



PERGAMON

Computers & Education 38 (2002) 241–252

**COMPUTERS &  
EDUCATION**

www.elsevier.com/locate/compedu

## Developing science activities through a networked peer assessment system

Chin-Chung Tsai<sup>a,\*</sup>, Sunny S.J. Lin<sup>a</sup>, Shyan-Ming Yuan<sup>b</sup>

<sup>a</sup>*Center for Teacher Education and Institute of Education, National Chiao Tung University,  
1001 Ta Hsueh Road, Hsinchu, 300, Taiwan*

<sup>b</sup>*Department of Computer and Information Science, National Chiao Tung University,  
Hsinchu, 300, Taiwan*

Accepted 13 November 2001

---

### Abstract

This paper described the use of a networked peer assessment system to facilitate the development of inquiry-oriented activities for secondary science education. Twenty-four preservice teachers in Taiwan participated in this study and experienced a three-round peer assessment for developing science activities. The findings suggested that teachers tended to develop more creative, theoretically relevant, and practical science activities as a result of the networked peer assessment. However, the peers' evaluations were not highly consistent with experts' (e.g., university professors) grades. This study also revealed that students who offered detailed and constructive comments on reviewing and criticizing other peers' work might help them improve their own work, especially in the beginning stage of revising their original work. © 2002 Elsevier Science Ltd. All rights reserved.

*Keywords:* Peer assessment; Science education; Networked learning; Taiwan

---

### 1. Introduction

Teachers are often expected to design instructional activities that integrate theoretical knowledge and promote students' creative thinking. Recent reform in science education also emphasizes an open-ended approach to instruction and student-centered inquiry activities (Tobin, 1993; Treagust, Duit, & Fraser, 1996; Tsai, 1999, 2000, 2001a). This instruction style opposes traditional teaching strategies that often present science knowledge as final-version facts (Duschl, 1990). Research on teachers, however, has repeatedly revealed that many teachers still use a traditional, fact-based, and textbook-oriented approach to instruction (Weiss, 1993). Their adher-

---

\* Corresponding author. Tel.: +886-3573-1671; fax: +886-3573-8083.

*E-mail address:* cctsai@cc.nctu.edu.tw (C.-C. Tsai).

ence to traditional instruction may stem from the fact that teachers do not have effective ways of developing open-ended inquiry activities. Successful inquiry activities may need a careful consideration among students' existing concepts, problem-solving skills and theoretical knowledge in the field, and they may also require teachers' creativity. Hence, if teachers can collaborate, the activities they develop may become more creative, practical and theoretically solid.

Peer assessment can be viewed as a form of collaboration for teachers (Tsai, 2001b). Teachers can propose some original thoughts for the instructional activities, and ask peer teachers to give some comments and suggestions, and then these activities can be refined into a more satisfactory level. Peer assessment may be particularly useful for preservice teachers, because they may not have developed mature conceptual frameworks and pedagogical content knowledge about students, instruction and subject matter.

Recent development in computer technology also reveals a new approach to implementing peer assessment. Networked or WWW-based environments where students remotely share their ideas and criticisms may provide a way of processing peer assessment. That is, we believe that peer assessment is most effectively implemented by using a networked system as a medium and control center, thereby eliminating communication restrictions such as time and place. The anonymity of networked environments may also help to build up a more objective manner of judging peers' work.

Because of the fast spread of Internet, the implementation of peer assessment on campus computer network via various Internet applications (e.g. Email or WWW) has recently become possible in Taiwan. In spring 1998, the Networked Peer Assessment System (NetPeas; Chiu, Wang, & Yuan, 1998) was designed to support peer assessment (either teachers or students) over the Internet. The study reported here was an attempt of using networked peer assessment to help a group of preservice teachers develop science activities for secondary school students.

## **2. Research about peer assessment**

Peer assessment is a widely used teaching strategy in higher education for very different fields, from writing, teaching, business, science, engineering, to medicine (Falchikov, 1995; Freeman, 1995; Strachan & Wilcox, 1996). Topping (1998) has provided an overall review of past studies of peer assessment and found it to be a reliable and valid evaluation method.

