

A Study of the Effects of Scaffolded Assessment on the Learning Effectiveness in Network Peer Assessment Activities

Chien-I Lee , Ya-Fei Yang, and Shin-Yi Mai
Department of Information and Learning Technology,
National University of Tainan,
Tainan, Taiwan

e-mail: {leeci@mail.nutn.edu.tw, cillia55@gmail.com, maitz3@hotmail.com}

Abstract—Peer assessment has been considered an important process for learning. However, students may not offer constructive feedback due to lack of expertise knowledge in network assessment activities. Scaffolding can be offered to students timely to assist in peer assessment. Therefore, to address this issue, this study proposed a scaffolded assessment approach accordingly. To evaluate the effectiveness of the proposed approach, the quasi-experimental design was employed to investigate the effects of scaffolded assessment for self-critiques and peer assessment on students' learning effectiveness in the network assessment activities. A total of ninety 7th graders participated in the experiment, and divided into three groups with or without the scaffolding critique. The results show that the validity in the network peer assessment has been well demonstrated, indicating scaffolded assessment could enhance the validity of the network peer assessment. Moreover, the participants' learning effectiveness is also enhanced. In addition, the participants showed a positive learning attitude toward the network assessment activities and agreed that the network assessment activities could enhance the participants' interactions between the peers and instructor. All in all, the proposed scaffolded assessment enables the participants to offer constructive feedback in the assessment activities.

Keywords—Network Peer Assessment; Scaffolding; Validity; Learning Effectiveness; Learning Attitude

I. INTRODUCTION

Research on network peer assessment had been extensively studied and indicated that network peer assessment has positive effects on learning [11]. Due to rapid development of network technologies, disadvantages of paper-pencil based peer assessment have been gradually corrected and then replaced by network assessment. Individuals who share common interests, feelings, or ideas over the Internet form a virtual community in which learners evaluate peers' works, receive feedback from peers, modify their original works based on peer assessment, and then make their works become better [14][18].

Peer assessment enables learners to gain diverse ideas and inspiration, enhances their higher order thinking skills, and offers peers opportunities to learn from each other, which can promote learners' learning motivation and achievements [7][16][17][22]. Network peer assessment defined as students are given the needed knowledge to tasks

review and feedback in peer assessment process through the Internet. Basically, network peer assessment is similar with peer assessment [12][13].

However, the lack of expertise may result in weak feedback and comments due to student incomprehension and misunderstanding of works during the peer assessment [5]. Thus, the appropriate support and scaffold given to student is needed, in which facilitate students to offer proper comments for their peers works [21]. Learners can observe peers' works, understand their learning progress, reflect on self-learning, and then provide feedback to peers for further improvement [1]. In addition, peers may gain more feedback from diverse backgrounds of peers understanding toward the works than that from teachers [2]. Previous research have proved the effectiveness of the scaffolded assessment. For instance, Cho, Schunn, and Wilson [3] stated that scaffolded assessment could facilitate evaluators to assess given tasks precisely.

Thus, this study proposes the scaffolded assessment for self-critiques and peer assessment by using self-critiques and peer feedback as an evaluation basis. To evaluate the effectiveness of the proposed approach, the quasi-experimental design was employed to investigate the effects of scaffolded assessment for self-critiques and peer assessment on students' learning effectiveness and learning attitude in the network assessment activities.

In this paper, a survey of related work will be studied in Section II. Section III is our research framework. The experimental process and activities will be given in Section IV. Section V describes the research tools. Section VI shows the experiment results. Finally, the conclusions and some future work are given in Section VII.

II. LITERATURE REVIEW

A. Peer Assessment

The concept of peer assessment is from Peer Assisted Learning (PAL) and multiple assessments. Types of PAL include peer tutoring, peer modeling, peer education, peer counseling, peer monitoring, and peer assessment [15]. Topping [14] stated that peer assessment is performed by students with similar degree of knowledge or background. Students not only have to learn knowledge and accomplish their assignments but also play a role as a tutor to observe and evaluate peers' works. Then, students receive peer

feedback and modify their original assignments. Isaacs [6] mentioned that peer assessment is an activity in which students demonstrate their own works to peers for assessment. Peers who participate in assessment activities must have the same educational background or be in the same class, and they play roles of being an author, evaluator, and evaluatee. Students have to observe, evaluate, and compare by playing different roles in multiple activities which can enhance their diverse thinking abilities and learning effectiveness.

