Effects of a Peer Assessment System Based on a Grid-based Knowledge Classification Approach on Computer Skills Training

Author(s): Ting-Chia Hsu (A.K.A. Ching-Kun Hsu)

Source: *Journal of Educational Technology & Society*, Vol. 19, No. 4 (October 2016), pp. 100-111

Published by: International Forum of Educational Technology & Society

Stable URL: https://www.jstor.org/stable/10.2307/jeductechsoci.19.4.100

Hsu, T. C. (2016). Effects of a Peer Assessment System Based on a Grid-based Knowledge Classification Approach on Computer Skills Training. *Educational Technology & Society, 19* (4), 100–111.

Effects of a Peer Assessment System Based on a Grid-based Knowledge Classification Approach on Computer Skills Training

Ting-Chia Hsu (A.K.A. Ching-Kun Hsu)
Department of Technology Application and Human Resource Development, National Taiwan Normal University // ckhsu@ntnu.edu.tw

(Submitted April 26, 2015; Revised September 27, 2015; Accepted February 25, 2016)

ABSTRACT

In this study, a peer assessment system using the grid-based knowledge classification approach was developed to improve students' performance during computer skills training. To evaluate the effectiveness of the proposed approach, an experiment was conducted in a computer skills certification course. The participants were divided into three groups: one that learned with traditional instruction, one given the conventional peer-assessment system, and one using the proposed peer-assessment system. The results showed that the learning achievements of the students using the proposed system were significantly better than those of the other two groups. Therefore, integrating the knowledge engineering approach with the peer-assessment process can benefit students' learning, and help them attain computer skills certification. The dynamic peer assessment with knowledge classification approach is not only useful, but can also be repeatedly applied to different question sets of the certificate of computer software application.

Keywords

Peer assessment, Grid-based knowledge classification, Knowledge engineering

Introduction

Presently, most vocational students have significant experience using computers, and endeavor to pass the certification of the office word processing software application in their first year of senior vocational school in Taiwan. The skill evaluation center of the workforce development agency of the Council of Labor Affairs (i.e., CLA) conducts such testing. San et al. (2006) indicated that the licensing rate is an important index for measuring labor quality; moreover, those who pass the licensing tests are deemed to be more professionally qualified and competitive in the labor market (San et al., 2006). There is no implementation subject which every vocational school student has to learn besides the certification of computer software application, because fostering such competence is one of the essential employability skills in the workplace. In other words, there is a particular population with an urgent need to learn this topic. The certification includes two stages. The first assesses the students' academic knowledge, while the second assesses their practical computing skills with real tasks which must be completed within two hours using the functions of the application software for managing, editing, and positioning the required fonts, tables, paragraphs, format styles, word art, visual material, columns, and layout size. After passing both the academic and practical tasks, CLA will certify the students.

Knowing how to operate software does not necessarily mean one can correctly use it for practical applications. Learning to apply the available information and software to solve problems is vital for vocational high school students (Valtonen, Hacklin, Dillon, Vesisenaho, Kukkonen, & Hietanen, 2012; Chang, Tseng, & Lou, 2012). Scholars have noted that vocational students have relatively low academic achievement and less positive attitudes toward learning, resulting in a lack of higher-order thinking (Lee, Shen, & Tsai, 2010; Yang, 2015). Yang (2015) warned that low engagement and outdated approaches to instruction are the two main challenges of vocational education. Repeated operation without supporting students' thinking may result in a lack of problem-solving capability in the workplace (Yang, 2015). Therefore, it is necessary to provide these students with higher-order thinking learning activities. Simply learning a software's functionality in class does not elicit the students' full potential because they may not be able to correctly apply, analyze or evaluate the functions they have learned to solve the problems or complete the tasks they have been assigned. To deal with this phenomenon, this study considers a peer assessment approach which lets the students learn not only from the teacher's lessons but also by finding their peers' implementation errors (Race, 2001; Hsia, Huang, & Hwang, 2015). Peer assessment is one of the higher-order thinking learning activities (Lorente-Catalán & Kirk, 2016). Moreover, performing the active role of a reviewer can have a positive impact on one's own learning (Chinn, 2005). Therefore, a peer assessment system with a grid-based knowledge classification approach based on knowledge engineering was developed.

Scholars have indicated that the peer assessment process is challenging for students because they have difficulty knowing what and how to explain their evaluation; however, providing them with specific assistance and conducting knowledge acquisition are two main strategies for overcoming these problems (Lawrence & Zollinger, 2015). Therefore, the current study hypothesized that using a repertory grid method, one of the knowledge acquisition approaches in knowledge engineering, would help the students conduct peer assessment and learn from the process. Knowledge engineering refers to the process of developing expert systems which are artificial intelligence programs capable of imitating human decision-making processes on the basis of knowledge derived from domain experts. Among the existing approaches to knowledge acquisition, the repertory grid method, originating from the personal construct theory (Kelly, 1955), provides valuable assistance in the organization and presentation of domain experience and knowledge (Hsu, Hwang, & Chang, 2010; 2013). In a standard computer skills certification test, students need to identify and implement the functions required for the assigned tasks in a limited amount of time; therefore, they must completely understand the evaluation criteria of the computer system plus its individual functions. Consequently, referring to the work of others and receiving suggestions from others during the learning process is likely to be helpful, further supporting the use of peer-assessment strategies in school settings (Reid, 2000).