Peer assessment in this study was a compulsory evaluation process for a course offered by a teacher education program in Taiwan (described later). At first, preservice teachers (or authors hereinafter) submitted assignments, that is, designing science activities, through a networked peer assessment system. Then, several anonymous peer reviewers were assigned to mutually assess peers' assignments and providing feedback. The scores and feedback were then sent back to the original authors. The authors had to revise the assignments based on peers' feedback. Thus, peer assessment was formative, anonymous, and asynchronous in nature. This peer assessment scheme has been modeled after the authentic journal publication process of an academic society (Roth, 1997; Rogoff, 1991). It should also be noted that in this study, each participant assumed the role of an author who handed in assignments as well as a reviewer who commented on peers' assignments. Thus participants constructed and refined knowledge through social interactions in a virtual community linked via the Internet.

In peer assessment, first, participants have more chances to view more peers' work than in customary teacher assessment settings (Lin, Liu, Chiu, & Yuan, 2001). They may get inspiration from concrete examples their peers offer. Second, when providing feedback, participants have more chances to engage in important cognitive activities, such as critical thinking (Graner, 1985) and may obtain a more accurate self-assessment (Topping, 1998). Third, receiving adequate feedback is correlated with more effective learning (Bangert-Drowns, Kulick, Kulick, & Morgan, 1991; Crooks, 1988; Kulik & Kulik, 1988). In receiving feedback, authors have more chances to execute reflective thinking (Falchikov, 1995). Also, receiving abundant and immediate peer feedback, some early errors can be detected and hints for further improvement may be easily obtained.

A pilot study to evaluate learning effects and students' perceptions about networked peer assessment was conducted (Lin, Liu, & Yuan, 2001). The result revealed that students not only performed better under networked peer assessment, but also displayed some important cognitive skills, such as critical thinking, planning, monitoring, and regulation. Students perceived peer assessment both as an effective learning process and an incentive to learn.

### **3. Research questions**

The main focus of this study was to investigate how a networked peer assessment system may help preservice teachers develop inquiry-oriented science activities for instruction. This study also concentrated on the correlation between peers' evaluation and experts' (e.g. university professors, experienced researchers) grading. The correlation may provide some hints about the validity of peer assessment. In addition, this study explored the relationship between preservice teachers' quality of reviewing peers' work and the performance of their own work. As previously proposed, providing feedback to peers may help the reviewers engage in higher-order cognitive activities, such as critical thinking. This analysis may suggest some ideas about how reviewing peers' work may help the authors develop their own science activities.

In sum, this study was conducted to explore the following three research questions:

1. Did the preservice teachers improve the quality of their assignment (designing science-oriented inquiry activity) as a result of peer assessment?
2. What were the relationships between the scores determined by the learning peers and those marked by experts?
3. What were the relationships between the preservice teachers' quality of commenting on other peers' assignments and the performance of their own assignment?

### **4. Methodology**

#### *4.1. Subjects*

The subjects of this study included 28 preservice teachers enrolled in a "science education" course offered by a national university in Taiwan. However, one student withdrew from the course and three students did not successfully submit all of their science activity designs for peer

assessment. Hence, their data were excluded from final analyses. Of the final sample of 24 preservice teachers most were college seniors or first-year students of relevant master’s programs and 11 of them were female.

These preservice teachers intended to become secondary school teachers of science, mathematics or engineering. In this course, these students were divided into the following three groups according to their specializations: science, mathematics and engineering. These groups included nine, ten, and nine preservice teachers, respectively. One course requirement was to ask these preservice teachers to design a science (or mathematics) activity suitable for secondary students to understand some mathematics, science or engineering concepts. The preservice teachers within the same group were asked to comment on their peers’ science activity design through a Vee heuristic networked peer assessment system (described later). Then, they were asked to revise their own science activity design after taking peers’ comments and suggestions. In other words, every preservice teacher acted as both an author and a reviewer throughout the peer assessment.

