Educational Technology & Society
An International Journal

Aims and Scope
Educational Technology & Society is a quarterly journal published in January, April, July and October. Educational Technology & Society seeks academic articles on the issues affecting the developers of educational systems and educators who implement and manage such systems. The articles should discuss the perspectives of both communities and their relation to each other:

- Educators aim to use technology to enhance individual learning as well as to achieve widespread education and expect the technology to blend with their individual approach to instruction. However, most educators are not fully aware of the benefits that may be obtained by proactively harnessing the available technologies and how they might be able to influence further developments through systematic feedback and suggestions.
- Educational system developers and artificial intelligence (AI) researchers are sometimes unaware of the needs and requirements of typical teachers, with a possible exception of those in the computer science domain. In transferring the notion of a “user” from the human-computer interaction studies and assigning it to the “student,” the educator’s role as the “implementer/ manager/ user” of the technology has been forgotten.

The aim of the journal is to help them better understand each other’s role in the overall process of education and how they may support each other. The articles should be original, unpublished, and not in consideration for publication elsewhere at the time of submission.

Educational Technology & Society and three months thereafter.

The scope of the journal is broad. Following list of topics is considered to be within the scope of the journal:


Founding Editor
Kynthia, University of North Texas, USA.

Editors
Nian-Shing Chen, National Sun Yat-sen University, Taiwan; Demetrios G Sampson, Curtin University, Australia.

(Term ended in January, 2019)

Advisory Board
Ignacio Aedo, Universidad Carlos III de Madrid, Spain; Mohamed Alby, Athabasca University, Canada; Luis Anido-Rifon, University of Vigo, Spain; Gautam Biswas, Vanderbilt University, USA; Rosa Maria Bottino, Consiglio Nazionale delle Ricerche, Italy; Mark Bullen, University of British Columbia, Canada; Tak-Wai Chan, National Central University, Taiwan; Kuo-En Chang, National Taiwan Normal University, Taiwan; Ni Chang, Indiana University South Bend, USA; Yam San Chee, Nanyang Technological University, Singapore; Sherry Chen, Brunel University, UK; Bridget Cooper, University of Sunderland, UK; Darina Dicheva, Winston-Salem State University, USA; Jon Dron, Athabasca University, Canada; Michael Eisenberg, University of Colorado; Boulder, USA; Robert Farrell, IBM Research, USA; Brian Garner, Deakin University, Australia; Tiong Goh, Victoria University of Wellington, New Zealand; Mark D. Gross, Carnegie Mellon University, USA; Roger Hartley, Leeds University, UK; J R Isaac, National Institute of Information Technology, India; Mohamed Jenni, University of Tunis, Tunisia; Mike Joy, University of Warwick, United Kingdom; Athanasis Karoulis, Hellenic Open University, Greece; Paul Kirshner, Open University of the Netherlands, The Netherlands; William Klemm, Texas A&M University, USA; Rob Koper, Open University of the Netherlands, The Netherlands; Jimmy Ho Man Lee, The Chinese University of Hong Kong, Hong Kong; Rudy Lelouche, Université Laval, Canada; Tzu-Chien Liu, National Central University, Taiwan; Rory McGreal, Athabasca University, Canada; David Merrill, Brigham Young University - Hawaii, USA; Marcelo Milrad, Växjö University, Sweden; Riichiro Mizoguchi, Osaka University, Japan; Permanand Mohan, The University of the West Indies, Trinidad and Tobago; Kiyoshi Nakabayashi, National Institute of Multimedia Education, Japan; Hiroaki Ogata, Tokushima University, Japan; Toshio Okamoto, the University of Electro-Communications, Japan; Jose A. Pino, University of Chile, Chile; Thomas C. Reeves, The University of Georgia, USA; Norbert M. Seel, Albert-Ludwigs-University of Freiburg, Germany; Timothy K. Shih, Tamkang University, Taiwan; Yoshiaki Shindo, Nippon Institute of Technology, Japan; Kevin Singley, IBM Research, USA; J. Michael Spector, Florida State University, USA; Slavi Stoyanov, Open University, The Netherlands; Timothy Teo, Nanyang Technological University, Singapore; Chin-Chung Tsai, National Taiwan Normal University, Taiwan; Jie Chi Yang, National Central University, Taiwan; Stephen J. H. Yang, National Central University, Taiwan; Yu-Mei Wang, University of Alabama at Birmingham, USA; Ashok Patel, CAL Research & Software Engineering Centre, UK; Reinhard Oppermann, Fraunhofer Institut Ange wandte Informationstechnik, Germany; Vladimir A Fomichev, K. E. Tsiolkovsky Russian State Tech Univ, Russia; Olga S Fomicheva, Studio “Culture, Ecology, and Foreign Languages,” Russia; Piet Kommers, University of Twente, The Netherlands; Chul-Hwan Lee, Inchon National University of Education, Korea; Brent Muirhead, University of Phoenix Online, USA; Erkki Sutinen, University of Joensuu, Finland; Vladimir Uskov, Bradley University, USA.

Editorial Assistant
She Wei (Sylvia) Chew, National Sun Yat-sen University, Taiwan.

Technical Manager
Jie Peng, National Taiwan Normal University, Taiwan

Executive Peer-Reviewers
see http://www.j-ets.net

Publisher
National Taiwan Normal University, Taiwan

Editorial Office
c/o Dr Nian-Shing Chen, Taiwan E-Learning and Digital Content Association, National Taiwan Normal University, 129, Sec. 1, Heping E. Rd., Education Building, Rm. 611, Taipei 10644, Taiwan.
Supporting Organizations

Center for Cognition and Gesture-based Computing, National Sun Yat-sen University, Taiwan
Taiwan E-Learning and Digital Content Association (TELDCA), Taiwan
Information Technologies Institute, Centre for Research & Technology Hellas (C.R.T.H), Greece
University of Piraeus, Greece

Subscription Prices and Ordering Information

For subscription information, please contact the editors at ets.editors@gmail.com

Advertisements

Educational Technology & Society accepts advertisement of products and services of direct interest and usefulness to the readers of the journal, those involved in education and educational technology. Contact the editors at ets.editors@gmail.com

Abstracting and Indexing


Guidelines for authors

Submissions are invited in the following categories:

- Peer reviewed publications: Full length articles (4000 - 7000 words)
- Book reviews
- Software reviews
- Website reviews

All peer review publications will be refereed in double-blind review process by at least two international reviewers with expertise in the relevant subject area. Book, Software and Website Reviews will not be reviewed, but the editors reserve the right to refuse or edit review.

For detailed information on how to format your submissions, please see: https://www.j-ets.net/ETS/guide.html

Submission procedure

Authors, submitting articles for a particular special issue, should send their submissions directly to the appropriate Guest Editor. Guest Editors will advise the authors regarding submission procedure for the final version.

All submissions should be in electronic form. The editors will acknowledge the receipt of submission as soon as possible.

The preferred formats for submission are Word document and RTF, but editors will try their best for other formats too. For figures, GIF and JPEG (JPG) are the preferred formats. Authors must supply separate figures in one of these formats besides embedding in text.

Please provide following details with each submission: • Author(s) full name(s) including title(s), • Name of corresponding author, • Job title(s), • Organisation(s), • Full contact details of ALL authors including email address, postal address, telephone and fax numbers.

In case of difficulties, please contact ets.editors@gmail.com (Subject: Submission for Educational Technology & Society journal).
### Table of contents

**Full Length Articles**

1. **Using Online Peer Feedback through Blogs to Promote Speaking Performance**  
   Hui-Chin Yeh, Sheng-Shiang Tseng and Yu-Sheng Chen  
   1–14

2. **Online Targeting Behavior of Peer-Assessors under Identity-Revealed, Nicknamed, and Concealed Modes**  
   Fu Yun Yu and Shannon Sung  
   15–27

3. **Partnership among Schools in E-Learning Implementation: Implications on Elements for Sustainable Development**  
   Siu Cheung Kong  
   28–43

4. **Self-Regulated Learning for Web-Enhanced Control Engineering Education**  
   Flavio Manganello, Carla Falsetti and Tommaso Leo  
   44–58

5. **Why and How Serious Games can Become Far More Effective: Accommodating Productive Learning Experiences, Learner Motivation and the Monitoring of Learning Gains**  
   Wim Westera  
   59–69

6. **Do Learning Styles Matter? Motivating Learners in an Augmented Geopark**  
   Tien-Chi Huang, Mu-Yen Chen and Wen-Pao Hsu  
   70–81

7. **Investigating Remote Access Laboratories for Increasing Pre-service Teachers’ STEM Capabilities**  
   Ting Wu and Peter Albion  
   82–93

8. **Fostering EFL teachers’ CALL Competencies Through Project-based Learning**  
   Sheng-Shiang Tseng and Hui-Chin Yeh  
   94–105

9. **Using a Bilingual Concordancer for Text Revisions in EFL Writing**  
   Yu-Fen Yang, Gwo-Haur Hwang and Ruey-Fen Harn  
   106–119

10. **Effects of Gamified Comparison on Sixth Graders’ Algebra Word Problem Solving and Learning Attitude**  
    Hong-Zheng Sun-Lin and Guey-Fa Chiou  
    120–130
Using Online Peer Feedback through Blogs to Promote Speaking Performance

Hui-Chin Yeh¹, Sheng-Shiang Tseng²* and Yu-Sheng Chen³

¹Graduate School of Applied Foreign Languages, National Yunlin University of Science & Technology, Yunlin, Taiwan // ²Graduate Institute of Curriculum and Instruction, Tamkang University, New Taipei City, Taiwan // ³National Jhuo-Lan Senior High School, Miaoli // hyeh@yuntech.edu.tw // u9241346@gmail.com // m10141008@yuntech.org.tw

*Corresponding author

(Submitted July 6, 2016; Revised November 22, 2016; Accepted September 13, 2017)

ABSTRACT

To extend the recent growth of literature on using peer feedback through blogs to enhance students’ speaking performance, this study investigated the effects of online peer feedback via blogs on the speaking performance of college students studying English as a Foreign Language (EFL). Participants comprised 45 EFL college students, from two classes English Conversation and English Listening and Speaking, who were required to practice English speaking by recording a series of video clips and giving/receiving peer feedback on their speaking performance. Students also reflected on their experiences at the end of the semester. The collected data included the students’ scores on their first and final video clips, their blog entries, and their self-reflection sheets. Based on the differences in scores on their first and final clips, the students were classified into groups who made more progress (MP) and less progress (LP) respectively. After receiving peer feedback through blogs, only the MP group showed significant progress in the development of the content of their videos, including introduction, supporting points and conclusions while both groups showed significant improvement in the delivery area except for vocabulary use and grammar. It was also found that those responding more actively to peers’ problematic feedback gained more progress in the revised clips. Several pedagogical implications are also discussed.

Keywords

Online peer feedback, Video Blogs, Speaking performance

Introduction

The use of peer feedback in language learning has received growing attention in student-centered and collaborative learning studies. Liu and Carless (2006) defined peer feedback as “a communication process through which learners enter into dialogues related to performance and standards” (p. 280). Whereas the sole reliance on teacher feedback might lead students into passive learning (Lee, 2008), peer feedback encourages learner participation and fosters communicative competence through two-way interaction (Hyland, 2003). When peer feedback is based on given criteria, it helps both parties recognize strengths and weaknesses in their performance and thus develop an awareness of the qualities of good performance.

During the past few decades, Computer-Mediated Communication (CMC) technologies have provided new affordances for peer feedback exchange. Indeed, computer-mediated peer feedback, also referred to as online feedback, has been identified as having advantages over traditional face-to-face oral or written peer feedback in that it (1) overcomes time-and-place constraints, (2) reduces students’ discomfiture arising from face-to-face critique, (3) alleviates the anxiety of giving immediate responses, and (4) gives teachers access to monitor all students’ discussion about peers’ language performance (DiGiovanni & Nagaswami, 2001; Ho & Savignon, 2007; Liou & Peng, 2009; Liu & Sadler, 2003; Tuzi, 2004; Warschauer & Ware, 2006). Scholars have further suggested that asynchronous CMC tools like blogs, which allow time for thoughtful review and constructive peer feedback, were more beneficial for language learning (e.g., Dippold, 2009; Liou & Peng, 2009; Pham & Usaha, 2016) than traditional face-to-face and synchronous CMC peer feedback.

Although scholars have highlighted the potential values of peer feedback using blogs in language learning, in particular by investigating its effects on students’ writing (e.g., Cifci & Kocoglu, 2012; Dippold, 2009; Liou & Peng, 2009; Pham & Usaha, 2016), the possibilities for enhancing students’ speaking performance via blogs remains underexplored. This study therefore investigated the effects on peer feedback that arose through the use of blogs on EFL college students’ speaking performance. Four research questions were addressed:

- What were the effects of blog-supported peer feedback on non-English major students’ speaking performance?
- How did the effects on content and delivery items differ between higher and lower performing non-English major students via blogs?
• How did peer feedback impact the students’ speaking performances of higher and lower performing non-English major students via blogs?
• What were non-English major students’ receptions of and follow-up responses to problematic feedback via blogs?

Literature review

Theoretical framework for peer feedback

The use of peer feedback draws upon the theoretical underpinning of social-constructivism. Instead of focusing solely on individuals’ learning processes, social constructivism asserts that knowledge is constructed by individuals when they interact with others socially (Swan, 2002). Through social interactions, individuals can gradually accumulate and internalize knowledge constructed collaboratively with other participants (Kanuka & Anderson, 2007; Bonk & Cunningham, 1998). To this end, having students provide feedback on peers’ work prompts them to generate knowledge through meaningful interactions. Peer feedback activities are also supported by Vygotsky’s (1978) zone of proximal development (ZPD), described as “the process distance between the actual development level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers” (p. 86). Similar to Vygotsky’s ZPD, scaffolding is usually construed as an interactive process by which less capable learners are guided by more capable ones (e.g., teachers or peers) to advance in knowledge or skills (Bull, Shuler, Overton, Kimball, Boykin, & Griffin, 1999; Ko, Schallert, & Walters, 2003). In peer feedback activities following the ZPD frame, more advanced students help less proficient peers by providing scaffolding, such as feedback and any appropriate assistance.

Online peer feedback

Advances in CMC technologies have propagated the emergence of computer-mediated peer feedback over the past few of decades. Warschauer and Ware (2006) describe online peer feedback as “the means by which human feedback, particularly peer response, can be provided via technology” (p. 109). Compared with traditional face-to-face peer feedback, online peer feedback offers students several learning benefits, such as overcoming time and place constraints, creating less threatening environments, and promoting textual exchange interactively. Students can choose to review peers’ work and give peer feedback at any time and place, as online peer feedback makes peer interaction viable beyond the limits of classroom walls (Liu & Sadler, 2003). In addition, giving feedback online helps students diminish their peer pressure, which in turn helps students respond to peers’ work more judicially and comfortably (MacLeod, 1999). Under this less threatening environment, online peer feedback represents one form of computer conversation as characterized by DiGiovanni and Nagaswami (2001), which “allows students to respond spontaneously, yet offers them the opportunity to reflect on their ideas, rehearse their responses, and respond at their own pace” (p. 269). Liu and Sadler (2003) also argued that students using technology-enhanced peer feedback could generate more revision-oriented comments as well as an increased quantity of comments compared with those using traditional peer feedback, and that the asynchronous nature of the discourse was even more effective in student interactions than conventional synchronous modes.

Studies on the use of peer feedback through blogs to enhance students’ speaking

There has been a recent growth in blog-related research on facilitating students’ speaking performance in the context of learning a second or foreign language. Researchers have investigated using audio blogs in language classes to evaluate students’ oral assignments (Hsu, Wang, & Comac, 2008), understand students’ voice blogging stages/strategies (Sun, 2009; Huang, 2015), and explore perspectives on video blogging to learn speaking (Huang, 2015; Hung, 2011; Shih, 2010), and examine the effectiveness of speaking performance (Hung & Huang, 2015; Hung & Huang, 2016). Hsu et al. (2008), for instance, found that the use of audio blogs provided an effective way to assist instructors in managing oral assignments and interacting with students. In addition, students perceived that voice blogging facilitated their English speaking and pronunciation skills. Sun’s (2009) study identified five stages in voice blogging, namely, conceptualizing, brainstorming, articulation, monitoring, and evaluating and found that students developed a number of strategies for tackling blogging difficulties. In brief, research so far has indicated that blogging is a promising approach to helping students develop speaking skills.
Although several studies concluded that students generally held positive perspectives toward video blogs to improve their speaking, only a few have explored how students perceived video blogs with the support of peer feedback to learn/improve English speaking (Hung, 2011; Hung & Huang, 2015; Hung & Huang, 2016; Shih, 2010). For example, Shih (2010) formed a blended learning environment by integrating video blogs with face-to-face instruction in an English public speaking course and investigated how this model would affect students’ learning and satisfaction with the process. Both the instructor and the students commented on uploaded video clips of the students’ blog entries. The students then uploaded their revised clips based on the comments received. The results of 5-point Likert scale survey questionnaires showed that the students were satisfied with peer feedback as a way to help them improve their public speaking performance.

Similarly, Hung (2011) explored students’ perceptions of using video blogs in an English for Specific Purposes (ESP) class. The students uploaded four video clips and provided feedback after watching clips on their peers’ blog entries. The findings indicated that students perceived that peer feedback provided them with diverse perspectives, and cooperating with peers in a learning community provided opportunities for speaking practice. In a more recent pilot study, Hung and Huang (2015) examined the effects of video blogs on students’ oral presentation performance. Throughout an 18-week semester, the students were required to upload four presentation files and give one another peer feedback. The results indicated that the students’ overall presentation performance significantly improved, especially in the areas of projection, intonation, posture, introduction, conclusion, and purpose. To sum up, the research to date has shown a relatively positive effect of video blogs coupled with peer feedback upon students’ learning of English speaking.

However, research on using peer feedback through blogs to improve students’ speaking performance is still in its infancy. More studies need to be carried out to investigate how peer feedback through blogs can lead to improved students’ speaking performance. Another problem is that participants in the previously reviewed studies (Hung, 2011; Hung & Huang, 2015; Hung & Huang, 2016; Shih, 2010) did not receive explicit peer feedback training, which has been identified as an important factor for students’ positive learning outcomes in video blog-mediated peer review activities (Dippold, 2009; Liou & Peng, 2009; Ware & O’Dowd, 2008). Without feedback training, students may fail to provide peers with feedback that is constructive enough for them to improve their speaking performance. The last problem is that students’ receptions of and responses to problematic peer feedback on blogs to learn speaking have not been thoroughly investigated as previous studies tended to focus on learning outcomes of blogs rather than on peer feedback. The absence of this information, which is vital to understanding students’ learning problems of peer feedback, may contribute to difficulties in providing immediate and proper scaffolding to help students benefit from peer feedback. Hence, the present study aimed to bridge the gap found in the literature by examining how the use of online peer feedback through blogs could promote EFL college students’ speaking performance.

**Methodology**

**Participants**

Two groups comprising 20 non-English and 25 English majors at a university of technology in Taiwan voluntarily participated in this study. The non-English majors were taking an 18-week elective course called English Conversation. They were assessed as intermediate level speaking ability by the General English Proficiency Test (GEPT), a nationwide English proficiency test developed by the Language Training and Testing Center (LTTC) in Taiwan. The English-major participants were enrolled in an 18-week required course called English Listening and Speaking, and their English speaking proficiency, also evaluated by the GEPT, ranged from upper-intermediate to advanced level. In accordance with the ZPD frame, therefore, the English majors represented high achievers whose feedback could help the non-English majors develop English speaking skills. All participants were native speakers of Mandarin Chinese and had been learning English for 10 to 12 years.

**Research design**

In the first week of the semester, the students took the GEPT speaking test to measure their English proficiency before participating in the study. After the test, the instructor introduced the students the course objectives and the video blogging project, which was one of the requirements for both classes. The rationale for choosing video over audio-only blogging was to enable students to authenticate their oral communication, an introduction to some aspect of their country, Taiwan, in English to foreigners, with non-verbal cues and visual aids. In the
following week, the instructor introduced the blog website (WordPress.com) and demonstrated its main functions to the students. After the tutorial, students registered and set up their individual blogs.

From the fourth week onwards, three 4-week cycles of blogging tasks were implemented to offer students additional speaking practice outside of class. In the first week, students individually video-recorded three-minute videos in which they introduced Taiwan in English, which they then uploaded to their individual blogs. During the following two weeks, the students were randomly assigned four peers’ blog sites (two from each class) by the researchers. For each blog entry, they provided peer feedback including strengths, weaknesses and suggestions for improvement (based on Hung, 2011; Hung & Huang, 2015; Hung & Huang, 2016). The recipient students were encouraged to respond to their peers’ feedback until it was clear to them. To reduce the anxiety of using English to provide peer feedback, the students were allowed to comment on peers’ video clips in either Chinese or English. In the last week of each blogging cycle, the students revised their video clips based on feedback given by their peers and then uploaded the revised clips to their blogs again.

After completing the three blogging cycles, each student had produced three initial and revised clips in total. In week 18, students were given self-reflection worksheets with four guiding questions (Appendix A) on which to reflect upon giving and receiving peer feedback via blogs in this course. To encourage in-depth reflection on their learning experience, they were allowed to complete the sheets in Chinese.

Peer feedback training

In the third week, prior to participating in peer feedback, the students were explicitly instructed about how to offer peers helpful feedback on comment boards. After they were given feedback sheets specifying the essential features of effective peer feedback and the scoring criteria adopted in the study (see discussion below), they watched a sample video clip and discussed the speaker’s strengths and weaknesses and formulated constructive feedback in small groups. To help students continue to improve the quality of their feedback, during the actual blogging cycles, the teaching assistant periodically chose examples of more and less effective peer feedback from the students’ blogs to discuss with the class.

Data collection and analysis

Both quantitative and qualitative data were collected, including (1) the scores of students’ video clips, (2) the students’ individual blog entries containing blogging videos and peer comments, and (3) self-reflection worksheets. Though these data were collected from both English and non-English major students, the analysis of the data was mainly based on non-English major students for the scope of the current paper. First, the scores of the students’ video clips were used to examine the 20 non-English majors’ progress in speaking performance after receiving online peer feedback (RQs 1 and 2). The rubric for scoring speaking performance was adopted from Brown’s (2004) oral presentation checklist and modified to better fit the context of the present study. The revised version comprised two main categories, content and delivery, with several respective subcategories. Specifically, the content area comprised five items: purpose, introduction, main idea, supporting points, and conclusion. The delivery area consisted of seven items: gesture and body language, eye contact, vocabulary use, volume of speech, fluency, pronunciation, and grammar. The score for each item ranged from 1 (poor) to 5 (excellent) for a range of 12 to 60 total points. To achieve scoring consistency between the researcher and the research assistant, two rater training sessions were conducted, during which the researcher first acquainted the research assistant with the study purpose and the rubric. Next, they individually rated the same five video clips, compared scores, and reconciled any scoring discrepancies. In the formal assessment process, the two raters scored each student’s first and final clips individually and then averaged the two scores on each video clip to represent the students’ speaking performance. The inter-rater reliability, measured by percentage agreement, reached .89.

Based on gains in scores between the first and final video clips, two groups were drawn from the 20 non-English majors, the More Progress (MP) Group, consisting of five students in the upper quartile of gain scores, and the Less Progress (LP) Group, which included five students in the lower quartile. The students’ speaking progress was analyzed using paired sample t-tests to investigate the two groups’ different degrees of improvement.

Secondly, the students’ individual blog entries were examined in order to explore the follow-up revisions and performances of higher and lower performing non-English majors after the use of online peer feedback (RQ 3). Two cases, Student 1 (S1) and Student 2 (S2) were then chosen from the MP and LP groups respectively based
on typical case sampling (Bamberger, Rugh, & Mabry, 2011). The criterion for selecting the two cases was that their gain scores were closest to the mean gain scores of their respective groups. The use of such purposeful sampling was to show how representative students from the two sub-groups were likely to perform relative to the entire group. Drawing on the archives of S1’s and S2’s blogs, their initial and revised videos, peer feedback received, and responses to peers were analyzed and compared.

Thirdly, the self-reflection sheets were analyzed to examine non-English majors’ perceptions of and responses to problematic peer feedback (RQ 4). Content analysis was used, including coding, categorization, description and interpretation (Patton, 2002). In the coding phase, the research team read students’ responses on the self-reflection sheets thoroughly and highlighted meaningful statements from which several codes were generated and then collapsed into categories. For example, concerns about peers’ language proficiency, the reliability of feedback, and incorrect feedback found were codes constituting one of the main categories, correctness of the students’ perceptions of peer feedback. In the description phase, the main ideas of the categorized units were summarized and documented. Finally, in the interpretation phase, the research team interpreted important factors by providing possible explanations, reaching conclusions, and drawing inferences from each main idea. Reliability was established through regular discussions between the two coders on emerging themes and the inter-rater reliability reached .86. The individual blog comment boards also served as explanatory data to show how the students attempted follow-up revisions and its impact on their revised videos.

Results

RQ 1: What were the effects of blog-supported peer feedback on non-English major students’ speaking performance?

A paired sample t-test was computed to examine the relationship between the mean scores of the students’ first and final video clips. As shown in Table 1, the results were significant (t = -16.40, p < .01), indicating that there was significant improvement in the students’ speaking between the first video clip (M = 35.95, SD = 2.34) and the final video clip (M = 43.40, SD = 3.60).

Table 1. Results of the t-test in the students’ first and final video clips (N = 20)

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>SD</th>
<th>t(19)</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>First clip</td>
<td>30.50</td>
<td>41.50</td>
<td>35.95</td>
<td>2.34</td>
<td>-16.40</td>
<td>.00**</td>
</tr>
<tr>
<td>Final clip</td>
<td>35.50</td>
<td>52.50</td>
<td>43.40</td>
<td>3.60</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: **p < .01.

RQ 2: How did the effects on content and delivery items differ between higher and lower performing non-English major students via blogs?

Two more paired-sample t-tests were conducted to determine the extent to which the students improved their speaking scores on the 12 related items under the two main speaking categories, content and delivery. As shown in Table 2, though improvements in purpose, main idea, vocabulary use, and grammar were not significant, the MP group’s gains in the other areas were statistically significant particularly in introduction (t = -5.70, p < .01), conclusion (t = -6.00, p < .01), and gesture and body language (t = -6.33, p < .01). This finding affirmed that the provision of online peer feedback helped the MP group enhance their speaking performance in most items in both content and delivery categories.

Table 2. Criteria results of students’ speaking performances in the first and the final video clips in the MP group (N = 5)

<table>
<thead>
<tr>
<th></th>
<th>First clip</th>
<th>Final clip</th>
<th>t(4)</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Content</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purpose</td>
<td>3.30</td>
<td>.48</td>
<td>3.70</td>
<td>.67</td>
</tr>
<tr>
<td>Introduction</td>
<td>2.80</td>
<td>.27</td>
<td>4.20</td>
<td>.53</td>
</tr>
<tr>
<td>Main idea</td>
<td>3.10</td>
<td>.65</td>
<td>3.30</td>
<td>.45</td>
</tr>
<tr>
<td>Supporting points</td>
<td>3.20</td>
<td>.45</td>
<td>4.30</td>
<td>.27</td>
</tr>
<tr>
<td>Conclusion</td>
<td>2.70</td>
<td>.76</td>
<td>3.90</td>
<td>.82</td>
</tr>
<tr>
<td><strong>Delivery</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gesture &amp; body language</td>
<td>3.00</td>
<td>.50</td>
<td>4.00</td>
<td>.61</td>
</tr>
</tbody>
</table>
Eye contact 3.40  .74  4.40  .41  -4.47  .01*  
Vocabulary use 3.00  .35  3.30  .27  -1.50  .21  
Volume of speech 3.40  .65  4.30  .57  -3.09  .04*  
Fluency 3.20  .27  4.50  .50  -3.83  .02*  
Pronunciation 2.70  .76  3.90  .42  -3.53  .02*  
Grammar 3.00  .35  3.10  .22  -4.1  .70  

Note: *p < .05; **p < .01.

Table 3, however, shows that the LP group made no significant gains in the subcategories of content, suggesting that the students in this group made less productive use of peer feedback when revising their video clips than those in the MP group. In the delivery component, apart from the items of vocabulary use ($t = -4.1, p = .70$), pronunciation ($t = -1.63, p = .18$) and grammar ($t = -2.45, p = .07$), the LP group showed improvements similar to those of the MP group: gesture and body language ($t = -4.81, p = .01$), eye contact ($t = -4.71, p = .01$), volume of speech ($t = -3.16, p = .03$), and fluency ($t = -6.00, p = .00$), indicating that the students in the LP group tended to improve largely in the delivery area, in which revisions might have been less cognitively demanding. Notably, in the content category, both groups did not statistically improve their scores in terms of purpose and main idea. In addition, neither group showed significant improvement in vocabulary use and grammar, suggesting that online peer feedback through blogs might have limited impact on these aspects of students’ speaking performance.

Table 3. Criteria results of students’ speaking performance in the first and the final video clips in the LP group ($N = 5$)

<table>
<thead>
<tr>
<th></th>
<th>First clip</th>
<th>Final clip</th>
<th>t(4)</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purpose</td>
<td>3.30</td>
<td>.27</td>
<td>3.40</td>
<td>.22</td>
</tr>
<tr>
<td>Introduction</td>
<td>2.80</td>
<td>.27</td>
<td>3.10</td>
<td>.22</td>
</tr>
<tr>
<td>Main idea</td>
<td>2.90</td>
<td>.22</td>
<td>3.00</td>
<td>.35</td>
</tr>
<tr>
<td>Supporting points</td>
<td>2.90</td>
<td>.42</td>
<td>3.00</td>
<td>.35</td>
</tr>
<tr>
<td>Conclusion</td>
<td>2.60</td>
<td>.42</td>
<td>2.90</td>
<td>.22</td>
</tr>
<tr>
<td>Delivery</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gesture &amp; body language</td>
<td>2.50</td>
<td>.35</td>
<td>3.40</td>
<td>.22</td>
</tr>
<tr>
<td>Eye contact</td>
<td>2.40</td>
<td>.65</td>
<td>3.60</td>
<td>.42</td>
</tr>
<tr>
<td>Vocabulary use</td>
<td>3.10</td>
<td>.42</td>
<td>3.20</td>
<td>.27</td>
</tr>
<tr>
<td>Volume of speech</td>
<td>3.20</td>
<td>.57</td>
<td>3.70</td>
<td>.45</td>
</tr>
<tr>
<td>Fluency</td>
<td>3.10</td>
<td>.74</td>
<td>3.70</td>
<td>.57</td>
</tr>
<tr>
<td>Pronunciation</td>
<td>3.10</td>
<td>.55</td>
<td>3.30</td>
<td>.27</td>
</tr>
<tr>
<td>Grammar</td>
<td>2.80</td>
<td>.27</td>
<td>3.10</td>
<td>.22</td>
</tr>
</tbody>
</table>

Note: *p < .05; **p < .01.