Although numerous studies have reported the effectiveness of peer assessment, some researchers have indicated that students might not benefit from it for several reasons (Chang & Tseng, 2011). For example, the assessment criteria and forms could be difficult to understand, causing them to spend much time deciphering the assessment items (Chang, Tseng, & Lou, 2012). Moreover, it could be difficult for them to evaluate and compare peers' work and summarize their findings without assistance or boundaries (Gijbels et al., 2005). Meanwhile, the scholars emphasized that finding ways to support and improve the question- or criteria-posing ability and performance of lower-achieving students is an important issue for instructors (Yu, Liu, & Chan, 2005). Additionally, there has been a call for support mechanisms to help make question-posing and peer assessment more manageable (Yu, Liu, & Chan, 2005). Because knowledge engineering helps people organize their acquired domain knowledge or expertise, it has been identified as a potential approach to help students organize what they have learned and experienced during their coursework (Liu & Lee, 2005). Therefore, in this study, the grid-based knowledge classification approach was used to develop a peer-assessment system for helping students evaluate their peers' performance. To appraise the effectiveness of this innovative approach, an experiment was conducted in a computer skills certification course at a vocational high school with three groups learning through different experimental treatments with or without a peer assessment tool. It was not certain that peer assessment would benefit the majority of students who have relatively low achievement without or with a supportive design (i.e., the grid-based knowledge classification approach). Therefore, three research questions were proposed as follows.

- Do the students make progress week by week after the experimental treatments?
- Do the students using the proposed approach outperform those using conventional approaches?
- Do the low-achievement students in the experimental group gain more benefits than those in the two control groups?

Literature review

Peer assessment

Falchikov (1995) defined peer assessment as mutual assessment of students in the same grade. Topping (1998) later proposed that it is a process by which students with similar levels of knowledge and learning backgrounds alternately assume the role of reviewers. By doing so, they can judge each other's learning outcomes. Topping further indicated that integrating technological advances into assessment has the potential to assist in carrying out peer assessments (Topping, 1998). Peer assessment was also conducted as part of the overall summative course assessment in an engineering education course (Hersam, Luna, & Light, 2004). It has been found to have high reliability and effectiveness (Ohland et al., 2005).

In the past decade, various web-based peer assessment systems have been applied in numerous domains, including computer science, science projects, and webpage design (Chen, 2010; Lin, Liu, & Yuan, 2001; Sung, Chang, Chiou, & Hou, 2005). Several studies have reported the benefits of peer assessment (Boud, Cohen, & Sampson, 2014; Kaufman & Schunn, 2011; Strijbos & Sluijsmans, 2010). For example, Alexander et al. (2008) compared the differences between traditional and peer assessment models and found that engaging students in peer-assessment activities improved their learning outcomes. Van Zundert, Sluijsmans, and van Merriënboer (2010) showed that peer assessment benefited the development of students' domain-specific skills via the evaluation, reflection, and revision cycle. Recently, one study showed that there was a positive correlation between the ratings given by teachers and peers, although the two groups interpreted the rubric criteria differently (De Grez, Valcke, & Roozen, 2012).

A peer assessment system entails consideration of peers' accomplishments, and requires students to use higher-order thinking skills such as analyzing, solving problems and evaluating phenomena rather than memorizing facts (Bloom, Englehart, Hill, & Krathwohl, 1956; Barak & Rafaeli, 2004; Sluijsmans et al., 2001; Zohar & Dori, 2003). Peer assessment is a higher-order thinking learning activity and was hypothesized to be a proper strategy for students to learn from checking others' work. However, researchers have indicated that without careful design of the assessment criteria and forms, it could be very difficult for students to understand, especially considering its purpose as a tool to enhance learning through clear peer-to-peer communication; hence peer assessment could in fact act as a learning barrier for students (Chang et al., 2012). Moreover, without assistance, students might be unable to organize and analyze their assessment results, implying the necessity and importance of developing effective peer-assessment strategies (Gijbels et al., 2005). This is why the current study proposes a dynamic repertory grid approach for a peer assessment system used in a computer skills class, allowing student reviewers to select personalized assessment criteria from a bank of items, causing them to reflect on the purpose of the assignment being reviewed and how to best characterize the performance of the student completing the assignment, as well as providing ample feedback for the students.

The grid-based knowledge classification approach in e-learning

In recent decades, researchers have proposed various knowledge engineering methods to help knowledge-based system developers acquire and organize knowledge from domain experts (Keynan, Ben-Zvi Assaraf, & Goldman, 2014; Siau, Tan, & Sheng, 2010). The repertory grid method proposed by Kelly (1955) is one of the most frequently adopted knowledge acquisition methods owing to its simplicity and clear structure (Malmström, Johansson, & Wincent, 2015; Kumar & Natarajan, 2010). Kelly (1955) suggested that people can create explanations of phenomena from their experiences, which can help them develop their understanding of underlying concepts and can be used as the basis of future judgments.

A repertory grid consists of three major components: elements, constructs, and ratings (Fransella, Bell, & Bannister, 2004). It can be viewed as a matrix with columns containing element labels and rows containing construct labels. Elements can be decisions to be made, objects requiring classification, or concepts to be learned (Chu et al., 2010). Constructs are features used to describe similarities or differences among elements. Each construct describes a trait and its opposite. A 5-point scale rating mechanism is commonly used to represent relationships between elements and constructs, where a score of 1 indicates that the element is very likely to have the trait, 2 indicates that it may have the trait, 3 corresponds to responses of "unknown" or "no relevance," 4 indicates that the element may possess the opposite trait, and 5 indicates that it very likely possesses the opposite trait (Fallman & Waterworth, 2010; Hsu et al., 2010; 2013).

Repertory grids have been employed to help researchers collect and organize domain knowledge for the development of expert systems, which, as previously defined, are artificial intelligence programs designed to simulate expert reasoning on the basis of knowledge elicited from domain experts (Keynan et al., 2014). Repertory grids have not only been used to integrate opinions from various experts to accumulate important data and preserve it in a database (Hoffman, Shadbolt, Burton, & Klein, 1995; Abdul-Gader & Kozar, 1990), but are also seen as an important mind-tool for e-learning (Hwang, Hung, Chen, & Liu, 2014). In our study, repertory grids are a knowledge classification approach for helping students organize their assessment results and illustrate their review judgements. This approach is introduced in the following section.