4.2. Vee heuristic WWW-based peer assessment system

Novak and Gowin (1984) proposed using a Vee graph as a heuristic strategy to explain the relationship between the structure of knowledge and the course of action in obtaining or expanding this structure. Preservice teachers can organize their own science activities through the Vee graph shown in Fig. 1. For example, a preservice teacher can click on “Focus Question” and then type in the major question related to his (or her) science activity (e.g. How to protect an egg if it is released from a height of 10 m?). The teacher can then click on the icons of “Theory”, “Principles” or “Concepts” to type in conceptual frameworks related to this science activity or

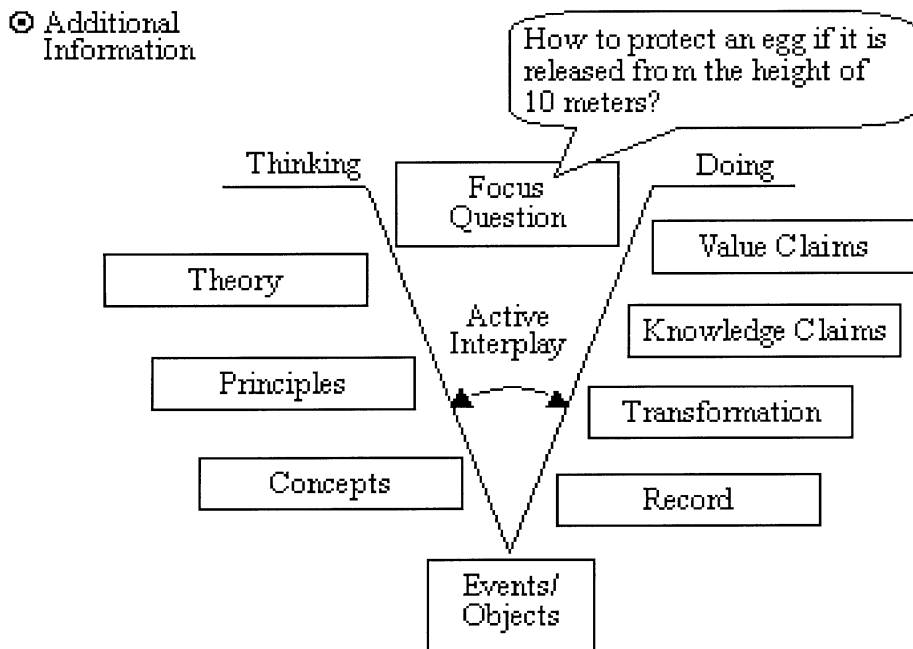


Fig. 1. Vee heuristic interface (adapted from Novak & Gowin, 1984).

related to the focus question (e.g. Newton's laws of motion, free fall motion, force, resistance). On the other hand, the methodological details of exploring the focus question are added to the icons of "Events/Objects", "Record" and "Transformation". The "Knowledge Claims" and "Value Claims" stemming from the focus question can also be added through this Vee interface. The "Additional Information" icon is designed for preservice teachers to describe some other important ideas related to this particular science activity. Hence, the Vee heuristic interface can easily display all relevant ideas of a science activity. Each science activity in this study is stored in this Vee-based format. Reviewers can also read each science activity design by simply clicking the icons of this interface (for details, please refer to Tsai, Liu, Lin, & Yuan, 2001).

The networked (or WWW-based) peer assessment system was performed by storing and retrieving DBMS's (Data Base Management System) information through a CGI program. These preservice teachers were asked to submit their science activity design over the Internet, which their peers read and then rated and commented on. These preservice teachers needed to modify their original assignments according to their peers' evaluations. Then, they were asked to submit revised science activity design on-line after taking peers' comments and suggestions. The peer assessment was conducted in three rounds, shown in Fig. 2. As shown, these preservice teachers needed to evaluate their peers' work three times, and they needed to revise their science activity design twice.

#### 4.3. Scores of peer assessment

Because the study subjects were a group of preservice teachers enrolled in a teacher preparation course, their performance that could be evaluated in quantitative terms might be quite important for the course and for the purpose of research analyses. Therefore, every preservice teacher's science activity design was quantitatively evaluated on the following three dimensions by his or her peers.

1. creativity: the extent to which the activity requires secondary students' creativity,
2. relevance: the extent to which the science activity is related to background knowledge in science or mathematics, and
3. feasibility: the extent to which the activity could be practically applied in real classrooms.