RQ 3: How did peer feedback impact the students’ speaking performances of higher and lower performing non-English major students via blogs?

To more closely examine how the students used peer feedback to improve their speaking performance, S1 from the MP group and S2 from the LP group were purposefully selected as representative cases in that their gain scores were the closest to the mean gain scores of their respective groups. As illustrated in Table 4, both S1 and S2 improved their speaking performances between the first clip and the final clip, with S1 obtaining 5.5 gain points more than S2. In the following discussions, the remaining non-English major students are consecutively identified as S3, S4, S5, etc. while the English-major students were labeled ES1, ES2, ES3, etc.

Table 4. Background of the two selected participants from the two groups

<table>
<thead>
<tr>
<th></th>
<th>Group</th>
<th>First clip</th>
<th>Final clip</th>
<th>Gain scores</th>
<th>Mean gain scores of the group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student 1</td>
<td>MP</td>
<td>35.5</td>
<td>45.5</td>
<td>10</td>
<td>10.1</td>
</tr>
<tr>
<td>Student 2</td>
<td>LP</td>
<td>34.5</td>
<td>39</td>
<td>4.5</td>
<td>4.7</td>
</tr>
</tbody>
</table>

6
Taking the second blogging cycle as an example, S1 introduced three well-known tourist spots in Taiwan in her video. In her initial clip, S1 exhibited the content problems with introductions and supporting ideas, and had the delivery problems with volume and fluency of speech, pronunciation, and eye contact. With regard to content, her peers pointed out that S1 failed to provide a strong introduction to her topic, and the supporting points for each tourist spot could have also been strengthened. For the delivery problems, throughout the video, S1’s voice was sometimes not loud enough to be heard clearly, and her speech was further obscured by frequent pauses and pronunciation errors. Three peers (S3, S4, and ES1) suggested that she maintained more eye contacts with her audience. However, her peers did not identify S1’s grammar as especially problematic, and only one peer mentioned an issue with vocabulary use.

After receiving peer feedback, S1’s revised video showed notable improvement in the coherence of content, pronunciation, vocabulary use, and grammar. Specifically, Table 5 compares parallel transcript excerpts from S1’s initial and revised clips. Words in bold font highlight the differences between S1’s initial and revised video clips. In terms of content, the information in S1’s revised clip was more complete and coherent than in her initial presentation. She not only followed peers’ suggestions to include an adequate introduction, but also extended supporting points to further explain each tourist spot as S3 and S4 had suggested (see excerpts 1 and 2). S1’s revised video showed that she adopted most of her peers’ feedback on her delivery. She maintained more eye contacts, spoke more audibly, corrected pronunciation errors, and achieved greater fluency by reducing the number of unnecessary pauses and voice hesitations. ES1’s and ES2’s feedback helped S1 reduce her pronunciation errors from twelve to five (which included two new errors in the revised clip). On the other hand, S1 made few changes in vocabulary use and grammar, which had not been the focus in the feedback she received. She adopted ES2’s grammar corrections by changing sit the boat into sit in the boat (excerpt 3), but she did not follow ES1’s suggestion to change the word large in relation to competition (excerpt 4).

<table>
<thead>
<tr>
<th>Excerpts from the initial clip</th>
<th>Highlights of peers’ feedback</th>
<th>Excerpts from the revised clip</th>
<th>Revising actions (category)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Hello, everyone. I’m Sandy. Taiwan is small in the (P) Earth, but it is really an amazing island for many beautiful scenery. If the foreigners (P) come to Taiwan, I would take them to travel...around the...Taiwan.</td>
<td>ES1: Introduction could be strengthened.</td>
<td>(1) Hi, everyone. Today I want to introduce about Taiwanese scenery. Taiwan is small in the (P) Earth, but it is an amazing island because it has many beautiful scenery. We should go there at least once in lifetime. For example, Taroko National Park, Alishan National Scenery Area, and Sun Moon Lake are famous around the world.</td>
<td>Adding an introduction. (content)</td>
</tr>
<tr>
<td>(2) I think that, that...natural landforms there (P) are the most amazing...to...attract many travelers coming Taiwan.</td>
<td>S3 and S4: You can explain more in each tourist spot to enrich your contents.</td>
<td>(2) There are many uncommon animals, natural landforms, peculiar (P) plants in the Taroko National Park. I think natural landforms there are the most amazing to...attract travelers coming Taiwan. If you like to go walking and hiking, Taroko...Gorge is the perfect place.</td>
<td>Adding more supporting points. (content)</td>
</tr>
<tr>
<td>(3) If the foreigners (P) get there (P), they can sit the boat to travel the lake and see so many beautiful...landscape around the lake.</td>
<td>ES2: You should say sit in the boat not sit the boat.</td>
<td>(3) If the foreigners get there, they can sit in the boat to travel the lake and see beautiful landscape</td>
<td>Fixing a grammar error. (delivery)</td>
</tr>
<tr>
<td>(4) In September, there (P) is a large competition for swimming through the lake.</td>
<td>ES2: It’s weird to use large to describe competition.</td>
<td>(4) In September, there is a large competition for swimming through the lake.</td>
<td>Sticking to the original word “large.” (delivery)</td>
</tr>
</tbody>
</table>

Note. (P): pronunciation error; …: unnecessary pause; uh/un: voiced hesitation sound; underlining: corrected pronunciation error.
The topic of S2’s speech was an introduction to popular festivals and food in Taiwan. His initial video showed a number of content problems with respect to supporting ideas, which his peers pointed out. S5 and ES3 commented that information about the population of Taiwan was not very relevant to the purpose of his topic. S5 and S6 also commented that the supporting points were not detailed enough to give viewers an integral understanding of the topic. S6 and ES4 also mentioned S2’s lack of a definite conclusion when summing up his speech.

Several pronunciation and grammar errors were also identified in S2’s delivery. For example, S2 mispronounced several words, such as dragon, lunar, and annually. He also paused frequently and voiced hesitation “um,” and his facial expression lacked animation, which made his speaking less professional and vivid. The peers suggested that he practiced more and preview the next recording before uploading it to the video blog. However, although grammatical errors could easily be found in S2’s initial transcript, as in S1’s case few peers provided feedback on these, suggesting that students attended less to grammar slip-ups than to other aspects when they were reviewing peers’ video clips.

With respect to the use of the comment board, S2 seldom utilized this function to interact with his peers, seek further support, or verify the nature of their feedback. He tended to merely acknowledge feedback by posting such phrases as “thank you” or simply pressing the “Like” button. As demonstrated in the comparison of his initial and revised transcript excerpts (Table 6), S2 made only slight modifications to the content of his speech in the revised clip. First, he simply removed the irrelevant statements from his introduction but failed to include new points pertaining to his topic (excerpt 1). Second, his supporting points were not sufficiently developed as only a few complete statements were included (excerpt 3). Third, the revised video clip still lacked the explicit conclusion his peers had suggested (excerpt 4). Therefore, there were few differences in content between the original and the revised clips. However, S2 made greater improvements in pronunciation, fluency, and gesture and body language in his delivery. First, he corrected his pronunciation errors such as dragon, annual, and lunar which were underlined as examples in excerpt 2, and 3. Second, he reduced the number of pauses and voiced hesitations. Third, he exhibited more appropriate body language and facial expressions. These adjustments resulted in moderate improvement of his overall speaking performance, though his grammatical errors remained uncorrected in his revised video.

**Comparison between S1 and S2’s speaking improvements**

The selected cases had different degrees of improvements in their speaking. First, S1 exerted more effort in revising and enriching her speaking content by adding introduction and supporting details while S2 made only minor adjustments on supporting ideas, so S1 made notably greater improvements in content area. Second, both S1 and S2 made notable progress in their delivery, especially in the aspect of fluency, which could be attributed to both peers’ feedback and additional practice before the second recordings. Third, both seldom corrected their grammatical errors or improved their vocabulary use in their revised presentations, perhaps because they received little feedback in these areas. Finally, the feedback that was given on vocabulary use, grammar, and pronunciation was provided by English majors (ES1, ES2, ES3, ES4), suggesting that non-English major students lacked the confidence or knowledge to address peers’ problems in these three areas.

### Table 6. Comparison between transcript excerpts from S2’s initial and revised video clips

<table>
<thead>
<tr>
<th>Excerpts from the initial clip</th>
<th>Highlights of peers’ feedback</th>
<th>Excerpts from the revised clip</th>
<th>Revising actions (category)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Hello, today I will introduce festivals and some food of Taiwan to you. There are many things you should know about Taiwan. Do you know we have over twenty, twenty three million (P) people live in Taiwan, and um...most of its population live (P) on the western plains (P).</td>
<td><strong>S5 and ES3:</strong> The beginning part is not related to your topic.</td>
<td><strong>(1) Hi guys, my name is Tom.</strong> Today I want to introduce festivals and some food of Taiwan to you.</td>
<td>Removing the old introduction without adding a new one. (content)</td>
</tr>
<tr>
<td>(2) First is the Dragon (P) Boat</td>
<td><strong>ES3 and ES4:</strong> You</td>
<td>(2) First <strong>one</strong> is the <strong>Dragon</strong> Boat</td>
<td>Correcting a...</td>
</tr>
</tbody>
</table>
Festival. This festival falls on May fifth on the lunar (P) calendar.

(3) This festival has its story of Qu Yuan. Uh...he jumped (P) into the river. People want to save him, so they throw rice dumplings into the river to keep fish from eating his body. Second is the Pingxi Sky Lantern Festival. This event is held annually. People write down their best wishes on the lanterns and release them. Third is the Moon festival. It take place on the fifteenth of August on the lunar (P) calendar.

(4) This is why people love Taiwan after this video. Bye.

Note. (P): pronunciation error; …: unnecessary pause; uh/um: voiced hesitation sound; underlining: corrected pronunciation error.

RQ 4: What were non-English major students’ receptions of and follow-up responses to problematic feedback via blogs?

Drawing on the data from the 20 students’ reflection sheets, a total of 44 statements were coded as students’ receptions of problematic peer feedback. The 44 statements of the problematic peer feedback were then categorized into three types: peer feedback lacking in clarity, doubting of its correctness, and having difficulties tackling inconsistencies arising from peers’ different feedback on the speaking content. As shown with examples in Table 7, the major problematic peer feedback type was clarity and this showed that most of the feedback the students received was not explicit enough for them to adopt in their revised speaking videos.

The students’ follow-up responses to problematic feedback were also identified. Table 8 presents the three major types: (1) ignoring peer feedback directly (58%), (2) seeking external help before adopting or ignoring the online peer feedback (23%), and (3) asking for clarification (19%). In Type 1, the majority of the responses indicated that students would disregard peer feedback which they did not think would help them improve their speaking performance. S8’s reflection represents this view: “If peers’ feedback was too vague, I would directly skip it because it was usually not valuable. I only considered adopting clear comments.” This type of follow-up response required less effort than asking their peers for clarification.

Most students who looked for external help (Type 2) either searched for related information on the Internet or asked others for opinions. For example, when S1 was not sure whether to adopt a peer’s feedback suggesting that big was a more common adjective than large to qualify competition (Figure 1), she consulted the Internet for verification before responding to her peer, “I checked it on Google. Both words work.”
Table 7. The students’ initial receptions of problematic peer feedback

<table>
<thead>
<tr>
<th>Types of reception (Total = 44 statements)</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clarity (22 statements)</td>
<td>“One peer told me my pronunciation was a bit weird, but he didn’t point out the specific problems, so I still did not know how to improve it.” (S8)</td>
</tr>
<tr>
<td></td>
<td>“Some comments were not clear enough. Some only commented that my introduction needed to be strength and without telling me how to do it” (S9)</td>
</tr>
<tr>
<td>Correctness (15 statements)</td>
<td>“I am not sure if I should trust peers’ feedback, especially those non-English majors, because I think some of them are at the same English proficiency level.” (S4)</td>
</tr>
<tr>
<td></td>
<td>“Classmates’ correction might be wrong…one peer suggested that I change the word from eat to drink the soup, but I knew eat is correct.” (S10)</td>
</tr>
<tr>
<td>Consistency (7 statements)</td>
<td>“Feedback on content is rather subjective. It’s hard to decide which one to adopt sometimes. But I like English-major peers’ feedback more.” (S11)</td>
</tr>
<tr>
<td></td>
<td>“I happened to have two peers’ different opinions on my speaking content. One said it’s good to use detailed examples, but the other said it is redundant.” (S12)</td>
</tr>
</tbody>
</table>

Table 8. The students’ major types of follow-up responses to problematic peer feedback (N = 20)

<table>
<thead>
<tr>
<th>Types of actions taken</th>
<th>Frequency (percentage)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Ignoring online peer feedback directly</td>
<td>25 (58%)</td>
</tr>
<tr>
<td>2. Seeking external help before adopting/ignoring online peer feedback</td>
<td>10 (23%)</td>
</tr>
<tr>
<td>3. Asking for clarification</td>
<td>8 (19%)</td>
</tr>
</tbody>
</table>

A transcript excerpt from S1’s initial video clips

In September, there (P) is a large competition for swimming through the lake.

<table>
<thead>
<tr>
<th>Providers</th>
<th>Highlights of peer feedback and responses on S1’s blog</th>
</tr>
</thead>
<tbody>
<tr>
<td>ES2</td>
<td>Peer feedback: It’s weird to use <em>large</em> to describe <em>competition</em>.</td>
</tr>
<tr>
<td>S1</td>
<td>Further response to ES2: …anything wrong with <em>large</em>? I remember that I’ve seen this usage somewhere before.</td>
</tr>
<tr>
<td>ES2</td>
<td>Further response to S1: Really!? I would use <em>big</em> instead of <em>large</em>. It’s just my opinion.</td>
</tr>
<tr>
<td>S1</td>
<td>Further response to ES2: Hi, I checked it on Google. Both words work.</td>
</tr>
</tbody>
</table>

Showing concerns for correctness of the peer feedback
Seeking external help from the Internet

A transcript excerpt from S1’s revised video clips

In September, there is a large competition for swimming through the lake.

Retention the same vocabulary “large”

Figure 1. An example of a student’s Type 2 follow-up attempt

The third most frequently mentioned follow-up attempt, Type 3, comprised requests for clarification from peers on blogs before revising their presentations. For instance, in response to the comment on her initial clip “introduction could be strengthened,” S1 asked ES1 to further clarify this feedback by providing more concrete suggestions for her revision (Figure 2). ES1 then responded, “maybe if you mention the names of the three places before introducing them, it will make your introduction more connected to the content.” After she adopted
ES1’s suggestion, S1’s introduction in the second video was more complete, and thus the feedback helped her improve the content of her presentation. This example shows that students can receive more support from peers if they directly ask about unclear feedback.

Among the three types of responses to feedback, it was found that Types 2 and 3 had better results because the students were more active in responding to peers’ feedback and expressed concerns regarding feedback they found problematic as well as further verified it before making modifications. That is, active involvement in using the comment board such as responding to peers’ unclear feedback or asking for more information from peers seemed to optimize the effect of peer feedback use.

<table>
<thead>
<tr>
<th>Providers</th>
<th>Highlights of peer feedback and responses on S1’s blog</th>
</tr>
</thead>
<tbody>
<tr>
<td>ES1</td>
<td>Peer feedback: Introduction could be strengthened.</td>
</tr>
<tr>
<td>S1</td>
<td>…in the beginning I’ve said that Taiwan is an amazing island and people should come to visit Taiwan. I think it’s not bad. Do you have any ideas on what else I can add?</td>
</tr>
<tr>
<td>ES1</td>
<td>Further elaborations on S1’s inquiry: ...the opening part is good, but maybe if you mention the names of the three places before introducing them, it will make your introduction more connected to the content.</td>
</tr>
<tr>
<td>S1</td>
<td>Your suggestion is very helpful. Thanks a lot.</td>
</tr>
</tbody>
</table>

A transcript excerpt from S1’s revised video clips

Hi, everyone. **Today I want to introduce about Taiwanese scenery**. Taiwan is small in the (P) Earth, but it is an amazing island because it has many beautiful scenery. **We should go there at least once in lifetime. For example, Taroko National Park, Alishan National Scenery Area, and Sun Moon Lake are famous around the world.**

Figure 2. An example of a student’s follow-up attempt in Type 3

**Discussion and conclusion**

This study has shown how the use of online peer feedback through blogs can enhance students’ speaking performance as captured in their blogging videos. A comparison of students’ scores on their first and final video presentations indicated that the students’ overall speaking performance significantly improved with the support of peer feedback. This finding corroborated prior research (Hung & Huang, 2015; Hung & Huang, 2016; Shih, 2010) which suggested that utilizing blogs as a learning tool for speaking practice combined with peer feedback had a positive impact on students’ speaking performance. In addition, this study can enrich the current literature (Hung, 2011; Hung & Huang, 2015; Hung & Huang, 2016; Shih, 2010), which merely evaluated students’
improvements either from their gain scores or perceived progress of speaking, by further dividing students’ speaking performance into the content and delivery components with several respective subcategories to discuss the influence of online peer feedback on students’ speaking performance specifically. The comparison of the results from Tables 2 and 3 indicated that for the content area, the MP group was the only group who made significant improvement in purpose, supporting ideas, and conclusion, and neither the MP or LP group made any significant progress in purpose and main idea. The lack of improvement in the content area especially in terms of purpose and main idea may possibly account for the fact that once the students settled down into the topics of their initial videos, most of them retained their original purpose or main idea in their revised clips in spite of doing extensive revisions to the introduction, supporting points and conclusion. The students’ receptions of feedback shown in Tables 7 and 8 supported this claim in which S11 indicated that “Feedback on content is rather subjective. It’s hard to decide which one to adopt sometimes. But I like English-major peers’ feedback more.” With respect to the delivery area, the MP and LP groups made significant progress in gesture and body language, fluency, and speech volume. The results revealed that the implementation of online peer feedback into video blogging primarily benefitted students’ expressive language skills such as gesture, speech volume, and the use of pauses and voice hesitations.

Both groups did not make significant improvements in their vocabulary use and grammar of the delivery component. Similarly, the comparison of S1 and S2’s transcripts showed that both seldom corrected grammar mistakes or improved their vocabulary use in the revised speaking videos, which might be attributed to the limited feedback on these two aspects which they received. While reviewing peers’ video clips, the students might not have paid much attention to these two types of errors as long as they could understand the contents of the presentations. Another possible explanation was that the students doubted the correctness of the peer feedback related to vocabulary use and grammar. The students’ receptions of feedback (Table 7) showed that the students did not trust the feedback of non-English majors. The students such as S1 and S2 benefited more from English major students’ feedback as they were able to provide more feedback on syntactical, morphological, or phonological problems than non-English major students, and their scaffolding played a major role in this blog-enhanced peer feedback activity, affirming the frame of the ZPD emphasis on the guidance of more knowledgeable peers (Ko et al., 2003).

The students’ receptions of feedback revealed in this study is in line with the researchers’ claims that since students are not considered knowledge experts, when compared with their teachers, their feedback to peers might contain problems and thus are less favored by students (Gielen, Peeters, Dochy, Ongena, & Struyven, 2010; Strijbos, Narciss, & Dünnebier, 2010). However, the interactive processes between feedback providers and receivers in this study were captured in students’ follow-up responses to problematic peer feedback on blog comment boards. It was found that after receiving peer feedback, those verifying comments received and taking the initiative in asking for further suggestions from peers made more progress in their revised clips. This finding echoes Yang, Badger, and Yu’s (2006) study, which suggested that students with doubts or reservations about peers’ feedback would, instead of solely relying on them, be encouraged to find related sources to confirm or correct received feedback, and such a self-correcting process was more likely to result in their better performance. In addition, students’ levels of engagement with responding to peers’ problematic feedback for clarification might also influence their learning outcomes. As Baggetun and Wasson (2006) stated, “in order to generate feedback one needs certain skills, in particular, participation skills… and [knowledge of] how to invite or [ask] questions so that someone feels tempted to reply” (p. 460).

Based on the findings in this study, several pedagogical implications may be drawn. First, because students made significant improvements in their overall speaking performance after receiving and responding to online peer feedback with teacher guidance and knowledgeable peer scaffolding, language teachers are recommended to provide students with explicit training in efficient feedback strategies and include more proficient learners in their instructional design so as to optimize the effects of blog-mediated feedback activities. Second, as seen in this study, although the students had received feedback training, sometimes they still provided unclear or implicit feedback. Hence, the teachers should encourage students who received the feedback to actively probe unclear feedback and negotiate problems through continuous interactions. Third, given that students’ vocabulary use and grammar did not significantly improve, teachers should focus online peer feedback on enhancing students’ speaking fluency. To develop students’ speaking accuracy through online peer feedback, teachers should set one specific goal such as commenting only on peers’ grammar or vocabulary use to help students attend to aspects to which they might otherwise pay little attention. In addition, students should be required to upload transcripts into their individual video blogs to supplement the audio portion of the videos.

As the present study may be regarded as an initial exploratory study, suggestions for further research include the following. First, larger sample sizes from different language proficiency levels or age groups can be recruited to
examine whether the use of online peer feedback via blogs can yield similar outcomes with a broader range of students. Second, this study did not explore the relationships between improvement of non-English major’s content and delivery items and EFL ability of English major students who provided feedback to those non-English majors. Future research may explore how the language proficiency of more capable majors’ content and delivery items impacts the improvement of less capable ones’ speaking performance. Third, researchers can further investigate whether posting transcripts as well as the video clips affects the quality of students’ feedback, which might in turn influence students’ speaking performance. Finally, more follow-up qualitative case analysis for documenting the experiences of individual students can also be conducted to generate different insights gained from peer feedback through blogs.

Acknowledgements

This project is supported by the Ministry of Science and Technology under the grant number 106-2628-S-224-001-MY3.

References


Appendix A

Self-reflection sheet

Please use the following guiding questions to elicit your learning in this course. You should write at least a 100-word response to each question.

1. What difficulties did you encounter while giving peer feedback on video blogs?
2. How did you solve each difficulty you encountered? Please explain in detail. For each difficulty, you may propose more than one solution.
3. What are the benefits of utilizing peer feedback through blogs?
4. What are your suggestions for improving this project especially in the aspect of utilizing peer feedback via blogs?
Online Targeting Behavior of Peer-Assessors under Identity-Revealed, Nicknamed, and Concealed Modes

Fu Yun Yu\(^1\) and Shannon Sung\(^2\)

\(^1\)Institute of Education, National Cheng Kung University, Taiwan // \(^2\)Education Studies Program, Spelman College, USA // fuyun.ncku@gmail.com // ssung@spelman.edu

\(^\ast\)Corresponding author

(Submitted August 4, 2016; Revised March 8, 2017; Accepted August 2, 2017)

ABSTRACT

This study examined whether different identity revelation conditions result in different online targeting behavior among peer-assessors through a pretest and posttest quasi-experimental research design. Students from six fifth-grade classes (\(N = 196\)) participated in online learning tasks where they generated and selected peer-generated questions to review and assess in their respective identity revelation modes—real-name, nickname, and anonymity. Several findings were obtained. First, there was a high redundancy rate for both high- and low-targeted assesses under the pre-treatment (control) and treatment conditions for all three identity revelation modes. Second, the non-significant results of the chi-square tests indicated that the identity revelation modes and assessee redundancy were neither considerably related for the high- nor low-targeted assesses groups. Third, the results from the ANCOVA on the number of peer-feedback messages each student received showed no significant differences among the three identity revelation treatments. Fourth, the Wilcoxon tests confirmed that there were no significant differences in the assessees rankings in terms of the number of times their generated questions were assessed between the pre-treatment and treatment conditions for all different identity revelation modes. Finally, the results from participants’ responses to the checkbox question revealed almost the same ranking pattern regarding factors dominating their targeting behavior (with the question-author ranked last among all factors), despite the various different identity revelation modes. In sum, identity revelation modes were not found to affect peer-assessors’ targeting behavior in an online peer-assessment activity.

Keywords

Identity, Online targeting behavior, Peer-assessment

Introduction

Assessment is one of the decisive factors affecting the processes and outcomes of teaching and learning. Conventional teacher-centered methods of assessment are now giving way to alternative and innovative assessments that enable the learners to actively share responsibility and initiative in assessing personal levels of understanding and growth (Cheng & Warren, 2005). Understandably, the potential of peer-assessment has attracted much attention since its emergence and has been found to promote critical thinking, cognitive restructuring, better quality work, learning motivation, self-efficacy, and attitudes toward learning (Cartney, 2010; Cheng & Warren, 2005; Kaufman & Schunn, 2011; Topping, 2009; van Gennip, Segers, & Tillema, 2010).

Recent developments in peer-assessment research

Recently, the effects of different types and arrangements of peer-assessment have been the focus of many studies, for example, anonymous vs. identified peer-assessment (Yu & Wu, 2011), concise vs. elaborated peer-feedback (Strijbos, Narciss, & Dünnebier, 2010), peer-feedback with vs. without justification (Bolzer, Strijbos, & Fischer, 2015), non-directive vs. specific peer-feedback (Cho & MacArthur, 2010), corrective peer-feedback with different degrees of directness (AbuSeileek & Abualsha’r, 2014), and peer-feedback with different levels of structure (Gielen & De Wever, 2015). Studies concerning the effects of the degree of anonymity of the assessor and assessed have special importance since researchers (e.g., Ballantyne, Hughes, & Mylonas, 2002; White, 2009) have questioned the objectiveness and fairness of peer-assessment. For instance, the outcomes of peer-assessment have been reported to be based more on the personal importance of the assessed rather than the work being examined, with more introvert individuals usually receiving higher scores than those who are more extroverted (Johnston & Miles, 2004). Friendship has also been found to influence assessments, where individuals belonging to the same inner circle receive inflated scores as opposed to those outside such circles (Chan, 2010) or when there is a high level of friendship between the assessors and assessees (Panadero, Romero, & Strijbos, 2013).

Because the disclosure and concealment of identities between interacting parties in a peer-assessment situation could be easily managed in online spaces (Yu, 2009), researchers have examined its effects on various
educational outcomes, such as the quality of the work produced, interpersonal relationships, perceptions toward peer-assessment, ratings and comments provided, professional behavior, and engagement (Ainsworth, Gelmini-Hornby, Threapleton, Crook, O’Malley, & Buda, 2011; Garner, McKendree, O’Sullivan, & Taylor, 2010; Raes, Vanderhoven, & Schellens, 2015; Li, 2017; Stepanyan, Mather, Jones, & Lusuardi, 2009; Thompson & McGregor, 2009; Wadhwa, Schulz & Mann, 2006). Motivated by the fact that few studies have investigated its effects on the actual interaction “process” or related behavior, the online targeting behavior of the assessor (i.e., decision-making as to which item to assess) was examined in one prior study using a one-group research design to understand the “process” of interaction in identified (pre-treatment) and anonymous (treatment) peer-assessment situations (Yu & Sung, 2016).

Given that the aforesaid design could have been weakened by several internal validity threats, such as history and maturation (Grinnell & Unrau, 2010), a more robust research design—a pretest-posttest quasi-experimental research method, was adopted in this study. Furthermore, since researchers suggest that different levels of user identifiability may lead to different interactions (Yu & Wu, 2011), and the use of nicknames is prevalent among online users, the online targeting behavior of the assessor using real-name (i.e., absolute revelation), nickname (i.e., partial revelation/concealment), and anonymity (i.e., complete concealment) modes is examined in this study.

Research questions

The assessors’ actual behavior when selecting a target for online peer-assessment (i.e., online targeting behavior) and their reasons for such behavior in three identity revelation modes were investigated. The two main research questions formulated for the study are as follows:

RQ1: Do different identity revelation modes (i.e., real-name, nickname, and anonymity) among the interacting parties affect the online interaction behavior and social dynamics of the group in online peer-assessment contexts? To address this, four sub-research questions are examined:

- Does the proportion of overlapping assesses among the high-targeted assesses (i.e., assesses who receive a comparatively high number of peer-feedback messages) vary significantly among the three different identity revelation modes?
- Does the proportion of overlapping assesses among the low-targeted assesses (i.e., assesses who receive a comparatively low number of peer-feedback messages) vary significantly among the three different identity revelation modes?
- Does the average number of peer-feedback messages each student receives differ among the three identity revelation treatments?
- Does ranking in terms of the total number of peer-feedback messages received in different identity revelation modes differ significantly between the pre-treatment (i.e., control) and treatment conditions?

RQ2: Do the major reasons selected by peer-assessors for their online targeting behavior vary significantly among the three different identity revelation modes?

Literature review

Face-to-face peer-assessment

Peer-assessment capitalizes on the idea of peers assisting one another by providing and receiving quantitative and/or qualitative feedback to and from their equivalent peers about their respective work or performance (van Gennip et al., 2010). Because peers generally have a similar learning status and are within each other’s zone of proximal development, peer-feedback is more easily perceived and understood, as compared to that given by teachers (Topping, 2009). Studies have found that students’ self-esteem and self-efficacy regarding academic competence tend to be more positively affected by evaluations from peers than from their teachers (Kaufman & Schunn, 2011), and the quality, validity, and reliability of peer-assessment has been attested in a number of works (e.g., Avery, 2014; Jones & Alcock, 2014).