B. Network Peer Assessment

Network peer assessment is also called as web-based peer assessment, in which students can facilitate contacts, assist in brainstorming and generate meaningful learning [10]. With the application of the Internet, students can observe and evaluate peers' assignments on a network teaching platform and receive feedback from peers, and then they can modify their original works based on the received feedback [8][14][18]. Knoy, Lin, Liu, and Yuan [9] stated that the three advantages of network peer assessment are: 1) to ensure anonymity and enhance peers' willingness to evaluate, 2) easy for teachers to monitor learning and evaluation process which reduce paper waste and copying cost, and 3) convenient for students to demonstrate their assignments for peer assessment.

C. Scaffolding

The concept of scaffolding is based on Vygotsky's learning theory. Scaffolding is the process of teaching children, which is similar to the construction of a house, by realizing children's characteristics, offering appropriate assistance, and ensuring that children have full support. When children are able to independently solve specific problems, support or assistance may be gradually reduced to develop their own problem solving ability. Once they can independently solve various problems without assistance, then scaffolding can be removed. Vygotsky [19] indicated development is the transformation of socially shared activities into internalized processes. The concept of scaffolding is similar to the Zone of Proximal Development (ZPD), in which instructors or peers with better academic performance can provide scaffoldings to assist students to develop their learning ability and achieve the goal of transferring learning. As learners' ability is enhanced, they can learn independently and construct their own knowledge which is the time to gradually reduce scaffolding. Thus, with scaffolded assessment, students can construct new knowledge and enhance their ability through peer feedback.

III. RESEARCH FRAMEWORK

This study aims to investigate the effects of scaffolded assessment on the students' learning effectiveness and validity of the network peer assessment. Since Modular Object-Oriented Dynamic Learning Environment (Moodle) [23] is a free software e-learning platform with three functions: management of website, learning, and course, it can run on Windows and Mac and is easy to access and manage. Thus, this study uses Moodle to construct a network

peer assessment system in which the instructor can upload and manage course content, and students can upload their assignments, conduct peer assessment, and view their scores. The students' assignments, scores, and behaviors are recorded in the database. After the network peer assessment activities are completed, this study uses the SPSS 17.0 statistical software [24] to analyze the data.

This study investigates the effects of different scaffolded assessment approaches on the learning effectiveness, learning attitudes, and validity of the network peer assessment. The variables include independent, dependent, and control ones, which are specified as follows:

1) The independent variable: Providing the scaffolded assessment.

a) The experimental group I is given the scaffolded assessment for self-critiques.

b) The experimental group II is given the scaffolded peer assessment.

c) The control group would not receive any scaffolded assessment when evaluating the other members' assignments in the network peer assessment activities.

2) The dependent variable:

a) Validity on the assessment: The validity refers to the consistency between the earned scores from the peer and instructor in this study.

b) Learning effectiveness: A student's final score on his or her assignment is the mean of the earned scores from the three network peer assessment activities.

c) Learning attitude: The leaning attitude indicates that the students' attitudes toward the network peer assessment activities.

3) The control variable:

a) Instructor: The instructor is a full-time computer instructor.

b) Course hour: The instructor would complete the teaching instruction in a 45-minute class in the first week of the network peer assessment activities.

c) Course content: The 5th and 6th chapters, which are "my creative photos" and "my creative photo frames", respectively, of the textbook.

d) Same softwares, PhotoCap 6 and Photo Magician, used to complete the course assignment in a computer lab.

The students of the two experimental groups and the control group would have the same course hours, learn the same content, and be taught by the same computer instructor in the same learning environment.

A. Participants

The participants of this study were from three classes of the 7th grade, with a total of 90 students who participated in the network peer assessment activities. They possess basic computer ability to implement the software, PhotoCap 6, and access computers, such as collection information, uploading and downloading data from the Internet. The participants are normally grouped by age in the junior high school. Three classes of the 7th graders are randomly selected and labeled as the experimental group I with 30 students conducted with self-critique for reference, the experimental group II with 29 students conducted with other peers' critique for reference,

and the control group with 31 students conducted without any reference.