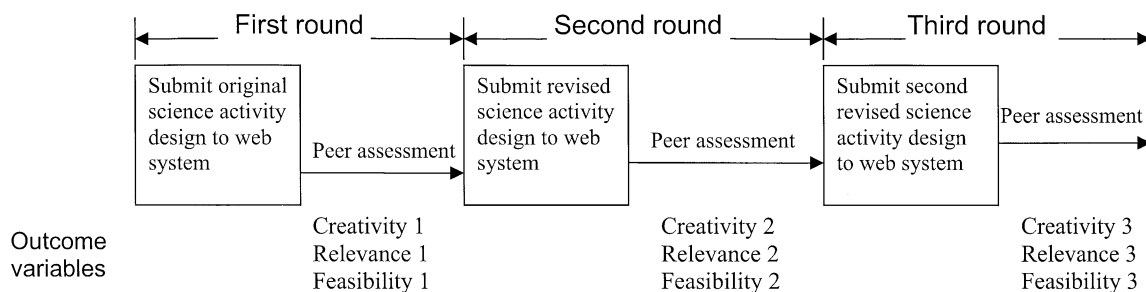


Fig. 2. The process of peer assessment used in this study.

The preservice teachers gave a score between 1 and 10 (with 1 point as unit) to every learning peer's science activity design on each dimension above. Such a scoring method was applied to the three-round peer assessment process, as shown in Fig. 2. This process produced nine outcome variables listed in Fig. 2. Each round of submitting activity design and peer assessment took about 3 weeks. The whole peer assessment in this study lasted about 2 months.

In addition to the quantitative evaluation above, every preservice teacher had to provide some qualitative comments to his or her peers. Their qualitative comments had to respond to their quantitative evaluation. For example, they needed to explain why he or she believed a piece of work should deserve for a lower grade for its creativity (say, 5 points) and another piece of work should deserve for a much higher grade (say, 10 points). A university professor and a research assistant examined the consistency between these subjects' quantitative grading and qualitative comments, and it was found that at least 85% of them were highly consistent. More importantly, these preservice teachers were encouraged to read their grades and comments reviewed by their peers and then to modify their original science activity design. On average, every student teacher had been commented on his or her assignment on each round by 6.68 peers (S.D. = 1.46), ranging from 5 to 9.

#### 4.4. Experts' assessment

Preservice teachers' science activity design was also evaluated by the following two experts: one was a university professor with a doctorate in science education and the other was a research assistant with a bachelor in physics and some years of research experiences in science education. These two experts evaluated preservice teachers' science activity design in all of the three rounds and then gave a score between 1 and 10 (with 0.5 as unit) on the dimensions of creativity, relevance and feasibility. Hence, experts also generated nine outcome variables for subjects' science activity design, as those created by their peers shown in Fig. 2. The agreement (Pearson's correlation coefficients) between two experts is presented in Table 1.

Table 1 shows that the correlation coefficients between two experts on each outcome variable ranged from 0.68 to 0.80. These two experts' scores on each outcome variable were significantly related. That is, these experts had similar perspectives in evaluating these teachers' work, indicating that experts' scores showed adequate reliability of assessing subjects' science activity

Table 1  
Pearson's correlation coefficients between two expert raters on every outcome variable ( $n = 24$ )

	Coefficient
Creativity (1st round)	0.70***
Creativity (2nd round)	0.68***
Creativity (3rd round)	0.78***
Relevance (1st round)	0.77***
Relevance (2nd round)	0.75***
Relevance (3rd round)	0.70***
Feasibility (1st round)	0.75***
Feasibility (2nd round)	0.79***
Feasibility (3rd round)	0.80***

design. Moreover, experts' scores were not revealed during the process of science activity design or peer assessment. That is, preservice teachers could only acquire their peers' comments and evaluation scores when designing and modifying their science activity design.

#### 4.5. Quality of comments

Preservice teachers' quality of comments was also considered as an important variable in the process of peer assessment. Preservice teachers' quality in making comments was also evaluated by the two experts described above. These experts read all of the (qualitative) comments throughout the whole process of peer assessment made by every individual. Then, the experts gave a score of between 1 and 10 (with 0.5 as unit) to represent the preservice teacher's quality of reviewing his or her peers' science activity design. Preservice teachers who provided more detailed, concrete and scientifically (or mathematically) sound suggestions on others' science activities would acquire higher scores. The correlation between two expert raters' scores on this variable was 0.64 ( $P < 0.01$ ).