Although the benefits of peer-assessment have been recognized by teachers, its adoption in classrooms has not been totally encouraging. Foremost, face-to-face peer-assessment is spatial- and physically-bound, as well as time-consuming and labor-intensive regarding material preparation and data compilation (DiPardo & Freedman, 1988). Peer-assessment may also be in conflict with the societal norms in certain cultures that implicitly
discourage confrontational discourse or critical feedback (Cartney, 2010), and this may result in inaccurate assessments or peer pressure especially in identified face-to-face contexts (Raes et al., 2015). Finally, students may not be equipped with the required capacity to provide constructive feedback to their peers without adequate scaffolds to ensure quality peer-assessment work (Liang, 2010). For all these reasons, many researchers have turned to computer-mediated peer-assessment.

**Computer-mediated peer-assessment**

The affordances of networked technologies are crucial in alleviating the challenges of face-to-face peer-assessment, while at the same time enabling authors to receive timely and constructive feedback. Many online peer-assessment systems have been developed in the past decade to help manage and implement peer-assessment effectively in classrooms (e.g., NetPeas, Vee heuristic, Web-SPA, and SWoRD). With unique features including high processing speed, immense storage space, learner control of sequencing, customizable scaffolds for feedback-provision, multimedia, simultaneity, instantaneity, and space-, time- and device-independence, various technological and pedagogical arrangements can be made available to online users of such systems (Yu & Wu, 2011). System features, such as automatic assignment of work to be assessed, a set of customized built-in peer-assessment criteria for reference, time-stamping, multiple peer reviews, asynchronous and synchronous interaction, anonymity, process display, history record, and real-time notification are frequently embedded in these systems to expedite and streamline tasks (Smaldino, Lowther, & Russell, 2008; Yu & Wu, 2011).

Although online peer-assessment generally benefits students cognitively and affectively, it has several shortcomings, including difficulties related to comprehension because it lacks contextual cues, which may compromise its effects (Jones & Alcock, 201; Liang, 2010). In an attempt to rectify this and further promote the learning effects of peer assessment, many researchers have focused on exploring different peer-assessment designs (e.g., anonymous/identified peer-assessment, concise/elaborated peer-feedback, peer-feedback with/without justification, non-directive/specific peer-feedback, and corrective peer-feedback with different degrees of directness), and examine their relative effects.

**The psychological functioning of anonymity during online interactions—social identity theory and sense of perceived psychological safety**

The social identity theory describes how people develop and form a sense of belongingness, membership, and social identity by associating themselves with particular affiliations (Pearce II, 2013; Tinson & Close, 2012). In addition to forming personal identities as individuals, people tend to search for and establish their social identities through group affiliation, which differentiates them from other groups (Hogg, 1987). This theory has been used to explain and predict how perceived status differences and personal characteristics in social groups govern an individual’s behavior and decision-making (Hogg, van Knippenberg, & Rast III, 2012; Tajfel & Turner, 1986). Accordingly, individuals perceived as having a prominent social identity (i.e., being superior or more favored) are believed to possess more reliable sources of information and thus will tend to exert more influence over the perceptions and behavior of other group members (Hogg et al., 2012). In contrast, those perceived to have a less prominent social identity are less likely to receive as much attention as those from more favored and prominent groups (Tinson & Close, 2012). Considering the potential for prejudiced interactions, it has been suggested that concealing the real identities of participants during group interactions could avoid judgments being formed on the sheer basis of the person’s affiliations and could help foster greater and more balanced participation among individuals of different backgrounds (Chester & Gwynne, 1998).

On the other hand, anonymity’s potential to remediate biased scoring and differential participation is believed to be associated with an enhanced sense of psychological safety (Yu, 2009; Miyazoe & Anderson, 2011; Roberts & Rajah-Kanagasabai, 2013) — a state where one feels secure and comfortable without excessive concerns over any possible effects on one’s well-being, image, and status (Zhang, Fang, Wei, & Chen, 2010). Identity concealment could help lessen negative emotions, such as apprehension, intensified pressure, and a sense of insecurity (Bullingham & Vasconcelos, 2013), which may be caused by social comparison, peer pressure, and self-validation (Franzoi, 2006; Moral-Toranzo, Canto-Ortiz, & Gómez-Jacinto, 2007; Raes et al., 2015). Similarly, excessive self-consciousness, which is linked closely with identifiability in group situations and may inadvertently affect assessors’ objective feedback, can be reduced by masking one’s identity (Bullingham & Vasconcelos, 2013; Wadhwa et al., 2006). Overall, psychologically safe learning contexts are generative and have been found to be associated with more active knowledge-sharing and genuine self-expression intention (Zhang et al., 2010), which can lead to improved work or more refined performance (van Gennip et al., 2010).
The psychological functioning of nicknames in online interaction situations—Identity construction

Nicknames disguise the real identity of participants, and may thus have similar liberating effects on the interacting parties as anonymous situations do. As such, the social identity theory and psychological safety serve well as the theoretical foundations for the inclusion of a nickname option in virtual worlds. Nonetheless, the use of nicknames may engender further benefits, as reported in identity construction studies.

Studies on identity construction stress the importance of one’s identity formation through reciprocal interaction with sociocultural factors (Bullingham & Vasconcelos, 2013; Hassa, 2012). Such studies mostly highlight the social constructive process of self-perception that is closely intertwined with language, culture, and society (Krämer & Haferkamp, 2011; Locher, 2008) and transcends conventional social categorizations (e.g., gender, Cornetto & Nowak, 2006), which is more easily attainable in online spaces.

In the cyberworld, interpersonal communications and conversations are extended (i.e., not limited to face-to-face contexts) (Tufekci, 2008), and it is possible for members to use nicknames, profile photos, personas, avatars, and usernames, where one’s identity can be explored and formed (Lindholm, 2013; Subrahmanym & Smahel, 2011). Indeed, people form impressions of others based on users’ online nicknames (Chester & Gwynne, 1998), which can trigger interpersonal interactions (Lindholm, 2013). As emphasized by personality development theorists (e.g., Marcia, 1966), it is essential to allow individuals to be able to actively explore options before reaching their final stage of identity development. Enabling users to recreate certain entities or characteristics enables them to avoid value-laden contradictions or confrontations or can even help them impress others and attract attention (Krämer & Haferkamp, 2011). In sum, the representation of oneself in the virtual world provides opportunities for identity exploration and construction in a risk-free fashion (Subrahmanym & Smahel, 2011).

To summarize, peer-assessment learning activities are frequently in the center of online teaching and learning designs. Most studies focus on the effects of different types and arrangements of peer-assessment; however, some issues remain unresolved. In particular, among the studies examining identity revelation, very few focus on the participants’ actual interaction behavior, and far fewer involve the use of nicknames. Since selecting the targets to be assessed is the first step in the peer-assessment process and may impact student experiences, attitudes, and/or learning outcomes, and because the allocation of assessed work is commonly practiced in peer-assessment situations, issues surrounding whether or not revealing the identity of the author may reduce the chances that the participants’ work will be assessed still need to be resolved before instructors can confidently apply self-selecting targets for peer-assessment. Thus, the effects of real-name, nickname, and anonymity on peer-assessors’ actual online targeting behavior and their accompanying reasons for targeting any specific items were investigated in this study.

Methods

Participants and the online learning system

Six fifth-grade classes (age 10-11, N = 196, male = 96, female = 100) were randomly selected from a primary school in Taiwan and introduced to an online question-generation and peer-assessment learning activity during their weekly computer literacy class. Computer literacy is a required course from third grade on, so the participants all had fundamental computer capabilities (e.g., keyboarding, web-searching, using a word-processor, inserting multi-media files into a project, and so on). An online learning environment—Question-Authoring and Reasoning Knowledge System (QuARKS) was adopted to support the focal activity, which was only accessible in class to ensure that all the participants had the same amount of time to use the system. Even though QuARKS is capable of customization (in terms of the specific sets of functions accessible to users in different contexts) (Yu, 2009), for the purpose of this study, with the exception of different identity revelation modes that were assigned to different treatment groups, the accessible functions and implementation procedures were kept identical for all treatment groups.

Specifically, the only difference in the treatment groups was whether the real-name or nickname of the question-author was revealed or concealed in the question-author field. Explicitly, when the questions to be assessed were listed in the peer-assessment list window (see Figure 1) and could be viewed in the top portion (i.e., targeted question) of the online peer-assessment form window (see Figure 2), the student’s full-name was retrieved automatically from the database and shown in the question-author field in the respective windows for the real-name mode (Figures 1a and 2a). In contrast, for the nickname mode, the nickname used by the question-author
was shown (Figures 1b and 2b), and for the anonymity mode, no information about the identity of the question-author was given, and the term “anonymity” was shown instead (Figures 1c and 2c).

Figure 1. Peer-assessment list window for the real-name (left, a), nickname (middle, b) and anonymity (right, c) modes

Figure 2. The top portion of the online peer-assessment form window for the real-name (left, a), nickname (middle, b) and anonymity (right, c) modes

Figure 3. The bottom portion of the online peer-assessment form window (for the real-name mode)

While the default item-sequencing is based on the question submission time (from the latest, marked as question No. 1, to the earliest), to assist viewing and selecting targets, participants can change the order of items that appear on the system, according to the number of assessors/peer-feedback received, question type, question, and question-author, by clicking on the menu bar of the peer-assessment list window (Figure 1). Once students
decide their assessment target and click on the item of their choice (Figure 1), they are directed to the online peer-assessment form window to view the targeted question (Figure 2). They then provide both quantitative and qualitative feedback in the online peer-assessment form (Figure 3).

**Experimental design and implementation procedures**

To examine the effects of different identity revelation modes (i.e., real-name, nickname, and anonymity) on participants’ targeting behavior in an online peer-assessment context, a pretest-posttest quasi-experimental research design was adopted. The six participating classes were randomly assigned to the three different treatment groups (two classes in each mode).

Before the study, the students were briefed about the purposes of the online learning activity — supporting science learning and promoting higher order thinking skills, and informed that their performance on the tasks would account for part of their computer literacy grades. They then received a training session in the computer lab, where the following topics were covered: basic principles for question item writing, the question-generation criteria and models, the peer-feedback criteria and models, and operational procedures for question-generation and peer-assessment on QuARKS.

Afterwards, in the actual study during the 40-minute instructional time, the students were instructed to first construct two multiple-choice questions on QuARKS related to the science content covered in the current week. They were then directed to assess four peer-generated questions of their choice on QuARKS, where their quantitative ratings of the quality of the assessed item as well as descriptive feedback would be given and directed to the question-author for consideration. The students were asked to focus on their individual tasks and to not talk to their classmates. Any questions were thus directed to the implementer.

To establish a baseline of the students’ online targeting behavior, the real-identity mode was set for all experimental groups as the pre-treatment (control) condition for the first session, and the respective treatment conditions were introduced for the following week. In other words, the learning tasks completed by the participants in the three treatment groups for the pre-treatment and treatment conditions were the same — generating two questions and assessing four peer-generated questions. The only difference was in the identity revelation mode used for the question-author during the treatment condition. Furthermore, to enhance question-generation and peer-feedback performance, the students’ work was reviewed by the implementer, and three pieces from the respective classes were selected and shown as exemplars at the beginning of the treatment condition week for whole-class feedback.

Finally, the students responded to one checkbox question after the activity, where possible reasons that might have affected their targeting behavior during the online peer-assessment phase of the activity were provided for selection. A flow diagram summarizing the experimental procedures of this study is given in Figure 4.

![Figure 4. Experimental procedures used in this study](image-url)
Instrumentation

One checkbox question was developed and disseminated for individual completion after the study to collect data on the participants’ accounts of their online targeting behavior. Based on informal, non-participant observation of the participants’ selection behavior during the activities, six responses were listed as options to choose from — question type, number of assessors already giving feedback on the item, question item itself, question submission time, who the question-author was, and the current average rating. Participants were instructed to select not more than three reasons that they felt often directed their online targeting behavior during the peer-assessment activities in this study.

Data preparation and analysis

Prior to a statistical analysis, careful preparation of the data by applying appropriate filtering techniques is essential for accurate data analysis (Hansen, Shneiderman, & Smith, 2010). Specifically, to ensure an unbiased comparison between the pre-treatment (i.e., real-name mode, 1st week) and respective treatment conditions (i.e., different identity revelation modes, 2nd week), only individuals who participated in both activities (i.e., represented by \( n' \)) were retained for data analysis for research question 1. As a result, the number of assesses decreased considerably (where \( n' \) ranges from 14 to 23, \( \bar{n'} = 19.2 \)) after deletion, as compared to the number of students in the original class, \( n \), ranging from 31 to 34, \( \bar{n} = 32.7 \) (see Table 1).

To answer research questions 1-1 and 1-2, firstly, high- and low-targeted assessees were operationally defined. With reference to the literature on item discrimination indices (e.g., Beuchert & Mendoza, 1979; Engelhart, 1965) and taking into consideration the small sample size of this study, 33% was chosen as the upper and lower cutoff percentage. The high-targeted subgroups consisted of participants whose total number of received peer-feedback messages in the pre-treatment control condition and the subsequent treatment condition, respectively, were in the upper 33% of their respective treatment groups, whereas the low-targeted subgroups were those whose total number of received peer-feedback messages were in the lower 33%. The recurring members found in both the pre-treatment and treatment conditions were considered redundant assessees. The proportions of redundant assessees for the three different treatment groups were then calculated, and the results were analyzed using a chi-square test. The assessors’ online targeting behavior in different identity modes could be revealed in this way by factoring in the patterns of redundant assessees in both the high- and low-targeted subgroups.

To answer research question 1-3, the two classes receiving the same treatment were combined for statistical tests (Table 1). An ANCOVA was performed to determine if there were significant differences in the average number of peer-feedback messages each student received among the three identity revelation groups. To control for pre-existing differences, data on the same variable in the pre-treatment were used as covariates.

To answer research question 1-4, the two classes receiving the same treatment were combined and then ranked based on the total number of peer-feedback messages received in the pre-treatment condition and the subsequent treatment condition, respectively. The Wilcoxon test was adopted to reveal any significant ranking differences as a result of changes in the three identity revelation modes.

Finally, to answer research question 2, the reasons that participants assigned to the same treatment group chose for their online targeting behavior were tallied together and compared among different identity revelation modes.

Results

Research question 1: Do different identity revelation modes among the interacting parties affect the online interaction behavior and social dynamics of the group in online peer-assessment contexts?

To answer research question 1-1, a two-way contingency table analysis was conducted. The proportions of redundant assessees for the high-targeted subgroup were calculated first, and 68.8%, 66.7%, and 72.7% redundancy rates were found for the anonymity, nickname, and real name conditions, respectively (see Table 1). The results of the chi-square test were not significant, \( \chi^2 (2) = .110, p = .946 \), Cramér’s \( V \) high = .051, indicating that the identity revelation mode and assesse redundancy were not significantly related for the high-targeted assessees subgroup.
Similarly, a two-way contingency table analysis was conducted to answer research question 1-2. The proportions of redundant assessees in the low-targeted subgroup were calculated first, for which a 66.7%, 63.6%, and 65.9% redundancy rate was found for the anonymity, nickname, and real name conditions, respectively (see Table 1). The results of the chi-square test were also not significant, \( \chi^2 (2) = .033, p = .984 \), Cramér’s V low = .028, meaning that the identity revelation mode and assesseee redundancy were not significantly related for the low-targeted assessees subgroup, either.

Furthermore, the assumption of the homogeneity of regression slopes was satisfied, \( F (2, 109) = 1.267, p = .286 \), before proceeding to the ANCOVA. The results of the chi-square test were also not significant, \( \chi^2 (2) = .033, p = .984 \), Cramér’s V low = .028, meaning that the identity revelation mode and assesseee redundancy were not significantly related for the low-targeted assessees subgroup, either.

Finally, for research question 1-4, the Wilcoxon test was performed, and the results showed non-significant differences in matched-pair signed rank between the pre-treatment and treatment conditions for all three different treatment groups (see Table 1).

### Table 1. Descriptive and inferential statistics of the three identity revelation modes

<table>
<thead>
<tr>
<th>Identity revelation modes</th>
<th>Class</th>
<th>( n(n^{1}) )</th>
<th>( m_{\text{pre}} ) (sd)</th>
<th>( m_{\text{post}} ) (sd)</th>
<th>( m_{\text{adj}} ) (se)</th>
<th>Wilcoxon test</th>
<th>Redundant</th>
<th>Pearson ( \chi^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anonymity</td>
<td>1</td>
<td>34 (23)</td>
<td>2.04 (1.41)</td>
<td>2.20 (1.99)</td>
<td>2.32 (0.38)</td>
<td>( Z = -.323 )</td>
<td>68.8</td>
<td>66.7</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>31 (22)</td>
<td>2.30 (2.30)</td>
<td>2.57 (2.71)</td>
<td>2.62 (0.40)</td>
<td>( Z = -.707 )</td>
<td>66.7</td>
<td>63.6</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>32 (16)</td>
<td>3.26 (2.18)</td>
<td>2.83 (2.92)</td>
<td>2.59 (0.45)</td>
<td>( Z = -1.482 )</td>
<td>72.7</td>
<td>65.9</td>
</tr>
</tbody>
</table>

Note. \(^{1}\) = the redundant participants; \(^{2}\) = combined mean and standard deviation for pre-treatment condition; \(^{3}\) = combined mean and standard deviation for treatment condition; \(^{4}\) = adjusted mean and standard error.

**Research question #2: Do the major reasons selected by peer-assessors for their online targeting behavior vary significantly among the three different identity revelation modes?**

As seen in Table 2 and Figure 5, the rankings for the reasons for peer-assessors’ online targeting behavior in the three identity revelation modes reflected very similar patterns. First, the rankings were identical for the real-name and anonymity groups. Second, “question type,” “number of assessors already given feedback on the item,” and “the current average rating” were the three main reasons for all three identity revelation modes, with an average of more than 50% of all participants highlighting their influence on targeting. Third, “who the question-author is” was found to be the factor least likely to affect targeting behavior among the six listed reasons, preceded by “question item itself” and “question submission time” in that order.

**Figure 5.** Rankings of reasons for peer-assessors’ online targeting behavior among different identity revelation modes.
Different interaction patterns and behavior, specific hypotheses regarding differentially ate for high and somewhat understood by the data collected suggesting. Moreover, the ANCOVA results on the number of peer-assessors’ decisions as to which questions to assess in online peer-assessment contexts were high (exceeding 60%) for all of the experimental groups. The non-significant chi-square tests results further demonstrated that the identity revelation mode and asseesee redundancy were not significantly related in the high- or low-targeted asseesees groups. Moreover, the ANCOVA results on the number of peer-feedback messages each student received among the three conditions were non-significant. Finally, the non-significant differences in the Wilcoxon matched-pair signed rank tests reinforced the finding suggesting that the identity revelation mode of each treatment group did not significantly affect their participant rankings in terms of the total number of feedback messages received from peers.

The findings from the quantitative results could be explained and somewhat understood by the data collected from the end-of-session checkbox question. As revealed in Table 2 and Figure 5, the assessors made their decision on which questions to assess based predominately on “question type,” “number of assessors already given feedback on the item,” and “the current average rating.” They paid little attention toward who authored the questions when it came to selecting their targets to complete the online peer-assessment task. The non-significant effect results from the chi-square, ANCOVA, and Wilcoxon tests of all treatment groups can thus be understood.

To conclude, this study did not find supportive, empirical data to validate that masking real identities or recreating identities in an online context would induce different interaction patterns and behavior, specifically, targeting behavior, on the part of participants. In other words, the suspected impacts gleaned from interpersonal relationships and psychological functioning did not undermine the peer-assessors’ targeting behavior in this study. This study provided some initial empirical evidence to support past findings of very limited existing studies that suggest that being in an identifiable or anonymous situations makes no difference on the actual interaction process (Roberts & Rajah-Kanagasabai, 2013; Yu & Sung, 2016; Yu & Wu, 2011).

### Discussion and conclusions

This work investigated the targeting behavior of peer-assessors in three different identity revelation modes (real-name, nickname, and anonymity) in an online peer-assessment context. The probable reasons for peer-assessors’ online targeting behavior were also examined. Both social identity theory and psychological safety studies provided the basis for our proposed hypotheses regarding differences in targeting behavior in the identified and unidentified (i.e., anonymity and nickname) modes. Identity construction studies further provided support for the use of nicknames in online spaces.

Contrary to the authors’ hypotheses, the results from the inferential analyses all found no evidence supporting our proposed hypotheses, and all results pointed to the conclusion that changes in identity revelation modes did not significantly affect peer-assessors’ decisions as to which questions to assess in online peer-assessment situations. To summarize the major findings, foremost, both high- and low-targeted assessesee redundancy rates were high (exceeding 60%) for all of the experimental groups. The non-significant chi-square tests results further demonstrated that the identity revelation mode and assessesee redundancy were not significantly related in the high- or low-targeted assessesee groups. Moreover, the ANCOVA results on the number of peer-feedback messages each student received among the three conditions were non-significant. Finally, the non-significant differences in the Wilcoxon matched-pair signed rank tests reinforced the finding suggesting that the identity revelation mode of each treatment group did not significantly affect their participant rankings in terms of the total number of feedback messages received from peers.

The findings from the quantitative results could be explained and somewhat understood by the data collected from the end-of-session checkbox question. As revealed in Table 2 and Figure 5, the assessors made their decision on which questions to assess based predominately on “question type,” “number of assessors already given feedback on the item,” and “the current average rating.” They paid little attention toward who authored the questions when it came to selecting their targets to complete the online peer-assessment task. The non-significant effect results from the chi-square, ANCOVA, and Wilcoxon tests of all treatment groups can thus be understood.

To conclude, this study did not find supportive, empirical data to validate that masking real identities or recreating identities in an online context would induce different interaction patterns and behavior, specifically, targeting behavior, on the part of participants. In other words, the suspected impacts gleaned from interpersonal relationships and psychological functioning did not undermine the peer-assessors’ targeting behavior in this study. This study provided some initial empirical evidence to support past findings of very limited existing studies that suggest that being in an identifiable or anonymous situations makes no difference on the actual interaction process (Roberts & Rajah-Kanagasabai, 2013; Yu & Sung, 2016; Yu & Wu, 2011).

### Implications of the study

This work has some implications for teaching practice and computer-supported learning system development. The non-significant results of the redundancy rate for high- and low-targeted assesseseees in this study provide some comfort to elementary school instructors interested in introducing peer-assessment tasks in their classes, yet having no access to systems supporting anonymity or nickname use. As found in this study, the revelation or concealment of identity is unlikely to have any bearing on the number of times a focal item is going to be assessed.
Second, the students’ targeting behavior may be directed by various factors. In contrast to the possible adverse effects frequently discussed under identity revelation versus the merit of anonymous assessment literature, this study approached the issue from a different perspective—an investigation of peer assessors’ decision-making process in terms of which question they chose to assess. The non-significant results in the students’ targeting behavior implied that instructors may consider allowing students to have some freedom in this regard because of their differing motivations, interests, needs, and preferences (Ediger, 2006). Along the same humanist line of thought, instructors with access to online learning systems that do support anonymity or nickname use are advised not to arbitrarily impose anonymity and nickname modes since doing so may have adverse effects on interpersonal relationships (Yu & Wu, 2011).

Finally, as reflected in the students’ selections in the end-of-session checkbox question, of the six available options, on average, more than 50% of the participants saw “question type,” “number of assessors already giving feedback on the item,” and “the current average rating,” as the top three factors directing their online targeting behavior. Designers of technology-enhanced learning systems should include details of these features in user interfaces in order to facilitate peer-assessment activities.

Limitations of this study

The present study has some limitations. First, this study took place within intact classes, in which the fifth-grade participants were already familiar with each other prior to the experiment. As assessments may be affected by the existing friendships among students (Chan, 2010), who may guess who the real person is based on the nickname chosen, the findings may not be applicable to online open forums or communities where participants who may or may not know each other or to other age groups.

Second, use of a person’s real-name was set as the pre-treatment condition in this study. It is possible that questions from the same author may exhibit similar writing styles during pre-treatment and treatment conditions in such a way that student identity is revealed, even if nicknames are used or identities are later concealed.

Third, a checkbox question was used in this study to collect data on the participants’ accounts of their online targeting behavior. Although the six responses listed as possible answers were based on informal, non-participant observation of the participants’ selection behavior during the activities, other possible reasons may have been overlooked.

Fourth, the time allocated for interacting with the online system was fixed and limited to in-class use. In addition, the actual study was conducted in two instructional sessions, which may not have been long enough to allow for the effects implied by identity construction theory to manifest. The generalizability of this study to contexts without time and access restrictions and with extended implementation time should thus be exercised with caution.

Finally, the study involved Taiwanese elementary school pupils. Studies from the field of social communication, as well as those examining social perceptions and emotional perceptions, have found that students with different backgrounds have different cultural values (e.g., independent self vs. interdependent self) (Masuda, Ellsworth, Mesquita, Leu, Tanida & Van de Veld, 2008) and exhibit different interaction patterns and behavior under identified and unidentified anonymity conditions in group situations (Chester & Gwynne, 1998; Hosack, 2004). Therefore, future studies involving different age groups with different cultural origins are thus needed to further extend the generalizability of the current work.

References


Partnership among Schools in E-Learning Implementation: Implications on Elements for Sustainable Development

Siu Cheung Kong

Department of Mathematics and Information Technology, The Education University of Hong Kong // sckong@eduhk.hk

(Submitted September 15, 2016; Revised June 18, 2017; Accepted August 8, 2017)

ABSTRACT

The study looked into how school partnership generates benefit. It aimed to identify the structures of partnership among collaborating schools and to examine elements that can contribute to sustainable e-Learning development. Six cluster project cases were purposefully selected from an e-Learning pilot scheme in Hong Kong to investigate how school partnership functions in e-Learning implementation through semi-structured focus group interviews. The findings identified five types of partnership structures that were adopted by the six e-Learning cluster projects, namely, a traditional leader-centered team leadership; a fusion of traditional leader-centered and distributed team leadership; a distributed-coordinated team leadership; an intermediate form of distributed-coordinated and distributed-fragmented team leadership; and a duplicated distributed team leadership structure. Elements including mutual benefit, active school engagement with dynamic communication and interaction, reasonable team size, and co-building of online sharing platform for channeling ideas and actions efficiently are critical to keep e-Learning school partnership sustainable.

Keywords

E-Learning, Partnership structures, School education, School partnership, Sustainable elements

Introduction

E-Learning refers not merely to the use of technology for learning and teaching (Stein, Shephard, & Harris, 2011), it concerns also the pedagogical issues, which focus on the ways of using digital resources, using digital communication tools and collecting learning data to effectively support learning, promote interaction and facilitate pedagogical decision making (Gebre, Saroyan, & Bracewell, 2014; Osborne, Dunne, & Farrand, 2013). E-Learning is multifaceted and requires talents of many parties in the formulation of a diversity of learning and teaching strategies, therefore collaboration is a key to carry out quality e-Learning implementation (Vandenhouten, Lepak, Reilly, & Berg, 2014).

Across-parties collaboration for e-Learning development is important, because different partners provide specialized supports and professional services for schools implementing e-Learning. Major stakeholders who play essential roles in e-Learning implementation include partnership schools, tertiary education sectors, business sectors and parents. This article focuses on the roles of partnership schools, which are crucial to schools that worked on a cluster basis and organized tenacious communities together in disseminating e-Learning practices and experience. Schools need to build a good partnership with different collaborating schools for the successful promotion of e-Learning. The collaborative initiatives and cooperation relationships can be delineated by studying the structures and related elements involved in partnership among cluster schools.

The distribution of leadership, division of labor and closeness of relationship are taken into consideration to investigate the structures of school partnership (Muijs, 2015; OECD, 2001). With partnership schools as an e-Learning stakeholder, teachers can benefit in various ways such as enhancing quality of lesson preparation and delivery, supporting fellows in resources sharing and management as well as reaching a larger pool of students (Wagner, Hassanein, & Head, 2008). The collaborative efforts generated from school clusters enable the schools to ease budget restriction for Information Technology (IT) infrastructure and human resources for e-Learning development, also to minimize the resistance from conservative teachers in adopting e-Learning strategies (Muijs, 2015; Wagner et al., 2008).

Research framework

Literature stresses the importance to identify the needs of school-based e-Learning in the first place; prior to schools’ decisions on when, why and with whom to partner; and hence set clear and common goals, objectives and expectations between the schools and their partners (Duffy & Gallagher, 2015; OECD, 2001). Schools can
be strategic in terms of forming school partnerships, by understanding the cultural and organizational differences as well as similarities of the motives and management styles among different e-Learning partners. It is suggested that there is a variety of stakeholder groups who work jointly with schools, each plays specific roles and exerts influences on the operation and effectiveness of e-Learning implementation (Duffy & Gallagher, 2015; Wagner et al., 2008). In order to build a good e-Learning partnership, it would be more manageable for schools to begin with homogeneous affiliated members, namely the partnership schools. This study focuses on partnership schools as one of the stakeholders, who account for a major part in developing e-Learning initiatives especially in cluster projects.

Mehra, Smith, Dixon, and Robertson (2006) identified four team leadership structures in business domains: (1) traditional leader-centered team leadership structure, of which leadership is centered on a single individual within the team; (2) distributed team leadership structure, of which leadership is dispersed widely across team members; (3) distributed-coordinated team leadership structure, which is one of the derived forms of distributed team leadership that there is a reciprocal tie between the individuals who had emerged as leaders for coordinating team network; and (4) distributed-fragmented team leadership structure, which is another derived form of distributed team leadership that leadership is distributed over multiple team members without an obvious coordination within the team.

The four identified structures serve as the groundwork for the investigation of school partnerships in this study (Mehra et al., 2006). Essential to the concept of partnership, as pointed out by Nehring and O’Brien (2012), is reciprocity, collaboration and the development of mutually beneficial relationships. In the vision of Haines, Gross, Blue-Banning, Francis and Turnbull (2015), reciprocity is able to maintain a dynamic partnership with boosted trust and active engagement among partners, which generates mutual benefit to create a win-win situation. Duffy and Gallagher (2015) and Haines et al. (2015) further asserted that effective communication involving all parties listening to each other, sharing ideas and resources as well as ensuring reciprocal benefit is pivotal to establish a strong and trusting partnership. Another factor that may influence performance in the partnership is the team size. As characterized by Markette (2013), bigger team size is not necessarily the better, but a team with more members is probable to increase the lines of communication exponentially. Besides, although teams in smaller size can more quickly build up team cohesion, larger teams are more likely to contribute more collective intelligence, which also helps to enhance the partnership efficiency (Thompson et al., 2015).

