The development of science activities via on-line peer assessment: the role of scientific epistemological views

Chin-Chung Tsai · Jyh-Chong Liang

Abstract   This study implemented an online peer assessment learning module to help 36 college students with the major of pre-school education to develop science activities for future instruction. Each student was asked to submit a science activity project for pre-school children, and then experienced three rounds of peer assessment. The effects of the online peer assessment module on student learning were examined, and the role of Scientific Epistemological Views (SEVs) in the learning process was carefully investigated. This study found that student peers displayed valid scoring that was consistent with an expert's marks. Through the online peer assessment, the students could enhance the design of science activities for future instruction; for instance, the science activities became more creative, science-embedded, feasible and more suitable for the developmental stage of pre-school children. More importantly, students with more sophisticated (constructivist-oriented) SEVs tended to progress significantly more for designing science activities with more fun, higher creativity and greater relevancy to scientific knowledge, implying that learners with constructivist-oriented SEVs might benefit more from the online peer assessment learning process. These students also tended to offer more feedback to their peers, and much of the peer feedback provided by these students was categorized as guiding or helping peers to carefully appraise and plan their science activity projects. This study finally suggested that an appropriate understanding regarding the constructivist epistemology may be a prerequisite for utilizing peer assessment learning activities in science education.

Keywords   Peer assessment · Scientific epistemological views · Online ·
Web · Science education · Pre-school · Constructivism

C.-C. Tsai (✉)
Graduate School of Technological and Vocational Education, National Taiwan University of Science and Technology, #43, Sec.4, Keelung Rd., Taipei 106, Taiwan, ROC
e-mail: cctsai@mail.ntust.edu.tw

J.-C. Liang
Department of Holistic Wellness, Chin Min Institute of Technology, No.110, Syuefu Rd, Toufen, Miaoli County 351, Taiwan

Introduction

Peer assessment has been increasingly implemented in educational settings, especially in higher education (Topping 1998; van den Berg et al. 2006). From a practical perspective, peer assessment can be used as an alternative method of assessment, and, to a certain extent, it can reduce the instructor's load for grading student work. From a learning perspective, relevant research has concluded that the use of peer assessment can help students achieve better learning outcomes, identify their own strengths and weaknesses, and enhance their problem-solving and metacognitive abilities involved in learning processes (e.g., Smith et al. 2002; Tsai et al. 2002). As peer assessment highlights peer interactions and feedback for learning, some educators believe that its practice is based on the theory of social constructivism (Lin et al. 2001; Yu et al. 2005). Although there are some concerns about peer assessment, for instance, the validity of peer grading (Cho et al. 2006; Falchikov and Goldfinch 2000), peer assessment has its positive impacts on learning (e.g., Barak and Rafaeli 2004; Venables Summit 2003) and it has been widely used by teachers in a variety of fields (Topping 1998).

Recently, online technology or Internet provides a potential avenue of executing peer assessment with enhanced effectiveness. Several studies regarding online peer assessment have clearly shown its advantages (e.g., Barak and Rafaeli 2004; Cho and Schunn 2007; Davies 2000; Tsai et al. 2001). For example, online peer assessment environments provide greater freedom of time and location for learners, and efficiently enhance peer interaction and feedback. The online environments, if used properly, may ensure a higher degree of anonymity and timely submission than traditional paper-and-pencil peer assessment. Moreover, the online environment or system can have sufficient capacity to record more thorough data about peer interactions and feedback, assisting teachers as well as researchers to acquire more information or electronic learning profiles from peer learners. Few studies related to peer assessment analyzed the peer feedback in great details, and deeply explore how different students may provide their comments to peers. By implementing an online peer assessment system, this study was an initial attempt to explore this.

Many will agree that science education should begin from early stage of life, as science requires observations, classifications, and manipulations, some basic skills developed even from infants. There is no doubt that developing certain science concepts or skills is quite important for pre-school pupils. Therefore, teaching science in pre-school stage is quite important, and developing suitable science activities for pre-school learners is critical for the success of early childhood science education. The sample of this study included a group of students with a major in early childhood and pre-school education. They were all preservice pre-school teachers. Therefore, these students were asked to take some science courses and they were also enrolled in a course for developing science activity (or game) for pre-school children. In this course, we utilized an on-line peer assessment system to help these students develop appropriate science activities for pre-school learners.

In recent decades, science education literatures have documented the importance of Scientific Epistemological Views (SEVs) for science teaching and learning (Lederman 1992; Songer and Linn 1991; Southerland et al. 2001; Tsai 1998, 2007; Windschitl and Andre 1998). The SEVs refer to people's views regarding the nature of science, including the assumptions, sources, certainty, justifications, consensus making and conceptual developments in science (Ryan and Aikenhead 1992). Research also indicated that there were basically two poles of SEVs held by students. One pole is constructivist-oriented, asserting that scientific knowledge is an invented (hence, tentative) reality, which is developed by agreed theories, shared forms of evidence and social negotiations in science

community. The other one is positivist-oriented, claiming that scientific knowledge comes from completely objective observations, totally neutral discoveries and interpretations, implying a relatively certain status about science. Students' SEVs can be roughly characterized by using a continuum between these two poles of different SEV aspects (Tsai 1999a; Tsai and Liu 2005). Relevant studies, in general, supported that students with constructivist-oriented SEVs tended to develop more integrated knowledge structures and employ more meaningful learning strategies. In particular, students with constructivist-oriented SEVs tended to perform better in peer-cooperative small-group or so-called "social constructivist" learning environments (such as laboratory) than those with positivist SEVs (Tsai 1999b, 2000; Wallace et al. 2003). Learning environments for offering student opportunities of peer assessment can also be perceived as those encouraging learner cooperation, as peer members are requested to evaluate peer learners' work and offer helpful suggestions to them. Therefore, this study hypothesized that student SEVs would be related to their learning in the development of science activities via the peer assessment process, and tried to gather empirical data to examine this hypothesis.