## 5. Findings

### 5.1. Did the preservice teachers improve the quality of their science activity design as a result of the processes of peer assessment?

As described earlier, preservice teachers' science activity design was evaluated by their peers and two experts on the dimensions of creativity, relevance and feasibility. Table 2 shows preservice teachers' scores of developing science activity throughout the whole process of peer assessment.

Table 2 shows that these preservice teachers' average scores in the first round of networked peer assessment as evaluated by their peers were found to be 7.69, 8.04 and 7.87 on the dimensions of

Table 2  
Preservice teachers' scores of science activity design both from peers' and from experts' perspectives

Variable	Peer scores			Expert scores		
	Mean	S.D.	Range	Mean	S.D.	Range
Creativity 1	7.69	0.51	6.00–8.57	7.59	0.53	6.25–8.75
Creativity 2	7.75	0.63	5.57–9.14	7.71	0.48	6.50–8.75
Creativity 3	8.11	0.59	7.33–9.40	7.82	0.45	6.50–8.75
Relevance 1	8.04	0.41	7.25–8.75	7.58	0.48	6.75–8.75
Relevance 2	8.14	0.59	5.86–9.00	7.93	0.47	7.00–8.75
Relevance 3	8.50	0.58	7.40–10.00	7.99	0.49	7.00–9.25
Feasibility 1	7.87	0.57	6.57–9.00	7.33	0.64	6.00–8.50
Feasibility 2	8.10	0.66	5.57–8.83	7.67	0.51	6.25–8.75
Feasibility 3	8.35	0.63	7.40–10.00	7.80	0.58	6.50–9.25

creativity, relevance and feasibility respectively. The same science activities that were rated by experts gained the average scores of 7.59, 7.58 and 7.33 on the above three dimensions respectively. However, the corresponding scores on the third round of peer assessment were 8.11, 8.50 and 8.35 as evaluated by peers, and 7.82, 7.99 and 7.80 as assessed by experts. An examination of the mean of subjects' scores on each dimension revealed that both from peers' and from experts' perspectives, these preservice teachers had an increasing average score on each dimension. A series of paired *t*-tests were further used to compare student score changes as a result of the networked peer assessment. The results are presented in Fig. 3.

Although preservice teachers' scores did not necessarily significantly increase in any two consecutive rounds (e.g. in the light of experts' evaluation, preservice teachers' scores on creativity did not statistically differ between the first round and the second round, nor did between the second round and the third round), their scores on each dimension did make a significant progress when comparing those of the final round with those of the first round ( $P < 0.01$ ). It is also important that such significant score increases could be observed both from peers' and from experts' perspectives. In other words, these preservice teachers, in the aspects of creativity, relevancy, and feasibility, improved the quality of their science activity design through the networked peer assessment system. We believe that the quantitative evaluation might help preservice teachers to feel dissatisfied with their original work (especially in the first round of the peer assessment), and their peers' qualitative comments provided constructive suggestions for further modification. Altogether, their work was significantly enhanced.