Based on the conventional idea, traditional leader-centered leadership structures are featured as “fixed top-down institutional hierarchical models of leadership and management” with leaders exercising dominating power over the subordinate followers (Jameson et al., 2006, p. 957). By contrast, frequent communication and close cooperation are associated with partnership structures with distributed leadership, in which “a mutual coordinated recognition of leadership authority and attributes by the formal and emergent leader(s) in teams” is emphasized (Jameson et al., 2006, p. 957). According to Mehr et al. (2006), distributed leadership structures in which leaders see each other as leaders within an atmosphere of trust and respect promote participation and information sharing among members, which in turn facilitate superior team performance and team outputs.

An increasing international interest in community of practice and knowledge management has developed along with the acknowledgement of significance of human networking and social capital (Duffy & Gallagher, 2015; Muijs, 2015). Jameson et al. (2006) endorse the view that without an obvious dominant position or hierarchical differentiation between team leaders in a cluster, school partnership structures with distributed leadership characteristics would be favorable for communities of e-Learning practice to function, in which partners share common values and beliefs for effective collaboration. Jameson et al. (2006) also recommended that clear accountability and authority in a cooperative relationship is imperative, especially for a time-limited e-Learning project aiming to achieve specific outcomes. When schools adopt a collaborative team-based approach to e-Learning management, it is anticipated that a healthy relationship in a partnership network relies on trust and reflexivity (Jameson et al., 2006). In other words, school partnership structures allow knowledge sharing and opinion exchange to take place among partner schools, based on respect and interdependence. There are two research questions guiding this investigation: (1) What are the structures of school partnership adopted by e-Learning schools in Hong Kong? (2) What elements pertaining to school partnership are beneficial to sustainable e-Learning development?
Methodology

The study was based on the data collected in the third year of a three-year e-Learning pilot scheme in Hong Kong schools about the processes and outcomes of e-Learning design and implementation, from which the school partnership on e-Learning is the focus of the current study. A multiple-case study approach was adopted to allow an intensive, holistic description and analysis of cases within each setting and across settings (Yin, 2003). Under the e-Learning pilot project, some schools worked individually as a discrete unit on a singleton basis, while some schools worked together with one or more collaborating schools on a cluster basis, the latter types of project schools were purposefully selected to address the scope of e-Learning partnership in the research questions. In total, six cluster project cases, which involved 32 pilot schools, i.e., 14 primary, 8 secondary, and 10 special schools, were selected for investigation. In each cluster case, the coordinating school worked collaboratively with one or more partner schools during the project. Each school was coded with a three-unit project case number: the first letter “C” stands for cluster project; the second digit indicates the number of school cluster; while the last one digit (except C610 which is the last two digits) refers to the individual partnership school within a particular cluster (see Table 1).

<table>
<thead>
<tr>
<th>Project case</th>
<th>Partnership schools</th>
<th>Number of interviewees</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>C11</td>
<td>C11, C15, C16</td>
<td>3</td>
<td>111 min</td>
</tr>
<tr>
<td></td>
<td>C12, C13, C14</td>
<td>5</td>
<td>103 min</td>
</tr>
<tr>
<td>C21</td>
<td>C21, C22</td>
<td>2</td>
<td>76 min</td>
</tr>
<tr>
<td>C31</td>
<td>C31, C32, C33, C34, C35, C36</td>
<td>7</td>
<td>130 min</td>
</tr>
<tr>
<td>C41</td>
<td>C41, C42, C43, C44</td>
<td>4</td>
<td>95 min</td>
</tr>
<tr>
<td>C51</td>
<td>C51, C52, C53, C54</td>
<td>4</td>
<td>118 min</td>
</tr>
<tr>
<td>C61</td>
<td>C61, C62, C63, C64, C65, C66, C67, C68, C69, C610</td>
<td>10</td>
<td>126 min</td>
</tr>
</tbody>
</table>

Total (Average): 35 (5) 759 min (108 min)

Note. C – Cluster case project.

Focus group interviews enable e-Learning stakeholders to make use of their experiences, knowledge and sensibilities to discuss different views of partnership issues (McPherson & Nunes, 2008). Data in this study was collected from semi-structured interviews with school senior management and leading teachers in both coordinating schools and partner schools (see Table 1). Project proposals in this pilot study were also used to examine the aim and objectives as well as the contexts for partnership and collaboration in these school clusters. Representatives from pilot schools commented on their collaborative model in scaling up e-Learning solutions and changes occurred across the cluster schools in the focus group interviews. Specific interview questions, such as, “How to describe the partnership among the e-Learning partnership schools in terms of responsibility, closeness and contribution?”; “What are the gains and advantages generated from the partnership for school e-Learning development?”; “What are the challenges and difficulties faced by the school cluster during the e-Learning implementation?”; were asked in order to collect views from the partnership schools on the (1) contexts, (2) beneficial outcomes, and (3) problems encountered during the collaboration in the implementation of the e-Learning projects. The audiotaped interview records were transcribed for analysis. Information on the three key features of partnership structure (i.e., context, beneficial outcomes and problems encountered) of the six project cases were systematically summarized. The research team with expertise in e-Learning development categorized the partnership structure of each project case by referring to the four team leadership structures empirically reviewed by Mehra et al. (2006), in order to ensure the validity and reliability of such categorization process.

Results and discussion

Five types of partnership structures from the six e-Learning project cases were identified. The categorization results were triangulated at two levels: a research team member who visited the representative schools in all school clusters for classroom observation and focus group interviews first agreed with the categorization results; and then another research team member who had no school visit but transcribed all focus group interview audio records agreed with the categorization results. The key features about the context, beneficial outcomes and problems encountered of the six e-Learning project cases were listed in Tables 2 to 7. Each project case is discussed with a brief description to elaborate the key features; and illustrated in a comprehensive figure (see Figures 1 to 6).
Partnership of project case C51

The school partnership of project case C51 (see Table 2) was initiated by the coordinating school C51 with C52, C53 and C54 as the partner schools in order to construct a Liberal Studies online learning platform with a question database for teaching resource management and students’ knowledge assessment.

Table 2. Details of context, beneficial outcomes and problems encountered of partnership structure in project case C51

<table>
<thead>
<tr>
<th>C51 Partnership</th>
<th>Traditional leader-centered team leadership</th>
</tr>
</thead>
</table>
| Context         | • C51 gathered schools to join efforts in constructing a Liberal Studies online learning platform with a question database for teaching resource management and students’ knowledge assessment; established the online learning platform; shared IT information and support; provided iPad training and e-Learning workshops for teacher development.  
• C52, C53, C54 made contributions to designing the online learning platform, producing Liberal Studies test questions and trying out e-Learning practice in different Liberal Studies topics.  
• Example of feedback - from C51 representative: “As the coordinating school, we focused on designing the online learning platform in the first year. Our partner schools at that stage mainly gave comments on the designed platform. We then provided training for partner schools to conduct trial teaching using the designed platform across different subject topics. We regularly gathered partner schools to exchange experiences and suggestions on e-Learning with the use of the designed platform.”  
• Example of feedback - from C53 representative: “We felt satisfied with the collaborative work on designing the online learning platform and exchanging e-Learning resources. There were a number of success factors for this cluster: first, the regular meeting and sharing arranged by the coordinating school for coordinating all member schools to collaboratively design tasks and collect resources for school-based e-Learning trials; second, the professional advice from the coordinating school on the possible challenges and feasible practices for school-based pedagogical integration of e-Learning resources in subject learning activities; and third, the opportunities among the member schools to make collaborative class observation and teaching reflection for valuable comments on improving e-Learning trial in each member school.”  
• Example of feedback - from C53 representative: “The partner schools were not active to start the assigned tasks in implementing the e-Learning initiative. We often waited for the start by the coordinating school, which tried hard to tackle challenges of insufficient time for an intensive technical and pedagogical guidance to the school-based e-Learning trial in each partner school.” |
| Beneficial outcomes | • An online learning platform with a Liberal Studies question databank established [Element 4];  
• E-Learning teaching materials and pedagogy enriched [Element 1];  
• E-Learning teaching experience shared [Element 2];  
• A division of labor [Element 3];  
• Project outcome of e-Learning pedagogical designs, experience and resources disseminated through sharing sessions to non-project schools [Elements 1 & 2].  
| Problems encountered | • Insufficient time in preparing e-Learning lessons due to limited work force [Element 3];  
• Partner schools not active enough to take initiative [Element 2];  
• In need of technology company to provide free technical support to build e-Learning infrastructure in schools [Element 4].  
• Example of feedback - from C53 representative: “The partner schools were not active to start the assigned tasks in implementing the e-Learning initiative. We often waited for the start by the coordinating school, which tried hard to tackle challenges of insufficient time for an intensive technical and pedagogical guidance to the school-based e-Learning trial in each partner school.” |

Note. Element 1 - Mutual benefit; Element 2 - Active school engagement with dynamic communication and interaction; Element 3 - Reasonable team size; Element 4 - Co-building of online sharing platform for channeling ideas and actions efficiently.

The model of partnership in this school cluster resembled a traditional leader-centered team leadership structure. C51 was the single school at the center of the leadership, which played a starter role and coordinated communication among the partner schools. It took on the major responsibility for designing the teaching framework and building the online learning platform. Under the leadership of C51, partner schools made collaborative efforts to refine the teaching resources on the online platform and to share their ideas and experience for the improvement of the e-Learning practice. All partner schools agreed that the sufficient assistance from the coordinating school C51 was helpful in achieving the project goals in their schools. The
collaborative relationship was considered successful as it accelerated the establishment of the online learning platform with a resourceful Liberal Studies question databank. Meetings, workshops and classroom observations relating to e-Learning were organized for teacher development. For improvement, teachers mainly reflected that there was insufficient time to prepare the main contents and supporting measures for the e-Learning lessons, which was largely resulted from limited work force. Partner schools recognized their lack of initiative, but were motivated by the project leader C51 to move on in the process. The partnership could be improved if more tasks were involved by the partner schools.

Based on the context and style of partnership observed, it is identified that the school cluster adopted a traditional leader-centered team leadership structure (Mehra et al., 2006) (see Figure 1). The coordinating school C51 was leading in technology and thus the partnership mainly benefited teacher development at the technical level.

![Figure 1. Traditional leader-centered team leadership structure in project case C51](image)

**Partnership of project case C41**

Under the same school sponsoring body, C41 initiated the partnership of project case C41 with three partner schools C42, C43 and C44 (see Table 3), aiming to encourage student-centered learning, build up students’ 21st century skills, and promote self-directed study along with the latest technological advances.

<table>
<thead>
<tr>
<th>Context</th>
<th>Beneficial outcomes</th>
<th>Problems encountered</th>
</tr>
</thead>
</table>
| C41 initiated the partnership for the common project goals; played the role as coordinator and decision-maker; sought outsourced support to develop teaching materials for the project; purchased IT infrastructure for partner schools; shared ideas and experience on e-Learning implementation; administrated the information hub on Facebook. | • The formation of an information hub on Facebook for information, ideas, experience sharing and teaching reflection among cluster schools [Elements 1 & 2];  
• Experience, resources and project outcome of e-Learning implementation disseminated via the information hub to project and non-project schools [Element 2];  
• A division of labor [Element 3].  
• Example of feedback - from C44 representative: “The coordinating school set up a Facebook group for the prompt uploading of e-Learning resources and the useful posting of pedagogical ideas and teaching examples related to the uploaded e-Learning resources. As a partner school, we valued these inputs for the convenience of experience sharing and feedback exchange among the four schools in the cluster.” | • Deficient human resources resulting in teachers’ heavy workload [Element 3];  
• Insufficient time for e-Learning development [Element 3].  
• Example of feedback - from C43 representative: “This collaborative e-Learning initiative was meaningful and successful. We noted the big challenge of the coordinating school to..." |

Table 3. Details of context, beneficial outcomes and problems encountered of partnership structure in project case C41
spend sufficient time and effort on giving each partner school intensive guidance and follow-up support for the existing implementation and future development of school-based e-Learning trial.”

Note. Element 1 - Mutual benefit; Element 2 - Active school engagement with dynamic communication and interaction; Element 3 - Reasonable team size; Element 4 - Co-building of online sharing platform for channeling ideas and actions efficiently.

As the project initiator, C41 coordinated tasks and communication among partner schools, made important decisions, sought outsourced support to develop teaching contents, and purchased IT facilities for partner schools. The three partner schools tried out e-Learning practice independently with advice and support from C41. E-Learning policies and pedagogical activities were different among the schools in this cluster. The cooperation mainly emphasized the sharing of ideas. C41 was highly appreciated by its cluster members for all the effort and information provided, especially the ideas and experience on e-Learning implementation with self-produced teaching materials.

In this cluster, classroom observations were held spontaneously and meetings were arranged when necessary, while WhatsApp and email were used as usual communication tools and video conferencing as the major means for “face-to-face” discussion. The attribute of this partnership was the formation of an information hub on Facebook for within-group sharing about teacher development courses, workshops, useful apps, description, reflection, photos and videos of e-Learning lessons. However, teachers found the workload heavy because of insufficient human resources and preparation time for e-Learning development.

With respect to partnership structure, on the one hand, C41 functioned as a traditional leader coordinating information dissemination, providing technical support and facilitating interactions. On the other hand, partner schools worked independently and automatically to design the appropriate pedagogy and teaching materials to adapt to their own school context, diffusing the leadership across the cluster members. With the aid of the information hub, school participation and information sharing were enhanced. This school cluster demonstrated a fusion of traditional leader-centered and distributed team leadership structure (see Figure 2), which is a derivative structure from the traditional leader-centered team leadership structure and the distributed team leadership structure (Mehra et al., 2006).

![Figure 2. Fusion of traditional leader-centered and distributed team leadership structure in project case C41](image)

**Partnership of project case C21**

The school partnership in project case C21 between C21 and C22 (see Table 4) was initiated for enhancing school learning and teaching with e-Learning resources and pedagogy as well as improving students’ self-directed learning and study motivation. As the initiator, C21 shared its IT knowledge and information with its partner school C22 that had less experience in e-Learning. It also took full responsibility for the coordinating and administrative work. Both C21 and C22 shared equal workload in developing e-Learning resources, which accounted for identical amount of different topics in Chinese Language and Mathematics for Primary 4 and 5.

Teachers in this cluster were in favor of the partnership because there was an equal division of labor for their work on preparing e-Learning lesson plans and teaching materials. Collaborative lesson planning and discussions, ideas and experience sharing, peer classroom observations, teaching reflection and evaluation, all provided inspiration and new insights into the pedagogical approach of e-Learning practice, which was of great significance for teacher development. Although the partnership between C21 and C22 was closely connected and in harmony in the cooperation, having only two members in the project gave teachers plentiful burden in academic workload due to insufficient work force. Both schools believed that if more schools engaged in the
partnership, more work force and ideas would be generated to elevate the quality and efficiency of the e-Learning implementation.

Table 4. Details of context, beneficial outcomes and problems encountered of partnership structure in project case C21

<table>
<thead>
<tr>
<th>C21 Partnership</th>
<th>Distributed-coordinated team leadership</th>
</tr>
</thead>
</table>
| **Context**     | - C21 initiated the partnership in the hope of enhancing school learning and teaching as well as improving students’ self-directed learning and study motivation; shared IT information with the partner school; took charge of the coordinating and administrative work.  
- C21 and C22 shared equal workload in developing e-Learning teaching materials; conducted collaborative lesson planning and provided feedback and advice for improvement; organized meetings and peer classroom observations beneficial to teacher development.  
- Example of feedback - from C21 representative: “As the coordinating school, we took charge of major administrative tasks in the cross-school e-Learning initiative, such as external liaisons and meeting arrangements. For the academic design of subject-specific e-Learning trials, we and the partner school equally shared the workload for both the Chinese Language learning stream and the Mathematics learning stream. Regular meetings were organized by the two streams for various partnership supports, including checks on trial progress, reminders on trial implementation, sharing of experiences and difficulties, collaborative preparation of e-Learning trial lessons.” |
| **Beneficial outcomes** | - An equal division of labor in lesson plans and teaching materials preparation [Element 1];  
- Teachers’ insight into e-Learning implementation enlarged [Element 1];  
- The co-building and sharing of an online platform with well-prepared teaching resources [Element 4];  
- E-Learning resources, experience and project outcome disseminated within and beyond the cluster schools through the closing ceremony [Elements 1 & 2].  
- Example of feedback - from C22 representative: “The collaborative efforts for school-based e-Learning implementation widened the horizons of both member schools in the cluster for e-Learning development. We benefited much from the teaching resources and lesson plans co-prepared by both member schools; and the open sharing and discussion of teaching ideas related to e-Learning trials on the online platform co-developed for the cluster.” |
| **Problems encountered** | - Heavy teacher workload due to insufficient human resources [Element 3];  
- More partners wanted for more ideas and work force to enhance the quality and efficiency of the e-Learning implementation [Element 3];  
- In need of more funding for the allocation of human resources, the maintenance and update of e-Learning facilities [Element 3].  
- Example of feedback - from C21 representative: “We were very satisfied with the co-development of teaching resources for e-Learning trials. However, the workload for teachers who took charge of this aspect became too heavy. We expected more peer support and less teaching workload for those teachers, with the maintenance and update of e-Learning infrastructure for efficient preparation of e-Learning resources.” |

Note. Element 1 - Mutual benefit; Element 2 - Active school engagement with dynamic communication and interaction; Element 3 - Reasonable team size; Element 4 - Co-building of online sharing platform for channeling ideas and actions efficiently.

Notwithstanding C21 took charge of the overall planning and external contact in this project, C21 and C22 had the same level of academic responsibility in e-Learning implementation. Dissimilar with the traditional leader-centered team leadership structure with only one single leader standing at the core of the leadership network, C21 and C22 recognized each other as the leader in preparing and designing the teaching materials in different topics with shared equal workload, and made consensual decisions to achieve concerted practices. Their leadership contributed to a harmonious and reciprocal relationship by synchronizing their leadership efforts in such a way that decisions and actions could channel more effectively within the partnership (Jameson et al., 2006; Mehra et al., 2006). However, the lack of school partners resulted in a partnership with restricted team size. Therefore, the partnership in the C21 project displayed a distributed-coordinated team leadership structure (Mehra et al., 2006) in a small scale (see Figure 3).
Partnership of project case C31

C31 initiated the partnership in project case C31 (see Table 5) with five other schools under the same school sponsoring body for developing e-Learning materials and establishing an effective learning management system for two key subjects to promote teaching efficiency and students’ self-directed learning. C31 was the partnership initiator and coordinator for making the overall administration, and providing its partner schools with technical backup to reduce their time and ease their financial burden of acquiring hardware groundwork. C33 was another leadership school. C31 led C32 and C35 as a subgroup for compiling e-Learning materials for the subject Putonghua as a Medium of Instruction in Teaching Chinese Language, while C33 directed C34 and C36 as another subgroup for producing e-Learning materials for the subject General Studies. The two subgroups exchanged their e-Learning materials, demonstrating a combination of intimate interaction and interdependence.

C31 and C33 directed and supported partner schools in their subgroups for collaborative lesson planning, classroom observations in turns, and follow-up meetings for discussion. The formation of “Study Circle,” in which members in this cluster interacted through both in-person meetings and virtual communications using email, instant messenger and cloud storage, was highly treasured for improving communication and ideas sharing, establishing a functioning and sustainable teacher network, and increasing confidence and willingness to try e-Learning practice. However, the participating teachers called for a better school policy to improve their working condition by cutting down their teaching time and increasing the workforce so that they could concentrate on experimenting e-Learning.

Table 5. Details of context, beneficial outcomes and problems encountered of partnership structure in project case C31

<table>
<thead>
<tr>
<th>C31 Partnership</th>
<th>Distributed-coordinated team leadership</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context</td>
<td>• C31 initiated the partnership for compiling e-Learning materials and establishing a learning management system for two subjects to promote teaching efficiency and students’ self-directed learning.</td>
</tr>
<tr>
<td></td>
<td>• C31 as the coordinating school and C33 as the leading partnership school each formed a subgroup responsible for e-Learning development in either of the subjects.</td>
</tr>
<tr>
<td></td>
<td>• C31 took charge of the overall administration, provided technical support, and contacted external partners for supporting the whole project.</td>
</tr>
<tr>
<td></td>
<td>• C31 and C33 provided administrative and operational backup, gave directions to designing e-Learning teaching materials and pedagogy within the subgroup, and organized sharing talks and sessions for teacher development.</td>
</tr>
<tr>
<td></td>
<td>• Partner schools engaged in subgroup-based collaborative lesson planning, took turns to organize classroom observations and arranged follow-up meetings for sharing and discussion.</td>
</tr>
<tr>
<td></td>
<td>• Schools modified the lesson plans and contacted their own external partners to ensure the teaching materials suitable for the school context before using.</td>
</tr>
<tr>
<td></td>
<td>• The adoption of a “Study Circle” improved interactions, knowledge exchange and experience sharing within the cluster.</td>
</tr>
<tr>
<td></td>
<td>• Example of feedback - from C31 representative: “The six member schools in this cluster shared the same objectives and coherent rationales for e-Learning development to facilitate students’ self-directed learning. As the coordinated school, we not only took charge of the administrative issues in the cluster, but also provided pedagogical and technical supports for each partner school to arrange school-based e-Learning trials. The strong link and extensive communication within the cluster fostered all six member schools to regularly organize joint-school professional development activities such as classroom observation and training seminars for increasing knowledge and sharing experience of e-learning development.”</td>
</tr>
<tr>
<td>Beneficial outcomes</td>
<td>• Teaching materials including lesson plans and technical devices shared [Element 1];</td>
</tr>
<tr>
<td></td>
<td>• A division of labor [Elements 1 &amp; 3];</td>
</tr>
<tr>
<td></td>
<td>• Teachers’ insight into e-Learning implementation broadened with increasing confidence in</td>
</tr>
</tbody>
</table>
e-Learning [Element 1];

- Communications improved and a well-functioning teacher network established by forming a “Study Circle” for knowledge and experience sharing [Elements 1 & 2].
- Example of feedback - from C35 representative: “The cluster formed a ‘Study Circle’ which facilitated the six member schools to engage in professional sharing of e-Learning development. Through the continuous observation and sharing of school-based e-Learning implementation within the cluster, even the teachers who at first felt reluctant to e-Learning in our school became adapted to integrate e-Learning into subject lessons for improving teaching and learning effectiveness.”

Problems encountered

- Heavy workload for teachers [Element 3];
- More funding needed for purchase, maintenance and update of e-Learning infrastructure, such as the tablets and an integrated self-sustaining e-Learning platform for the cluster [Element 4].
- Example of feedback - from C34 representative: ”We concern most the provision of tablets for e-Learning implementation. The total number of tablets in our school now is insufficient for every student in one class to use in an e-Learning trial lesson. Moreover, the existing model of tablets cannot fully support the smooth operation of the designed e-Learning activities. We need funding support for purchasing the updated model of tablets for maintaining the e-Learning initiative.”

Note. Element 1 - Mutual benefit; Element 2 - Active school engagement with dynamic communication and interaction; Element 3 - Reasonable team size; Element 4 - Co-building of online sharing platform for channeling ideas and actions efficiently.

With C31 and C33 as two leaders each directing a subgroup to compile e-Learning materials for either of the two subjects, the leadership of this cluster was distributed. The successful cooperation and intimate interaction relied not only on the leadership of the two leaders within and between the subgroups, but also the active teamwork performance of the partner schools (Jameson et al., 2006; Mehra et al., 2006). Being both interdependent and autonomous in preparing the suitable e-Learning materials, each school in this cluster had made its own contributions in the course of e-Learning development. With a six-member-school composition, the partnership in the C31 project was of distributed-coordinated team leadership structure (Mehra et al., 2006) with a relatively reasonable team size (see Figure 4).

Figure 4. Distributed-coordinated team leadership structure in project case C31

Partnership of project case C11

Similar with project case C31, the coordinating school C11 started the partnership with five partner schools for trying out e-Learning in four subjects and nurturing students’ capacity for self-directed learning (see Table 6). The work in this cluster was equally assigned to two subgroups led by C11 and C12 respectively. According to school expertise, C11 led C15 and C16 for developing Mathematics and English Language e-Learning materials, while C12 led C13 and C14 for developing Chinese Language and General Studies e-Learning materials. The two subgroups worked separately with limited interactions between each other, though they shared their well-prepared lesson plans and teaching materials. The overall administrative and operational support was provided by the coordinating school C11.

Different from project case C31, project case C11 first assigned the two subgroups leaders to independently design lesson plans and teaching materials. Partner schools in each subgroup then practiced the lesson plans and teaching materials in class, with follow-up feedback and advice for amendment. Each subgroup conducted classroom observations and discussions, and on-site technology training and e-Learning workshops. The cluster
schools maintained a positive attitude towards the partnership with a clear division of work. Teachers in each subgroup found they enriched pedagogical experience, gained important visions and built teaching confidence in e-Learning implementation. However, teachers expected more preparation time and less teaching duty for designing e-Learning lessons.

Table 6. Details of context, beneficial outcomes and problems encountered of partnership structure in project case C11

<table>
<thead>
<tr>
<th>Partnership</th>
<th>Intermediate form of distributed-coordinated and distributed-fragmented team leadership</th>
</tr>
</thead>
</table>
| Context     | • C11 initiated the partnership with 5 partner schools for the common goals to try out school e-Learning in 4 subjects and enhance students’ self-directed learning.  
• C11 as the coordinating school and C12 as the leading partnership school each formed a subgroup accountable for e-Learning development in 2 different subjects based on school expertise.  
• The 2 subgroups worked separately with limited interactions but shared the teaching materials when well-prepared.  
• C11 provided overall administrative and operational support for the cluster.  
• C11 and C12 designed the lesson plans and teaching materials independently, invited their partner schools for classroom observations and discussions, and offered assistance in teacher development programs within the subgroup.  
• Partner schools practiced the lesson plans and teaching materials designed by their leading schools on a subgroup basis, gave feedback and advice for refinement.  
• Partner schools in C11 subgroup began to develop the teaching materials on their own in the third year while those in C12 subgroup kept on following and practicing the lesson plans developed by C12.  
• C13 in C12 subgroup only cooperated to develop the teaching materials of one subject due to the use of different textbook of the other subject from its subgroup members.  
• Example of feedback - from C11 representative: “The member schools in the cluster all shared the similar rationale for e-Learning development and aimed to collaboratively develop suitable teaching and learning resources for e-Learning trials according to school needs. As the coordinating school, we tended to be an initiator and then a supporter, but not a director; to stimulate the planning and support the implementation of school-based e-Learning trials among the partner schools. We focused on organizing classroom observation activities with post-observation meetings for the partner schools to concretely gain insights into the design and implementation of school-based e-Learning trials.”  
| Beneficial outcomes | • A division of labor [Element 1];  
• A school network established [Element 1];  
• Information, knowledge and experience in e-Learning obtained [Element 1];  
• Lesson plans and teaching materials shared [Elements 1 & 2];  
• Teachers’ insight into e-Learning implementation expanded [Element 1].  
• Example of feedback - from C15 representative: “The greatest benefit in this cluster project is widening the horizons of e-Learning development. At the level of school planning, we gain insights into the desirable directions of e-Learning implementation in primary school classrooms. At the level of teacher development, we gain knowledge of effective e-Learning pedagogies for addressing students’ learning difficulties. We value the sharing of lesson plans and teaching resources co-developed by the member schools.”  
| Problems encountered | • Insufficient time and human resources for e-Learning development [Element 3];  
• Deficient school infrastructure for e-Learning lesson implementation [Element 2].  
• Example of feedback - from C16 representative: “We faced two major challenges: there were insufficient tablets for classroom use in e-Learning trial lessons; and the teachers were overloaded to prepare for e-Learning pedagogies and resources.”  

Note. Element 1 - Mutual benefit; Element 2 - Active school engagement with dynamic communication and interaction; Element 3 - Reasonable team size; Element 4 - Co-building of online sharing platform for channeling ideas and actions efficiently.

The project case C11 adopted a similar partnership structure with similar team size as in project case C31, that is, the distributed-coordinated team leadership structure, with two teams devoting to subjects of expertise. Despite the fact that the two subgroups equally shared the workload in designing the teaching plans and materials which would be exchanged within the cluster when well-prepared, the bonding between the two subgroups in project case C11 was not strong to form a concrete distributed-coordinated team leadership structure owing to rare inter-
group communication and interaction (Haines et al., 2015; Nehring & O’Brien, 2012). Correspondingly, project case C11 could be regarded as a halfway structure from distributed-coordinated to distributed-fragmented team leadership structure (see Figure 5) derived from the structures termed by Mehra et al. (2006).

![Figure 5. Intermediate form of distributed-coordinated and distributed-fragmented team leadership structure in project case C11](image)

**Partnership of project case C61**

The school partnership in the last project case directed by the coordinating school C61 was established among 10 special schools with a partnership background under a previous project (see Table 7). It aimed to build an online platform with suitable teaching plans and materials to support teachers in implementing e-Learning in schools for students with physical or intellectual disabilities, and to improve these students’ learning efficiency by adopting student-centered approach assisted with e-Learning pedagogy.

This project case focused on two learning areas, namely Personal, Social and Health Education (PSHE) and Chinese Language. All cluster schools arranged teachers and staff to engage in the e-Learning preparation process in both learning areas. C62 and C63 were the two leading partnership schools coordinating the efforts with the other nine schools to develop e-resources of PSHE and Chinese Language respectively. Meanwhile, C61 was the overall coordinator and invigilator of the whole project for the support of cluster communication, project logistics, project finance, system operation, and teacher development programs.