B. Network peer assessment

Moodle is an open Course Management System (CMS) based on the theory of social construction. Moodle is supported by a large international community of educators because of the three main functions: management of website, learning, and course. It has an easy to learn user interface which enables instructors to manage their courses, and students can learn course content independently by opening a web browser and linking to the teaching platform. The proposed network peer assessment system is constructed by Moodle 2.4.1 and AppServ software. AppServ can configure Apache, PHP, and MySQL to form an integrated environment for the experimental activities. MySQL is used to construct the backend database.

The interfaces of the proposed network peer assessment system shows in Figure 1. The network peer assessment system is installed on the serve host system. The instructor and students have to access the Internet, link to the server, and enter the network peer assessment system. Student account and password required based on the course enrollment by teachers. Figure 2 shows the peer assessment activities announce and assessing records in the system. Students can observe and evaluate peers' assignments on a network teaching platform and receive feedback from peers.



Figure 1. Student entry interface of the network peer assessment system

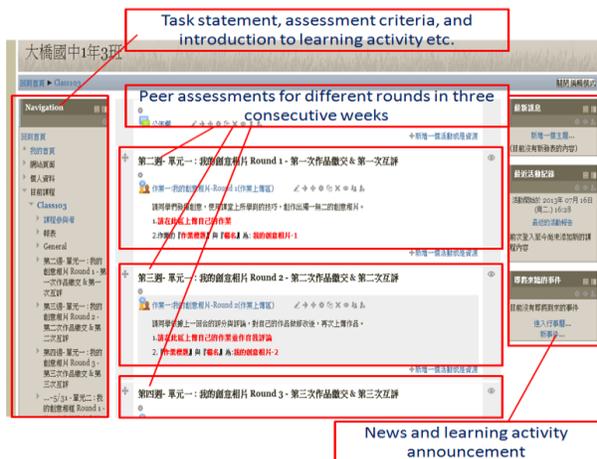


Figure 2. Peer assessment activity announced in the System

IV. THE DESIGN OF THE EXPERIMENTAL ACTIVITIES

A. The experimental implementation phase

After the preparation phase of the first four weeks, the experimental activities would be carried out for seven weeks. The details of the experimental activities are described in Table I.

TABLE I. THE DETAILS OF THE EXPERIMENTAL ACTIVITIES

1 st week	Experimental groups I and II. The control group
	1. Explanation of the course assignment and grading policy 2. The participants are required to upload their assignments to the system after the instructor delivered instruction
2 nd week	Experimental groups I and II. The control group
	Completed the assignments, evaluated the other peers' assignments, offered feedback.
3 rd week	The experimental group I
	Modified their original assignments based on the received feedback from their peers, provided self-critiques on their modified assignments, and then uploaded them to the proposed system in the first activity
3 rd week	Experimental group II. The control group
	Modified their original assignments based on the received feedback from their peers and then uploaded them to the proposed system in the first activity
4 th week	Experimental group I
	Read the evaluatees' self-critiques and evaluated the assignments in the second activity.
4 th week	Experimental group II
	Read the other members' feedback, which was offered in the first network peer assessment activity, and then evaluated the assignments in the second activity.
4 th week	The control group
	Evaluated the peers' assignments in the second activity.
5 th week	Experimental groups I
	Modified their assignments again based on the received feedback from their peers, provided self-critiques on their modified assignments, and uploaded them to the system
5 th week	Experimental groups II. The control group
	Read the other members' feedback, modified their assignments again, and uploaded them to the system
6 th week	Experimental group I
	Read the evaluatees' self-critiques and evaluated the assignments in the third activity.
6 th week	Experimental group II
	Read the other members' feedback on the assignments, which was offered in the first and second assessment activities, and then evaluated the third activity.
6 th week	The control group
	Evaluated the peers' assignments.
7 th week	Experimental groups I and II. The control group
	After completing all the three assessment activities, the participants are required to fill in the network peer assessment attitude scale in the system

B. The result analysis phase

After the experimental activities are completed, the instructor and the teaching assistant would also evaluate the participants' assignments. The mean of the two given scores from the instructor and teaching assistant would be the earned score from the instructor on each participant's assignment. The scores from the instructor and peers would be incorporated for further analysis and investigation. The instructor has been a computer instructor for five year and fully familiar with the course content. The teaching assistant is also familiar with the necessary knowledge and skills required in this computer course.