Method

Participants

A total of 224 tenth graders whose average age was 16, from six classes, participated in this study. They studied computer software applications for two hours a week in a senior vocational school. This was the first semester they had taken the computer software application certificate course. They were therefore beginners preparing for the Council of Labor Affairs certification. One class was assigned to be the experimental group, while the other two classes were assigned to be control groups 1 and 2. The experimental group did peer evaluation with the online peer assessment system based on the grid-based knowledge classification approach, while control group 2 used the conventional peer-assessment system. Every student in these two groups acted as a practitioner and a reviewer, and so had to assess one set of implementation work done by another student. Control group 1 did not perform peer assessment. All of the work of the groups was evaluated by the teacher each week. The six classes were taught by the same instructor who did the certification evaluation with another expert. Both of these teachers had more than ten years' vocational teaching experience.

The conventional peer-assessment system

Control group 2 was issued the assessment form used by the CLA as their peer-assessment form. Figure 1 shows the peer-assessment criteria used in the conventional CLA peer-assessment approach. The rows in the assessment form are fixed, preventing students from revising them. Thus, the peer-reviewers were only allowed to fill in the number of errors without any input into designating the assessment criteria.

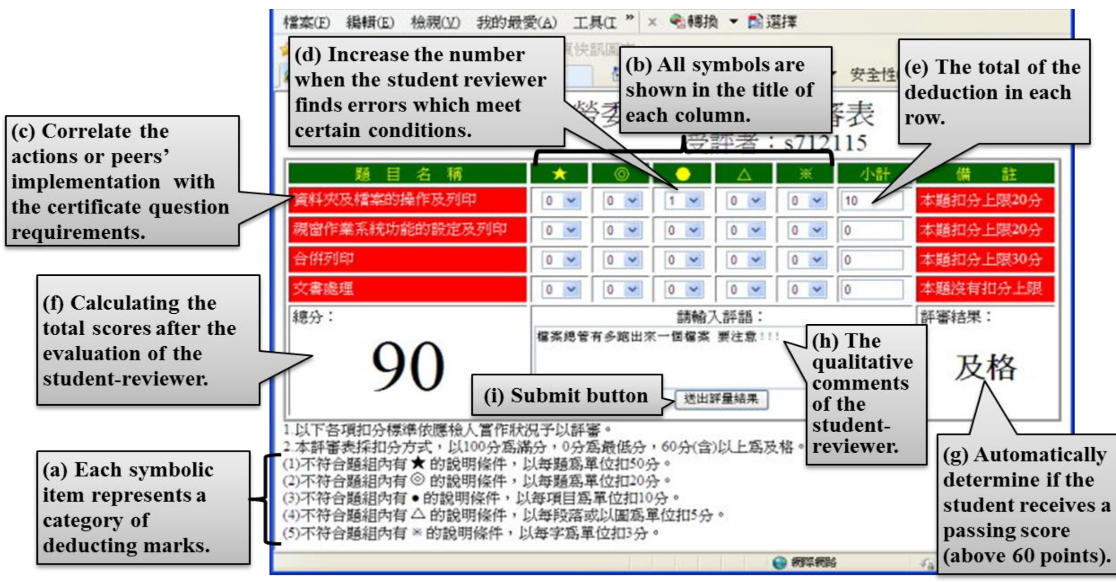

Figure 1. Assessment form used by control group 2

The computer application certificate includes the most popular Office suites, information processing, word-processing, and document formatting. Some users may think they know how to format a document; however, they may not use appropriate, correct or efficient functions. Are they aware of the mistakes they make during the test? This study conducted experiments to compare the effectiveness of the peer assessment systems and traditional instruction.

A peer assessment system based on a grid-based knowledge classification approach

An alternative system which is a repertory grid-based peer assessment tool, was developed for the experimental group. Two instructors specializing in computer software certification were invited to design an assessment form based on the repertory grid approach. Table 1 presents an illustrative example of a repertory grid in which the 10 elements (i.e., P_1, P_2, …, P_{10}) represent 10 paragraphs in an article edited by a student, and the 19 categories (C_1, C_2, …, C_{19}) represent a basic structure of peer-assessment assistance. In each category (e.g., paragraph format), a set of sub-items (e.g., indent in the first line of the paragraph and space between the lines in the paragraph) could be selected for assessing the paragraphs via a drop-down menu. The students were required to identify the relevance of each sub-item to the assignment they were assessing, and choose the items themselves when assessing their peers' work. In this way, a partially self-constructed assessment form was used during the assessment process. In this study, we attempted to provide the students with assistance when they conducted the higher-order thinking learning activity of peer assessment. Engaging them in identifying the importance of the sub-items and constructing their own assessment forms was intended to support their higher-order thinking, as suggested by several researchers (Barak & Dori, 2009; Van den Berg, 2004).

In the repertory grid-based assessment form, a rating of 1 means the student correctly employed or developed the function, showing that he/she had fully developed the computer skill described in the construct (e.g., data processing skill); 2 means that he/she is equipped with a sufficient but incomplete computer skill; 3 means that he/she has learned part of the skill; 4 means that he/she seems to have failed to use the appropriate computer skill to fulfill the requirements; 5 means that he/she has completely failed to learn the skill. Note that the peer-assessment process was conducted anonymously. Figure 2 presents an illustration of the repertory grid assessment form. Using the repertory grid developed for this study, the development procedure of the peer assessment activity is given as follows.