In addition, previous studies also found that students' SEVs would shape their perceptions toward the nature of learning tasks and learning environments (Tsai 1999b, 2000; Wallace et al. 2003). Requesting students to grade or comment on peers' work may be quite a different task from traditional learning they experienced before, and online peer assessment learning environments are likely innovative ones for them. Students' SEVs may guide their perceptions toward the nature of learning involved in peer assessment as well as their views about the online learning environments, and then influence their learning or progression in such instructional context. Certainly, the development of adequate science activities requires an understanding of relevant scientific knowledge, process and skill; therefore, the role of students' SEVs, that is, their views toward the nature of science, may be particularly important for designing and evaluating these activities. All of arguments above strengthen the need of exploring the relationships between students' SEVs and their learning involved in online peer assessment of developing science activities. Almost none previous research has tried to explore how students' epistemological views may play a role in their learning involved in peer assessment.

In sum, the implementation of peer assessment has been viewed as effective for enhancing student learning, especially with the assistance of online technology. This study set up online peer assessment system to help students (preservice teachers) with the major of pre-school education to develop science activities for pre-school children. Moreover, how students' SEVs may play a role in their gains and learning involved in on-line peer assessment processes would be investigated.

Research purposes

This study was conducted with the following research purposes. First, the validity of peer assessment, a fundamental issue for peer assessment research (Falchikov and Goldfinch 2000; Topping 1998), was examined by the correlations between the scores marked by peers and those graded by expert. Then, we would examine whether these students design more appropriate science activities for pre-school children along the implementation of on-line peer assessment. That is, the effect of the on-line peer assessment system on students' science activity development would be evaluated. As stated previously, student SEVs were considered as an essential factor for their development of science activities. This study explored the role of students' SEVs on their gains or development derived from the on-line

peer assessment process. Moreover, we would analyze the content of students' peer comments and explore the kinds and trends of peer feedback provided throughout different rounds of online peer assessment. Finally, this study examined the relationships between students' SEVs and their offering of peer comments. That is, we would investigate whether students of different SEVs provided different kinds of peer comments along the peer assessment learning activities. To state more specifically, the study, through gathering research data from 36 preservice pre-school teachers, explored the following five research questions:

- What are the correlations between peer scores and expert scores?
- What are the effects of utilizing online peer assessment on the development of science activities?
- What are the relationships between the participants' SEVs and their gains from the online peer assessment process?
- By a series of content analyses of peer comments recorded in the online system, what are the types and trends of peer feedback offered by the participants for different rounds of peer assessment?
- What are the relationships between the participants' SEVs and their generation of peer feedback?

Method

Participants

The participants of this study included 36 college students in Taiwan, who all majored in early childhood and pre-school education. All of them were preservice pre-school teachers, and enrolled in a course called "science activities for pre-school children." One course requirement was to ask these students to design a science activity suitable for pre-school learners (3–6 year-olds) to explore some science concepts or natural phenomena.

Online peer assessment module in the course

Each student was asked to design a science activity and then submit it to an on-line system for peer assessment. These students were, then, assigned to comment on their peers' science activity design also via the on-line peer assessment system. Each activity was assessed by five peers. (In other words, each student assessed five peers' work). Then, they were asked to revise their own science activity design after taking peers' comments and suggestions. The peer assessment was conducted in three rounds. That is, these students needed to evaluate their peers' work three times, and they needed to revise their science activity designs twice (that is, initial submission, first peer assessment, revision submission, second peer assessment, second revision submission, and third peer assessment). This online system and peer assessment procedures were used by some previous studies (e.g., Tseng and Tsai 2007; Wen and Tsai in press). The whole peer assessment process took about 2 months, and it was undertaken in an anonymous way. It is expected that the use of on-line system can facilitate greater anonymity and timely submission.

Peer and expert scores (marks)

For each peer assessment round, every student's science activity design was quantitatively evaluated on the following five dimensions by his or her peers.

1. Developmental suitableness (called "Development" later): the extent to which the science activity is developmentally suitable for 3–6-year-olds to play.
2. Science fun: the extent to which the activity leads to some fun about science.
3. Scientific relevancy: the extent to which the science activity is related to some background knowledge in science.
4. Creativity: the extent to which the activity requires pre-school students' creativity.
5. Feasibility: the extent to which the activity could be practically applied to real pre-school classrooms.

The five dimensions above were decided by three experts in pre-school education or science education. They collaboratively contributed to their ideas regarding "what counts as 'good' science activities for pre-school children." By expert agreement, they believed that these five dimensions were the most important for evaluating the activities. The students gave a score between 1 and 7 (with 1 point as unit) to every learning peer's science activity design on each dimension above, similar to that utilized by Tseng and Tsai (2007). The seven-point scale, rather than 1–100 scale, was employed, as it could more possibly avoid the situation of arbitrary scoring by the students. These scores can represent the quality of each student's science activity design for each peer assessment round. In a similar manner, the instructor of the course also marked each student's science activity design in each round by the same scoring method. These marks were viewed as expert's scores. In addition to quantitative marks, the students were requested to provide qualitative comments to each peer's work.