V. RESEARCH TOOLS

This study uses the peer assessment score sheet as a tool for the participants to score the other members' assignments during the network peer assessment activities. The instructor has to explain the scoring criteria to the participants in the first week of the experimental activities. The scoring guidelines are developed based on the discussion between the teacher and participants. The peer assessment score sheet is used to evaluate the participants' assignments based on accuracy, expression, and completeness. The lowest score is 1, and the highest score is 100 on the peer assessment score sheet.

This study modifies the Wen and Tsai's [20] questionnaire, which originally had 34 items with five-point Likert scale (1: strongly disagree, 5: strongly agree), with an aim to investigate students' perceptions of network peer assessment. The instrument contains four subscales: positive attitude (PAS), online attitude (OAS), understanding-and-action (UAS), and negative attitude (NAS). The analysis results showed that the reliability of the four individual subscales took values more than 0.63 and 0.8 for the overall reliability. Thus, the questionnaire has high reliability which can be used to evaluate the students' attitudes toward network peer assessment.

The modified questionnaire contains 20 items. One of the three classes was randomly selected to evaluate the reliability and validity of the modified questionnaire. The results show that the Alpha coefficient takes a value of 0.6 for the four individual subscales and 0.86 for all the four subscales, indicating that the modified questionnaire is adequately reliable [4].

VI. RESULTS

The participants are required to complete the three assessment activities in this experiment. The valid sample size of the experimental group I is 25, making the completing rate of 83.3%. The valid sample size of the experimental group II is 26, making the completing rate of 89.7%. The valid sample size of the control group is 25, making the completing rate of 90.3%.

A. The reliability analysis

This study investigates the reliability between the scaffolded and non-scaffolded assessment activities. The experimental group I is given the scaffolded assessment for

self-critiques. The experimental group II is given the scaffolded peer assessment. The control group would not receive any scaffolded assessment when evaluating the other members' assignments in the network peer assessment activities. This study uses the Pearson's correlation coefficient to measure the reliability between the scores from the instructor and teaching assistant. The results show that the reliability between the scores from the instructor and teaching assistant on the participants' assignments are positively correlated in the three assessment activities. The Pearson's correlation coefficient takes values more than 0.8, indicating that the scores from both the instructor and teaching assistant are adequately reliable and can be an external criterion for the validation of the network peer assessment.

B. The validity analysis of the scaffolded assessment for self-critiques in the experimental group I.

This study uses the scores from both the instructor and teaching assistant as the external criterion to examine whether the validity of the scaffolded assessment for self-critiques is enhanced in the network assessment activities. As can be seen in Table II, the results show that the scores from the instructor and peers are significantly positively correlated in the three network assessment activities, and the correlation coefficients, which take values between 0.5 to 0.7, gradually increase.

TABLE II. THE PEARSON CORRELATION COEFFICIENT BETWEEN THE SCORES FROM THE INSTRUCTOR AND PEERS IN THE EXPERIMENTAL GROUP I

Assessment activities	Evaluator	n	Mean	SD	r
The 1 st	Instructor	25	73.44	8.905	.525**
	Peers		76.68	12.284	
The 2 nd	Instructor	25	77.14	5.467	.624**
	Peers		79.04	7.738	
The 3 rd	Instructor	25	84.25	5.592	.695**
	Peers		83.96	6.465	

** p<0.01

The consistency between the scores from the instructor and peers are enhanced after the second network assessment activity.

C. The validity analysis of the scaffolded peer assessment in the experimental group II

This study uses the scores from both the instructor and teaching assistant as the external criterion to examine whether the validity of the scaffolded peer assessment is enhanced in the network assessment activities. As it can be seen in Table III, the results show that the scores from the instructor and peers are significantly positively correlated in the first network assessment activity when $p < 0.05$.

The scores from the instructor and peers are significantly positively correlated in the second network assessment activity when $p < 0.01$, and their correlation coefficient takes values between 0.4 to 0.8. Thus, the consistency between the scores from the instructor and peers is enhanced.