A steering committee consisted of 10 school principals was authorized to oversee and review the work progress for a well-structured organization and so the smooth operation of the project among the 10 cluster schools in two writing teams for e-Learning development of PSHE and Chinese Language. An accelerator team composed of representatives from C61 (IT coordinator), C62 (PSHE teacher), C63 (Chinese teacher) and C68 (speech therapist) was set up to formulate the work schedule and monitor the progress in teaching plans and materials development by regulating the tasks and communication between the two writing teams. A technical team constituted of teachers with higher IT competency from each school was formed to solve the problem hindering e-Learning particularly due to students’ physical disability by designing special devices, such as installing a panel with buttons on the wheelchairs. Furthermore, schools catering for similar special needs of students paired up and formed small groups with varied assigned duties. Tasks were distributed to different schools in accordance with specialty of fellows in the schools to guarantee optimal quality of work.

**Table 7. Details of context, beneficial outcomes and problems encountered of partnership structure in project case C61**

<table>
<thead>
<tr>
<th>C61 Partnership</th>
<th>Duplicated distributed team leadership</th>
</tr>
</thead>
</table>
| Context         | • The partnership network was established under a previous project with 10 special education schools aiming to build an online platform to support teachers in e-Learning implementation, and to enhance students’ learning efficiency by promoting student-centered pedagogy with the aid of IT devices.  
• C61 as the overall coordinator and invigilator provided assistance to all the details in the events.  
• C62 and C63 each formed a writing team leading the other 9 partner schools to develop e-Learning resources on one of the two key learning areas respectively.  
• A steering committee consisted of 10 school principals to oversee and review project progress.  
• An accelerator team was set up as a middle management to formulate the work schedule and coordinate tasks between the writing teams.  
• A technical team constituted of teachers with higher IT competency from each school was responsible for system maintenance and IT support. |
• A learning platform with lesson plans and teaching materials was constructed by joint efforts of all schools in the cluster.
• Example of feedback - from C61 representative: “This cluster consisted of 10 special schools of which all had a common motivation that there was a great need of an online learning platform tailored for special schools to efficiently share e-Learning resources for different subjects. Each of the 10 schools assigned a teacher with high IT competency to join the technical team in the cluster to contribute expert knowledge for integrating accessible technological devices with suitable e-Learning pedagogies for teaching students with different special educational needs.”

<table>
<thead>
<tr>
<th>Beneficial outcomes</th>
<th>• An online learning platform with scheme of work and e-Learning resources shared [Element 4];</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• E-Learning implementation efficiency improved due to the clear division of labor [Element 2];</td>
</tr>
<tr>
<td></td>
<td>• An upgrade of teachers’ ability of IT usage [Element 1];</td>
</tr>
<tr>
<td></td>
<td>• An enhancement of IT application in schools [Element 1];</td>
</tr>
<tr>
<td></td>
<td>• An excellent environment created for e-Learning implementation with teamwork spirit [Elements 1 &amp; 2];</td>
</tr>
<tr>
<td></td>
<td>• Insight into e-Learning implementation gained in schools for students with physical or intellectual disabilities [Element 1];</td>
</tr>
<tr>
<td></td>
<td>• A change of mindset of teachers and parents on the impact of e-Learning [Element 1].</td>
</tr>
<tr>
<td>Problems encountered</td>
<td>• Insufficient time and resources for e-Learning development - desire for more funding to better allocate human resources and to develop tailor-made software to cater for learner diversity [Element 3];</td>
</tr>
<tr>
<td></td>
<td>• Heavy workload for teachers - desire for more technical support to solve the difficulties in integrating pedagogy with technology and technical problems in implementation [Element 3].</td>
</tr>
<tr>
<td></td>
<td>• Example of feedback - from C63 representative: “We collaborated to prepare e-Learning lesson plans, develop and try the online learning platform, and conduct e-Learning trial teaching. This cluster consisted of the hierarchies of steering committee with school principals, accelerator team with middle management teachers, and technical team with teachers with high IT competency. The strong team spirit fostered all member schools to actively use the online learning platform and regularly organize cluster meetings to report implementation progress, exchange pedagogical experience, and share teaching resources in school-based e-learning trials. All member schools worked closely for the common goal of benefitting students in special schools to receive e-Learning opportunities in lesson time.”</td>
</tr>
</tbody>
</table>

Note. Element 1 - Mutual benefit; Element 2 - Active school engagement with dynamic communication and interaction; Element 3 - Reasonable team size; Element 4 - Co-building of online sharing platform for channeling ideas and actions efficiently.

This cluster constructed a platform with scheme of work able to be accessed and modified according to school-based teaching plans. Due to the distinct expertise of the cluster schools, this platform accumulated a wide range of resources systematically to serve as a central management system to facilitate lesson preparation and increase teaching quality. The most delighting part in the partnership was the creation of an exceptional environment for e-Learning through teamwork, in which information and ideas on e-Learning pedagogy and practices were shared and exchanged. Teachers gained insight into e-Learning in schools for students with physical and intellectual disabilities in the hope of transferring e-Learning to the other learning areas.

This cluster indicated the room for improvement by putting adequate funding in order to have a better arrangement of human resources for e-Learning development and to design tailor-made software to meet the needs of different students. The teachers expected sufficient time and resources for school e-Learning development. They also expected more technical support not only to help them integrate pedagogy with technology for special school students, but also to fix the IT problems happened during e-Learning lessons. The teachers also hope for preventing the assignment of several roles at the same time, e.g. an operation manager of the writing team for the e-Learning development of either PSHE or Chinese Language could simultaneously be a member of the accelerator team.
The project case C61 displayed a duplicated distributed team leadership structure, a derivative of the distributed team leadership structure identified by Mehra et al. (2006), in which two structurally identical writing teams were formed under the leadership of C62 and C63 respectively (see Figure 6). Each writing team was responsible for the e-Learning development of one key learning area, with C61 as the overall coordinator to assist and facilitate the communication and interaction between the writing teams. All schools within the cluster sent teachers and experts to engage in both of the teams. Joint efforts of members were much more emphasized instead of great reliance of work on a dominant leader. Since more participants played a more active part, or even took up multiple roles in knowledge, experience and resource sharing, collaboration was reinforced with boosted mutual benefit (Haines et al., 2015; Jameson et al., 2006; Markette, 2013; Thompson et al., 2015). The partnership in project case C61 exerted constructive influence on e-Learning development.

![Figure 6. Duplicated distributed team leadership structure in project case C61](image)

**Implication**

From the five types of school partnership structures identified from the six e-Learning cluster projects, this study reveals four elements critical for school partnership to effectively maintain e-Learning implementation. They are mutual benefit, communication and interaction, team size, and the online sharing platform.

**Mutual benefit**

Mutual benefit among the participating schools served as the most essential reason to practice and continue the partnership (Nehring & O’Brien, 2012). Results from the six project cases C11, C21, C31, C41, C51 and C61 indicated that they did share mutual benefits. Based on the specific needs of different schools and cluster projects, leadership schools emerged by virtue of their knowledge and experience in implementing e-Learning, which were able to offer partner schools advice and guidance (OECD, 2001). In return, partner schools engaged in lesson planning, teaching material preparation, trial classes, co-building of the e-Learning online platform, helping to improve the e-Learning pedagogy and teaching quality and to accelerate the development progress. Reciprocal benefit generated by school partnership thus performed as a bonding tie to draw these schools into a collaborative network (Haines et al., 2015; Nehring & O’Brien, 2012). Findings indicated that mutual benefit in these project cases included the sharing of IT information, ideas and experience on e-Learning implementation, the exchange of lesson plans and teaching materials, and the division of work to ease the workload of teachers.

**Communication and interaction as well as team size**

In spite of the aforementioned mutual benefit, the communication and interaction among the associated partners as well as the team size of the partnership structure also significantly affect the practice and effectiveness of school partnership for e-Learning development. This explains why the following project cases employed similar partnership structures but resulted in various outcomes.
Case C41 and C51 were similar in team size and model. However, case C41 was in a more dynamic structure; compared with case C51 in which a single school C51 was at the center of the leadership network with superior performance in directing and propelling the collaboration with its less technologically experienced partners due to its competency in technology use. In case C41, the leadership was distributed when the partner schools took charge of designing tailor-made teaching materials based on their own school context with the technical support from the coordinating school C41. This kind of fusion of traditional leader-centered and distributed team leadership structure enhanced team performance when being compared with the pure traditional leader-centered team leadership structure in project case C41. It is because partner schools took more initiative and involved in more active team communication and interaction, which, in turn, expedited the process of information, experience and reflection sharing about e-Learning practice (Jameson et al., 2006; Mehra et al., 2006).

The comparison of case C11, case C31 and case C21 further highlights the importance of these two critical elements. The partnership structures of these three cases all belonged to distributed team leadership patterns. Each of the three cases could be regarded as a combination of two subgroups sharing the equal division of academic workload.

Case C11 and case C31 adopted a similar partnership structure with similar team size. Case C11 operated in a mechanism with two subgroup leaders each directing two partner schools to develop e-Learning teaching materials for two diverse subjects respectively. Although partner schools in this cluster cooperated closely within the subgroup, there was scarce inter-group communication and interaction, which contrasted strikingly with the case C31. Even if the e-Learning lesson plans and teaching materials would be exchanged, as a whole the link to support the partnership between the two subgroups was weak and likely to be torn apart since actions and efforts were hard to be synchronized (Mehra et al., 2006). This made the partnership structure of case C11, which could be considered as an intermediate form of distributed-coordinated and distributed-fragmented team leadership structure, not effective to maintain or promote the reciprocal influence and thus not a very desirable type of partnership for e-Learning development in the long run.

Case C21 and case C31 demonstrated features of the distributed-coordinated team leadership structure. These two cases were different in team size, but both comprised two teams working simultaneously with an exchange of information and developed resources in the e-Learning project. Though the two cases involved active engagement of the team members and systematic division of labor, the approach of work allocation were different. In project case C21, academic tasks were assigned to the two teams with topics equally distributed. In project case C31, two subgroups were formed each responsible for one subject. The divergent team size in these two cases also brought a noticeable difference between the two projects on e-Learning development. In project case C31, each team was made up of three school members; and there were in total six schools in the cluster making joint efforts to develop e-Learning resources. In project case C21, the team size was small with only one member in each team. The work force was therefore weakened, which restricted not only the circulation of information and knowledge, but also the efficiency of school partnership (Markette, 2013; Thompson et al., 2015).

Co-building of online sharing platform for channeling ideas and actions efficiently

Compared with the partnership structures of five previously mentioned project cases, the last structure identified in project case C61 might have more possibility to maintain sustainability. Case C61 was in a complicated structure, which consisted of two identically structured writing teams led by two partnership schools C62 and C63 respectively. With C61 as the overall coordinator, the partnership functioned systematically with 10 participating schools all dedicating themselves to both of the writing teams, each of which was working on the e-Learning resources of one key learning area. The relatively large team size and the clear division of labor according to expertise were the apparent characteristics of this structure, which provided more work force and knowledge to operate the partnership (Jameson et al., 2006; Markette, 2013; Thompson et al., 2015). However, what made it more productive and energetic was the dynamic communication and interaction among the cluster members via both face-to-face meetings with support from the school principals and an online sharing platform, so that ideas and actions could be channeled within the cluster efficiently (Haines et al., 2015). Dissimilar with case C51, in which the online learning platform was incomplete and being constructed under the leadership of the coordinating school C51 to inspire the three partner schools in the cooperation, the online learning platform in project case C61 was mature, well-structured and resourceful, which was an effective outcome of joint efforts from a larger scale of member schools with more enthusiastic participation in case C61. Meanwhile, the boosted mutual benefit including the enrichment of teaching resources, the enhancement of pedagogical quality and the improvement of implementation efficiency through the partnership structure in turn motivated more teachers’
engagement, making the partnership a well-run ecology (Haines et al., 2015; Nehring & O’Brien, 2012). Therefore, the duplicated distributed team partnership structure in project case C61 was regarded as an ideal partnership structure for schools to practice long-term e-Learning collaboratively.

Conclusion

Although some previous researches have tested the relationship between different team-level leadership structures and have discussed the structure-related elements affecting the team performance, seldom of them were in connection with the promotion of sustainable school e-Learning development. This study contributes to the empirical literature in this area by examining the network of leadership and the collaborative relationship of the school partnership structures in an e-Learning pilot project in Hong Kong. This study gained insights from the business-domain perspectives for an innovative analysis of the school partnership of the six project cases in an e-Learning pilot scheme in Hong Kong, in order to inform the employment of school e-Learning partnership which meets the needs and achieves the goals of individual school in long-term e-Learning implementation. This study identified five partnership structures for school e-Learning development, and four critical elements associated with sustainable e-Learning implementation in school partnership.

The duplicated distributed team partnership structure is found favorable to realize all four critical elements for the sustainable school partnership for e-Learning development. The project case C61 in this study, of which had an insightful co-building of online platform for sharing the use of tangible products among the member schools in the cluster, can best illustrate this research conclusion. First, partnership schools share common goals, objectives and expectations in e-Learning collaborate for mutual benefit on technical, resource and pedagogical supports for e-Learning implementation. Next, partnership schools actively engage in dynamic communication and interaction in the cluster to facilitate exchange flow of experience, resources and project outcome of e-Learning implementation within the partnership. Furthermore, school clusters in reasonable team size balance the work force in preparing and trialing e-Learning materials in the collaborative relationship for a significant impact to maintain mutual benefit. Finally, cluster-based online sharing platforms serve as the fourth critical element to synergize the joint efforts of the partnership schools for realizing the first three critical elements through providing a channel for communication and sharing tangible outcomes.

Results from this study reveal that policy support on funding school clusters to co-build online sharing platforms will be beneficial to encourage teachers to continue their give-and-take actions to keep the dynamic ecology and sustainable environment for e-Learning development. It is noted that the generalization of the above research conclusion is limited due to the small sample size and the self-reported data nature in this study. Further research is recommended to look into more project cases on school partnership in e-Learning development for a concrete generalization of the success models and factors for schools to collaboratively and sustainably develop e-Learning initiatives in cluster projects.

Acknowledgements

The author would like to acknowledge the support of the “Provision of Services for Conducting the Research Study on the Pilot Scheme on e-Learning in Schools (Part 2)” project of the Education Bureau of the Government of the Hong Kong Special Administrative Region, PRC.

Reference


Self-Regulated Learning for Web-Enhanced Control Engineering Education

Flavio Manganello1*, Carla Falsetti2 and Tommaso Leo2
1Institute for Educational Technologies, Italian National Research Council, Genoa, Italy // 2Università Politecnica delle Marche, Ancona, Italy // manganello@itd.cnr.it // c.falsetti@univpm.it // tommaso.leo@univpm.it

*Corresponding author

(Submitted October 29, 2018; Revised December 4, 2018; Accepted December 12, 2018)

ABSTRACT

Web enhanced active learning has been demonstrated to be an effective approach in Engineering Higher Education, as it provides students with more flexibility in dealing with the development of skills related to professional knowledge. However, students require a sufficient level of self-efficacy and control over their own learning, which might impact negatively on their effort and academic performance. Therefore, promoting self-regulated learning among students would help them to develop effective strategies they could adopt in planning, monitoring and evaluating their learning process. In this paper, a web-enhanced active learning approach is proposed which integrates a self-regulated learning strategy that supports Control Engineering students in managing their learning process. In order to evaluate the effectiveness of the proposed approach, a three-year quasi-experimental study was performed in the context of an undergraduate Automatic Control course at the Università Politecnica delle Marche, Italy. This involved 418 students and 4 teachers. Both quantitative and qualitative measurement tools were used for the evaluation. The results of the study confirmed the effectiveness of a learning design specifically tailored to implement self-regulated learning features in a web-enhanced active learning approach for undergraduate engineering students. Moreover, the qualitative-quantitative evaluation model proved to be effective in capturing and gauging a more comprehensive and detailed picture of the triggered self-regulated learning dynamics.

Keywords

Engineering education, Material knowledge, Self-regulated learning, Active learning, Web-enhanced learning

Introduction

Over the last few years, researchers and educators in the field of Engineering Education have investigated strategies for actively promoting development of professional knowledge among university students (Llorens, Berbegal-Mirabent, & Llínás-Audet, 2017; Vallim, Farines, & Cury, 2006). It is recommended that engineering students should be trained to master complex physical phenomena, to acquire competence in the analysis of complex systems in several application domains, and acquire the capacity to develop proper design processes (Robinson, Sparrow, Clegg, & Birdi, 2005). The focus of attention should be on the development of professional skills of an expert designer. One of these is Material Knowledge (MK). According to Leo, Pagliarecci, and Spalazzi (2009), MK is knowledge related to real operation conditions in the engineering of complex systems. MK is possessed by experts; it is compiled, implicit, and can be described as intuitive; it is empirical. In a previous study (Leo, Falsetti, Manganello, & Pistoia, 2010a), it was demonstrated that teaching MK effectively demands active, experiential and collaborative learning processes enacted in supervised learning settings; here, educators act as experts providing students with appropriate scaffolding and facilitation. Nevertheless, such an approach requires teachers to devise and implement relevant learning strategies and give their students more autonomy (Coto, Mora, & Alfaro, 2013); it requires students to mobilize their personal desire to acquire knowledge or proficiency in the competence domain (Leo, Manganello, & Chen, 2010b).

Online technology and web-enhanced learning have demonstrated considerable potential for optimizing the learning process in Engineering Higher Education. They also help students to acquire professional knowledge through active learning strategies, as online activities provide students with more flexibility than traditional campus-based courses (Cabrera, Villalon, & Chavez, 2017; Hoic-Bozic, Mornar, & Boticci, 2009; Garousi, 2010; Lundquist, Skoglund, Granström, & Glad, 2013). However, it is not always evident that use of such flexible technology-enhanced active learning settings leads to the development of the necessary skills. The critical issues can be traced back to the functional characteristics of the environments themselves, the way they are configured and used by the teacher, as well as students’ ability to use them. Within such learning settings, students might be challenged by the lack of pedagogical guidance or exposed to information retrieval issues (Dron & Anderson, 2014). This might affect their motivation, effort and academic performance, as students not familiar with such approaches and not sufficiently self-motivated are most likely to experience a decrease in effort and give up. In this light, students using technology-enhanced learning environments should first be
engaged in scaffolded practice and then in fading of the guidance (Azevedo & Hadwin, 2005; Beishuizen & Steffens, 2011). It becomes crucial for students to be supported so as to gain confidence within such learning settings, entailing an adequate level of self-efficacy, as well as of control over their own learning (Hartnett, George, & Dron, 2011). Therefore, it is important to promote self-regulatory processes in learning within computer-based learning settings (Azevedo, 2007). Previous studies have demonstrated the positive effects of computer supported self-regulation on learning performance, specifically with reference to mathematics (Cho & Heron, 2015; Kramarski & Gutman, 2006; Lai & Hwang, 2016) and Technology (Santhanam, Sasidharan, & Webster, 2008; Whipp & Chiarelli, 2004), as well as on learning processes, learning outcomes and self-efficacy in the field of Science (Lai, Hwang, & Tu, 2018; Sha, Looi, Chen, Seow, & Wong, 2012).

With specific reference to Engineering Education, few research studies have evaluated self-regulated learning in the context of online and web-enhanced learning. Notably, a previous study investigating engineering students’ self-regulated learning skills in web-intensive learning environments demonstrated that students may respond differently with respect to self-regulated learning, thus having an impact on their level of engagement (Lawanto, Santos, Lawanto, & Goodridge, 2017). Moreover, few research studies have focused on gauging the impact of designing and embedding self-regulated learning prompts in web-enhanced learning settings, or on evaluating the process. As far as the implementation framework is concerned, self-regulated learning has been considered cumulative with active learning in web-enhanced settings (Debnath, Rahman, & Hossain, 2014), although design in particular has been investigated as part of efforts to implement motivational features in reactive blended learning applied to an introductory Control Engineering course (Mendez & Gonzalez, 2011). Both qualitative and quantitative approaches were considered in terms of evaluation methods for describing and analyzing the practice and development of self-regulated learning specifically within digital learning environments; in the former case, a set of self-regulated learning indicators were adopted for spotting clues of self-regulated events within students’ messages (Debnath et al., 2014); in the latter case, learning analytics were used to identify patterns in students’ decision making about learning with respect to the conditions that might have influenced them (Gašević, Dawson, Rogers, & Gasevic, 2016).

So with the aim of reinforcing the effectiveness of active learning in promoting professional knowledge (in particular MK) among engineering students, and fostering their ability to self-regulate the learning process, in this study a web-enhanced active learning approach was developed. This integrated a self-regulated learning strategy to support students in managing their own learning pace, monitoring their effort investment and their performance, and making decisions about the strategies they adopt when seeking help and information. It was expected that this approach would support students in reflecting on and evaluating their learning performance. Moreover, a specific quali-quantitative evaluation method was implemented for recognizing and gauging the practice and development of self-regulated learning within the approach.

**Literature review**

**Active learning**

Active learning is grounded in constructivist learning (Anthony, 1996) and is tightly linked to learning through discovery (Bruner, 1961) and learning through meaningful reception (Ausubel, 1978). In an active learning setting, students actively engage in a knowledge building process within lessons, which are student-centred rather than teacher-based. Moreover, the process of knowledge creation relies on students’ prior knowledge, where students are self-aware of the process at a cognitive level and can control and regulate it themselves. That is, activation and recognition of prior knowledge are crucial tasks when designing an active learning scenario. A typical learning strategy for fostering students’ active learning is problem solving (Kapur, 2010). Within Science, Technology, Engineering, and Mathematics (STEM), active learning has been shown to be more effective than traditional lectures in promoting both comprehension and memory, and in enhancing student performance (Freeman et al., 2014). As far as Engineering Education is concerned, a study conducted by Prince (2004) defined common, relevant forms of active learning and found that there is broad support for the core elements of active, collaborative, cooperative and problem-based learning.

Technology tools and multimedia adopted within active learning settings have proven effective in engaging and motivating students, as well as in improving their learning processes (Huizenga, Admiraal, Akkerman, & Dam, 2009; Moreno & Mayer, 2000). However, not all students are necessarily comfortable with an active learning approach, especially those with weaker control of their learning process (Margolis & McCabe, 2004). Hence, further investigation is needed to understand the way students self-regulate in such active learning settings, as well as the most appropriate mechanisms or tools for promoting self-regulation strategies.
Self-regulated learning

Self-regulated learning (SRL) is considered “the self-directive process by which learners transform their mental abilities into academic skills” (Zimmerman, 2002, p. 65). SRL occurs in a pro-active way and consists in controlling the learning process not only at the cognitive, metacognitive and behavioural levels, but also in terms of motivation and emotions. According to Zimmerman (2000; 2002), SRL is a recursive process characterized by three different phases, as shown in Figure 1: forethought, performance, and self-reflection. In the forethought phase, students set their learning goals, select the suitable learning strategies, and make an initial assessment of their capacity to reach the goals (self-efficacy) and to activate their prior knowledge. In the performance phase, students monitor their learning process in terms of maintaining attention, use of self-learning strategies, management of time, management and/or configuration of the study environment, and possible search for help (help-seeking). In the third phase, students evaluate their learning process, particularly the achievement of learning outcomes with respect to the goals initially set and the strategies selected.

In recent years, advancements in technologies have impacted individual learning strategies, and technology-based learning environments play an ambivalent role with respect to SRL skills. While using technologies might sometimes be challenging, at the same time they can offer tools and functionalities allowing or even stimulating the development of SRL skills. With this in mind, several studies have investigated the relationship between SRL and technologies, aiming to understand the potential technologies offer for SRL and the conditions that allow optimal exploitation of such potential. Many of the studies have demonstrated the potential support granted to SRL by student-centered web-based learning environments (Azevedo, Behnagh, Duffy, Harley, & Trevors, 2012; Dabbagh & Kitsantas, 2004), by personal learning environment and social media (Dabbagh & Kitsantas, 2012; Manganello, Falsetti, Spalazzi, & Leo, 2013).

According to the literature, active learning has proven to be an effective approach in Engineering Education, promoting student comprehension and memory, and improving their learning performance (Freeman et al., 2014). Problem-based learning and collaboration strategies have proven effective in fostering active learning among engineering students (Prince, 2004). Furthermore, researchers have demonstrated that students need to be self-regulating within active learning settings, in order to better control their learning process (Margolis & McCabe, 2004). With this in mind, in this study a web-enhanced active learning approach was proposed for fostering SRL among engineering students, aiming at supporting them in managing their own learning pace, in monitoring their effort investment and their performance, and in making decisions about the strategies they adopt for seeking help and information. To this end, the following research questions were investigated so as to evaluate the effectiveness of the proposed approach, in the context of Engineering Higher Education:

- **RQ.1a** - Can SRL be effectively promoted online among students?
- **RQ.1b** - Can SRL development online be evaluated and measured?
- **RQ.2** - Can SRL be instructionally defined in the actual learning process?

### A self-regulated web-enhanced active learning approach

In this study, a self-regulated web-enhanced active learning approach for undergraduate engineering students was developed featuring the Moodle e-learning platform. As shown in Figure 2, the system consists of a learning dashboard, a set of SRL monitoring tools, a teaching dashboard and a database. The learning dashboard provides...
the students with the learning contents and the learning assignments uploaded by the teacher(s). From the same dashboard, the students can access the course syllabus and the contextualized learning activities (i.e., Blog/Forum/Wiki). The SRL monitoring tools are prompts explicitly and implicitly embedded in the e-learning platform that provide the students with specific affordances to foster the SRL strategy in terms of the following processes: knowledge activation, time management, effort investment, help seeking, and self-evaluation. By means of the teaching dashboard, the teacher(s) can upload the learning material and the syllabus, as well as provide the students with comments and feedback based on their learning process. Finally, the platform’s database tracks the students’ learning logs and stores their grades in the gradebook.

Figure 2. Flowchart of the self-regulated web-enhanced active learning approach

Figure 3 shows the learning procedure of the self-regulated active learning web-enhanced approach. The students engage with learning activities that – implicitly or explicitly – provide them with SRL prompts. SRL features are integrated in the learning path from the beginning, and are embedded into the online learning environment through learning activities and materials. Four teachers are involved: two teachers for the face-to-face classroom activities, one teacher for the laboratory applications, and one teacher for the online activities.

Figure 3. Learning procedure of the self-regulated web-enhanced active learning approach

At the begin of the course, the teachers introduce the syllabus and the learning model to the students, who are left free to decide whether to follow the course in the traditional way or in web-enhanced mode. Presentation of the syllabus is intended to provide students with all the information about the course and help them in taking more responsible decisions about the organization of their learning process (time management, organization, effort investment, etc.).

The students following the web-enhanced approach are asked to log on to the e-learning platform and access the learning contents. These are organized in units, whose weekly scheduling is presented in the syllabus and is aligned to the topics of the traditional course. These learning materials present the declarative knowledge of the discipline and introduce preliminary aspects of the MK (operative instructions on how to manage professional tools - i.e., Immersive Telelaboratory and/or MATLAB). At the beginning and end of each unit, the students are provided with self-assessment tests. Those at the begin of each unit are designed to help students activate prior
knowledge, while those at the end are intended for self-evaluating the learning process. After that, the students are led to the supervised learning activities, which consist of weekly homework and resource hunting.

The Homework activities are implemented in Moodle by means of the Assignment activity. Each homework assignment involves solving a problem related to the topics presented at the current time, and hence its complexity increases with time. To promote MK, laboratory experiences are as set as homework. Students are proposed Immersive Telelaboratory and/or MATLAB-based learning activities with the aim of fostering ‘learning by doing’ experiences. The students are asked to work in groups and to schedule their activities week by week. Those interested are invited to form small groups of three people and assume specific roles (i.e., role-playing) (Wenger, McDermott, & Snyder, 2010). Each group is free to take decisions about the roles during the activities, but they were asked to implement a weekly turnover so as to share the workload. Each group is provided with a discussion forum, supervised by the teachers. The aim of running such supervised, well-structured, group-based activities is to promote the SRL strategy (especially in terms of time management and effort investment) among the students in a ‘scaffold and fade-away’ fashion.

In addition to Homework, the students engage in a Resource Hunting activity aimed at familiarizing them with the ‘contextual’ dimension of the domain knowledge and eliciting pre-existing knowledge (Eraut, 2000). The students are asked to use specific Moodle features (i.e., blog, forum, and Wiki) to collect meaningful and contextualized Web resources, share them with others, and provide explanations and comments and/or collaboratively write glossary entries. The aim is to explain key concepts related to the course contents, and to link the entries to meaningful and contextualized Web resources featuring explanations and comments. Moreover, the Resource Hunting activity attempts to promote students’ help-seeking and effort investment in a more autonomous and less structured fashion than in the Homework activity.

After completing both the Homework and Resource Hunting activities, students receive comprehensive feedback from the teachers. Participation in these activities, as well as completion of the learning contents and related self-assessment tests, is not mandatory and not subject to summative evaluation. However, learning contents and Homework are graded as a formative assessment activity, with the aim of helping the students fine tune and reflect on their learning process.

Method

To evaluate the effectiveness of the self-regulated web-enhanced active learning approach in terms of the research questions, a three-year quasi-experimental study was carried out. This was performed in the context of three sequential editions of an Automatic Control course, a compulsory unit of the Information Engineering undergraduate program at the Università Politecnica delle Marche (UNIVPM), Italy. In total, 418 students and 4 teachers were involved. The course design was an iterative process. Each edition comprised a suitable mix of face-to-face lectures/numerical case study presentations and online activities, both supervised and not. Several methods were used to measure and evaluate the approach, including ex-ante survey, ex-post survey, tracking of the learning activities, forum discussion analysis, and interviews with the teachers.