In order to fully understand how peer assessment would play a role for student learning, expert scores were not revealed during the peer assessment process. In other words, students could only acquire their peers' scores and comments when modifying their science education projects.

Instrument probing students' SEVs

To assess the students' SEVs, this study used the instrument developed by Tsai and Liu (2005), which suggested a multi-dimensional framework of representing student SEVs. By adopting multi-dimensional framework of SEVs, it was anticipated to describe students' different aspects of SEVs in more details. The five subscales (dimensions) of the instrument, with sample item provided, are shown in Table 1.

The dimensions of the instrument were based on the conceptual framework developed in previous studies (Tsai 1999a, 2002) with student interviews and experts' content validity. These dimensions were also verified by exploratory factor analysis (Tsai and Liu 2005; Liu and Tsai in press). These dimensions, in essence, embrace the issues of the epistemology of science proposed by Ryan and Aikenhead (1992) and Lederman et al. (2002), which mainly include the conceptual inventions in science, consensus making in scientific communities, and tentative and cultural-embedded features of scientific knowledge.

Table 1 The subscales of SEV questionnaire used in the study

Subscale	Description	Sample item(s)
Invented and creative reality of science	Assessing the extent of agreement toward the idea that science is created by human	The development of scientific theories requires scientists' imagination and creativity (constructivist-oriented view)
Theory-laden exploration	Investigating the extent of agreement toward the view that the exploration in science is theory-laden	Scientists' research activities will be affected by their existing theories (constructivist-oriented view)
		Scientists can make totally objective observations, which are not influenced by other factors (positivist-oriented view, scored in reverse)
Changing and tentative feature of science knowledge	Exploring the extent of agreement toward the perspective that scientific knowledge is always changing and its status is tentative	Contemporary scientific knowledge provides tentative explanations for natural phenomena (constructivist-oriented view)
Social negotiation in community	Assessing the extent of agreement toward the view that social negotiation in science community plays an important role in the development of scientific knowledge	Through the discussion and debates among scientists, the scientific theories become better (constructivist-oriented view)
Cultural impacts	Measuring the extent of agreement toward the idea that science is culture-dependent	Different cultural groups have different ways of gaining knowledge about nature (constructivist-oriented view)
		Scientific knowledge is the same in various cultures (positivist-oriented view, scored in reverse)

Each SEV subscale contained four to five items. All of the instrument items were presented in 1–5 Likert scale. Students' responses were scored below to represent their SEVs. For the constructivist-oriented perspective items (e.g., the sample items of the first four subscales), a "strongly agree" response was assigned a score of 5 and a "strongly disagree" response assigned a score of 1, whereas the items stated in a positivist-aligned view (e.g., the sample item of the last subscale) were scored in a reverse manner. Students having stronger beliefs regarding the constructivist view for a certain dimension (i.e., subscale) thus attained higher scores on the subscale; on the other hand, students with positivist-aligned SEVs for a certain subscale would have lower scores. Tsai and Liu (2005) reported the alpha coefficients of this SEV instrument to be around 0.70 for each subscale. The same instrument was used in another study in Taiwan for assessing Taiwan college students' SEVs with different majors (Liu and Tsai in press). The alpha coefficients calculated from the students in this study for each subscale were around 0.80, showing adequate reliability for representation of student SEVs.

Some science educators have expressed concerns for assessing students' SEVs by using instruments composed of forced-choice items such as agreement/disagreement or Likert-type, because of the reservations regarding the validity and the ways of labeling student SEVs as "adequate" or "inadequate" by total scores (e.g., Lederman et al. 2002). First, the SEV instrument used by this study was developed by a series of experts' justifications and

factor analyses and further validated by students' interviews or written responses (Tsai and Liu 2005; Liu and Tsai in press), ensuring its satisfactory validity. Moreover, in this study, we did not label students' SEVs by total questionnaire scores; rather, we used their scores on each subscale, representing their extent of agreement toward one specific SEV construct. Also, students were not labeled as "adequate" or "inadequate"; we used the characterization, such as "constructivist-oriented" or "relatively positivist" to more neutrally represent their SEVs in various dimensions. The use of this SEV instrument, clearly, was convenient and equipped to the quantitative analysis for this study.

Content analysis of peer feedback

As described previously, in addition to quantitative scores, assessors needed to provide qualitative comments or feedback via the online system to each student who designed the science activity. In order to acquire a better understanding about the role of peer feedback, all of the qualitative feedback given by students was categorized by the framework in Table 2. The same as the definition proposed by Tseng and Tsai (2007), a piece of feedback or comment was defined as the one which expressed a complete thought that might contain one or several sentences. The framework in Table 2 was modified from Guan et al. (2006) and Henri (1992). By the analysis framework, peer feedback was classified into three major dimensions: affective, cognitive and metacognitive. Affective dimension included students' praise or emotional responses toward peers' work. The cognitive comments consisted of the correction, the expression of personal opinion (without giving more information), and the guidance for peers' projects of science activities. The peer feedback classified into the metacognitive dimension, which was perceived as higher-order comments, helped peers to deeply evaluate, plan, regulate and reflect their own work. As shown in Table 2, each dimension included two or three specific categories. The definition and example of student feedback for each category are also illustrated in Table 2. For the metacognitive dimension, we combined "evaluating" and "planning," and also "regulating" and "reflecting" as individual category because the occurrence of these types of feedback was not very frequent.