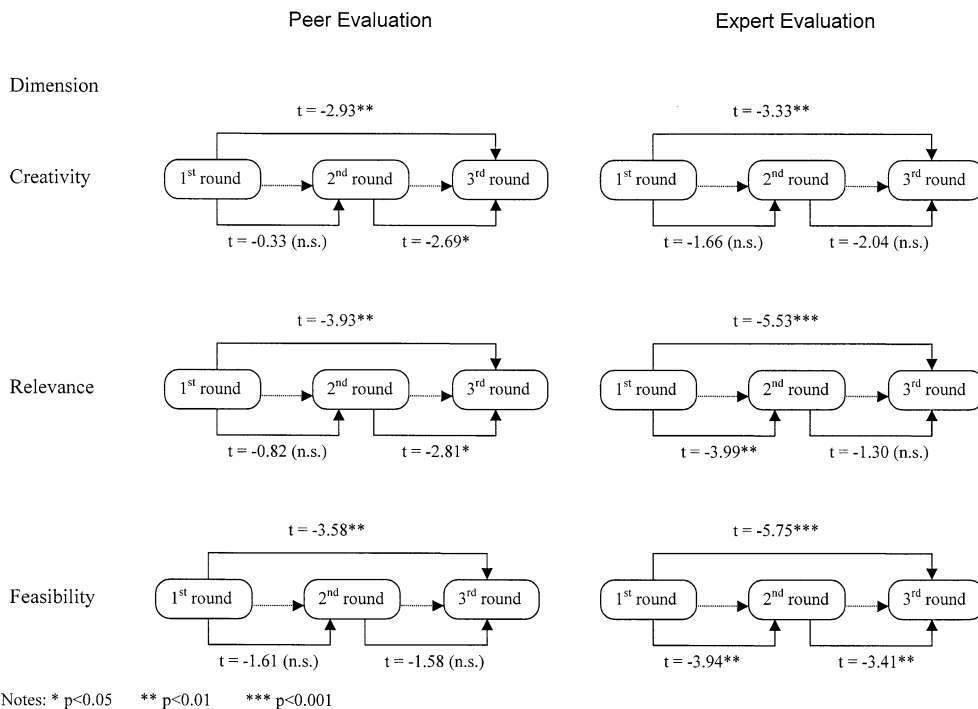


Fig. 3. Paired *t* tests on preservice teachers' score changes through networked peer assessment rounds.



The current study asked preservice teachers to engage in a three-round peer assessment; however, as in the case of Lin, Liu, and Yuan (2001), computer majors mutually assessed peers' work for only two rounds. They found still many students did not improve over two rounds. The present study also showed similar findings. For instance, based on peers' evaluations, these preservice teachers' scores on creativity, relevance and feasibility were not significantly different between the first round and the second round. From the results of current study and Lin, Liu, and Yuan (2001), it is very likely that improvement of peer assessment can be more easily observed through three rounds. Thus, effects of peer assessment take time to be fully revealed.

### 5.2. Were the scores determined by the learning peers correlated with those marked by experts?

Table 3 presents the correlation coefficients between peers' scores and experts' scores on every outcome variable. Although the correlation is positive for all of these variables, only two of them (i.e. creativity 3 and feasibility 1) reveal a statistically significant correlation. That is, peers and experts may share some agreement of evaluating the science activity design; however, such an agreement was not sufficiently high. This implies that peers' evaluation may not always produce a valid indicator of representing students' performance. It deserves to know that Topping (1998) reviewed the application of peer assessment in various subjects of higher education from 1980 to 1996. Twenty-five of the 31 papers conferred on a high correlation between peer assessment and teacher grading.

It is likely that the process of peer assessment may involve complicated interactions among peers' background knowledge, their preferences, and interrelations of others. These interactions may reduce the validity of using peer assessment. For example, Lin, Liu, and Yuan (2001) reported several limitations of peer assessment. Participants regarded that (1) peer assessment was time and effort consuming, (2) peers might not have adequate knowledge to evaluate others' assignments, (3) although peer assessment was anonymous, students still hesitated to criticize their peers, (4) peers' feedback was often ambiguous or not relevant for further modification and (5) some students tended to give extremely low scores to others to keep his or her achievement at a relatively high level. All of these limitations may reduce the validity of peer assessment.

It should also be noted that, although this study did not show a high consistency between peers' evaluations and experts' grades, these preservice teachers did improve their science activity design by reading peers' quantitative marks and qualitative comments, as shown in the analyses of the previous section. Educators can not ignore the important role of peer assessment on facilitating these teachers' development of better instructional activities.

Table 3

The correlation coefficient between expert and peer scores on every outcome variable

Creativity 1	Creativity 2	Creativity 3	Relevance 1	Relevance 2	Relevance 3	Feasibility 1	Feasibility 2	Feasibility 3
0.16	0.33	0.55**	0.16	0.08	0.13	0.45*	0.37	0.32

\*  $P < 0.05$ .

\*\*  $P < 0.01$ .

### 5.3. Were the preservice teachers' quality of reviewing other peers' science activities related to the performance of their own science activity design?