Table 1. Illustrative example of a repertory grid

Positive constructs	Elements					Opposite constructs
	P_1	P_2	P_3	…	P_{10}	
C1. Good paragraph format (e.g., first line of the paragraph is indented by 2 characters)	1	1	5	…	5	*C1'. Poor paragraph format* (e.g., first line of the paragraph is not indented by 2 characters)
C2, Good font format (e.g., using the standard font, such as Times New Roman)	5	1	2	…	4	*C2', Poor font format* (e.g., using various fonts or incorrect fonts)
:	:	:	:	…	:	*:*
C19. Good list and bullet format (e.g., the list and bullet format are used when necessary)	5	5	1	…	1	*C19'. Poor list and bullet format* (e.g., abusing the list and bullet format)

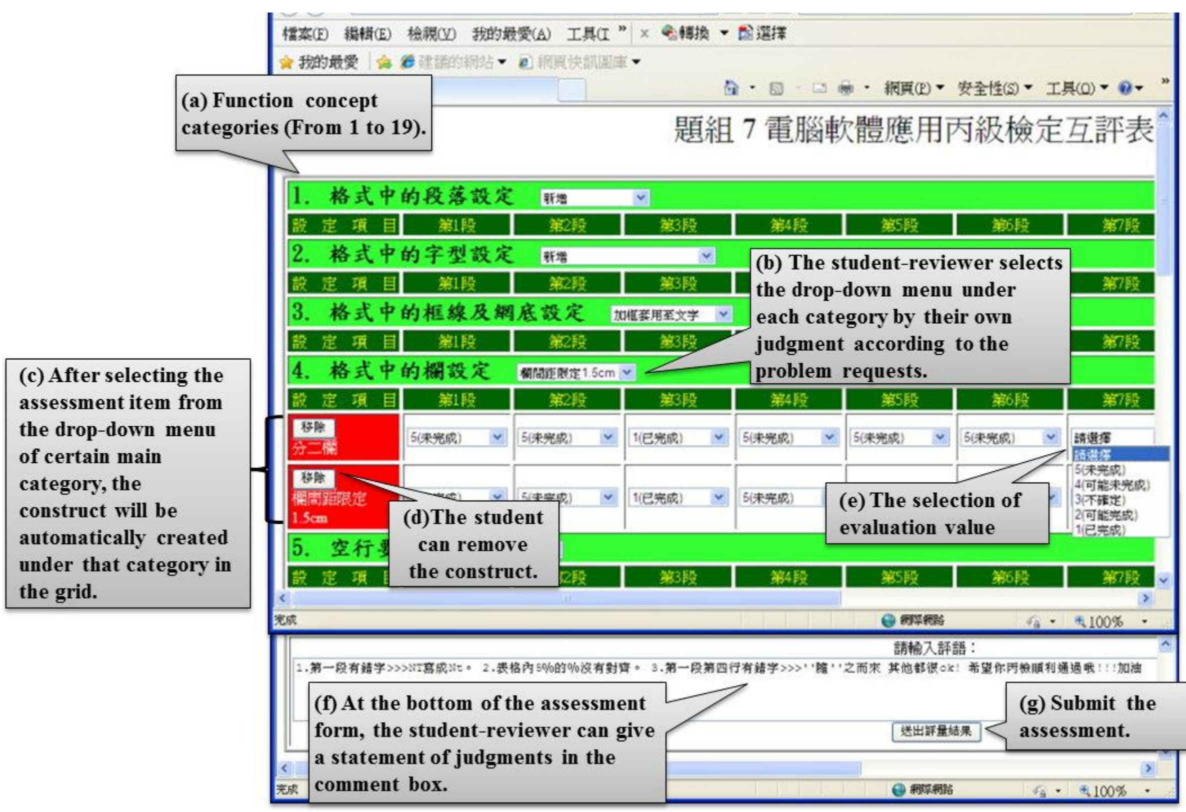

Figure 2. Experimental group assessment form

Step 1: Choose sub-items from the drop-down menu in each available category on the basis of the requirements of the certificate questions. After choosing one sub-item, a construct for evaluating each paragraph is produced, with the construct's description and its opposite shown on the left and right sides of the repertory grid, respectively. It is shown from (a) to (e) in Figure 2.

Step 2: Check the criteria of the learning task and select additional sub-items from the categories if needed.

Step 3: Evaluate each paragraph on the basis of the selected sub-items by filling in the rating, ranging from 1 to 5 in order of best to worst, for each <item, sub-item> entry.

Step 4: Provide qualitative comments at the end of the assessment form (shown as (f) in Figure 2).

Step 5: Submit the assessment results (shown as (g) in Figure 2).

Experimental design

Students in control group 2 and the experimental group participated in peer assessments every week, whereas students in control group 1 were assessed only by the instructor. The experiment was conducted over a period of six weeks. Each week the students did a certificate item set and then handed over their implementation. The first week named week 0 was the pretest, then the results were obtained in four weekly stages named week 1 to week 4, while the last week named week 5 was the posttest. The students in each class were further grouped into three achievement levels according to their achievement on the pretest.

Two independent variables were included in the experiment design: pretest achievement and experiment treatments (control and experimental groups). The covariate was the pretest scores representing the students' prior achievements. The experimental procedure is outlined in Figure 3. The dependent variable was the students' learning outcomes.

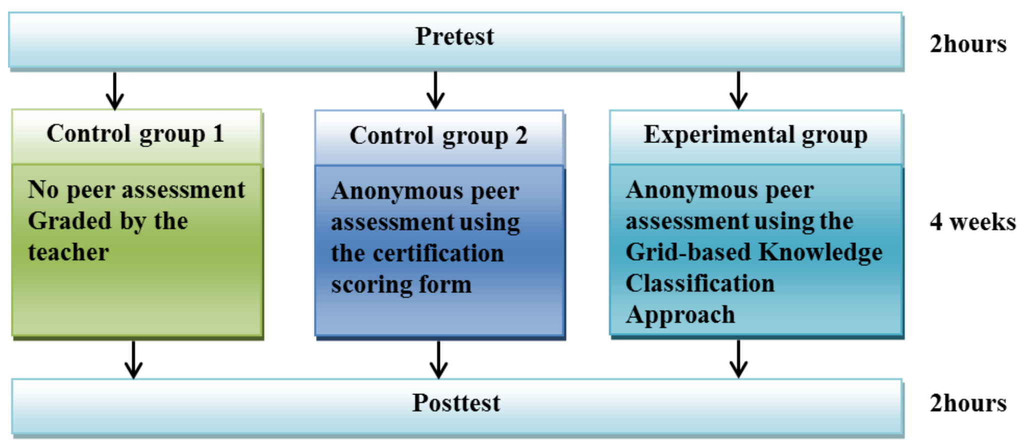

Figure 3. Experimental procedure

Data collection

The government certificate database includes 15 data processing projects (tests), all of which have been proved to have equal degrees of difficulty and discrimination. The 15 tests are comparable, so every student taking the official certification test is randomly assigned only one. The 15 tests involve similar required functions from the application software, although their implementation appearances are different. The six tests in this study including the pre-test at the beginning, the 4 tests during the following 4 weeks, and the post-test in the final week, were randomly selected from the 15 projects (tests).