By conducting the content analysis of peer feedback, we retrieved the comments provided by each participant from the online peer assessment database. First, we counted the total pieces of feedback for each round by the participant. For each piece of feedback, we categorized it first into one dimension (e.g., cognitive) and then one specific category under the dimension (e.g., direct correction, personal opinion).

In each round of peer assessment process, the frequency (i.e., the number of pieces of feedback) in which each student provided to their peers for each category of peer feedback was counted for analyses. For example, a student, in the first round of peer assessment, might give his/her peers ten pieces of comments categorized in "supporting," three in "emotional response," two in "direct correction," nine in "personal opinion," four in "guiding," eight in "evaluating and planning" and four in "regulating and reflecting." By this way of analysis, this study would explore the trends of peer feedback provided by the participating students. Moreover, each student's sum frequency of all rounds for each feedback category was utilized to relate to his/her SEVs to examine how his/her SEVs may play a role on the generation of peer feedback.

The categorization process was conducted by one of the authors and a panel of graduate students. The author analyzed all of the feedbacks. The online records, which thoroughly documented student peer comments, helped the completion of the content analysis. Students'

Table 2 The framework of analyzing peer feedback or comments

Dimension	Category	Definition	Example
Affective comments	Supporting	Statements referring to clear support and praise (i.e. showing praise, and supporting one's ideas)	• Good activity, suitable for pre-school children • Very creative and feasible science activity for children
	Emotional response	Simply express one's feeling about the work	• I don't like this activity plan • I like it!
Cognitive comments	Direct correction	Focus on the correctness of work	• This activity is too dangerous. You cannot do this for pre-school children • The concept involved in this activity is not scientifically correct
	Personal opinion	Offering general advice or personal perspectives without providing solid evidences or concrete directions.	• I think your design should be more complete • It would be better if you extend your activity
	Guiding	Questioning for direction, or introducing concepts, contents or frameworks	• I think this activity is quite difficult. Is it suitable for pre-school pupils? • This activity looks simple, but can pre-school children finish it by their own? • The activity should better introduce the concept of weight for pupils

Table 2 continued

Dimension	Category	Definition	Example
Metacognitive comments	Evaluating and planning	Evaluating: assessment, appraisal or verification of one's knowledge and skills, and the efficacy of a chosen strategy, for instance,	• I would suggest you to simplify your activity, and try to allow more opportunities for the children to have really hands-on experiences
		–Asking whether one's statement is true	• You can see if the pre-school like this part of the activity or not. And then, you can decide how to do for the rest part of the activity
		–Commenting on one's manner of accomplishing a task.	• I would suggest you to divide the activity into two major parts. Otherwise, the children will get confused
		–Raising doubts and querying	
		Planning: selecting, predicting, and ordering an action or strategy necessary to the accomplishment of an action, for example	
		–Predicting the consequences of an action	
		–Organizing aims by breaking them down into sub-objectives	
		–Providing clues to foster in-depth thinking	
	Regulating and reflecting	Regulating: setting up, maintenance and supervision of the overall cognitive task.	• You need to involve the introduction of more scientific knowledge for children. Please note that the main purpose of the activity is to acquire some science concepts for students
		Reflecting: challenging learners to think deeply and stimulating constructive reflection by learners as they conjecture, probe and contemplate.	• Do you really think whether your activity brings more science for children?
			• I think your ability is not so limited. You can have a better design. You can think more deeply about your activity and have a better deign

on-line feedback for the first round was selected for the graduate students to perform the same categorization process. In general, the agreement between coders' categorization is over 85%, indicating sufficient reliability for analyzing students' peer feedback.

Results and findings

The correlation between peer and expert scores

In this study, each student's science activity was scored by their peers between 1 and 7 points on the dimensions of development, science fun, scientific relevancy, creativity, and feasibility. Similarly, an expert, the course instructor, also evaluated students' work by the same method. It is important to examine whether student peers displayed valid scoring that was consistent with the expert's marks. Table 3 shows the correlation between expert and peer scores for each dimension of each peer assessment round.

According to Table 3, except the correlation for the "development" of the first round, all of the correlation was positively significant, indicating that expert and peer scores were statistically consistent. Peer scores in this study could be deemed as valid measurements. It was also interesting to find that the correlation coefficients between peer and expert scores have increased along the peer assessment learning module. For instance, for the first round, the coefficients ranged from 0.22 to 0.51, and for the final round, almost all of them were around 0.60. In other words, when these peer students acquired more experiences about peer assessment, they would have better capability to judge their peers' work, which in turn was more correlated with their teacher's marks.

The effects of online peer assessment on the development of science activity

The descriptive data of students' project scores evaluated by peers on the five assessment dimensions for each round are presented in Table 4. Table 4 shows that students' average scores in the first round of on-line peer assessment as evaluated by their peers were found to be 4.90, 4.95, 4.96, 4.67, and 4.69 on these dimensions respectively. The scores on the third round of peer assessment were 5.65, 5.53, 5.61, 5.41, and 5.59 for the five dimensions as assessed by peers. An observation of the mean of students' scores on each dimension in Table 4 revealed that these students had an increasing average score on each dimension. Table 5 shows the scores evaluated by the expert (the course instructor), and a similar increasing trend was revealed across different peer assessment rounds.