Context of the study

At the UNIVPM, Automatic Control is the only undergraduate course of controls and systems throughout the degree program in Information Engineering. The objective of the course is to prepare students to become skilled practitioners in the field of engineering. The discipline is organized around three knowledge building blocks: how to model the objects to be controlled and then those that will be designed for control; once the model is clarified and the formal definition of the desired performances are explained, the next step is the synthesis of the control law for attaining the desired performance of the whole system; finally, how to operate the control in real or realistic cases so as to acquire the MK through experience using an Immersive Telelaboratory (Leo et al., 2010a) or by simulation using MATLAB. With the teacher’s help, the students are expected not only to acquire basic knowledge of the structured aspects of the discipline, but also to face critical issues drawing on minimal information provided by the teacher. For this purpose, it is necessary to clearly define the problem to map the problem data with regards to the available tools, and to use those tools properly. To achieve this, the course needs to comprise specific learning activities based on collaboration and role-playing. Within this educational setting, learning design and formative assessment become crucial. To this end, a web-enhanced active learning approach has been defined, as illustrated in Figure 4.
The implementation model combines traditional classroom methods with an online learning environment; the instructional model integrates direct instruction and collaborative strategies; the team-teaching model reckons on the active participation of multiple teachers, with specific role and tasks, at various stages. The directive approach is mainly used to support the teaching and learning of declarative knowledge through a mix of traditional activities (face-to-face lectures and numerical case-study presentations), access to structured e-learning contents (Battistini et al., 2009), and formative and summative assessment. This approach does not seem to be effective for supporting the acquisition of procedural knowledge in the domain (Spiro & DeSchryver, 2009). Moreover, students to be educated as design engineers of Automatic Control should be engaged in practical activities such as laboratory experiences (blended with theoretical explanations) to attain a more solid grounding in the subject matter (Leo et al., 2010b). Here students are engaged in supervised learning activities in the online environment, where a mix of collaborative and contextualized learning strategies is adopted.
Within such a course, it is necessary to adopt a team-teaching approach to ensure that students have access to an appropriate mix of authentic competencies. The students are in touch with two teachers, one explaining theory via traditional classroom lectures and the other introducing them to the exercise practice (numerical case-studies). Consistency between the two teachers is a key requisite. When laboratory experiences are planned, a third teacher manages the laboratory equipment. In addition, a fourth teacher is involved as an online facilitator to train and assist the students during the online activities. Final examinations are carried out by various examiners, who consider the arguments independently and then work jointly on collegial and specialized evaluation. Coordination among the teachers and the leadership is critical. Behavioral aspects of collaboration and communication are essential (Leo et al., 2010a; Manganello, Falsetti, & Leo, 2009). The learning procedure of the self-regulated web-enhanced active learning approach, along with its study design, is shown in Figure 5.

Participants

Within the three-year period of the study, four teachers were involved: one for the face-to-face lectures, one for the face-to-face numerical case-study presentations, one for the laboratory experiences/MATLAB activities, and one for the online learning activities. A total of 418 students were involved. The number per year was respectively: \( n = 127 \) in the first year, \( n = 138 \) the second year, and \( n = 153 \) the third year.

Measurement tools

Several measurement tools were used to collect data from the three-year period of study and to evaluate the approach. Quantitative tools included ex-ante and ex-post surveys (for each of the three years of the study), and tracking logs of the supervised learning activities (only the Homework activities were considered for this study) on the online platform (for each of the three years of the study); qualitative tools included interaction analysis of student group forum discussions (for each of the three years of the study) and semi-structured interviews with two teachers (one teacher was interviewed at the beginning of the first year study, the other at the end of the third). The evaluation model is presented in Table 1.

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Indicators</th>
<th>Measurement tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRL-Forethought</td>
<td>Self-motivation</td>
<td>EAS (1 item)</td>
</tr>
<tr>
<td></td>
<td>Knowledge activation</td>
<td>FDA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSI</td>
</tr>
<tr>
<td>SRL-Performance</td>
<td>Time management</td>
<td>EPS (1 item)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TSLA</td>
</tr>
<tr>
<td></td>
<td>Help seeking</td>
<td>FDA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSI</td>
</tr>
<tr>
<td></td>
<td>Effort investment</td>
<td>EPS (1 item)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TSLA</td>
</tr>
<tr>
<td>SRL-Self-reflection</td>
<td>Self-evaluation</td>
<td>EPS (1 item)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FDA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSI</td>
</tr>
</tbody>
</table>

Table 1. The evaluation model

Note. EAS = Ex-ante survey, EPS = Ex-post survey, TSLA = Tracking of the supervised learning activities, FDA = Forum discussions analysis, SSI = Semi-structured interviews.

The ex-ante survey was composed of eight items. With respect to the aims of the study, this survey contributed to collecting students’ self-reported quantitative data on the following SRL dimension:

- SRL-Forethought: “self-motivation” (1 item) and “knowledge activation” (1 item), to be evaluated on a four-point scale (1=strongly disagree, 2=disagree, 3=agree, and 4=strongly agree).
The ex-ante survey was composed of 15 items evaluated on the same four-point scale. With respect to the aims of the study, this survey contributed to collecting students’ self-reported quantitative data on the following SRL dimensions:

- SRL-Performance: “time management” (1 item), “help seeking” (3 items), “effort investment” (1 item);
- SRL-Self-reflection: “self-evaluation” (1 item).

The Homework group-based activities were tracked and analyzed, based on the logging patterns of the groups that actively participated in the proposed activities within the e-learning platform, week by week, in terms of “submitted Homework” and “role-play.” By tracking the related activities in Moodle, quantitative data were collected about SRL events regarding time management and effort investment.

Analysis of interactions in the forum discussions was aimed at collecting qualitative data useful for tracking and measuring self-regulated events within the student groups. The online forum discussions were thematically analyzed according to an interaction analysis model specifically tailored for detecting SRL in online communities (Dettori & Persico, 2008). However, the model for coding the text messages was re-designed according to the three SRL dimensions (and their related indicators) considered in the study, as shown in Table 1. All the messages in the student’ group forum discussions for each of the three editions of the course were considered for the analysis. All the messages were collected, and each message was classified by two coders according to 19 coding labels (i.e., nodes) associated to observable SRL-related events. The Nvivo software application was used for data coding and analysis. The coding labels, implemented in Nvivo as “Nodes,” are presented in Table 2 below.

The interviews with the teachers were aimed at collecting qualitative data about the six SRL-related indicators presented in Table 1. One teacher was interviewed at the beginning of the study to gauge expectation, while the other was interviewed at the end of the study to gauge outcomes. The interviews were audio-recorded and transcribed verbatim. The analysis was organized as follows: definition of a taxonomy of elements to be used as coding labels; organization of these coding labels in four macro-categories (teaching context, teacher profile, course implementation, and course evaluation); coding of the interviews by two coders according to the defined coding labels. The labels used for coding SRL-related expressions within the two interviews are the same as those used for the forum discussion analysis, presented below in Table 2.

<table>
<thead>
<tr>
<th>SRL dimensions</th>
<th>SRL indicators</th>
<th>Coding labels (Nodes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forethought</td>
<td>Self-motivation</td>
<td>• Target goal setting</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Goal orientation adoption</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Efficacy judgments</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Task value activation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Interest activation</td>
</tr>
<tr>
<td></td>
<td>Knowledge activation</td>
<td>• Prior content knowledge activation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Metacognitive knowledge activation</td>
</tr>
<tr>
<td>Performance</td>
<td>Time management</td>
<td>• Time planning</td>
</tr>
<tr>
<td></td>
<td>Help seeking</td>
<td>• Awareness and monitoring of need for help</td>
</tr>
<tr>
<td></td>
<td>Effort investment</td>
<td>• Help seeking behavior</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Effort planning</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Increase/decrease effort</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Change or renegotiate task</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Persist, give up</td>
</tr>
<tr>
<td>Self-reflection</td>
<td>Self-evaluation</td>
<td>• Cognitive judgments</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Affective reactions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Attributions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Evaluation of task</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Evaluation of context</td>
</tr>
</tbody>
</table>

Results

The complete dataset compiled over the three-year period of the study comprised:
• n = 179 ex-ante surveys filled out by the students
• n = 121 ex-post surveys filled out by the students
• n = 56 active groups (n = 168 active students) tracked in the supervised learning activities
• n = 1 audio-recorded and transcribed interview with 1 teacher (face-to-face lectures)
• n = 1 audio-recorded and transcribed interview with 1 teacher (numerical case-study presentations)
• n = 220 forum messages collected from the student group discussions forum.

Both the ex-ante survey and the ex-post survey were filled out on a voluntary basis by students who opted for the web-enhanced approach and then enrolled themselves in the online course. On the aggregate for the three years, results were collected from both the ex-ante survey (n = 179) and the ex-post survey (n = 121). The following percentages present the aggregate data for the three years.

The main results of the ex-ante survey analysis are shown in Table 3. With respect to the “SRL-Forethought” dimension, most of the students self-reported to have intrinsic motivation toward the course topics (75.2%); prior knowledge was considered adequate for addressing the activities of the course by just under half of the students (48.8%).

### Table 3. Results of Ex-ante survey by SRL dimensions of three years aggregation

<table>
<thead>
<tr>
<th>SRL-Forethought</th>
<th>S.A.</th>
<th>A.</th>
<th>D.</th>
<th>S.D.</th>
<th>N.A.</th>
</tr>
</thead>
<tbody>
<tr>
<td>I am interested in the topics of this course</td>
<td>57%</td>
<td>18.2%</td>
<td>2.5%</td>
<td>2.5%</td>
<td>19.8%</td>
</tr>
<tr>
<td>I consider my prior knowledge of the subject adequate for successfully carrying out the activities of this course</td>
<td>23.1%</td>
<td>25.6%</td>
<td>18.2%</td>
<td>13.2%</td>
<td>19.8%</td>
</tr>
</tbody>
</table>

**Note.** Total respondents = 179. S.A. = Strongly Agree, A. = Agree, D. = Disagree, S.D. = Strongly Disagree, N.A. = Not Answered.

The main results of the ex-post survey analysis are shown in Table 4. With respect to the “SRL-Performance” dimension, most of the students self-reported that scheduling and managing the online activities of the course with the rest of the curricular activities was not an issue (62.8%). Help-seeking was practiced by most of the students, who self-reported that the Homework and Resource Hunting activities made it easier to decide whether to seek assistance (68.8%) and that they looked for peers and/or the online tutor to discuss issues (68.8%). Most of the students enjoyed the proposed group-based activities because they helped each other (69.4%). In terms of effort investment, only 38.8% of the students deemed role-playing to be an effective strategy for self-monitoring the learning process within the group-based activities. With respect to the ‘SRL-Self-reflection’ dimension, most of the students declared to have exploited the self-assessment tests provided with the learning materials as a benchmark for self-evaluating their progress (62%).

### Table 4. Results of Ex-post survey by SRL dimensions of three years aggregation

<table>
<thead>
<tr>
<th>SRL-Performance</th>
<th>S.A.</th>
<th>A.</th>
<th>D.</th>
<th>S.D.</th>
<th>N.A.</th>
</tr>
</thead>
<tbody>
<tr>
<td>I had no difficulty managing my study programs with the learning activities in this course</td>
<td>40.5%</td>
<td>22.3%</td>
<td>11.6%</td>
<td>5.8%</td>
<td>19.8%</td>
</tr>
<tr>
<td>To find the information I needed to do Homework activities, I used the learning contents and / or asked help from others.</td>
<td>47.9%</td>
<td>20.7%</td>
<td>4.1%</td>
<td>3.3%</td>
<td>24.0%</td>
</tr>
<tr>
<td>I looked for a peer and / or for the online tutor to discuss issues.</td>
<td>37.2%</td>
<td>25.6%</td>
<td>9.1%</td>
<td>0.8%</td>
<td>23.1%</td>
</tr>
<tr>
<td>I enjoyed the group-based activities because we helped each other.</td>
<td>53.7%</td>
<td>15.7%</td>
<td>5.0%</td>
<td>3.3%</td>
<td>22.3%</td>
</tr>
<tr>
<td>My active participation in group-based activities was facilitated by role-playing.</td>
<td>13.2%</td>
<td>25.6%</td>
<td>18.2%</td>
<td>19.8%</td>
<td>23.1%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SRL-Self-reflection</th>
<th>S.A.</th>
<th>A.</th>
<th>D.</th>
<th>S.D.</th>
<th>N.A.</th>
</tr>
</thead>
<tbody>
<tr>
<td>The self-assessment tests and Homework were useful for monitoring my progress</td>
<td>27.3%</td>
<td>34.7%</td>
<td>12.4%</td>
<td>4.1%</td>
<td>21.5%</td>
</tr>
</tbody>
</table>

**Note.** Total respondents: 179. S.A. = Strongly Agree, A. = Agree, D. = Disagree, S.D. = Strongly Disagree, N.A. = Not Answered.

As shown below in Table 5, the students who actively carried out the supervised learning activities were 68.3% (n = 168) of the total number of students who started the activities (n = 246). In the first year this percentage was 30.7% (n = 39), in the second year 45.7% (n = 63), and in the third year 43.1% (n = 66). It should be noted that “active groups” are those that completed at least half of the number of the Homework activities proposed during the current edition of the course. The number of Homework activities assigned was eight during the first year, 10 during the second year, and 12 during the third year.
Table 5. Number of active students and active groups

<table>
<thead>
<tr>
<th></th>
<th>First year</th>
<th>Second year</th>
<th>Third year</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students self-enrolled in the online course</td>
<td>127</td>
<td>138</td>
<td>153</td>
<td>418</td>
</tr>
<tr>
<td>Students who started the activities</td>
<td>69</td>
<td>78</td>
<td>99</td>
<td>246</td>
</tr>
<tr>
<td>Students who completed the activities</td>
<td>39</td>
<td>63</td>
<td>66</td>
<td>168</td>
</tr>
<tr>
<td>Groups that started the activities</td>
<td>23</td>
<td>26</td>
<td>33</td>
<td>82</td>
</tr>
<tr>
<td>Groups that completed the activities</td>
<td>13</td>
<td>21</td>
<td>22</td>
<td>56</td>
</tr>
</tbody>
</table>

As shown below in Figure 6, the year-long response to the Homework activity was similar in each of the three years: in each case, the number of Homework outputs submitted was particularly high in the first half of the year, then decreased significantly in the second half.

The diagrams shown in Figure 7, on the other hand, visualize role-playing as a proportion of the various Homework activities. Except for the first year, there was consistency between the group activities and the management of the same through the proposed roles. The data presented in Table 6 show that students actively participated in the group forum discussions. It should be noted that the highest total number of messages \((n = 97)\) was observed in the second year, whereas the highest number of active groups with respect to the group-based online activities was observed in the third year. Although the total number of messages per year was heterogeneous, the percentage of messages containing SRL indicators and the average number of indicators per SRL-related message were consistent over the three years.

Table 6. Basic data about students’ groups forum discussions

<table>
<thead>
<tr>
<th></th>
<th>First year</th>
<th>Second year</th>
<th>Third year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of messages</td>
<td>42</td>
<td>97</td>
<td>81</td>
</tr>
<tr>
<td>Total number of student messages (N and %)</td>
<td>35 (83.3%)</td>
<td>92 (94.8%)</td>
<td>63 (77.8%)</td>
</tr>
<tr>
<td>Total number of teacher messages (N and %)</td>
<td>7 (16.7%)</td>
<td>5 (5.2%)</td>
<td>18 (22.2%)</td>
</tr>
<tr>
<td>Messages containing SRL indicators (N and %)</td>
<td>16 (38.1%)</td>
<td>35 (36.1%)</td>
<td>29 (35.8%)</td>
</tr>
<tr>
<td>Total number of SRL indicators</td>
<td>21</td>
<td>42</td>
<td>38</td>
</tr>
<tr>
<td>Average number of indicators per SRL-related message</td>
<td>1.3</td>
<td>1.2</td>
<td>1.3</td>
</tr>
<tr>
<td>Kappa value ((k))</td>
<td>0.92</td>
<td>0.91</td>
<td>0.89</td>
</tr>
</tbody>
</table>
Table 7 shows the breakdown for SRL-related expressions detected by the two coders with respect to the six SRL-indicators presented in Table 2. Concurrency on these figures was gained following comparison and discussion between the two coders regarding the selected expressions within the messages (k values for each year are presented in Table 6). More detailed examination reveals that out of the total number of SRL indicators detected by the coders (N = 101) within the messages, 28.7% were related to “help seeking” (n = 29), 18.8% to “self-evaluation” (n = 19), 16.8% to “time management” (n = 17), 14.9% to “self-motivation” (n = 15), 10.9% to “knowledge activation” (n = 11), and 9.9% to “effort investment” (n = 10).

<table>
<thead>
<tr>
<th>SRL Indicators</th>
<th>First Year</th>
<th>Second Year</th>
<th>Third Year</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-motivation</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>Knowledge activation</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>Time management</td>
<td>3</td>
<td>8</td>
<td>6</td>
<td>17</td>
</tr>
<tr>
<td>Help seeking</td>
<td>6</td>
<td>11</td>
<td>12</td>
<td>29</td>
</tr>
<tr>
<td>Effort investment</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>Self-evaluation</td>
<td>4</td>
<td>7</td>
<td>8</td>
<td>19</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>42</td>
<td>38</td>
<td>101</td>
</tr>
</tbody>
</table>

Table 8 shows basic data from the semi-structured interviews. Once again, concordance was gained following comparison and discussion between the two coders (k values for each interview are presented in Table 8). Table 9 shows the breakdown of SRL-related expressions detected by the two coders with respect to the six SRL-indicators presented in Table 2. Once again, concordance was gained following comparison and discussion between the two coders (k values for each interview are presented in Table 8). More detailed examination reveals that out of the total number of SRL indicators detected by the coders (N = 17) within the interviews, 24% was related to “help seeking” (n = 4), 18% to “time management” (n = 3), 18% to “effort investment” (n = 3), 18% to “self-evaluation” (n = 3), 12% to “self-motivation” (n = 2), and 12% to “knowledge activation” (n = 2).

<table>
<thead>
<tr>
<th>Interview 1</th>
<th>Interview 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expressions containing SRL indicators</td>
<td>5</td>
</tr>
<tr>
<td>Total number of SRL indicators</td>
<td>8</td>
</tr>
<tr>
<td>Average number of indicators per SRL-related expression</td>
<td>1.6</td>
</tr>
<tr>
<td>Kappa value (k)</td>
<td>0.89</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SRL Indicators</th>
<th>Interview 1</th>
<th>Interview 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-motivation</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Knowledge activation</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Time management</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Help seeking</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Effort investment</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Self-evaluation</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>8</td>
<td>9</td>
<td>17</td>
</tr>
</tbody>
</table>

Discussion and conclusions

Recently, the implementation of web-enhanced active learning in the context of Engineering Education has been gaining increasing attention from scholars. This approach has proven to be particularly effective in providing engineering students with suitable flexibility and contextualization for acquiring professional knowledge. Nonetheless, several studies have recommended that appropriate pedagogical guidance should be provided within such flexible learning settings. With the aim of promoting active learning and SRL in a supervised manner, in this study, a web-enhanced active learning approach was developed and tested. This implemented an SRL-based strategy to support students in managing their own learning pace and in monitoring their effort investment and performance, as well as offering support for decision making about the strategies they adopt to seek help and information. In order to evaluate the proposed approach, a three-year quasi-experimental study was employed in the context of an undergraduate Automatic Control course. In the following, the research questions presented at the beginning of the paper are discussed. It should be noted that RQ.1a and RQ.1b are discussed together for convenience.
SRL - Promotion online (RQ.1a) and measurement (RQ.1b). The results of this study show that the proposed approach significantly benefited the students in terms of promoting and developing SRL strategies. Analysis of the ex-ante survey and the ex-post survey reveals that significant effects were achieved regarding several aspects of SRL, namely self-motivation, time-management, help seeking and self-evaluation. However, it should be noted that no significant effects were generated regarding prior knowledge activation and effort investment. These results were confirmed by analysis of the student group forum discussions, where results align perfectly with those from the surveys: self-motivation, time-management, help seeking, and self-evaluation were the SRL indicators most evident within the forum discussion messages. Analysis of the interview results show that the SRL factors most significantly effected (in terms of teachers’ expectations and observations) were, once again, self-motivation, time-management, help seeking, and self-evaluation; however, the teachers also considered that prior knowledge activation and effort investment were also effected. Analysis of the supervised learning activities (which was, as already stated, limited to the Homework activities) clearly showed that the number of the students participating in the Homework activities decreased over time in the same course: as curricular activities comprise various courses running in parallel, effort investment may have been affected. Moreover, the interested and motivated students seemed to have regulated the workload for accomplishing a personal objective - i.e., completing at least half of the Homework assigned to obtain a benefit in terms of formative assessment, while optimizing their resources in terms of time management and effort investment. The Homework activity probably helped the students in monitoring their time use and changing/renegotiating tasks. Therefore, we conclude that engineering students can improve their awareness of self-regulation when learning under the web-enhanced active learning approach, with respect to three areas for regulation:
- cognition, regarding complex problem solving (the capacity to deal with complex problems, finding connections and relations among concepts) and activation of reflective learning processes;
- motivation, regarding self-motivation (towards the subject); and
- behavior, regarding autonomy (ability to perform both individual and collaborative work), time management, and effort investment (in the course activity).

As far as evaluation and measurement of SRL development online, the quali-quantitative analysis adopted for the evaluation was in some ways innovative compared to what is reported in the literature concerning Control Engineering Education. The effectiveness of this approach can be measured by means of the narrative that the qualitative data (collected through thematic analysis of student group forum discussions and teacher interviews) offered in terms of keys for the interpretation of quantitative data (collected through self-reported student surveys and tracking of online activities). The data thus obtained can be read in an interrelated way to obtain - at a higher and more comprehensive level of detail - information on the triggered SRL dynamics, which then become observable and usable for a more authentic evaluation by the different stakeholders involved in the learning process.

SRL – Instructional representation (RQ.2). With respect to the implementation framework, the approach confirmed the effectiveness of the specific design strategy to implement SRL features in a web-enhanced active learning setting. It was important to deploy the course and, more in general, the learning activities within a learning environment capable of supporting flexible course design. Judging by the level of student satisfaction in the online environment, Moodle appeared to offer sufficient flexibility for the purpose. Moreover, in line with what is suggested by the reviewed literature, the online component was fundamental for providing students with a “walled garden” where they could nurture their own learning strategies, or familiarize with the new ones proposed, by means of implicit and explicit SRL prompts. The online environment proved to be effective in supporting the team of teachers in the implementation of multiple learning activities, each related to the appropriate area of the subject. In particular, the approach was effective in boosting some aspects of students’ professional training: the declarative and procedural aspects of the knowledge domain could be dealt with by means of appropriately designed learning materials and self-assessment, while the contextual aspects of the knowledge were handled by means of supervised learning activities, some of which were formal tasks (i.e., Homework), while others were non-formal tasks (i.e., Resource Hunting). It should be noted that the proposed semi-structured tasks attracted less student participation compared to the more formal and structured tasks. This is a point to ponder more thoroughly so as to determine whether the incentive to carry them out should be strengthened or not. Finally, the encouraging initial results from evaluation of the approach might have broader implications for Engineering Higher Education in terms of seeking flexible learning models for integrating SRL into actual learning process within a formal curriculum.

There is one significant limitation to this study that should be noted, namely that related to student learning and measurement of the actual evidence of student achievement. As student achievement in mastering the requisite knowledge and skills should be considered the most relevant criterion in the learning process, students’ ability to self-regulate their learning process should be considered important as this is instrumental in achieving such
mastery. Discussion of this point, however, falls outside the scope of this work. This could be done, for example, by tracking more data from the actual student learning activities, such as the results of the self-assessment tests, and analyzing them with respect to meaningful SRL factors. In order to make the study findings generalizable to other engineering courses, it is suggested that future studies explore such scenario by adopting the evaluation method proposed in this study.

Acknowledgements

Professor Tommaso Leo, passed away on June 16, 2014, aged 70. He received the M.S. degree in electronic engineering from the Università La Sapienza, Rome, Italy, in 1968, where he was in touch with the prestigious school of Automatic Control held by prof. Antonio Ruberti. Since 1981, he was a Full Professor of Automatic Control with the Faculty of Engineering, Università Politecnica delle Marche, Ancona, Italy. With the same University, he was Director of the Department of Electronic and Automation from 1982 to 1988, Dean of Engineering Faculty from 1990 to 1996, Director of the Department of Information Engineering (formerly Department of Informatics, Management and Automation Engineering) from 2005 to 2011. From 2003 to 2013, he was rector’s delegate for e-learning. He was chair of research programs and projects in robotics, automation, human movement analysis and e-learning. In the same research fields, Prof. Leo was author and co-author of more than 250 scientific papers, editor of scientific books and of special issues of International Scientific Journals. He was IEEE senior member. The life of Professor Leo was dedicated to science, education and engineering. His attention to teaching and to the students’ needs was a key feature that has distinguished his teaching activity and that has seen him very active in technology enhanced learning. Prof. Leo had made significant intellectual contributions to the study presented in this manuscript and had been involved in the initial stages of its preparation.

Reference


Why and How Serious Games can Become Far More Effective: Accommodating Productive Learning Experiences, Learner Motivation and the Monitoring of Learning Gains

Wim Westera
Open University of the Netherlands, Valkenburgerweg, The Netherlands // wim.westera@ou.nl

(Submitted October 18, 2016; Revised May 29, 2017; Accepted July 26, 2017)

ABSTRACT
This paper aims to improve the design methods for serious games (games for learning) by identifying a set of well-established pedagogical misconceptions and presenting design guidelines to avoid these. It analyses the pedagogical principles and models that are commonly used in serious game design, and contrasts these with evidence and advances in instructional psychology and instructional design research. The paper particularly focuses on (1) the concept of experience-based learning, which many serious games comply with, (2) the concept of learner motivation, which most games strongly claim to support, and (3) the score systems that many games use to track and display progress. Structural design weaknesses are exposed and countered with a large body of research evidence from the literature. A set of practicable design guidelines are presented that help to avoid the pedagogical flaws and contribute to improving the design methods for serious games.

Keywords
Serious games, Applied games, Game design, Learning effectiveness

Introduction
Games for learning are persistently gaining popularity. Many scholars and practitioners have recognised and embraced serious games because of their great potential for learning by predominantly referring to the dynamic, responsive and visualised nature of games, which produces high motivations, strong user involvement and penetrating learning experiences (Westera, 2015). Game development studios are a branch of so-called “creative industries,” a term that emerged in the 1990s to connect the arts and other cultural activities with emerging digital technologies and the associated knowledge economy (O’Connor, 2010). The expected impact of the creative industries on society is substantial as the creative industry product is innovative rather than routine, and is generally characterised by originality, technical professional skill, uniqueness and quality (Caves, 2000). Creative industry product design goes with large degrees of freedom, while routine recipes are not available and in many cases not desirable. This certainly holds for the leisure game industry, but to some extent also for serious games. The appreciation of serious games is largely determined by their gaming properties, that is, the quality of game play, the fidelity of the environment and the fun they offer to the end-users. Serious games that lack the manifest features of commercial video games are easily disqualified by learners (Van Rosmalen & Westera, 2014). Less prominent seems the criterion of learning effectiveness. However, the effectiveness of games for learning and their supposed superiority over other teaching approaches are not self-evident. Various reviews studies (Connolly, Boyle, MacArthur, Hainey & Boyle, 2013; Linehan, Kirman, Lawson & Chan, 2011) show that among many thousands of serious game studies only very few try to evaluate the educational effectiveness of serious games in a rigorous manner by a sound quantitative or qualitative analysis of educational outcomes. Just asking end-users what they think about the game they played is less convincing a method for collecting evidence of effective learning than arranging a randomised controlled trial with calibrated pre-tests and post-tests. On the positive side, quite an increasing body of evidence is becoming available supporting the effectiveness of serious games for learning (Boyle et al., 2016; Connolly et al., 2013; Sitzmann, 2011).

This paper explores the main pedagogical concepts and principles that are relevant for serious game design, in particular those games that rely on an experience-based approach, and it critically evaluates these against available research evidence. It argues that serious game design may be substantially improved by taking advantage of the large body of empirical findings in half a century of teaching and learning research. In particular it identifies potential weaknesses in the ways serious game design deals with instructional guidance, motivation and the assessment of learning progress, respectively, and it formulates recommendations to address these for improving the effectiveness of serious games. First, a brief analysis is presented of the principle conflict between play and learning that arises when games are used for serious purposes.
The presumed conflict between learning and play

Abt (1970) introduced the term “serious game” as to contrast the purposeful application of games with leisure games. In practice, the term serious games is used as a container allowing for a wide variety of styles and approaches, including e.g., quizzes, interactive stories, sandbox virtual worlds and simulations. However, the term “serious game” is an oxymoron: a figure of speech combining two concepts that are contradictory (e.g., “old news”). In his seminal book “Homo Ludens” Huizinga (1938) describes game play exclusively as a leisure activity: the quintessence of play is in its voluntariness and openness, the freedom it offers and having no other purpose than playing. Play takes place in what Huizinga defines as the magic circle: a playground, which is a temporary world detached from real life, and which allows for fantasy and pretending, while special rules apply. However, these principles are readily undermined when a game is re-positioned as a serious educational endeavour with its obligations, regulations, imperative classes, homework and examinations. Still, play is not without effect. Whether pursued or not, playing a game involves experiences that influence the individuals’ mental states at affective and cognitive levels. Play is generally considered a natural way of human learning (Blanchard & Cheska, 1985; Huizinga, 1938). Exactly because of their engaging and absorbing capacities games are increasingly being used in schools as they are supposed to support and enrich the curricula. Whether or not these serious games are capable of inducing play in its purest sense remains a philosophical question. A question of more relevance would be to what extent games are capable of supporting learners to master new knowledge and skills in effective, efficient and pleasant ways. A growing body of evidence supports a positive answer to this question (Connolly et al., 2013). Apparently fun and seriousness need not necessarily be in conflict, as fun doesn’t necessarily mean “easy.” The best fun is “hard fun” (Papert, 1980): people like to be challenged by difficult tasks, which require seriousness and dedication. Above all they are eager to see how they can stretch their own abilities. It is well recognised that games are capable of hooking and absorbing their players in such a way that they can hardly stop playing (Aldrich, 2005; Dickey, 2005; Gee, 2003). In extreme cases excessive gaming is known to display compulsive and addictive properties very similar to drugs and alcohol and producing similar damaging effects in individuals’ mental health (Griffiths, Király, Pontes & Demetrovics, 2015). Since motivation of school children and students is generally known to be low, how nice it would be if a touch of the compulsive properties of games could be exploited to induce wholesome levels of engagement and the ideality of learners eager for extra challenges, extra homework, extra lessons and even staying after school hours (Westera, 2015).