Results

Weekly improvement

There was no significant difference among the three groups in the pre-test (F =0.229, p > .05). The experimental results show that the integration of peer assessment and knowledge engineering guided the students to focus on constructing solutions from the concepts they had learned and integrating their computer software skills to complete the tasks they were assigned. The achievements of the participants in each group are presented in Figure 4. The students in the first control group did not make progress in some weeks, and their performance in week 3 was not significantly better than in the pretest. The second control group made slow progress, indicating that peer assessment with an inappropriate assessment tool does not necessarily produce ideal results.

A 3 x 5 mixed ANCOVA with experimental treatments as a between-subjects variable, time-point (weeks 1-4 and posttest) as a within-subjects variable, and the pretest as the covariate was employed. The analysis allowed concise description of whether there was an effect of the experimental conditions across the time-points, or possibly interaction between the two. The results showed that there was no interaction between the two variables (F = 0.919, p > .05). The main effects of the tests across different weeks had significant differences (F = 10.775^{***}, p < .001). In other words, there was remarkable improvement among the weeks based on the evidence from the within-subjects effect test. After multiple comparisons, it was found that the effectiveness in the second and fourth week was significantly higher than that in the first week when the pre-test was the covariance. The means and standard deviations can also be seen in Table 2. Overall, the post-tests of each group were all remarkably higher than their pre-test.

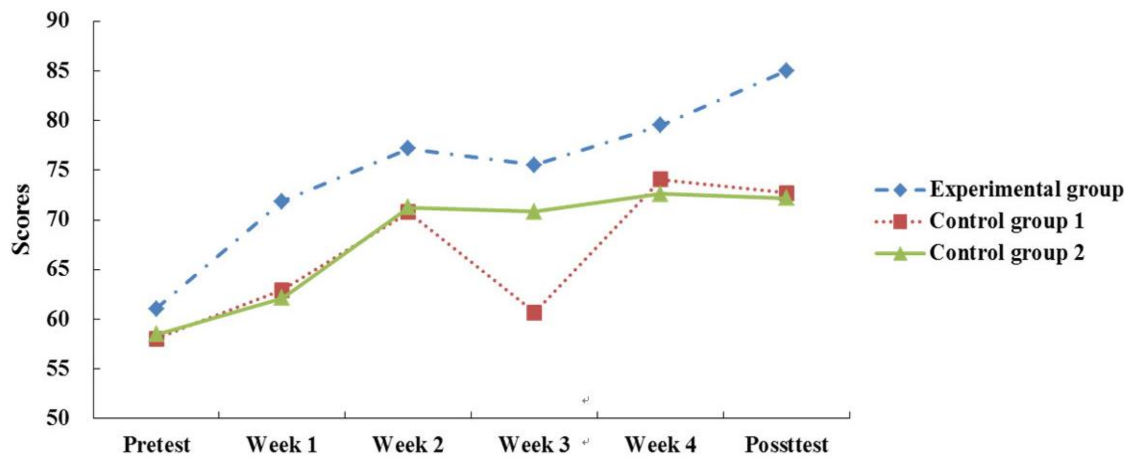

Figure 4. Weekly learning achievements of each group

Multiple comparisons among groups

Based on the results of the 3 x 5 mixed ANCOVA, the main effect of the experimental treatment was found to be significant (F = 11.552***; p < .001). In other words, the experimental group performed significantly better overall than control groups 1 and 2 after the paired comparisons. To explore the weekly progress and the difference among groups, the means, standard deviations and adjusted means are shown in Table 2.

From the tests of the between-group effects, listed in Table 2, there are significant differences in the post-test results of the three groups (F = 7.90***; p < .001). This made it necessary to conduct multiple comparisons, which revealed that the learning results of the experimental group were significantly better than those of the two control groups. However, no significant difference was observed between the learning achievements of the two control groups. In other words, the conventional peer assessment strategy did not result in better learning achievements in comparison with the traditional teaching method. In week 3, both the experimental group and control group 2 significantly outperformed control group 1 (F = 6.34*; p < .05). Moreover, it was found that the learning achievements of control group 1 approached those of control group 2 in the week 4 and post-test, shown as Table 2.

Table 2. The weekly adjusted mean after ANCOVA analysis

Time	Group	*N*	Mean	*SD*	Adjusted mean	*SE*	*F*	Pairwise comparisons
Week 1	Control group 1 (a)	76	62.92	30.28	63.72	2.78	2.34	
	Control group 2 (b)	69	62.13	30.02	62.64	2.92		
	Experimental group (c)	79	71.71	27.43	70.50	2.73		
Week 2	Control group 1 (a)	76	70.79	30.71	71.37	2.60	1.15	
	Control group 2 (b)	69	71.25	27.42	71.62	2.72		
	Experimental group (c)	79	77.16	16.85	76.28	2.55		
Week 3	Control group 1 (a)	76	60.67	31.03	61.16	2.82	6.34*	b > a*
	Control group 2 (b)	69	70.81	23.22	71.12	2.96		c > a*
	Experimental group (c)	79	75.56	24.23	74.82	2.77		
Week 4	Control group 1 (a)	76	74.08	20.54	74.54	2.07	2.13	c > b*
	Control group 2 (b)	69	72.64	23.66	72.94	2.18		
	Experimental group (c)	79	79.53	16.88	78.83	2.04		
Post-test	Control group 1 (a)	76	72.72	29.53	73.20	2.36	7.90***	c > a*
	Control group 2 (b)	69	72.20	22.84	72.51	2.47		c > b*
	Experimental group (c)	79	85.01	13.02	84.28	2.31		

Note. $^{*}p$ < .05; $^{***}p$ < .001.