A series of F-tests were further used to compare student score changes as a result of the on-line peer assessment. The results are also shown in Table 4 for peer scores and in Table 5 for expert scores. It was found that students significantly progressed their performance from the on-line peer assessment process. That is, in almost all situations, their

Table 3 The correlation between expert and peer scores for each outcome variable

	Development	Science fun	Scientific relevancy	Creativity	Feasibility
First round	0.22	0.43**	0.40*	0.51**	0.40*
Second round	0.39*	0.46**	0.52**	0.62***	0.58***
Third round	0.41*	0.60***	0.61***	0.58***	0.58***

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

Table 4 Students' scores of science activities from peers' perspective and their progression ($n = 36$)

	(1) First round (mean, SD)	(2) Second round (mean, SD)	(3) Third round (mean, SD)	F-value	Pair test
Development	4.90 (0.63)	5.28 (0.65)	5.65 (0.63)	22.17[*]	3 > 2 > 1
Science fun	4.95 (0.59)	5.22 (0.73)	5.53 (0.68)	11.43[*]	3 > 2 > 1
Scientific relevancy	4.96 (0.69)	5.28 (0.66)	5.61 (0.65)	18.01[*]	3 > 2 > 1
Creativity	4.67 (0.72)	5.06 (0.65)	5.41 (0.65)	16.93[*]	3 > 2 > 1
Feasibility	4.69 (0.66)	5.10 (0.72)	5.59 (0.62)	38.88[*]	3 > 2 > 1

[*] $P < 0.001$

Table 5 Students' scores of science activities from expert perspective and their progression ($n = 36$)

	(1) First round (mean, SD)	(2) Second round (mean, SD)	(3) Third round (mean, SD)	F-value	Pair test
Development	4.72 (0.61)	5.33 (0.53)	5.69 (0.58)	27.73[**]	3 > 2 > 1
Science fun	4.36 (0.90)	4.89 (1.06)	5.11 (0.89)	11.23[**]	3 > 1; 2 > 1
Scientific relevancy	4.72 (0.94)	4.89 (0.89)	5.19 (0.86)	4.84[*]	3 > 1; 3 > 2
Creativity	4.25 (1.11)	4.94 (1.15)	5.08 (1.05)	16.21[**]	3 > 1; 2 > 1
Feasibility	4.58 (0.81)	5.03 (0.70)	5.22 (0.64)	10.55[**]	3 > 1; 2 > 1

[*] $P < 0.05$; [**] $P < 0.001$

scores for each dimension were statistically higher in a later round than those in former round. Students significantly improved their science activities as involving the peer assessment process in all evaluation dimensions. The process of online peer assessment helped these preservice teachers design science activities that were more fun, creative, feasible, and more related to the acquisition of scientific knowledge as well as more suitable for the developmental stage of pre-school children. In particular, the dimension of "development" showed much progression in both peer and expert perspectives ($F = 22.17$ in Table 4, $F = 27.73$ in Table 5), suggesting that peer assessment was very effective to improve the design of science activities to fit the development of pre-school children.

The correlation between student SEVs and their gains from peer assessment module

Tables 4 and 5 supported that students, on average, had significant gains along the implementation of on-line peer assessment. Then, the next research question of this study was to investigate the relationships between students' SEVs and their gains in the peer assessment. Each student's gains (progressions) were defined by his/her score differences between the first round of peer assessment and the third round for each of the five scoring dimensions.

Table 6 shows the correlations between students' SEVs (assessed by Tsai and Liu 2005) and their gains of peer scores. It was found that for the dimension of "science fun," students with more constructivist-oriented SEVs in the subscales of "invented and creative reality of science," "theory-laden exploration," "changing and tentative feature" and "social negotiation" tended to have significantly more progressions ($r = 0.33$–0.44, $P < 0.05$). Students expressing more constructivist-oriented SEVs in the subscale of "invented and creative reality" tended to gain significantly more in the dimensions of

Table 6 The Pearson correlations between peer assessment gains and SEVs

	Development gain	Science fun Gain	Scientific relevancy gain	Creativity gain	Feasibility gain
Invented and creative	0.20	0.38*	0.37*	0.34*	0.17
Theory-laden	0.20	0.44**	0.25	0.28	0.03
Changing and tentative	0.17	0.34*	0.25	0.25	0.09
Social negotiation	0.11	0.33*	0.23	0.24	0.09
Cultural impacts	0.19	0.32	0.28	0.37*	0.08

* $P < 0.05$; ** $P < 0.01$

"scientific relevancy" ($r = 0.37$) and "creativity" ($r = 0.34$) than those holding positivist-aligned SEVs. A significantly positive correlation was also revealed between students' SEVs of "cultural impacts" subscale and their gains in the "creativity" dimension of designing science activities ($r = 0.37$).

Similarly, Table 7 presents the relationships between students' SEVs and their gains of expert (teacher) scores. It was found that the SEV subscales were significantly related to their gains in the "science fun" dimension ($r = 0.37$–0.45), similar to those revealed by Table 6. Students with more constructivist-oriented SEVs in "invented and creative" and "changing and tentative" tended to progress more in the dimension of "scientific relevancy." These findings, in general, supported that students with more constructivist-oriented SEVs might benefit more from the peer assessment process, especially for helping the students to design science activities with more fun, higher creativity and greater relevancy to scientific knowledge.

