this system emphasizes on letting students assess the best and poorest works, so that they get a sense of quality and awareness about making judgment.

The Target Users

The web-based peer assessment systems are gradually used in education, ranging from universities, junior high schools, to elementary schools, etc. About the target users of the systems explored in this study, some of them are applied in computer science course. NetPeas (Liu et al., 2002) was used at a university, and students are asked to do some writing assignments to verify their understanding of operating systems. Web-SPA (Sung et al., 2005) is used

in a junior high school and the participants are required to design multimedia web-pages. GSS (Kwok & Ma, 1999) educate students to become Information Systems professionals through the application of knowledge and skills required by the course to a specific business problem. And OASYS (Bhalerao & Ward, 2001) is used for students taking a programming class at university.

Besides computer science, SWoRD (Cho & Schunn, 2007) and Yang's system (Yang et al., 2005) are both used to improve writing performance in the native language through peer-assessing, where students assess peers' works and can make some reflection about their own work. Most of the assignments submitted in these two systems require students' imagination, creativity, and thinking skills. However, with respect to writing performance, the authors easily have blind spots upon doing this kind of work or the works are limited by their own creativity. Since a problem of students in writing is the awareness of its audience, peer assessment can especially help in this matter.
Furthermore, also other subjects are addressed through web-based peer assessment systems. Vee heuristic (Tsai et al., 2001) facilitates the development of inquiry-oriented activities for secondary science education. SPARK (Freeman & Mckenzie, 2002) aims at improving team work. Therefore, through peer assessment university students are asked to evaluate their peers about their contribution in team work. CAP (Davies, 2000) is also used at university level and focuses on the issue of plagiarism and asks assessors to examine whether their peers' work copies some materials from internet or not. Yeh's system (Yeh, 2001) is designed in a subject independent way, supporting junior high school students to create their own questions and a peer-review procedure is applied to evaluate others' questions.

The Role of Instructors and Students

The web-based peer assessment systems obviously are integrated into a course. In this section, we discuss what kind of roles instructors and students going to take when using web-based peer assessment systems.

In conventional classroom assessments teachers usually play a major role. Students work on an assigned topic within a limited period of time according to the teacher's guidance. Therefore, students generally receive feedback only from the teacher. Executing the web-based peer assessment, students partially undertake the teachers' role as evaluators and feedback providers in addition to their conventional role as learners (Lin et al., 2001). Students must put more effort on works than in a normal setting, and through the mechanism, it is possible to make students communicate and collaborate rather than just working alone. However, there are situations in which that kind of role causes drawbacks. Some students disliked peer assessment because raters were also competitors, and that can affect their impartiality on evaluations (Liu et al., 2002). Furthermore, web-technology can provide students with more opportunities of peer interaction beyond the constraints from time and locations since the activities can be conducted either within the classroom or in after-class situations (Sung et al., 2005).

Conclusions

In this paper, we showed how web-based peer assessment systems were developed and applied to assist in assessing peers' works in several school/university activities. The paper aimed at providing an overview of web-based peer assessment systems, discussing their characteristics, differences and similarities. Furthermore, the target users and the changing role of teachers and students were discussed.

By applying peer assessment, students are encouraged to deeply reflect on peers' work, evaluate it, and provide their peers with useful comments. While technology has improved the traditional way of peer assessment already significantly, there are still more possibilities to use technology in order to enhance the effectiveness, convenience, and success of peer assessment. For example, the quality of students' feedback has still potential to be improved by the use of technology. For example, more work can be done on incorporating students' individual characteristics and abilities such as proposed, implemented, and evaluated by Lan et al. (2007) with respect to learning styles. In addition, future work can deal with enhancing the validity and reliability of web-based peer assessment. Another direction of future research can be to focus on how to assist students in providing useful feedback in the process of peer assessing. On the one hand, this can be done from a pedagogical point of view through giving them some

guidance how to assess their peers work and teach them how to do assessment. Another possibility would be to implement an agent-based environment, where agents can detect students' difficulties in assessing peers' work and provide useful hints and guidance such as helpful information or a reference to similar assessments.

References

Bhalerao, A., & Ward, A. (2001). Towards electronically assisted peer assessment: a case study. *Alt-J| Association for Learning Technology Journal, 9*(1), 26-37.

Cho, K., & Schunn, C. D. (2007). Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer assessment system. *Computers & Education, 48*(3), 409-426.

Chung Hsien Lan, Sabine Graf, K. Robert Lai, and Kinshuk (2007). Effect of Learning Styles on Peer Assessment in an Agent-based Collaborative Learning Environment, Proceedings of the IEEE International Conference on Advanced Learning Technologies (ICALT 2007), Niigata, Japan, July 2007, pp. 421-423.

Davies, P. (2000). Computerized Peer Assessment. *Innovations in Education and Teaching International, 37*(4), 346-355.

Freeman, M., & McKenzie, J. (2002). SPARK, a confidential web-based template for self and peer assessment of student teamwork: benefits of evaluating across different subjects. *British Journal of Educational Technology, 33*(5), 551-569.

Kwok, R. C. W., & Ma, J. (1999). Use of a group support system for collaborative assessment. *Computers & Education, 32*(2), 109-125.

Lin, S. S. J., Liu, E. Z. F., & Yuan, S. M. (2001). Web-based peer assessment: feedback for students with various thinking-styles. *Journal of Computer Assisted Learning, 17*(4), 420-432.

Liu, E. Z. F., Lin, S. S. J., & Yuan, S. M. (2002). Alternatives to Instructor Assessment: A Case Study of Comparing Self and Peer Assessment with Instructor Assessment under a Networked Innovative Assessment Procedures. *International Journal of Instructional Media, 29*(4), 395-405.

Sung, Y. T., Chang, K. E., Chiou, S. K., & Hou, H. T. (2005). The design and application of a web-based self-and peer-assessment system. *Computers & Education, 45*(2), 187-202.

Tsai, C. C., Lin, S. S. J., & Yuan, S. M. (2002). Developing science activities through a networked peer assessment system. *Computers & Education, 38*(1-3), 241-252.

Tsai, C. C., Liu, E. Z. F., Lin, S., & Yuan, S. M. (2001). A Networked Peer Assessment System Based on a Vee Heuristic. *Innovations in Education & Teaching International, 38*(3), 220-230.

Wen, M. L., & Tsai, C. C. (2006). University Students' Perceptions of and Attitudes Toward (Online) Peer Assessment. *Higher Education, 51*(1), 27-44.

Yang, J. C., Ko, H. W., & Chung, I. L. (2005) Web-based Interactive Writing Environment: Development and Evaluation. *Educational Technology & Society, 8*(2), 214-229.

Yeh, C. C. (2001). A System to Support Designing and Peer-Assessing Activities of Items for Students.