The improvements of the low-achievement students

From the above-mentioned results, there was no significant difference among the three groups in the pre-test ($F = 0.229$, $p > .05$), but there was a remarkable main effect of the tests across the weeks. Overall, the students made progress in the post-test compared with the pre-test. The benefits which the low-achievement group gained with the assistance of the Grid-based Knowledge Classification Approach for higher-order thinking activities were further explored. Prior to the experiment, the students in each group were divided into three achievement levels on the basis of their pretest scores. In addition to no significant difference being found among the three groups in the pre-test overall ($F = 0.229$, $p > .05$), the low-achieving students ($F = 1.799$, $p > .05$) also displayed no significant difference across the three groups. After the experimental treatments, the low-achievement students in the experimental group gained significantly more than the low-achievement students in control group 1 did ($F = 4.696^{*}$, $p < .05$). However, there was no significant difference between the improvement of the low-achievement students in control groups 1 and 2. The low-achieving students in the experimental group showed the greatest progress over the course of the learning activity.

Table 3. ANOVA results of improvements in scores among the low-achievement students in the three groups

Tests	*N*	Mean of gain scores (Posttest-Pretest)	*F*	Comparisons
Control group 1 (a)	26	22.96	4.696^{*}	$c > a^{*}$
Control group 2 (b)	23	29.74		
Experimental group (c)	26	46.65		

Note. $^{*}p < .05$.

Discussion and conclusions

The results of this study indicated that the assessment form used by the second control group was unable to provide the same benefits as the grid-based knowledge classification approach used in the experimental group. Based on the study of Tseng and Tsai (2007), who conducted an experiment in a computer course in a senior high school and found that the students made significant progress after three rounds of peer-assessment for the same project, it was expected that every group in this study would make similar progress after several iterations of peer assessment. However, it was confirmed that the students who learned with the peer assessment system in control group 2 did not improve as much as those who learned with the peer-assessment approach in the experimental group. The conventional assessment system is the traditional tool which did not provide any further assistance. Both teachers and students had used this traditional assessment form before. After comparing the grid-based knowledge classification approach with the conventional approach, it was found that the former supported the students in conducting higher-order thinking learning activities and in making greater advances as a result of the peer assessment. It is reasonable to infer that the assessment form used in the conventional assessment system was too straightforward, because little thought is required to fill it out; thus, the students did not need to reflect much on what they had learned, resulting in the limited development of their higher-order thinking ability. Despite the fact that the repertory grid can produce peer assessment forms that include partially self-selected items, the students were shown to learn a great deal through construct selection for the evaluation of others' work. From the evidence of the system logs, it was found that the students in the experimental group selected almost all the possible assessment items (mean = 122) from the selection pools in the top-down menu (i.e., totally 143 choices from 19 categories) in the first week. However, from the third week, they had more confidence and selected only necessary assessment items based on the task requirements. Therefore, the number of assessment items selected stabilized at around half that of the first week (mean = 67). It is thus inferred that the students in the experimental group developed related evaluation thinking ability and learned from their experiences as well as from the mistakes of others week by week. In contrast, the students in the first control group had to accumulate experience by themselves.

The students in the experimental group did not necessarily have the ability to appropriately use the assessment form. Due to the design of the experimental system, when the students in the experimental group were unsure of their judgments, they were able to choose "may" (i.e., "2") or "may not" (i.e., "4") as a reference for their peers. As the data remain in the system, the teacher was able to identify the students' weaknesses. As the system used by the experimental group left indistinct choices for the learners, those choices would remind them of the areas where their peer reviewer was not certain, and the student-reviewers would have to strengthen their own knowledge and skills related to that target. Consequently, the grid-based knowledge classification approach in this study provided the students with a structured and reflective peer assessment system. Although considerable research on the application of repertory grids in peer assessment has been performed (Liu & Lee, 2005), this study proposes a more flexible instrument than those previously presented. The repertory grid used in this study differs from prior instruments (Liu & Lee, 2005; Liu & Tsai, 2005; Tsai, Lin & Yuan, 2002) which offered a fixed or dynamic grid without classification assistance. The novel contribution of this study is the use of a dynamic repertory grid assessment form to aid learning activities which prepare students for the computer software application certificate. The term "dynamic" here means the students independently chose evaluation items from the drop-down menu under each fixed category. However, the limitation of this study is that the assessment content was designed for one particular certificate, although it is the one which most vocational school students aim to acquire. Another innovation of this study is that it takes the students' ability into consideration. The instructors designed 19 fixed categories in advance, and offered a pool of several sub-items in the drop-down menu in each category on the basis of the function of the computer application software. The students selected evaluation items provided by the experts from the drop-down menu, which caused them to think analytically about how to best evaluate another's work and better comprehend the requirements of the set problem to be able to choose proper evaluation items. Boundaries for their choices are thus provided; otherwise, novices may have difficulty establishing the evaluation items independently.