Content analysis of peer feedback

As concluded previously, students in this study progressed significantly derived from the online peer assessment module. Similar to the assertion made by Liu and Carless (2006), this study hypothesized that one of the major sources for their progression came from peer feedback. Therefore, an analysis of students' peer feedback is quite essential. Each participating student provided various types or categories of feedback to their peers on each round of peer assessment. This study analyzed the content of peer feedback provided by

Table 7 The Pearson correlations between gains from teacher's perspective and SEVs

	Development gain	Science fun gain	Scientific relevancy gain	Creativity gain	Feasibility gain
Invented and creative	0.17	0.39*	0.39*	0.30	0.11
Theory-laden	0.14	0.45**	0.24	0.21	0.02
Changing and tentative	0.19	0.44**	0.33*	0.24	0.14
Social negotiation	0.10	0.37*	0.24	0.20	0.06
Cultural impacts	0.13	0.43**	0.23	0.29	0.03

* $P < 0.05$; ** $P < 0.01$

Table 8 Content analysis of peer feedback by dimensions

	First round Mean (SD)	Second round Mean (SD)	Third round Mean (SD)
Affective	14.91 (7.00)	19.53 (7.32)	22.93 (6.88)
Cognitive	15.40 (6.05)	10.00 (5.33)	12.13 (7.30)
Metacognitive	10.00 (6.89)	11.09 (7.44)	6.53 (4.73)
Total	40.31 (14.03)	40.62 (13.87)	41.60 (12.96)

each student in terms of the dimensions of affective, cognitive and metacognitive by the framework in Table 2. Table 8 shows the results across three rounds of peer assessment.

It was found that on average, for each round of peer assessment, each student provided about 40 comments to their peers. (In this study, each student rated five peers' work each round). By and large, more comments were categorized as "affective." However, if making cross-round comparisons by each dimension, the frequency of cognitive comments achieved highest in the first round, metacognitive attained highest on the second round, and affective occurred most frequently in the final round. These results implied a peer assessment learning model that students might engage in cognitive processing for evaluating peers' work first, then they were enhanced to some metacognitive thoughts, and then they finally involved some affective responses toward peers' work. Table 9 further presents the analysis of peer comments by more specific categories (shown Table 2).

According to Table 9, "supporting" was the most frequent feedback provided by the students. Then, still many comments were categorized as "personal opinion." Again, by cross-round comparisons by each category, the frequency of "guiding" and "evaluating and planning" achieved highest in the first round, "regulating and reflecting" attained highest in the second round, while numerous "supporting" comments elicited in the final round.

The relationships between SEVs and generation of peer feedback

This study further addressed another important research question: What is the role of student SEVs on the comments in which they provided to their peers' science activity projects? Table 10 shows a correlational analysis between student SEVs and the peer feedback provided. First, it was revealed that students with more constructivist-oriented SEVs tended to give more comments to their peers along the online peer assessment

Table 9 Content analysis of peer feedback by categories

	First round Mean (SD)	Second round Mean (SD)	Third round Mean (SD)
Supporting	13.06 (5.98)	19.00 (7.14)	22.50 (6.89)
Emotional response	1.86 (2.18)	0.53 (0.96)	0.43 (0.68)
Direct correction	1.11 (1.66)	0.35 (0.81)	0.33 (0.76)
Personal opinion	11.20 (5.01)	8.24 (5.10)	10.47 (7.18)
Guiding	3.09 (2.63)	1.41 (1.86)	1.33 (1.71)
Evaluating and planning	7.71 (5.84)	6.35 (4.72)	3.03 (2.58)
Regulating and reflecting	2.29 (2.74)	4.74 (4.26)	3.50 (3.52)

Table 10 The correlation between students' SEVs and their providing of peer feedback

	IC	TL	CT	SN	CU
Supporting	0.43**	0.39*	0.29	0.28	0.42*
Emotional response	0.05	−0.01	0.10	−0.02	−0.13
Direct correction	−0.12	−0.02	−0.04	−0.08	−0.03
Personal opinion	0.32	0.35*	0.32	0.32	0.38*
Guiding	0.33*	0.43**	0.44**	0.38*	0.46**
Evaluating and planning	0.33*	0.38*	0.42*	0.38*	0.42*
Regulating and reflecting	0.08	0.16	0.21	0.15	0.18
Total	0.42*	0.45**	0.43**	0.38*	0.49**

* $P < 0.05$; ** $P < 0.01$

IC: Invented and creative; TL: Theory-laden exploration; CT: Changing and tentative; SN: Social negotiations; CU: Cultural impact

module ($r = 0.38$–0.49 for the total). Moreover, the students with more sophisticated SEVs in the subscales of "invented and creative," "theory-laden exploration" and "cultural impacts" tended to present more feedback for supporting. The students holding more constructivist-oriented SEVs (in each SEV scale) tended to disclose more comments to guide peers' work (guiding, $r = 0.33$–0.46). Moreover, these students also afforded significantly more feedback to help their peers to engage in metacognitive thinking for developing science activity (e.g., evaluating and reflecting, $r = 0.33$–0.42). In general, these results suggested that students with advanced (i.e., constructivist-oriented) SEVs tended to offer more feedback to their peers, and much of the feedback was categorized as guiding or helping peers to critically evaluate and plan their designs of science activities. These findings were consistent with those of some studies that the having of more sophisticated SEVs was the prerequisite for implementing peer-supported collaborative learning (e.g., Tsai 1999a, b).