TABLE III. THE PEARSON CORRELATION COEFFICIENT BETWEEN THE SCORES IN THE EXPERIMENTAL GROUP II

Assessment activities	Evaluator	n	Mean	SD	r
The 1 st	Instructor	2	67.23	8.823	0.478*
	Peers	6	67.35	10.737	
The 2 nd	Instructor	2	71.43	7.407	0.694**
	Peers	6	71.35	11.788	
The 3 rd	Instructor	2	82.60	8.404	0.713**
	Peers	6	83.31	9.388	

*p<0.05; ** p<0.01

D. The validity analysis of the control group

This study uses the scores from both the teacher and teaching assistant as the external criterion to investigate the consistency between the earned scores from the instructor and peers on the same assignment in the network assessment activities. As it can be seen in Table IV, the results show that the scores from the instructors and peers are significantly positively correlated in the first and second network assessment activities.

TABLE IV. THE PEARSON CORRELATION COEFFICIENT BETWEEN THE SCORES IN THE CONTROL GROUP

Assessment activities	Evaluator	n	Mean	SD	r
The 1 st	Instructor	28	70.82	8.147	0.684**
	Peers		79.36	11.656	
The 2 nd	Instructor	28	76.11	6.632	0.417*
	Peers		80.86	10.302	
The 3 rd	Instructor	28	87.19	4.865	0.117
	Peers		82.18	10.890	

*p<0.05 ; **p<0.01

However, the results show that the scores from the instructor and peers are not significantly correlated in the third network assessment activity. The process of repeatedly turning in, evaluating, and modifying assignments reduce the validity of peer assessment due to the lack of patience and motivation.

E. The comparison between the validity analyses of the two experimental groups and control group

As it can be seen in Table V, in the condition of no scaffolded assessment, the Pearson correlation coefficients between the scores from the instructor and peers are significantly positively correlated in the experimental group I and control group when $p < 0.01$ in the first assessment activity. The Pearson correlation coefficient between the scores from the instructor and peers is significantly positively correlated in the experimental group II when $p < 0.05$. The results show that the validity of the network assessment activities is significantly enhanced in the two experimental groups, and the validity of the network assessment activities is significantly reduced in the control group. Moreover, the validity of the network assessment activities in the experimental group I is higher than that in the experimental group II, indicating that the consistency of the scores from the instructor and peers in the experimental group II is higher than that in the experimental group I.

F. The analysis of learning effectiveness of the three groups

This study investigates the effects of scaffolded assessment on the students' learning effectiveness by comparing the results of the three assessment activities from the three groups. Since the three groups are not provided with the scaffolded assessment in the first assessment activity, this study performs the one-way ANCOVA where the students' scores in the first assessment activity are treated as a covariate to investigate the effects of the scaffolded assessment on the students' learning effectiveness in the second and third assessment activities. With regard to the second network assessment activity, since the assumption of the homogeneity of regression coefficient within groups was satisfied, the ANCOVA test was then performed. As it can be seen in Table VI, there is no significant difference among the three groups in the second assessment activity ($F = 1.603$ and $p = 0.208 > 0.05$), indicating that the students' learning effectiveness does not have a significant difference among the three groups in the second assessment activity.

TABLE V. THE PEARSON CORRELATION COEFFICIENT BETWEEN THE SCORES IN THE THREE GROUPS

Assessment activities	Group	n	r
The 1 st	Experimental groups I	25	0.525**
	Experimental groups II	26	0.478*
	Control group	28	0.684**
The 2 nd	Experimental groups I	25	0.624**
	Experimental groups II	26	0.694**
	Control group	28	0.417*
The 3 rd	Experimental groups I	25	0.695**
	Experimental groups II	26	0.713**
	Control group	28	0.117

*p<0.05; **p<0.01

The results also show that the effects of providing the scaffolded assessment on the students' learning effectiveness in the second assessment activity are not significant. As it can be seen in Table VII, there is no significant difference among the three groups in the third assessment activity ($F = 2.625$ and $p = 0.079 > 0.05$), indicating that the students' learning effectiveness does not have a significant difference among the three groups in the third assessment activity. The results also show that the effects of providing scaffolded assessment on the students' learning effectiveness in the third assessment activity are not significant. Therefore, the students' learning effectiveness is enhanced in the three assessment activities, but the effects of providing the scaffolded assessment on the students' learning effectiveness are not significant in the three assessment activities.