Table 4 presents the relationships between preservice teachers' quality of comments and their performance of science activity design. The first part of the table shows the correlation coefficients between preservice teachers' scores of science activity design as viewed by their peers and their quality of comments (marked by experts). The second part of the table presents the coefficients between preservice teachers' science activity design scores as determined by experts and their quality of comments.

Preservice teachers' quality of comments was more likely correlated with their performance of science activity design (from either peers' or experts' perspectives) of the second round and third round, but not of the first round. This finding is plausible due to the fact that students did not comment any other's science activity when submitting the first-round activity design. That is, the performance of the first round was mainly determined from the preservice teachers' original ideas, which was independent of the peer assessment process. However, in the second round and third round, preservice teachers who provided more detailed and constructive comments on reviewing others' science activity design, in many cases, tended to design better science activities. In particular, preservice teachers' quality of reviewing others' work was related more to the second round performance ( $r=0.43$ ,  $0.34$  and  $0.45$  for the three dimensions from peers' perspectives, and  $r=0.48$ ,  $0.54$  and  $0.38$  for the dimensions from experts' perspectives). The second round's activity design was produced by preservice teachers' first attempt to modify their initial design according to peers' comments. This implies that students who spend more efforts on reviewing and criticizing other peers' work could help them acquire better and richer ideas in improving their own work, especially in the beginning stage of revising their original work.

Table 4

The relationships between preservice teachers' quality of comments and their own science activity design

	Creativity 1 <sup>a</sup>	Creativity 2	Creativity 3	Relevance 1 <sup>a</sup>	Relevance 2	Relevance 3	Feasibility 1 <sup>a</sup>	Feasibility 2	Feasibility 3
(1) Outcome variables (Creativity, Relevance and Feasibility) determined by peers									
Quality of comments	0.22	0.43*	0.28	0.20	0.34	0.05	0.04	0.45*	-0.09
(2) Outcome variables (Creativity, Relevance and Feasibility) determined by experts									
Quality of comments	0.26	0.48*	0.51*	0.12	0.54**	0.49*	0.05	0.38	0.32

<sup>a</sup> This correlation may not be worth interpreting, as the performance on the outcome variable was mainly determined by the preservice teachers' original ideas, which was independent of the peer assessment process.

\*  $P < 0.05$ .

\*\*  $P < 0.01$ .

## 6. Implications

This study revealed that preservice teachers could improve their work through a networked peer assessment system. It is hypothesized that participants may benefit more from peer assessment when they are required to complete a task in an open-ended task format, for instance, designing science activity used in this study. Also, we found that peers' assessment may not be necessarily coherent with experts' rating, though we rather suspect the small sample size ( $n = 24$ ) caused the difficulty of showing the real effects. However, future implementation of peer assessment may need more careful management to ensure its effect. For example, Topping (1998) stressed announcing and clarifying goals of peer assessment to promote a trusting environment and recommended training assessors to question, prompt, and scaffold rather than merely supplying a simple corrective answer. We suggest teachers' closer monitor in peer assessment process. Teachers need to motivate students if anyone reluctant to hand in homework, to encourage comprehensive discussion about evaluation criteria or scoring standards before peer assessment taking place, and to intervene when any inadequate attack is found in feedback or marking. It is important to note that the sample is rather small in this study. More research work on analyzing the processes and conceptual development involved in networked peer assessment with a much larger sample (or even different countries and cultural backgrounds) is necessary to reach more definitive conclusions.

In sum, this study provided initial evidence that networked or WWW-based environments may provide a potential avenue for participants to share others' ideas, criticize others' work and then achieve the goal of peer assessment. Internet-based environments that free the time and space constraints could allow participants remotely cooperatively to complete relevant work (either synchronously or asynchronously). The environments could also record all of students' ideas and comments throughout the whole processes of peer assessment, allowing participants to review their own ideas and relevant comments and allowing researchers to document participants' knowledge and ideational development.