Discussion and conclusions

This study implemented an on-line peer assessment system to help a group of preservice pre-school teachers design science activities for children. First, the correlations between expert and peer scores were examined, and positively medium to high correlations were revealed. In light of these results, the peer scores could be viewed as a valid measurement for assessing student work. This conclusion concurred with the findings presented by previous studies (e.g., Cho et al. 2006; Topping 1998). This study further proposed that when students had more experiences of peer assessment or after more rounds of peer assessment, the validity of peer scores was likely enhanced. Sufficient experiences in peer assessment, particular with appropriate guidance or scaffolds (Cho et al. 2006), may greatly help students to provide valid assessment scores to their peers. It is suggested for future studies in peer assessment to include some trial practices before actual implementation, if peer score will be considered as an indicator for representing student performance.

Moreover, it was found that through the online peer assessment, these students (or preservice teachers) could enhance the design of science activities for future instruction.

The positive effects on student performance were clearly illustrated, which were consistent with numerous studies showing the favorable impacts of peer assessment on student work (Barak and Rafaeli 2004; Cho and Schunn 2007; Tsai et al. 2002). In addition, instructors in any area of teacher education, no matter at preservice or inservice level, can use a similar online system to effectively help teachers share, comment, and collaboratively design some instructional plans for students. Future research may be undertaken to evaluate the practical effects of using the designed activities or plans on students in real classrooms.

This study considered students' SEVs as an essential factor for developing science activities via the online peer assessment system. In terms of the content involved in student work, the development of science activities requires a better understanding toward science; thus SEVs may guide their thinking and evaluation about "what constitutes a good science activity?" In terms of the instructional method utilized in this study, that is, online peer assessment, the students were engaged in inquiry peer-supported, social constructivist learning environments, where previous research revealed the high importance of SEVs for student learning (Tsai 1999b; Wallace et al. 2003; Windschitl and Andre 1998). This study presented evidence that the participants' SEVs were related to their progressions in the peer assessment environments. In particular, students with more sophisticated (i.e., constructivist-oriented) SEVs tended to progress significantly more for designing science activities with more fun, and possibly with higher creativity and relevancy to scientific knowledge. These findings implied that learners with constructivist-oriented SEVs might benefit more from the online peer assessment learning process. Educators should be highly aware learners' SEVs when implementing peer assessment activities in science education. An appropriate understanding regarding the constructivist epistemology may be a prerequisite for utilizing peer assessment learning activities in science education.

The content analyses of peer feedback found that peer comments were more frequent in the "affective" dimension (e.g., supporting, emotional responses), rather than in the "cognitive" (e.g., guiding) or "metacognitive" (e.g., evaluating, regulating) dimension. Educators may find more ways to encourage students' cognitive as well as metacognitive feedback when engaging in peer assessment learning. Peer modeling, exemplar sharing, explicit discussion or instruction on this before the implementation of peer assessment may be helpful. Furthermore, by cross-round comparisons of peer feedback, this study suggested an initial model for describing student learning involved in the peer assessment process. The peer assessment learning model has suggested that students may engage in cognitive activities first, then they are moved to more metacognitive thinking toward peers' work, and at the end of the peer assessment, likely an indication of the maturation of peer work, they express more affective responses about it. Certainly, this model requires more empirical data to verify its validity.

Students' SEVs may also be related to the comments they provide to their peers. This study found that students with SEVs more oriented to the constructivist philosophy tended to offer more feedback to their peers along the online peer assessment module. And more importantly, by a series of content analyses, it was found that much of the peer feedback provided by these students was categorized as guiding or helping peers to carefully appraise and plan their science activity projects. Such feedback is perceived as profitably valuable for promoting higher-order reflective thinking. This finding, again, strengthens an aforementioned argument that an adequate understanding about the constructivist epistemology may be a prerequisite (though not sufficient condition) for utilizing peer assessment learning activities in science education, as these students may more likely offer high-quality feedback or comments to their peers. Therefore, how to help students acquire

relatively constructivist-oriented SEVs may become an important issue for educators. This perspective is also consistent with that proposed by Lederman et al. (2002), suggesting a focus on classroom interventions for enhancing SEVs is much more important than mass evaluations aimed at depicting students' SEVs. Abd-El-Khalick and Lederman (2000) proposed major approaches of changing people's SEVs such as implementing science-based inquiry activities or utilizing elements from the history and philosophy of science for instructional interventions. This study further hypothesized that online peer assessment might be a vehicle of changing participants' epistemological beliefs, as the peer assessment process would involve in-depth inquiry for each peer project and different perspectives are induced in the process, which are often perceived as beneficial to epistemological development (Tsai 2001a, b, 2004). In fact, the study completed by Tsai (in press) has indicated that some online inquiry activities are quite effective for students' acquisition about constructivist-oriented SEVs. Nevertheless, how online peer assessment may contribute to participants' epistemological development needs additional research to explore.

Finally, based on the implementation experiences of this study, online technology is quite helpful to facilitate the implementation of peer assessment. The students can submit and review the course projects in a more effective way, with possibly higher anonymity. Also, course instructor and researchers can gain more detailed records about the peer assessment process involved, because all of the records are stored in online database. The teacher and researcher can access the data for either practical or academic purposes. Computer-supported or Web-assisted peer assessment is a contemporary trend for research and practice (Barak and Rafaeli 2004; Wen and Tsai 2006). Educators are encouraged to implement more online peer assessment environments to facilitate student learning.

There were still some limitations for current study. For example, some other ways of probing students' SEVs, such as open-ended questionnaire and interviews, might be utilized to acquire more complete representations (Lederman 1992; Lederman et al. 2002). Similarly, some interviews with selected participants might be quite helpful to know more about how they progress from the online peer assessment module, how they evaluate and provide comments to peers' work and how they perceive these comments. The psychological factors that are involved in peer assessment process clearly need more research.