TABLE VI. SUMMARY OF ANCOVA FOR THE THREE GROUPS IN THE SECOND ASSESSMENT ACTIVITY

Source of variation	Sum of Squares	df	Mean Square	F	Sig.
Covariates	1770.434	1	1770.43	22.109	0.000
Within Groups	256.790	2	128.395	1.603	0.208
Error	6005.839	75	80.078		

TABLE VII. SUMMARY OF ANCOVA FOR THE THREE GROUPS IN THE THIRD ASSESSMENT ACTIVITY

Source of variation	Sum of Squares	df	Mean Square	F	Sig.
Covariates	1384.810	1	1384.810	20.674	0.000
Within Groups	351.730	2	175.865	2.625	0.079
Error	5023.795	75	66.984		

G. The analysis of learning attitude of the three groups

To investigate the degree to which an individual agree or disagree the network assessment activity in the study, ANOVA was employed to analyze four dimensions of the attitude questionnaire. The analyzed result was shown in Table VIII. Table VIII shows both dimensions "positive attitude" and "online attitude" take values of significant difference among three groups ($F = 4.95, p < .05$ and $F = 3.76, p < .05$, respectively). The post hoc of positive attitude shows the mean of experimental group II significantly higher than that of control group, indicating that students of experimental group II have positively higher level of degree to the network peer assessment than those of control group. In addition, The post hoc of online attitude shows the mean of experimental group II significantly higher than that of control group, indicating that students of experimental group II have higher level of acceptance to the network peer assessment than those of control group.

TABLE VIII. ANOVA OF LEARNING ATTITUDE TOWARD NETWORK PEER ASSESSMENT FOR THREE GROUPS

Dimension	Group	SD	Mean	F	Post hoc
Positive Attitude (PAS)	(1)	0.61	3.90	4.95	*(2)>(3)
	(2)	0.57	4.32		
	(3)	0.70	3.55		
Negative Attitude (NAS)	(1)	0.98	2.36	0.78	
	(2)	0.66	2.59		
	(3)	0.58	2.59		
Online Attitude (OAS)	(1)	0.75	3.94	3.76	*(2)>(3)
	(2)	0.74	4.29		
	(3)	0.66	3.57		
Understanding and Action (UAS)	(1)	0.67	4.06	0.98	
	(2)	0.76	3.92		
	(3)	0.57	3.80		

* $p < 0.05$, (1): Experimental groups I, (2): Experimental groups II, (3): Control group

VII. CONCLUSION AND FUTURE WORK

Scaffolding provided by teachers or peers can assist students to develop their learning ability and achieve the goal of transferring learning. This study uses the scaffolded assessment which provides scoring guidelines to investigate whether the students are benefited from the peer assessment activities. The results show that the learning effectiveness of the participants in the two experimental and control groups is not significantly enhanced during the first and second network peer assessment activities. However, the learning effectiveness of the participants in the two experimental groups is significantly enhanced, but the learning

effectiveness of the participants in the control group did not show a significant difference.

The use of the scaffolded assessment in the network peer assessment activities did not show a significant difference in the students' learning effectiveness. Even though the results show that learning effectiveness of the participants in the two experimental groups is significantly enhanced, there is no significant effect of providing the scaffolded assessment on the participants' learning effectiveness in the three groups. Thus, the effects of the use of the scaffolded assessment on students' learning effectiveness should be further investigated.

The participants all have a positive attitude toward the network peer assessment activities. The results of the questionnaires which address the participants' perceptions of network peer assessment show that the participants agree that network peer assessment is beneficial for learning, enhances the sense of participation and motivation, increases interactions among peers and improves learning effectiveness. The participants with other peers' evaluation highly agree the network assessment in compared to those without scaffolded assessment in the network peer assessment activities. In addition, the scaffolded assessment provides a solid basis for the participants when evaluating the peers' assignments. The participants could also gain more diverse ideas and insights in the network peer assessment activities. In the future, to enhance the validity of the research, the number of the subjects would be increased. Furthermore, learning retention can be considered to investigate the effectiveness of the proposed approach.