Students are guided to reflect deeply on their learning when they add an assessment item (i.e., a construct in the repertory grid), revise or delete a construct, and compare different portions of a task to assess their peer's implementation. Previous research indicates that awareness and reflection can help develop students' metacognition, enhancing their learning and creative abilities (Burleson, 2005). Therefore, in the experimental group, the students were able to construct assessment items to enhance their knowledge and experience reflective thinking through peer assessment. Scholars have remarked that technology can be used as a kind of mind-tool to support students in experiencing reflective thinking that is necessary for meaningful learning (Jonassen & Carr, 2000). The experimental approach left an irresolute space for the students, and the indistinct choices caught the attention of both the students and the student-reviewers. Peer assessment in the experimental group promoted the students' self-awareness and active role when they selected assessment items specific to the evaluation of each peer's project that they reviewed, and used these items to evaluate their peers' implementation, enhancing their reflective thinking through this semi-self-constructed assessment. This research not only helped the students prepare for the certification of computer software application, but also cultivated their higher-order thinking skills (Forehand, 2010) which may be transferred to other implementation practices in the future (Yang, 2015).

The current results show that the achievements of the experimental group were significantly better than those of the two control groups. The low-achieving students in the experimental group benefited the most in this study and achieved the same target as that of a previous study (Yu, Liu, & Chan, 2005). The peer assessment model could be repeatedly used with many question sets like the projects of the computer software application certificate, which differs from the previous experiment which only used peer assessment for one project in the computer course (Tseng & Tsai, 2007). Collaborating with peers may achieve a level which studying alone cannot reach; this improvement is often referred to as the Zone of Proximal Development (ZPD) (Hung & Chen, 2001). Therefore, future studies could extend these methods to include collaborative peer assessment. When students are grouped heterogeneously to use the peer assessment system, they would have a partner to negotiate and discuss which assessment items to choose based on the requirements of the task when evaluating others. It is worth exploring whether such collaborative peer assessment would be more effective.

In the current study, the repeated iteration of doing peer assessment for one task per week accumulated and transferred the skills which the students learned and reflected on to the following tasks. We are planning to apply the dynamic repertory grid method to the peer-assessment activities of other subjects. We also hope to use serial analysis to code and analyze the collected qualitative data (i.e., the students' comments to peers) to determine whether different peer assessment patterns can be shown to cause different levels of thinking or diverse feedback styles in the innovative assessment mechanism proposed by this study. Finally, allowing students to take part in discussing and setting the criteria for peer assessment may be a feasible way to enhance their learning in computer-related courses in the future (Lai & Hwang, 2015).

Acknowledgements

This study is supported in part by the Ministry of Science and Technology in Taiwan under contract numbers: MOST 103-2628-S-003-003-MY2 and MOST 104-2511-S-003-034.

References

Abdul-Gader, A. H., & Kozar, K. A. (1990). Discourse analysis for knowledge acquisition: The Coherence method. *Journal of Management Information Systems*, 61-82.

Alexander, P., Van Loggerenberg, J., Lotriet, H., & Phahlamohlaka, J. (2010). The Use of the repertory grid for collaboration and reflection in a research context. *Group Decision and Negotiation, 19*(5), 479-504.

Barak, M., & Dori, Y. J. (2009). Enhancing higher order thinking skills among inservice science teachers via embedded assessment. *Journal of Science Teacher Education, 20*(5), 459-474.

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*(1), 84-103.

Bloom, B., Englehart, M. F. E., Hill, W., & Krathwohl, D. (1956). *Taxonomy of educational objectives: The Classification of educational goals. Handbook I: Cognitive domain*. New York, NY: Longmans Green.

Boud, D., Cohen, R., & Sampson, J. (Eds.). (2014). *Peer learning in higher education: Learning from and with each other*. New York, NY: Routledge.

Burleson, W. (2005). Developing creativity, motivation, and self-actualization with learning systems. *International Journal of Human-Computer Studies, 63*, 436–451.

Chang, C. C., & Tseng, K. H. (2011). Using a web-based portfolio assessment system to elevate project-based learning performances. *Interactive Learning Environments, 19*(3), 211-230.

Chang, C. C., Tseng, K. H., & Lou, S. J. (2012). A Comparative analysis of the consistency and difference among teacher-assessment, student self-assessment and peer-assessment in a Web-based portfolio assessment environment for high school students. *Computers & Education, 58*(1), 303-320.

Chen, C. H. (2010). The Implementation and evaluation of a mobile self-and peer-assessment system. *Computers & Education, 55*, 229–236.

Chinn, D. (2005). Peer assessment in the algorithms course. *ACM Special Interest Group on Computer Science Education (SIGCSE) Bulletin, 37*(3), 69-73.

Chu, H. C., Hwang, G. J., & Tseng, Judy C. R. (2010). An Innovative approach for developing and employing electronic libraries to support context-aware ubiquitous learning. *The Electronic Library, 28*(6), 873-890.

De Grez, L., Valcke, M., & Roozen, I. (2012). How effective are self-and peer assessment of oral presentation skills compared with teachers' assessments? *Active Learning in Higher Education, 13*(2), 129-142.

Falchikov, N. (1995). Peer feedback marking: developing peer assessment. *Programmed Learning, 32*(2), 175-187.

Fallman, D., & Waterworth, J. (2010). Capturing user experiences of mobile information technology with the repertory grid technique. *Human Technology, 6*(2), 250-268.

Forehand, M. (2010). Bloom's taxonomy. In M. Orey (Ed.), *Emerging perspectives on learning, teaching, and technology* (pp. 41-47).

Fransella, F., Bell, R., & Bannister, D. (2004). *A Manual for repertory grid technique*. West Sussex, UK: John Wiley & Sons.

Gijbels, D., van de Watering, G., & Dochy, F. (2005). Integrating assessment tasks in a problem-based learning environment. *Assessment & Evaluation in Higher Education, 30*(1), 73-86.

Hersam, M. C., Luna, M. & Light, G. (2004). Implementation of interdisciplinary group learning and peer assessment in a nanotechnology engineering course. *Journal of Engineering Education, 93*(1), 49-57.

Hsia, L. H., Huang, I., & Hwang, G. J. (2015). A Web-based peer-assessment approach to improving junior high school students' performance, self-efficacy and motivation in performing arts courses. *British Journal of Educational Technology, 47*(4), 618-632. doi:10.1111/bjet.12248

Hsu, C. K., Hwang, G. J., & Chang, C. K. (2010). Development of a reading material recommendation system based on a knowledge engineering approach. *Computers & Education, 55*(1), 76-83.