Learning from experience as the basis of game-based learning

Learning from experience is the dominant pedagogical paradigm in serious game design (Gosen & Washbush, 2004; Canhoto & Murphy, 2016; Connolly et al., 2013; Aldrich, 2005; Gala, 2014; Reese, 2011; Schank, Berman & Macpherson, 1999; Catalano, Luccini & Mortara, 2014; Arnab, Lim, Carvalho, Bellotti, De Freitas, Louchart, Suttie, Berta & De Gloria, 2015; Rooney, 2012). It refers to learning by active exploration and self-direction rather than learning from instruction. Many related terms are used to indicate comparable approaches, such as discovery learning (Bruner, 1961), problem-based learning (Barrows & Tamblyn, 1980), inquiry learning (Papert, 1980), experiential learning (Kolb, 1984), constructivism (Jonassen, 1991), situated learning (Lave & Wenger, 1991) and learning by doing (Schank, 1995; Aldrich, 2005), respectively. Well before the wave of constructivism Dewey (1916) stated that learning should be connected with some real world context in order to allow the learner to relate symbolic content (e.g., concepts and principles) to real-world referents. The same holds for its recent descendents. Such model of learning from experiences that result directly from one’s own actions is often contrasted with the information transfer model, which describes learning from e.g., reading instructions or listening to lecturers. The knowledge transfer model and the experience-based learning model essentially deal with distinct types of knowledge that is learned: explicit knowledge versus implicit knowledge, respectively. Explicit knowledge refers to all knowledge that can be codified, that is, written down in texts and formulas, and that thereby allows for being transferred directly from one person to another. Alternative terms are declarative knowledge, formal knowledge and articulate knowledge. In contrast, implicit knowledge is the knowledge that is hidden in the action. One may readily learn explicit knowledge - say about bicycles - from a textbook, e.g., about the different parts and components, but riding a bicycle requires experiencing the essential subtleties of timing, balance, power and sudden disturbances, which just cannot be explained in words, but can only be learned by doing. Such implicit knowledge is also referred to as procedural knowledge or tacit knowledge (Polyani, 1966). The difference between explicit and implicit knowledge is also referred to as “knowing that” and “knowing how” (Ryle, 1949).

As serious games are readily positioned as the attractive alternative of traditional teaching practice many games tend to avoid the explicit knowledge transfer model and promote experience-based learning to achieve learning
goals (Schank et al., 1999). Various authors consider serious games as pedagogically-driven games while referring to such experience-oriented models as a theoretical foundation (Catalana et al., 2014; Canhoto & Murphy, 2016; Rooney, 2012). A recent survey among game developers revealed the same preference, while guidance and instruction were reported to be seldom used during the game (Saveski et al., 2015). Many games rely on offering the “experience” and stay aloof of learner guidance (Canhoto & Murphy, 2016).

**Issues with experience-based learning**

**Rote learning rather than deep understanding**

As experience-based learning focuses on contextualised actions, many serious games in e.g., math, geography or language teaching are based on drill and practice approaches, which may support the successful reproduction of knowledge and automation of operational skills (rote learning), but fail to support deep level insights and generalised understandings that are based on conceptual background facts, principles and theories. Although the engaging experiences of serious games are valuable as such, a lack of explicit framing may impede deep understanding of the patterns and relationships of the content at hand (Mayer, 2004). Occasionally, these limitations are recognised and removed by adding a debriefing, which include reflective evaluation of the events in the game (Crookall, 1995; Saveski et al., 2015; Garris, Ahlers & Driskell, 2002). This is in accordance with Dewey’s claim that learning is the product of experience and reflection (Dewey, 1938). Still it means that instructional events are either neglected or positioned outside the game rather than inside the game, possibly to allow the game to be shielded from any reference to traditional teaching practice (Malone, 1981; Gee, 2003; Aldrich, 2005). The strive for fun, however, may have taken its toll when the game just produces routines without the required deep understanding: a missed opportunity that can be easily restored.

**Neglecting the ineffectiveness of minimal guidance approaches**

Many serious games offer an experience-based learning environment focused on exploration, discovery or problem solving that is devoid of instructional guidance (e.g., Canhoto & Murphy, 2016). However, evidence for the effectiveness of such minimal guidance approaches is almost non-existent: the far majority of research studies report that learner guidance approaches produce superior results as compared to approaches without or with minimal guidance (Mayer, 2004; Kirschner, Sweller & Clark, 2006). Kirschner et al. (2006) attribute the failure of minimal guidance approaches to the properties of human cognitive architecture: when learners are novel to certain learning content and thereby do not have detailed mental schemas that reflect the conceptual structure of the content domain, they are prompted to explore large search spaces that generate heavy loads on working memory, which is detrimental to learning. Worst case, the learner will adopt a thoughtless trial and error strategy, which eventually may lead to task completion be it without any learning gains. Moreover, research shows that exactly less able students suffer mostly from minimal guidance approaches.tragically, they still favour those approaches although they learn less from those (Clark, 1989; Kyllonen & Lajoie, 2003; Kirschner et al., 2006). Mayer (2004) notes that in spite of the overwhelming evidence against minimal guidance approaches, the tendency to reduce or even eliminate learner guidance can be seen in many subsequent teaching alternatives that aim to abandon traditional teacher-lead instruction, e.g., discovery learning (Bruner, 1961), problem-based learning (Barrows & Tamblyn, 1980), inquiry learning (Papert,1980), experiential learning (Kolb, 1984), constructivism (Jonassen, 1991) and learning by doing (Schank, 1995; Aldrich, 2005), respectively: every decade the same set of minimal guidance principles are covered by a new label. Each generation of minimal guidance advocates seem to be unaware of the refutations of the previous approaches, and naively starts repeating the same mantra (Mayer, 2004). Serious gaming may readily become the next wave of counterproductive pedagogical innovation.

Altogether, it seems that in many experience-based serious games the pedagogical foundation is shallow, ill-articulated, erroneous, if not tendentious. It may be attributed to a lack of understanding of associated pedagogical labels (e.g., constructivism, experiential learning) or to the neglect of a vast body of empirical evidence in learning research that has been made available over the past 50 years. The neglect of scaffolding and other forms of instructional guidance seems to have become a goal in its own right as to avoid any contamination with traditional teaching models (Kirschner et al., 2006). The mantra of experience-based learning may be easily abused to stay away from traditional instruction and to focus on the gameplay rather than on the learning. Yet, the rich multimedia properties of digital games would allow for both challenging gameplay and dedicated instruction. Serious games could thus become far more effective.
Issues with motivation

Determinants of motivation

Games are valued for their motivational power. Players of a game are challenged to actively engage in e.g., problem solving, exploration, goal formation, critical analysis, strategic thinking and enhanced creativity (Westera, 2015). Motivation is a complex psychological construct that is related to energy, intention, direction, and persistence, which all bring individuals into action (Ryan & Deci, 2000). It is widely considered as a main determinant of effective learning (Keller, 1987; Keller, 2008; Ryan & Deci, 2000; Cordova & Lepper, 1996). Research into motivation has identified various favourable outcomes for learning, such as increased attention, enjoyment, engagement, depth of involvement, task persistence and cognitive flow (Cordova & Lepper, 1996; Garris et al., 2002; Csikszentmihalyi, 1991). A variety of theoretical models of motivation have been proposed (e.g., Keller, 2008; Malone, 1981; Malone & Lepper, 1987; Ryan & Deci, 2000), all recognising the distinction between intrinsic motivation and extrinsic motivation. Intrinsic motivation is the inherent tendency to seek out novelty and challenges, to extend and exercise one's capacities, to explore, and to learn (Ryan & Deci, 2000). It is the personal drive to engage in an activity because of the activity itself. Extrinsic motivation in contrast refers to performing an activity driven by external factors, e.g., by external pressure or pursued outcomes such as rewards, prestige, diplomas or salary (Westera, 2015). Both intrinsic and extrinsic motivation are important determinants of learner behaviour. However, extrinsic motives are known to be often less productive than intrinsic motives: extrinsic motives are readily associated with shallow learning rather than deep processing (Habgood & Ainsworth, 2011). A large number of studies have demonstrated that intrinsic motivation as compared with extrinsic motivation leads to more interest, excitement, and confidence, which in turn contributes to enhanced performance, persistence, creativity, vitality, self-esteem, and general well-being (Ryan & Deci, 2000).

In their cognitive evaluation theory and self-determination theory Ryan and Deci (2000) identify three determinants of intrinsic motivation: competence, autonomy and relatedness. Intrinsic motivation has been demonstrated to increase when people feel “they can do it” (competence), provided that they “did it by themselves” (autonomy). The third factor “relatedness” is less prominent: it refers to the support from the social environment in terms of security and confirmation. The absence of relatedness, however, e.g., imposed tasks, directions, surveillance or deadlines, is known to diminish intrinsic motivation and to lead to poor performance or failure to complete. In games a number of motivation drivers can be distinguished: presentation style, gameplay scenario and reward systems. They should be used properly.

Style elements

The rich soundscapes and dynamic visual sceneries in games unescapably trigger the senses and help to effect penetrating user experiences (Dickey, 2005). The associated motivation effects remain largely extrinsic in nature: players are hooked, if not seduced, by the appealing, dynamic sceneries as is the case in sales promotion and commercial advertising. As a side-effect, the rich, dynamic user interfaces may readily undermine favourable conditions for learning, because of unwanted distractions and excessive cognitive load. Likewise, the strive for realism in games may be counterproductive for learning. Although realism to some degree is needed to promote the credibility of practical skills training assignments (Westera, Nadolski, Hummel & Wopereis, 2008), reducing the representation of the synthesised world to the essential functional properties would favourably reduce the cognitive load for learners. Experiments based on media equation theory (Reeves & Nass, 1996) have demonstrated that human individuals respond socially and naturally to a variety of non-human objects such as robots, computers, games, avatars, or symbols. By the same mechanism simple, non-realistic game representations (e.g., the ghost characters in Pacman) may still raise exciting, dramatic, if not thrilling experiences. The human brain is simply not capable of suppressing natural interpersonal responses, as it has no neural circuits or anatomic regions that distinguish between inter-human and mediated or symbolic interactions (Reeves & Nash, 1996). Games displaying rich, realistic or dynamic visualisation styles may stimulate extrinsic motivation, at the expense of unfavourable distractions and cognitive overloads. A justified balance is needed.

Gameplay scenarios

In contrast with the manifest game sceneries and features the underlying gameplay scenarios are the true carriers of intrinsic motivation as the players are active participants in their stories, having adopted goals, roles, responsibilities (autonomy) and powers (competence) that put them at the centre of the action: By engaging in
the gameplay scenario the player gets involved in a sense making process. This refers to the search for appropriate relations of causality, space and time, which allows for the construction of the game’s story (narrative or fabula) (Bordwell, 1989). This ongoing mental process of creating meaningful narratives from events and experiences is the main carrier of experience-based learning. As Gee (2003) explained, the secret of a videogame is “not the high quality, immersive 3-D graphics but in the underlying architecture, which balances the challenges offered to the player with the players’ abilities seeking at every point to be hard enough to be just doable”. If the game is fun, well-balanced and content-wise interesting, the players are likely to play the game because of the activity itself: intrinsic motivation. Scenarios that support the player’s competence and autonomy would make serious games far more effective.

**Reward systems**

Motivation in games is also driven by reward systems, which may include scores, permissions, property, reputation and more. Three different reward classes can be distinguished:

**Post-practice rewards**

Post-practice rewards are obtained after successful game completion. These rewards enhance extrinsic motivation driven by a future promise related to the outcomes of the game rather than the game activity itself, for instance, a diploma or a certificate, enhanced reputation e.g., by a favourable appearance on a public leader board ranking, a monetary prize, or privileges such as access to a new level: playing the game is no more than a means to an end (Vallerand, Fortier & Guay, 1997).

**Dynamic in-game scores**

Dynamic in-game scores indicate progress of achievement, usually expressed as a number, a percentage, a set of stars, or similar. Such in-game scores are frequently updated, which re-establishes extrinsic motivation again and again. Most in-game scores are not directly fed back into the gameplay, but they are just used as monitors of achievements. It means that the rewards are oriented on results rather than the activities themselves.

**Affordance-oriented rewards**

Affordance-oriented rewards provide incentives or privileges that directly support and enhance gameplay. Rewards resulting from successful actions open up new powers, resources and opportunities for intensifying gameplay: the rewards are part of the narrative that the player is constructing, for instance, gaining a high profit in a management game would allow for more investment options or take-overs. This comes close to receiving natural feedback: being confronted with the direct or indirect consequences of decisions made during gameplay. Because their focus is on enhanced activity, affordance-oriented rewards are likely to support intrinsic motivation.

Most rewards systems, however, function as external drivers and thereby they trigger extrinsic motivation, which may be effective as such at activating people, be it not without limitations. A lot of games, for instance in math, spelling or vocabulary learning, offer a dressed up drill-and-practice, which uses the behaviourist notions of repetition and reinforcement (and punishment) to condition learners to routine tasks. As opposed to contemporary constructivist approaches these behaviourist approaches focus on automation, while deep thinking and reasoning are discouraged. Such operant conditioning is generally considered a truncated, low-level mode of learning most suited for teaching tricks to animals. Games thus run the risk of being reduced to skinner boxes, which are “incentive dispensers that dole out rewards for attention” (Bogost, 2007). By using the right reward systems serious games could become far more effective.

**Extrinsic causes for intrinsic motivation**

External reward systems are prone to undermine intrinsic motivation because they are easily perceived as external controllers of behaviours, associated with force, pressure or surveillance, which all affect the player’s sense of autonomy. Nevertheless, Ryan and Deci (2000) have recognised the practical importance of extrinsic
motivation - our behaviours are for a large part externally driven -, and based on their self-determination theory they suggest that external drivers may amplify intrinsic motivation provided that there is sufficient room for perceived competence, autonomy and relatedness. For instance, children doing their homework may be externally motivated by the desire to satisfy their parents’ expectations (social relatedness), to receive a diploma (post-practice reward) or to satisfy requirements for a future profession. Despite this external pressure they can be highly committed provided that the requested behaviours have been fully integrated and internalised so that it is perceived and accepted as a personal striving. This can be promoted by allowing for self-selected, not imposed, goals, which assumes the opportunity for choice and freedom of movement (autonomy). When in addition self-confidence and self-esteem are established and re-established (competence), personal goals will develop and motivation may become internalised. In games similar mechanisms apply. If players are allowed to pursue self-selected goals and if they feel competent, that is, they feel in control rather than being controlled, the activity and the goals get internalised and integrated in the self.

In sum, serious games have the inherent capacity to address both intrinsic and extrinsic motivation. Motivation should preferably be intrinsic, that is, associated with the activity, the gameplay and the game’s content rather than the outcomes. Style elements may contribute to attractive gameplay, but they should not be distracting. Reward systems in games tend to be output-based but should instead be affordance oriented. Choice, freedom of movement, control and competence are crucial factors to support the intrinsic sense of achievement and self-fulfilment in the game. The best reward for an achievement is in the achievement itself. Many serious games violate these principles.

**Issues with progress monitoring and assessment**

The many traces that players leave during gameplay offer great opportunities for the detailed monitoring and assessment of players’ behaviours, progress and achievements. Therefore game scores are not just included so often for motivation purposes only (see above), but also for informing the players how well they are performing. While in leisure games the score systems are self-establishing, viz., they are defined as part of the rules of play, in serious games the scores may be expected to reflect the progress toward the games’ learning goals. Unfortunately, the score systems in serious games are likely to conform to gameplay standards rather than educational standards and to focus on events rather than the underlying skills, knowledge or competence frameworks. Consequently, game score systems seldom comply with the strict requirements of validity (whether or not the assessment measures what it claims to measure), reliability (whether the assessment produces consistent, reproducible outcomes) and fairness (whether the assessment is free from bias, e.g., racial, religious, gender, age) of educational assessments. A number of assessment issues are addressed below.

**Assessment as a by-product**

Assessment has been identified as the most powerful determinant of learning behaviours (Hattie, 2009; Brown & Glasner, 2003). Metaphorically this was expressed as “the assessment tail that wags the curriculum dog” (Hargreaves, 1989). It means that the assessment should not be considered as a final add-on to the game design, but instead should be the design’s starting point. Serious game design, however, is likely to start at the wrong end by designing a (nice) game around some relevant content, and then adding some score mechanisms. Studies in evidence-centred design (Mislevy, Steinberg & Almond, 2003) have suggested that educational design should start with a detailed task analysis, identifying the right activities and composing favourable situations that would allow players to exert and demonstrate certain well-specified behaviours. In serious games, the assessments are likely a by-product, which is detrimental for their quality. Serious game design should start with an assessment analysis.

**Complexity of covert, dynamic assessment**

As one of the promises of learning analytics the continual stream of player interaction data can be used to extract player’s progress indicators in a covert way. Such stealth assessment occurs on the fly without any specific testing assignments or other interruptions of the game (Shute, 2011; Shute & Ventura, 2013). Moreover, stealth assessment is based on sequences of highly dependent interactions in the game, which represent a large evidence base, while traditional testing approaches would only offer single, independent data points. However, as opposed to common game score systems, the implementation of stealth assessment is complex, laborious and time-consuming, leaving aside the tendency to view game scores just as a motivational add-on. In addition, game
designers, even experienced ones, may not have gained the level of assessment expertise that educational professionals, e.g., teachers, have. Nevertheless, stealth assessment could make serious games far more effective.

**Performance versus learning progress**

Most game scores are indicators of the player’s performances, but they do not necessarily indicate learning progress. Learning and performance are conflicting concepts that often require opposite attitudes. Various authors (VandeWalle, Brown, Cron & Slocum, 1999; Fisher & Ford, 1998) explain the difference between a performance orientation and a learning task orientation. Performance is linked with achieving milestones (in many cases under time constraints), the swift completion of tasks, avoiding errors, and the use of proven methods for reducing risks. Aiming at high performance scores draws players toward activities that they are good at already: they are reluctant to try our new approaches because they are anxious for running into penalty points. Effective learning, however, requires spending sufficient time for in-depth understanding, having sufficient opportunities for reflection, revision, and self-evaluation, and being prepared to make mistakes (Westera, 2015). Errors and failure are productive sources of learning (Mory, 2003; Mathan & Koedinger, 2005), be it that the score mechanisms used should not discourage these by imposed penalties. In practice, game score systems often enforce the achievement of performance goals, which stimulates learners to demonstrate high ability and to avoid poor performance. In such contexts failure becomes a threat to success and thereby it affects self-esteem, self-confidence, and motivation. The resulting self-defence reactions (Mory, 2003) include discounting (Kelley, 1973), task avoidance, feigning boredom, and task-irrelevant actions to bolster self-image (Dweck & Legget, 1988), and learned helplessness (Seligman, Maier & Geer, 1968). To support a learning orientation serious games should lower the price of failure (Gee, 2003). Game score systems should not primarily be based on time-constraints and penalty points, but should focus on learning progress, while allowing players to make mistakes, to spend sufficient time and effort, to try and retry, to reflect on attainments and to decide upon their own strategies (autonomy). By promoting a learning orientation and applying learning-oriented score systems serious games could become far more effective.

**Conclusion**

Research in instructional sciences has produced a huge body of evidence over the past 50 years. In serious game design, however, many established insights in teaching and learning seem to be neglected. This neglect may be partially explained by the ambition to contrast serious games with existing teaching approaches, by emphasising their fun properties and excluding manifest references to school and classroom practices. Also, the high visual, acoustic and narrative qualities of entertainment games inevitably frame the expectations that learners and teachers may have and thereby urge serious game designers to focus on the gaming part at the expense of the teaching part.

Still, the ambition to make learning more attractive and joyful should never go at the expense of learning effectiveness. Although many serious game designers would concur with the paramount importance of pedagogy, they are readily in default when it comes to activity, either by misconceptions or ignorance. This paper has identified a variety of weaknesses in serious game design, which needs further attention from the field. Main conclusion is that serious games can become more effective if the considerations about experience-based learning, motivation and assessment, respectively, would be taken into account. Based on our analysis we present the following guidelines.

With regard to experience-based learning, serious game design should:

- explicitly base its design on advances and evidence in learning sciences research
- include instructions and explanations about conceptual background principles and theories, even if when these would require the interruption of play
- include adaptive scaffolding mechanisms and other forms of instructional guidance and learner support in the design.
- provide feedback on learning rather than feedback on performance
- include opportunities for reflection and metacognitive activity, even if when these would require the interruption of play.

With regard to motivation serious game design should:

- avoid distracting presentation style elements because of the unwanted extraneous cognitive load they induce
• start with an assessment analysis that identifies relevant tasks and content to be covered by the game, and base the gameplay scenarios on the assessment analysis
• offer gameplay scenarios that allow for sufficient freedom of movement and player responsibilities to enhance the player’s sense of autonomy
• offer gameplay scenarios that pose cautiously balanced and just doable challenges for enhancing the player’s sense of competence
• avoid external pressure, surveillance or time locks as to preserve the player’s sense of relatedness.

With regard to scoring and assessment serious game design should:
• include affordance-oriented incentives that directly support and enhance gameplay rather than output-oriented rewards
• use dynamic in-game scores that provide a reliable monitor of learning progress
• benefit from stealth assessment based on evidence-centred design or similar, which allows for monitoring learning progress
• promote a learning attitude, which allows for failure, retrials, reflection and well-considered strategy development without time constraints or penalty scores.

For practical reasons this paper only looked into potential issues in instructional support, motivation and the assessment of learning progress, respectively. Still, various other aspects of serious games pedagogy deserve similar analysis, for instance the transfer of the knowledge learned in games to a variety of different operational contexts, the role of reflection, self-regulation and other metacognitive capabilities in games, the design of scaffolding, feedback mechanisms, and personalisation, the role of affective and emotional elements for learning, and the integration of serious games in the wider curriculum context. On all of these topics a vast body of research evidence is available, which only sparsely permeated the domain of serious game design. By putting more efforts in their pedagogical dimensions, serious games can become far more effective.

Acknowledgements

This research was partially funded by the European Union’s Horizon 2020 research and innovation programme under grant agreement No 644187, the RAGE project (www.rageproject.eu).

Reference


Do Learning Styles Matter? Motivating Learners in an Augmented Geopark

Tien-Chi Huang, Mu-Yen Chen* and Wen-Pao Hsu

Department of Information Management, National Taichung University of Science and Technology, Taiwan
// tchuang@nutc.edu.tw // mychen.academy@gmail.com // andrew780110@gmail.com

*Corresponding author

(Submitted December 5, 2016; Revised September 13, 2017 Accepted October 10, 2017)

ABSTRACT

Augmented reality (AR) technology has recently been applied to outdoor learning in an attempt to overcome the drawbacks associated with traditional teaching environments. This study conducted an experiment designed to examine how augmented reality (AR) technology in mobile devices can be used to generate virtual objects to create a context-aware, AR-enabled guided tour application for outdoor learning. The participants were 70 elementary school students (average age: 11 years), who were randomly divided into control and experiment groups. The results showed the proposed system provided learners with a friendly, interactive interface and rich, engaging media to improve learning performance and stimulate the students’ internal motivation to learn. The system’s quantification of the learning motivations noted in Keller’s ARCS model and Kolb’s learning style theory can be used to improve the design of the learning materials. In conclusion: (1) The proposed system and activity helps stimulate learning intention via the pursuit of outdoor learning objectives, (2) the AR technology provides learners with contextual information related to the outdoor learning environment, and (3) the benefits of the proposed model do not differ for students with different learning styles.

Keywords

Augmented reality, Ubiquitous learning, ARCS, Motivation, Learning styles

Introduction

Authentic learning activities have been proven to be an effective method by which to improve students’ problem-solving skills and learning achievements (Chen & Lin, 2016; Herrington & Oliver, 2000; Hwang et al., 2010). Fieldwork in geology—a form of authentic learning—has become one of the main activities to focus on the natural environment in today’s school education systems (de Barros et al., 2012). Among school subjects, geology places the most emphasis on observation and authentic learning, in which rich geological natural resources can stimulate students’ intention to learn. However, the traditional one-guiding-many method of teaching geology outdoors usually fails to arouse the individual’s motivation to learn, and does not provide content that facilitates individual exploration by the student.

Recent studies have stated that integrating innovative digital technology with the learning of geology in an outdoor setting can enable students to explore actual geology in depth (Di Serio et al., 2013; Moutinho et al., 2015). Augmented reality (AR) technology, which was developed over the last few decades and has become popular in educational technology research, is able to overcome the key weakness of authentic learning: the lack of adaptive systematic material (Huang et al., 2016a). Given the ability to create a ubiquitous learning environment, the approach can be an efficient approach to increase student intention, letting education to arise in authentic contexts outside the classroom.

Disseminating digital content via educational technology makes it possible to arouse students’ motivation to learn. The ARCS model of motivational design emphasizes the four key elements of motivation: Attention, Relevance, Confidence and Satisfaction (Keller, 1987a; Keller, 1987b). The model elaborates the content of motivation, which allows us to improve educational projects and teaching quality (Keller 1999). Besides learning motivation, learning style is another issue of concern in the learning process (Kolb, 1976). Although digital learning can support adaptive learning, the extent to which individual differences influence the learning outcome and even the motivation remains a question (Terrell, 2002). According to Kolb (1976), individual differences in information perception and information processing form four types of learning styles: diverging, assimilating, converging and accommodating. Different learning styles may bring about different outcomes for e-learning activities, particularly in terms of e-learning motivation.

In an authentic Geopark, an AR system might be able to improve learning effectiveness and motivation for learners who have different learning styles. Thus, this study develops an AR-based ubiquitous learning system and applies it in Taiwan’s National Geopark. In other words, this study attempts to realize systematic,
personalized learning in an authentic learning environment using AR technology. In view of the above research purpose, the following research questions are addressed in this study:

- Does the AR u-Geopark learning system improve learning achievement?
- Do students with different learning styles achieve the same level of learning when using the AR u-Geopark learning system?
- Did the AR u-Geopark learning system augmented by AR motivate students?

**Literature review**

**AR for ubiquitous learning**

Ubiquitous learning (u-learning) is “anywhere and anytime learning.” U-learning systems are designed and implemented followed by situated learning theory, which states that effective learning is best achieved by settling questions and obtaining knowledge from authentic experiences (Wang et al., 2015). The u-learning process has four advantages over traditional teaching methods. First, it allows access to information and sharable content via a mobile device (Hwang et al., 2011). Second, allows learning to happen at any location, indoors or outdoors (Klopfer et al., 2002). Third, the learning content is presented in authentic contexts (Liu et al., 2009). Fourth, the approach allows for the accumulation of a portfolio showing the individual’s learning progress (Glaserfeld, 1990). Situated learning theory also states that learning motivation comes from the stimuli within the authentic environment. However, once students’ learning motivation is aroused, an instructor faces the disadvantage of the lack of systematic information in the authentic environment.

To overcome the problem, augmented reality (AR) technology, which can combine virtual objects with the real world on a mobile device, is usually adopted in authentic learning environments. While the student is immersed in the real environment, virtual objects are produced via mobile devices or cameras and computers. Students can learn the study subject by using virtual objects on the equipment (Okita, 2014; Huang et al., 2016a; Huang et al., 2016b). Since mobile devices are context-aware in the real environment, learners can use them to understand the information available in the real world, and form a better understanding of the surrounding environment (Yang 2015). Therefore, the application of AR in this study of u-learning is summarized as follows:

- **Integration of the real-world and AR learning**: The use of AR technology allows learning materials to be more diverse and rich. With an integrated information system, students can relate the learning materials directly to the real world (Tsai et al., 2011), and the context-aware environment helps facilitate their understanding of the subject (Gutiérrez et al., 2010).

- **Development of an adaptive learning strategy with AR technology**: When students encounter a learning system with a complicated interface, their good intentions (deeper learning) are often dampened by the complexity of operating the system. Creating learning objectives based on the student’s current location allows the learner to use the system with ease, increasing learning effectiveness (Chen & Lin, 2016).

- **Creation of an innovative u-learning application**: Improving on the traditional learning model by using mature AR and multimedia streaming technologies, this innovative learning model allows students to extend their education via self-learning in context (Zhu & Jin, 2012). Students do not need to memorize texts; their study subjects are generated by a combination of the real world and virtual objects, allowing learners to deepen their prior knowledge (Ng, 2012).

**AR and ARCS motivation model**

In 1983, Keller proposed the ARCS model of motivational design which helps us to understand learners’ motivations and how to stimulate their desire for knowledge. The model is divided into four categories, namely Attention (A), Relevance (R), Confidence (C), and Satisfaction (S) (Keller, 1987a; Keller, 1987b). The first step is to understand what catches students’ attention in this motivation process, and to explore those triggers in depth. The subject matter and curricula are then tailored to relate closely to the individual student’s motivations. As they complete these learning tasks, their confidence builds. The learning process has to be both educational and interesting in order for students to concentrate on the subject materials (Gutiérrez et al., 2010). Students’ motivation is assessed via a questionnaire.

Recently, Hassan et al. (2012) adopted the ARCS model to investigate learning performance in distance learning. Their results indicate that AR or context-aware learning objectives increase students’ willingness to learn continuously. The adoption of AR technology motivates students by drawing their attention with augmented information, providing relevant content according to their current location, building learning confidence, and
satisfying them with rich information. Therefore, compared to traditional learning models, the AR assisted model proved more effective in enhancing both the learning and the willingness to learn (Cheng & Su, 2012; Di Serio et al., 2013).

**Learning style and individual differences**

Learning style is one of the main concepts involving individual differences in learning. In a u-learning environment, multimedia is adopted in the learning process, and differences in preference regarding information processing may influence the students’ achievement.