REFERENCES

- [1] C.-C. Chang and K.-H Tseng, "Using a web-based portfolio assessment system to elevate project-based learning performances," *Interactive Learning Environments*, vol. 19, no. 3, 2011, pp. 211-230.
- [2] C.-H. Chen, "The implementation and evaluation of a mobile self-and peer-assessment system," *Computers & Education*, vol. 55, no.1, 2010, pp. 229-236.
- [3] K. Cho, C. D. Schunn, and R. W. Wilson, "Validity and reliability of scaffolded peer assessment of writing from instructor and student perspectives," *Journal of Educational Psychology*, vol. 98, no.4, 2006, pp. 891.
- [4] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.
- [5] M. Freeman, "Peer assessment by groups of group work," *Assessment and Evaluation in Higher Education*, vol. 20, 1995, pp. 289-300.
- [6] G. Isaacs, "Brief briefing: Peer and self assessment," *The Effective Assessment Conference*, University of Queensland, 1998.
- [7] A. Jaillet, "Can online peer assessment be trusted?," *Journal of Educational Technology & Society*, vol. 12, no. 4, 2009.
- [8] R. L. Johnson, F. McDaniel and M. J. Willeke, "Using portfolios in program evaluation: An investigation of

- interrater reliability," *American Journal of Evaluation*, vol. 21, no. 1, 2000, pp. 65-80.
- [9] T. Knoy, S.-J. Lin, Z.-F. Liu, and S.-M. Yuan, "Networked peer assessment in writing: Copyediting skills instruction in an ESL technical writing course," 2001, Unpublished dissertation, National Chiao Tung University, Taiwan.
- [10] E. Z. F. Lin, C. H. Chiu, S. S. J. Lin, and S. M. Yuan, "Student participation in computer science courses via the Networked Peer Assessment System (NetPeas)," *Proceedings of the ICCE' 99*, 1999, pp. 774-777.
- [11] C. Rushton, P. Ramsey, and R. Rada, "Peer assessment in a collaborative hypermedia environment: A case-study," *Journal of Computer-Based Instruction*, vol. 20, 1993, pp. 75-80.
- [12] P. M. Sadler and E. Good, "The impact of self-and peer-grading on student learning," *Educational assessment*, vol. 11, no. 1, 2006, pp. 1-31.
- [13] Y.-T. Sung, K.-E. Chang, S.-K. Chiou, and H.-T. Hou, "The design and application of a web-based self-and peer-assessment system," *Computers & Education*, vol. 45, no. 2, 2005, pp. 187-202.
- [14] K. Topping, "Peer assessment between students in colleges and universities," *Review of Educational Research*, vol. 68, no. 3, 1998, pp. 249-276.
- [15] K. J. Topping and S. W. Ehly, "Peer assisted learning: A framework for consultation," *Journal of Educational and Psychological Consultation*, vol. 12, no. 2, 2001, pp. 113-132.
- [16] C.-C. Tsai and J.-C. Liang, "The development of science activities via on-line peer assessment: The role of scientific epistemological views," *Instructional Science*, vol. 37, no. 3, 2009, pp. 293-310.
- [17] C.-C. Tsai, S. S. Lin, and S.-M. Yuan, "Developing science activities through a networked peer assessment system," *Computers & Education*, vol. 38, no. 1, 2002, pp. 241-252.
- [18] S.-C. Tseng and C.-C. Tsai, "On-line peer assessment and the role of the peer feedback: A study of high school computer course," *Computers & Education*, vol. 49, no. 4, 2007, pp. 1161-1174.
- [19] L. S. Vygotsky, *Thought and language*: MIT press, 2012.
- [20] M. L. Wen and C.-C. Tsai, "University students' perceptions of and attitudes toward (online) peer assessment," *Higher Education*, vol. 51, no. 1, 2006, pp. 27-44.
- [21] D. Wood, J. S. Bruner, and G. Ross, "The role of tutoring in problem solving," *Journal of child psychology and psychiatry*, vol. 17, no. 2, 1976, pp. 89-100.
- [22] Y.-F. Yang and C.-C. Tsai, "Conceptions of and approaches to learning through online peer assessment," *Learning and Instruction*, vol. 20, no. 1, 2010, pp. 72-83.
- [23] Dougiamas, M., & Taylor, P. (2003). Moodle: Using learning communities to create an open source course management system. In *World conference on educational multimedia, hypermedia and telecommunications* (Vol. 2003, No. 1, pp. 171-178).
- [24] George, D. & Mallery, P. (2010). *SPSS for Windows step by step: A sample guide and reference 17.0 update*. Boston, Mass. : Allyn & Bacon.