Hsu, C. K., Hwang, G. J., & Chang, C. K. (2013). A Personalized recommendation-based mobile learning approach to improving the reading performance of EFL students. *Computers & Education, 63*, 327-336.

Hoffman, R. R., Shadbolt, N. R., Burton, A. M., & Klein, G. (1995). Eliciting knowledge from experts: A Methodological analysis. *Organizational behavior and human decision processes, 62*(2), 129-158.

Hung, D. W., & Chen, D. T. (2001). Situated cognition, Vygotskian thought and learning from the communities of practice perspective: Implications for the design of web-based e-learning. *Educational Media International*, *38*(1), 3-12.

Hwang, G. J., Hung, P. H., Chen, N. S., & Liu, G. Z. (2014). Mindtool-Assisted In-Field Learning (MAIL): An Advanced ubiquitous learning project in Taiwan. *Journal of Educational Technology & Society, 17*(2), 4-16.

Jonassen, D. H. & Carr, C. S. (2000). Mindtools: Affording multiple knowledge representations for learning. In S. P. Lajoie (Ed.), *Computers as cognitive tools. No morewalls* (Vol. 2, pp. 165–196). Hillsdale, NJ: Lawlence Erlbaum Associates, Inc.

Kaufman, J. H., & Schunn, C. D. (2011). Students' perceptions about peer assessment for writing: Their origin and impact on revision work. *Instructional Science, 39*(3), 387-406.

Keynan, A., Ben-Zvi Assaraf, O., & Goldman, D. (2014). The Repertory grid as a tool for evaluating the development of students' ecological system thinking abilities. *Studies in Educational Evaluation, 41*, 90–105.

Kelly, G. A. (1955). *The Psychology of personal construct*. New York, NY: Norton.

Kumar, S., & Natarajan, B. R. (2010). Facilitating knowledge acquisition methodologies for a knowledge-based decision support system. *Artificial Intelligent Systems and Machine Learning, 2*(2), 27-32.

Lai, C. L., & Hwang, G. J. (2015). An Interactive peer-assessment criteria development approach to improving students' art design performance using handheld devices. *Computers & Education, 85*, 149-159.

Lawrence, J., & Zollinger, S. W. (2015). Assessment matters: Enriching design education through online peer critique. *The Journal of Effective Teaching*, *15*(2), 78-86.

Lee, T. H., Shen, P. D., & Tsai, C. W. (2010). Enhance low-achieving students' learning involvement in Taiwan's higher education: an approach via e-learning with problem based learning and self-regulated learning. *Teaching in Higher Education, 15*(5), 553-565.

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

Liu, C. C., & Lee, J. H. (2005). Prompting conceptual understanding with computer-mediated peer discourse and knowledge acquisition techniques. *British Journal of Educational Technology, 36*(5), 821-837.

Liu, C. C., & Tsai, C. M. (2005). Peer assessment through web-based knowledge acquisition: tools to support conceptual awareness. *Innovations in Education and Teaching International, 42*(1), 43-59.

Lorente-Catalán, E., & Kirk, D. (2016). Student teachers' understanding and application of assessment for learning during a physical education teacher education course. *European Physical Education Review, 22*(1), 65-81.

Malmström, M., Johansson, J., & Wincent, J. (2015). Cognitive constructions of low-profit and high-profit business models: A Repertory grid study of serial entrepreneurs. *Entrepreneurship Theory and Practice, 39*(5), 1083-1109.

Ohland, M. W., Layton, R. A., Loughry, M. L., & Yuhasz, A. G. (2005). Effects of behavioral anchors on peer evaluation reliability. *Journal of Engineering Education*, *94*(3), 319-326.

Race, P. (2001). *A Briefing on self, peer & group assessment* (Vol. 9). York, UK: Learning and Teaching Support Network.

Reid, J. (2000). A Task-oriented non-interactive evaluation methodology for information retrieval systems. *Information Retrieval, 2*(1), 115-129.

San, G., Huang, T. C., & Huang, L. H. (2006). The Establishment and application of a labor quality index: The Case of Taiwan's manufacturing industry. *Social Indicators Research, 79*(1), 61-96.

Siau, K., Tan, X., & Sheng, H. (2010). Important characteristics of software development team members: An Empirical investigation using repertory grid. *Information Systems Journal, 20*(6), 563-580.

Sluijsmans, D., Moerkerke, G., van-Merrienboer, J.,& Dochy, F., (2001). Peer assessment in problem based learning. *Studies in Educational Evaluation, 27,* 153–173.

Strijbos, J. W., & Sluijsmans, D. (2010). Unravelling peer assessment: Methodological, functional, and conceptual developments. *Learning and Instruction, 20*(4), 265-269.

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

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

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

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*(4), 1161-1174.

Valtonen, T., Hacklin, S., Dillon, P., Vesisenaho, M., Kukkonen, J., & Hietanen, A. (2012). Perspectives on personal learning environments held by vocational students. *Computers & Education, 58*(2), 732-739.

Van den Berg, G. (2004). The Use of assessment in the development of higher-order thinking skills. *Africa Education Review, 1*(2), 279-294. doi:10.1080/18146620408566285

Van Zundert, M., Sluijsmans, D., & Van Merriënboer, J. (2010). Effective peer assessment processes: Research findings and future directions. *Learning and Instruction, 20*(4), 270-279.

Yang, Y. T. C. (2015). Virtual CEOs: A Blended approach to digital gaming for enhancing higher order thinking and academic achievement among vocational high school students. *Computers & Education, 81*, 281-295.

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.

Zohar, A., & Dori, Y. J., (2003). Higher order thinking skills and low achieving students: are they mutually exclusive? *The Journal of the Learning Sciences, 12*(2), 145–182.