References

Abd-El-Khalick, F., & Lederman, N. G. (2000). Improving science teachers' conceptions of nature of science: A critical review of the literature. *International Journal of Science Education, 22*, 665–701.

Barak, M., & Rafaeli, S. (2004). On-line question-posing and peer-assessment as means for web-based knowledge sharing in learning. *International Journal of Human-Computer Studies, 61*, 84–103.

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

Cho, K., Schunn, C. D., & Wilson, R. W. (2006). Validity and reliability of scaffolded peer assessment of writing from instructor and student perspectives. *Journal of Educational Psychology, 98*, 891–901.

Davies, P. (2000). Computerized peer assessment. *Innovations in Education and Training International, 37*, 346–355.

Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research, 70*, 287–322.

Guan, Y.-H., Tsai, C.-C., & Hwang, F.-K. (2006). Content analysis of online discussion on a senior-high-school discussion forum of a virtual physics laboratory. *Instructional Science, 34*, 279–311.

Henri, F. (1992). Computer conferencing and content analysis. In A. R. Kaye (Ed.), *Collaborative learning through computer conferencing: The Najaden papers* (pp. 115–136). New York: Springer.

Lederman, N. G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching, 29*, 331–359.

Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching, 39*, 497–521.

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*, 420–432.

Liu, N. F., & Carless, D. (2006). Peer feedback: The learning element of peer assessment. *Teaching in Higher Education, 11*, 279–290.

Liu, S.-Y., & Tsai, C.-C. Differences in the scientific epistemological views of undergraduate students. *International Journal of Science Education*, in press.

Ryan, A. G., & Aikenhead, G. S. (1992). Students' preconceptions about the epistemology of science. *Science Education, 76*, 559–580.

Smith, H., Cooper, A., & Lancaster, L. (2002). Improving the quality of undergraduate peer assessment: A case for student and staff development. *Innovations in Education and Teaching International, 39*, 71–81.

Songer, N. B., & Linn, M. C. (1991) How do students' views of science influence knowledge integration? *Journal of Research in Science Teaching, 28*, 761–784.

Southerland, S. A., Sinatra, G. M., Matthews, M. R. (2001). Belief, knowledge, and science education. *Educational Psychology Review, 13*, 325–351.

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

Tsai, C.-C. (1998). An analysis of scientific epistemological beliefs and learning orientations of Taiwanese eighth graders. *Science Education, 82*, 473–489.

Tsai, C.-C. (1999a). The progression toward constructivist epistemological views of science: A case study of the STS instruction of Taiwanese high school female students. *International Journal of Science Education, 21*, 1201–1222.

Tsai, C.-C. (1999b). "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). 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. (2001b). A review and discussion of epistemological commitments, metacognition, and critical thinking with suggestions on their enhancement in Internet-assisted chemistry classrooms. *Journal of Chemical Education, 78*, 970–974.

Tsai, C.-C. (2002). A science teacher's reflections and knowledge growth about STS instruction after actual implementation. *Science Education, 86*, 23–41.

Tsai, C.-C. (2004). Beyond cognitive and metacognitive tools: The use of the Internet as an "epistemological" tool for instruction. *British Journal of Educational Technology, 35*, 525–536.

Tsai, C.-C. (2007). Teachers' scientific epistemological views: The coherence with instruction and students' views. *Science Education, 91*, 222–243.

Tsai, C.-C. The use of Internet-based instruction for the development of epistemological beliefs: A case study in Taiwan. In M. S. Khine (Eds.), *Knowing, knowledge and beliefs: Epistemological studies across diverse cultures*. Dordrecht: Springer, in press.

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

Tsai, C.-C., & Liu, S.-Y. (2005). Developing a multi-dimensional instrument for assessing students' epistemological views toward science. *International Journal of Science Education, 27*, 1621–1638.

Tsai, C.-C., Liu, E. Z.-F., Lin, S. S. J., & Yuan, S.-M. (2001). A networked peer assessment system based on a Vee heuristic. *Innovations in Education and Teaching International, 38*, 220–230.

Tseng, S.-C., & Tsai, C.-C. (2007) On-line peer assessment and the role of the peer feedback: A study of high school computer course. *Computers & Education, 49*, 1161–1174.

van den Berg, I., Admiraal, W., & Pilot, A. (2006). Design principles and outcomes of peer assessment in higher education. *Studies in Higher Education, 31*, 341–356.

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

Wallace, C. S., Tsoi, M. Y., Calkin, J., & Darley, M. (2003). Learning from inquiry-based laboratories in nonmajor biology: An interpretive study of the relationships among inquiry experience, epistemologies, and conceptual growth. *Journal of Research in Science Teaching, 40*, 986–1024.

Wen, M. L., & Tsai, C.-C. (2006). University students' perceptions of and attitudes toward (online) peer assessment. *Higher Education, 51*, 27–44.

Wen, L. M. C. & Tsai, C.-C. Online peer assessment in an inservice science and mathematics teacher education course. *Teaching in Higher Education*, in press.

Windschitl, M., & Andre, T. (1998) Using computer simulations to enhance conceptual change: The roles of constructivist instruction and student epistemological beliefs. *Journal of Research in Science Teaching, 35*, 145–160.

Yu, F.-Y., Liu, Y.-H., & Chan, T.-W. (2005). A web-based learning system for question-posing and peer assessment. *Innovations in Education and Teaching International, 42*, 337–348.