Although the networked peer assessment system studied here was limited to preservice teachers, practicing teachers can also implement it. Through the system, teachers may develop their instructional activities in a more efficient way. Certainly, when practicing teachers use this system the qualitative comments may be more useful than the quantitative evaluation. For practicing teachers, peer assessment can be better viewed as peer cooperation, facilitating teachers' development of relevant activities through sharing and providing experienced suggestions. Some attempts that gathered a group of practicing science teachers of different schools to collaboratively develop instructional activities through the networked system have just begun (e.g. Tsai, 2001a). These attempts may enlarge the scope of the present study, and illuminate more potential for the improvement of science education.

## Acknowledgements

The funding of this research was supported by the National Science Council, Taiwan, under Contract Nos. NSC88-2511-S-009-006, NSC 89-2520-S-009-004 and NSC 89-2520-S-009-004. This paper was presented at the Computer Assisted Learning Conference, University of Warwick, Coventry, UK, April, 2001.

## References

- Bangert-Drowns, R. L., Kulick, C. L. C., Kulick, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61, 213–238.
- Chiu, C. H., Wang, W. R., & Yuan, S. M. (1998). Web-based collaborative homework review system. *Proceedings of ICCE'98*, 2, 474–477.
- Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58, 45–56.
- Duschl, R. A. (1990). *Restructuring science education*. New York: Teachers College Press.
- Falchikov, N. (1995). Peer feedback marking: developing peer assessment. *Innovations in Education & Training International*, 32, 175–187.
- Freeman, M. (1995). Peer assessment by groups of group work. *Assessment & Evaluation in Higher Education*, 20, 289–300.
- Graner, M. H. (1985). Revision techniques: peer editing and the revision workshop. *Dissertation Abstracts International*, 47, 109.
- Kulik, J. A., & Kulik, C. L. C. (1988). Timing of feedback and verbal learning. *Review of Educational Research*, 58, 79–97.
- Lin, S. S. J., Liu, E. Z., Chiu, C. H., & Yuan, S.-M. (2001). Peer review: an effective web-learning strategy with the learner as both adapter and reviewer. *IEEE Transactions on Education*, 44, 346–251.
- Lin, S. S. J., Liu, E. Z., & Yuan, S.-M. (2001). Learning effects of Web-based peer assessment for students with various thinking styles. *Journal of Computer Assisted Learning*, 17, 420–432.
- Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. Cambridge: Cambridge University Press.
- Strachan, I. B., & Wilcox, S. (1996). Peer and self assessment of group work: developing an effective response to increased enrollment in a third year course in microclimatology. *Journal of Geography in Higher Education*, 20(3), 343–353.
- Roth, W.-M. (1997). From everyday science to science education: how science and technology studies inspired curriculum design and classroom research. *Science & Education*, 6, 296–372.
- Rogoff, B. (1991). Social interaction as apprenticeship in thinking: guidance and participation in spatial planning. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 349–383). Washington, DC: APA.
- Tobin, K. (1993). *The practice of constructivism in science education*. Washington, DC: AAAS.
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68, 249–276.
- Treagust, D. F., Duit, R., & Fraser, B. J. (1996). *Improving teaching and learning in science and mathematics*. New York: Teachers College Press.
- Tsai, C.-C. (1999). “Laboratory exercises help me memorize the scientific truths”: a study of eighth graders’ scientific epistemological views and learning in laboratory activities. *Science Education*, 83, 654–674.
- Tsai, C.-C. (2000). Relationships between student scientific epistemological beliefs and perceptions of constructivist learning environments. *Educational Research*, 42, 193–205.
- Tsai, C.-C. (2001a). Collaboratively developing instructional activities of conceptual change through the Internet: science teachers’ perspectives. *British Journal of Educational Technology*, 32(5), 619–622.
- Tsai, C.-C. (2001b). The interpretation construction design model for teaching science and its applications to Internet-based instruction in Taiwan. *International Journal of Educational Development*, 21, 401–415.
- Tsai, C.-C., Liu, E. Z.-F., Lin, S. S. J., & Yuan, S.-M. (2001). A V-based network peer assessment system: student design of science activity supported by a Vee-heuristic. *Innovations in Education and Training International*, 38, 220–230.
- Weiss, I. R. (1993). Science teachers rely on the textbook. In R. E. Yager (Ed.), *What research says to the science teacher: The science, technology, society movement* (pp. 35–41). Washington, DC: National Science Teachers Association.