Kolb proposed the theory of learning style (Kolb, 1976) which was classified into the two dimensions of data grasping and data transformation, and these dimensions encompass four independent learning stages. Data grasping includes concrete experience (feeling) and reflective observation (watching). Data transformation includes abstract conceptualization (thinking), and active experimentation (doing). These four learning stages are also correlated to four learning types. Throughout the four stages, each of the four different types of learners choose the learning model which they find most favorable (see Table 1).

<table>
<thead>
<tr>
<th>Learning style</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diverging</td>
<td>(Concrete, Reflective)</td>
</tr>
<tr>
<td></td>
<td>• Emphasizes innovative and imaginative approaches to doing</td>
</tr>
<tr>
<td></td>
<td>• Interested in people and tends to be feeling-oriented</td>
</tr>
<tr>
<td></td>
<td>• Richly imaginative; will increase focus when the subject is found to be interesting</td>
</tr>
<tr>
<td>Assimilating</td>
<td>(Reflective, Abstract)</td>
</tr>
<tr>
<td></td>
<td>• Good at independent inductive reasoning and prefers theories and logic</td>
</tr>
<tr>
<td></td>
<td>• Cares more about thoughts and abstract theoretical concepts</td>
</tr>
<tr>
<td>Converging</td>
<td>(Abstract, Active)</td>
</tr>
<tr>
<td></td>
<td>• Emphasizes problem solving and the practical application of ideas</td>
</tr>
<tr>
<td></td>
<td>• Performs best when solving problems that have a single ideal answer</td>
</tr>
<tr>
<td>Accommodating</td>
<td>(Active, Concrete)</td>
</tr>
<tr>
<td></td>
<td>• Good at hands-on learning, learning by doing</td>
</tr>
<tr>
<td></td>
<td>• Comfortable with new learning models and will not eschew new learning methods</td>
</tr>
</tbody>
</table>

The effectiveness of a given learning method depends on the student’s learning style. When the student is guided in a way that matches his or her learning style, learning is less burdensome. For example, diverging and accommodating learners find it difficult to sit alone in their seats to learn. A gaming scenario stimulates these types of students’ motivation to learn, because they understand the message and gain knowledge via doing (Girvan et al., 2013). In contrast, learning in a solitary environment may help some converging and assimilating learners. These types of learners can easily focus in a classroom. When doing field studies outside of the classroom, they may feel the need for a separate space for reading in order to digest the content (Konak et al., 2014).

To investigate the relationship between e-learning and learning performance, Konak et al. (2014) adopted Kolb’s Experiential Learning Cycle as a framework to design hands-on activities in virtual computer laboratories. They found that such activities enhance student learning outcomes. In this study, Kolb’s learning style is adopted to investigate individual differences in u-learning motivation and achievement.

**System design and architecture**

This research integrates a context-aware ubiquitous learning environment and AR techniques, using content based on the Yehliu Geological Park. The proposed u-learning system incorporates radio frequency identification (RFID) and augmented reality (AR) technologies in mobile devices, allowing the learner to immediately access supplementary information regarding the geological objects observed in the park. This section describes the hardware and software requirements of the proposed system.

In this research, context-aware technology was used to allow users to follow learning activities along a predetermined route, using the augmented reality features to engage in image memory-based learning to increase comprehension and retention. Figure 1 shows the system architecture.
Teachers can use a classroom version of the AR system to introduce students to the system and its content in a classroom setting. In traditional classrooms, students learn about nature through text and pictures only, making it difficult for them to truly understand the essence of the natural sciences. However, using the proposed system, the teacher in the classroom can lead students through the supplemental teaching material at various difficulty levels. The teaching materials module allows teachers to select an appropriate level of subject difficulty. The teaching materials are classified according to level of difficulty. It is considered an aggregate of teaching units, including a topic name and its geology content, and the degree of difficulty in association with other units. The teaching units are also designed by teachers and are based on the concepts of geology knowledge. Teachers also constructed the tests that assess the student’s mastery of each teaching unit. In the system development phase, these tests were executed and obtained the degree of difficulty parameters. When learners finish Teaching Unit 1 of the instruction, they subsequently experience their first assessment using their tablets. The system then calculates their test scores. If a student fails to pass, the system will recommend other material regarding the same concept at the appropriate degree of difficulty. Using wireless technology, the system will transmit the suitable materials directly to the student’s tablet. On the other hand, if the student passes the exam in Teaching Unit 1, the student can continue the enrichment activities with additional related topics. On the other hand, the content management module allows teachers to easily swap different activities in or out of the course. Students can also use ancillary and supplemental teaching materials in the AR environment, using geographical location information to increase their learning efficiency.

From the learner’s perspective, the system logs all student activity, automatically adjusts the scope of subjects, selects an appropriate learning difficulty level, and updates parameters according to the student’s search results. While navigating the learning path, students can engage with one topic after another. In the traditional classroom, learners often have difficulty keeping pace with a teacher’s presentation because of a lack of understanding. Using the proposed u-learning system, students can repeatedly engage with the instructional material at their own pace, thus developing a familiarity of the topic.

**Augmented reality interface and operations**

On the system entry screen, students see four options. They can choose the option that best suits their needs, as shown in Figure 2. The upper left option is (1) “Personal Information,” (2) the upper right is “Learning Portfolio,” (3) the lower left is “Augmented Reality,” and (4) the lower right is the “Natural Encyclopedia.” “Learning Portfolio” is selected to show whether the goal of individual learning has been completed.

Selecting “Augmented Reality” and the corresponding subject 1 “Queen’s Head Rock” allows the user to select D’Fusion to display the AR model corresponding to the labels, as shown in Figure 3(a). Clicking on the side buttons reveals popup information, as shown in Figure 3(b).

If students are not clear about the observation requirements of the learning objective during the process, they can select “Natural Encyclopedia.” The popups offer more information regarding key features of the object being observed.
Research design and evaluation

A three-phase quasi-experiment was performed on students in the 4th \((n = 25)\), 5th \((n = 23)\) and 6th \((n = 22)\) grades \((N = 70; \text{male } = 39, \text{female } = 31)\), using the proposed AR u-Geopark system. The participants were from a public elementary school in northern Taiwan. All students had 1 to 2 years of experience using tablets at school or at home. Figure 4 shows the research chronology.

The 70 students were randomly divided these into experimental and control groups. Both groups had students from each of the three grade levels. First, 70 students were required to take pre-test to test their basic knowledge of these patterns and to assess their learning style. This study adopted the “Kolb learning style inventory” (KLSI,
V. 3.1) to determine students’ learning style of the experimental group. The Cronbach’s alpha of the four dimensions are .82, .73, .83 and .78 (Kolb, 1985). After “Kolb learning style inventory” was fulfilled, there were 13, 2, 5, and 15 learners belong to the diverging, assimilating, converging, and accommodating learning styles.

In Phase 1, teachers adopted traditional teaching approach in the classroom to carry on geology concepts, referring to some of Yeliu’s specific geological features: “Queen’s Head Rock,” “Candle Shaped Rock,” and “Bean Curd Rock.” This research was conducted in Yehliu Geopark which is comprised of many different kinds of landscapes and rocks.

<table>
<thead>
<tr>
<th>Learning materials</th>
<th>Experimental Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td><img src="image1" alt="Introduction" /></td>
<td><img src="image2" alt="Introduction" /></td>
</tr>
<tr>
<td>Queen’s Head Rock</td>
<td><img src="image3" alt="Queen's Head Rock" /></td>
<td><img src="image4" alt="Queen's Head Rock" /></td>
</tr>
</tbody>
</table>

*Figure 5. Learning materials for the two groups*

In Phase 2, the two groups of student proceeded with the experiment separately, and two experienced teachers instructed the two groups regarding what to observe and where to go. The students in the experimental group used the system to observe the objects. When they reached the observation site, they could observe the authentic objects and use the device with RFID reader to capture augmented learning information via sensing RFID tags that were built on the site. The learning information was accompanied by AR objects. The control group used the traditional commentator-guided method to observe the same objects. The learning information for the control group was presented on paper-based materials (e.g., brochures, information sheets). The two kinds of learning materials are shown in Figure 5. This activity spent 60 minutes for the two groups. For the control group, they tried to discover the objective geological subjects adopting textbook and a guide map. In the meantime, the experimental group engaged the learning actives by adopting the u-Geopark learning system. The learning subjects were “Queen’s Head Rock,” “Candle Shaped Rock,” and “Bean Curd Rock” and the experimental infrastructure was the Yehliu Geological Park, as shown in Figure 6.

In Phase 3, both the experimental and control groups come back to the classroom and fulfilled a post-test. In addition, the experimental group filled in the five-point-Likert-type ARCS questionnaire to assess how effectively the system increased learning motivation and performance (Appendix A) (10 minutes). In general, the purpose of the post-test was to access students’ comprehension of the basic geology of Yehliu Geopark and their knowledge of “Queen’s Head Rock,” “Candle Shaped Rock,” and “Bean Curd Rock.” And the ARCS questionnaire of this study was designed based on Keller (2010), Hassan et al. (2012), Cheng & Su (2012), and Di Serio et al. (2013). We compared these two groups in terms of learning motivation and learning effectiveness, and analyzed the learning styles of different learners to see how such information could help increase learning effectiveness.
Results

The AR u-Geopark learning system developed in this study facilitates student learning at the Yehliu Geopark. After the learning experiment was completed, the learning styles and motivation level of the students were analyzed to explore the impact of different learning styles on learning effectiveness and the willingness to learn. The pre-test and post-test scores of the samples are shown in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>Numbers</th>
<th>Pre-test</th>
<th>Post-test</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experimental Group</strong></td>
<td>35</td>
<td>54.64</td>
<td>74.86</td>
<td>-7.092*</td>
</tr>
<tr>
<td><strong>Control Group</strong></td>
<td>35</td>
<td>55.71</td>
<td>68.00</td>
<td>-3.34**</td>
</tr>
<tr>
<td>t</td>
<td>0.25</td>
<td>-2.707</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. **p < .01.

The difference in the pre-test scores of the two groups is not significant, while the t test between these two groups in the posttest (t = -2.707, p < .01) were significant. Meanwhile, the difference between pre-test and post-test in the experimental group is significant (t = -7.092, p < .005), so is the control group (t = -3.34, p < .005). That means the devised learning activities are exactly helpful to stimulate students’ learning achievements.

A pair-sample t-test was used to analyze the significance of the pre-test and post-test scores for each of the four learning style groups as shown in Table 2. The results show that the differences in pre- and post-test scores for diverging and accommodating learners are both statistically significant, and the p-value for each is below 0.05. The results of the Kruskal-Wallis test for pre-achievement and post-achievement test showed no significant difference in four learning style groups (p > .05), respectively.

According to past literature (Terrell & Dringus, 2000), we further divided four group (learning style) into two sets: information-perceiving and information-processing. The first set is concrete (Diverging and Accommodation) vs. abstract approach (Assimilating and Converging). The second set is reflective (Diverging and Assimilating) vs. active approach (Accommodation and Converging). The independent sample t-test results show that there are no differences in both sets in Table 3.
Table 2. Difference tests for achievement testing for the four Learning Styles

<table>
<thead>
<tr>
<th>Kolb’s Learning Style</th>
<th>Paired variation</th>
<th>Pair-sample</th>
<th>Independent-sample Ach. Kruskal-Wallis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>SEM</td>
</tr>
<tr>
<td>Diverging</td>
<td>-20.62</td>
<td>16.81</td>
<td>4.66</td>
</tr>
<tr>
<td>Assimilating</td>
<td>-40.00</td>
<td>28.28</td>
<td>20.00</td>
</tr>
<tr>
<td>Converging</td>
<td>-14.20</td>
<td>19.87</td>
<td>8.89</td>
</tr>
</tbody>
</table>

Note. $^*$p < .05.

Table 3. t-test of achievement performance for two participant sets

<table>
<thead>
<tr>
<th>Kolb’s Learning Style</th>
<th>Mean</th>
<th>SD</th>
<th>t</th>
<th>df</th>
<th>Sig (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>First set: Concrete (Div &amp; Acc)</td>
<td>54.02</td>
<td>17.4</td>
<td>-.46</td>
<td>33</td>
<td>.66</td>
</tr>
<tr>
<td>Pre-test: Abstract (Ass &amp; Con)</td>
<td>57.14</td>
<td>15.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First set: Concrete (Div &amp; Acc)</td>
<td>75.36</td>
<td>11.4</td>
<td>.70</td>
<td>33</td>
<td>.50</td>
</tr>
<tr>
<td>Post-test: Abstract (Ass &amp; Con)</td>
<td>72.86</td>
<td>7.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second set: Reflective (Div &amp; Ass)</td>
<td>53.3</td>
<td>18.0</td>
<td>-.39</td>
<td>33</td>
<td>.70</td>
</tr>
<tr>
<td>Pre-test: Active (Acc &amp; Con)</td>
<td>55.6</td>
<td>16.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second set: Reflective (Div &amp; Ass)</td>
<td>74.0</td>
<td>11.8</td>
<td>-.40</td>
<td>33</td>
<td>.70</td>
</tr>
<tr>
<td>Post-test: Active (Acc &amp; Con)</td>
<td>75.5</td>
<td>10.0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In Table 4, the outcome of ARCS scale presents that students’ motivation towards the “augmented Geopark learning activity” is high, the average mean value is between 4.21 and 4.39. Then, the students’ motivations were analyzed separately according to their learning styles. The degree of satisfaction was crucial to their motivation in three learning style groups: the diverging, the assimilating, and the accommodating groups. The converging group scored highest in attention ($m = 4.54$) which is the second high among the four learning styles in average score.

Table 4. Mean grades of four perspectives for the ARCS questionnaire

<table>
<thead>
<tr>
<th>Learning style</th>
<th>ARCS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Attention</td>
</tr>
<tr>
<td>All</td>
<td>4.34</td>
</tr>
<tr>
<td>Diverging</td>
<td>4.21</td>
</tr>
<tr>
<td>Assimilating</td>
<td>4.79</td>
</tr>
<tr>
<td>Converging</td>
<td>4.54</td>
</tr>
<tr>
<td>Accommodating</td>
<td>4.32</td>
</tr>
</tbody>
</table>

Discussion

Learning achievement of AR u-Geopark learning system

In this study, the control group adopted a relatively traditional strategy like reading instruction paper, while the experimental group learned with the ubiquitous AR system. An achievement test is a test of developed skill or knowledge. The most common type of achievement test is a standardized test developed to measure skills and knowledge learned in a given grade level, usually through planned instruction, such as training or classroom instruction (Koretz, 2002). In advance, achievement test scores are often used in an educational system to determine what level of instruction for which a student is prepared (Chen & Huang, 2012; Chen & Lin, 2016).

In this research, we used the pre-test and post-test for both groups, the average testing score is improving from 54.64 to 74.86 of the experimental group, in the meantime, the average testing score is improving from 55.71 to 68.00 of the control group. The descriptive statistics shows there is a significant increase for the testing score between pre-test and post-test. In advance, the difference between pre-test and post-test in the experimental group is significant ($t = -7.092, p < .005$), so is the control group ($t = -3.34, p < .005$). That means the devised learning activities are exactly helpful to stimulate students’ learning achievements.
Learning performance of different learning styles

In Table 2, the results show that the “AR u-Geopark system” improved students’ learning achievement, especially for those with diverging and accommodating learning styles, both of which are concrete-oriented style. The outcome shows that the real environment activity is more attractive for students who prefer concrete experiences, and the system provides dynamic, interactive information for their education. The differences of pre-test and post-test scores among the four groups did not reach the level of significance. In other words, when immersed in the same process in an AR-assisted ubiquitous learning environment, learning achievement levels do not differ among the different learning styles.

However, the result is different from the related literature. Past studies have examined learners’ achievement while using hypermedia and online learning systems according to their learning style (Di Serio et al., 2013; Terrell & Dringus, 2000; Terrell, 2002; Wei et al., 2015). They stated that in a digital learning environment (non-face-to-face), the “abstract” approach learners performed better than the “concrete” approach learners. In this study, the “concrete” approach learners (diverging and accommodating) improved significantly after the u-learning experience. This indicates that while all the learners are exploring in an authentic environment, the concepts can be concretize by presenting images of different angles on the mobile interface (Chen & Chen, 2017). Therefore, we suggest that AR may provide more benefits to learners with a concrete approach than to those who learn with an abstract approach. Besides, we conclude that the AR possesses great different features comparing to other types of e-learning technology, it helps learner to connect the concrete environment and abstract concept. Comparing to the environment, we think to assess students’ characteristic and give them the most proper e-learning assistance are the key issues of this study.

Motivation in AR U-learning environment

In Table 4, all four groups have a high average level (lowest mean = 4.21) of motivation in all four aspects (A, R, C, and S). The degree of learning satisfaction is critical to their motivation while learning with the AR system in a ubiquitous environment, and this result is in line with Huang et al. (2015).

ARCS model is a process of elements which are able to stimulate learners’ motivation. The converging learners enjoy problem solving in which observation with attention is the primary step. For the converging learners, this explains why attention is the key incentive of learning with AR system. Otherwise, the diverging, assimilating, and accommodating learners have higher levels of satisfaction regarding the activity, compared to the other three elements. Interestingly, although learners possess different learning preferences (diverging learners like creative materials, assimilating learners enjoy logical thinking, and accommodating learners prefer real experiences), they all self-report high levels of satisfaction in regards to the AR system. In this study, we found out that the 3D presentation in the system satisfied the diverging learners’ innovative need, the detailed explanation about the Geopark sites in the system fulfilled what assimilating learners want, and the activity was conducted in an authentic environment where accommodating learners could explore. The AR system and the ubiquitous learning environment aroused the motivation to learn, even among students whose learning style orientations are different. In other words, the diverse needs of the students were satisfied by the characteristics embedded in the AR which assisted the learning activities.

As a form of digital learning, a key feature of the AR system differs significantly from other kinds of digital learning: the connection with a situated, authentic environment. To be more specific, the present learning environment is a ubiquitous environment with AR technology installed. The AR system augmented the authentic environment, allowing the co-existence of real interactions and abstract information, which simultaneously satisfy students with different learning styles, enhancing their learning achievement. Future studies could further focus on exploring the relationship between learning style and u-learning achievement in different subjects.

Conclusion

This research integrates a context-aware ubiquitous learning environment and AR techniques, using content based on the Yehliu Geological Park. The proposed AR u-Geopark learning system allows the learner to immediately access supplementary information regarding the geological objects observed in the park. After the learning experiment was completed, the learning styles and motives of learners were analyzed to explore the impact of different learning styles on learning effectiveness and the willingness to learn. The results show that the proposed system can improve learning performance and stimulate students’ internal motivation to learn.
In this research, there are several limitations and can be considered improving in the future works. First, sampling was carried out in the Yehliu Elementary School of New Taipei City, an area in north Taiwan nearby the Yehliu National GeoPark. Future researches can be enlarged to involve more samples or other grade schools to differ from our research sample. Second, we recommended to apply the proposed u-learning approach to other outdoor courses and evaluate the learning performance of the teaching activities. Third, we divided the research sample into experimental group and control group. Besides, we used the pair-sample t-test, Kruskal-Wallis test, and independent sample t-test to assess the learning outcomes. Therefore, more statistical analysis or structural equation modeling can be considered. Finally, we used the ARCS model to understand the learning motivation under the u-learning environment. However, other motivation models or theories, such as achievement motivation theory, attribution theory, and self-efficacy theory can be considered adopting to extend the explanatory capacity of this research.

Acknowledgements

The authors wish to thank the Ministry of Science and Technology of the Republic of China for financially supporting this research under Contracts No. MOST 106-2511-S-025-003-MY3, MOST 105-2410-H-025-015-MY2, and MOST 104-2511-S-025-002-MY3.

Reference


Appendix A

ARCS Questionnaire

A
1. What interests me is the learning program using the 3D model system.
2. The content of the learning program of the 3D model system is very appealing to me.
3. The content of the learning program of the 3D model system is vivid and it makes me more focused in class.
4. The interpretation of the 3D model system helps me stay focused for longer periods of time.
5. Personally operating the digital teaching aids stimulates my curiosity.
6. Using the 3D model system to learn the course knowledge surprises me.
7. The use of 3D model system helps me become more focused on learning.

R
1. The goal is consistent with what I learn in class; I believe I can understand the learning content.
2. The comprehensive teaching materials help build my confidence.
3. I felt fulfilled and satisfied when I completed this learning program.
4. The diversity of learning activities makes me feel more interested.
5. The diagrams and examples in the 3D model system help me understand more about the learning content.
6. The content is worth understanding when studying the subject using the teaching material from the 3D model system.
7. The goal is consistent with what I learn in class; I can learn in my daily ordinary life.
8. The content of this course will be useful to me.

C
1. When I first saw the materials, they made an impression on me.
2. Good teaching materials help me build confidence.
3. After using the system, I have more confidence in my learning ability.
4. After using the teaching materials of the 3D model system, I am more confident that I will pass the exam.
5. With the diversified 3D model system for learning, I am more confident that I can learn comprehensively.

S
1. Upon completion of this program, I felt fulfilled and satisfied.
2. I am pleased to participate in this program, and I would like to continue to use the system for after-school learning.
3. I really liked learning this subject.
4. I was satisfied and had a sense of accomplishment when I completed each subject.
5. The curriculum of this 3D model system is vivid, and I found it pretty novel.
Investigating Remote Access Laboratories for Increasing Pre-service Teachers’ STEM Capabilities

Ting Wu1* and Peter Albion2

1Southwest Jiaotong University, China // 2University of Southern Queensland, Australia // tingwoo@126.com // peter.albion@usq.edu.au

(Submitted November 21, 2016; Revised October 21, 2017; Accepted January 18, 2018)

ABSTRACT

Facing calls for greater emphasis on STEM education in primary school classrooms, teachers may be anxious because of limited exposure to STEM in their own education. The Australian Curriculum: Technologies is new and many teachers are not familiar with its content. Hence both in-service and pre-service teachers (PSTs) require preparation. This research used a case study method to investigate factors influencing PSTs’ use of Remote Access Laboratories (RAL) with activities intended to develop their capacity to teach STEM in primary schools. Results highlighted the importance of PSTs’ experience of STEM in their own education and showed the benefits of hands-on learning and scaffolding to support preparation of PSTs for teaching STEM subjects.

Keywords
Remote access laboratories, STEM education, Pre-service teachers

Introduction: STEM education

Achieving a productive and progressive future for Australia will require a workforce with high levels of scientific and digital literacy developed through studies of STEM (Science, Technology, Engineering and Mathematics) subjects (Office of the Chief Scientist, 2013). However, the Australian workforce has shortages of STEM graduates, including STEM teachers, caused by a decline in STEM study at the tertiary level, which flows on from a high dropout rate from STEM courses in high school (Freeman, 2013). In turn, the decline in STEM interest in high school has been attributed to inadequate time spent on STEM teaching and learning in primary schools (Office of the Chief Scientist, 2013).

STEM as a merged area of study, and especially the technology component, is new in primary schools but, based on the “scientific” nature of its content, likely to be similar to science which has been studied extensively. Primary teachers’ reasons for lack of attention to science include limited exposure to science in their own education (Westerlund, Radcliffe, Smith, Lemke, & West, 2011), limited access to relevant teaching resources, and low confidence in their ability to teach science and technology effectively (Ping et al., 2011). Hence, it is important to support teacher professional learning to increase teachers’ confidence to teach science and provide teaching resources for teaching science. Similar arguments can be advanced for technology and STEM.

Australian curriculum: Technologies

The Australian Curriculum: Technologies (ACARA, 2015) was developed to ensure all students benefit from learning about technologies that shape the world they live in (Falkner & Vivian, 2015). It consists of two distinct but related subjects: Design and Technologies and Digital Technologies (ACARA, 2015). It provides opportunities for students from Foundation (F) to Year 10 to develop their design thinking, computational thinking and related skills.

For most in-service and pre-service teachers, many of the elements in the Australian Curriculum: Technologies were not part of their own schooling or teacher preparation. They will be unsure about the relevant knowledge and skills and will lack the repertoire of teaching ideas that they have for more familiar subjects. Many primary school teachers are unfamiliar with the concepts of computational thinking and design thinking and consequently may be anxious about teaching the Technologies curriculum (Albion et al., 2016). To ensure the successful implementation of the curriculum, high-quality learning resources and activities are needed (Falkner & Vivian, 2015). Therefore, it is urgent to provide professional learning opportunities to build up primary school teachers’ capacity and confidence to be able to teach the Australian Curriculum: Technologies.
Remote Access Laboratories (RALs)

Remote Access Laboratories (RAL) are well established in universities for providing students with more flexible access to experiments, especially in electrical and computer control engineering (Maiti, Maxwell, & Kist, 2013). RAL systems enable students to view and control equipment at a distance using cameras and sensors, and download real-time data in a computer laboratory, a classroom or at home (Tho & Yeung, 2016). They have been used effectively in secondary schools (Lowe, Newcombe, & Stumpers, 2013) and may also offer benefits for primary schools through sharing of equipment that is expensive to acquire and maintain. RAL has been used widely in engineering and computer science but there is little research on its application in other disciplines such as teacher education (Kist et al., 2014).

RALfie

The Remote Access Laboratories for Fun, Innovation and Education (RALfie) project was a joint effort between academics in Engineering and Education at University of Southern Queensland. Engineers were responsible for the technical aspects and educators provided pedagogical support (Kist et al., 2011). In traditional client-server RAL systems a university or other organisation hosts the RAL system and manages user access (Maiti, Kist, & Maxwell, 2015). In contrast, RALfie is a distributed RAL system with a modular design that allows participants to create and house experiments at distributed locations. The distributable feature makes RALfie more flexible for users (Kist et al., 2014). It expands the one-to-many approach, where one central laboratory serves many users, to a many-to-many approach, with many users using multiple equipment installations shared by various providers (Maiti et al., 2015). This project is unique because it provided both hands-on and remote activities that were incorporated into classes for preparing pre-service teachers to teach the Technologies curriculum as a medium for increasing their confidence with STEM activities. Pedagogical support provided by educators was important for the design of activities.

Hands-on learning

Piaget’s developmental stage theory suggests that learning by children starts from the concrete and moves to the abstract (Piaget, 1974). Many educators believe that adults pass through similar stages when learning in new areas and learn best, especially when beginning in a new field, with concrete or hands-on experiences (Jacobs, 2001). Tinkering and making are powerful ways to learn because they allow makers to try out ideas, make adjustments and experiment with new things (Martinez & Stager, 2013). The tinkering approach is characterized by “a playful, experimental, iterative style of engagement, in which makers are continually reassessing their goals, exploring new paths, and imagining new possibilities” (Resnick & Rosenbaum, 2013, p. 164).

Scaffolding

Constructivists argue that learning is an active process (Vygotsky, 1978). Students construct their understanding and knowledge based on their existing knowledge (Bryant et al., 2013). Scaffolding is a metaphor to explain guiding learning and development paths. It describes how teachers or peers supply students with the assistance they need to learn and then slowly withdraw help as students are capable of doing more independently (Jacobs, 2001). Scaffolded professional development is significantly superior to professional learning through self-study in terms of teacher beliefs and motivation, student performance and quality of instruction, and evidence shows that expert scaffolding in professional learning based on science curriculum has an advantage for primary teacher preparation for teaching science (Kleickmann, Tröbst, Jonen, Vehmeyer, & Möller, 2016).

STEM anxiety and its measurement

Science anxiety has been defined as a fear of, or aversion towards, science concepts, scientists, and science related learning activities (Sahin, Caliskan, & Dilek, 2015). Anxiety leads to panic, tension, and loss of ability to concentrate (Idowu, 2013). A large number of teachers required to teach STEM potentially have STEM anxiety (Bryant et al., 2013) which is detrimental to the effective teaching of STEM subjects in the classroom. The Positive and Negative Affect Schedule (PANAS) is a self-report measure assessing adult experiences of positive and negative affect (Watson, Clark, & Tellegen, 1988). There are twenty items, 10 each for positive (PA) and negative (NA) affect, related to various affective items which are adjectives describing mood states. A five-point
Likert-type scale is used by respondents to rate their mood against each item. The PA and NA have been identified as two dominant and relatively independent dimensions of the structure of affect (Watson et al., 1988), which can be used and analysed separately because they are two independent constructs (Hughes & Kendall, 2009).

Research context and problem

This research was conducted with PSTs enrolled in a final year course designed to prepare them for teaching the Australian Curriculum: Technologies in primary schools. The availability of RALfie offered an opportunity to provide them with enjoyable activities that had potential to alleviate anxiety about STEM through successful experiences. The study also enabled preliminary exploration of the potential of RAL for teacher development and use in primary schools, which is of particular interest in Australia where population centres and schools are often separated by significant distances.

Because the research was conducted in the context of an existing course, the time available for RAL activities was limited. Moreover, only 25 of the 168 students in the course were enrolled on the campus where they had direct access to the RALfie activities. The remainder were able to access the activities remotely via web browser. The study was conducted in a single semester during which all students had access to two online RALfie activities and those enrolled on campus could also participate in two face-to-face sessions.

The study explored PSTs’ responses to the experience of working with the RALfie activities. This paper reports on the following research questions:

- What factors make a difference in pre-service teachers’ experience of RAL?
- What can we learn from the RALfie experience to guide future use of RAL in primary education?

Research method

Quantitative methods have dominated analysis of anxiety since very early research. However, quantitative surveys are unable to provide specific reasons for changes in participants’ emotional status. A solely quantitative approach is inadequate to explore the relationship between pre-service teachers’ capacities for teaching STEM and engagement with hands-on and remote RAL activities. Additionally, the context of this study was complex because RAL was not the only intervention in the class. It was important to use a qualitative approach to attribute reasons to changes in pre-service teachers’ emotional status.

The study used mixed methods. PANAS was used to assess PSTs’ positive and negative affect scores before and after participating in RALfie activities. PSTs were also invited to volunteer for semi-structured interviews following completion of the RALfie activities. Difference scores calculated separately for PA and NA were used to identify PSTs’ who had been interviewed for inclusion in case studies for deeper investigation of changes in their emotional states.

Case study allows deep analysis of individuals, especially with consideration of the individual’s background, preconceptions and attitude. For this study, case study allowed deep analysis of the changes of confidence and emotional state resulting from the interaction with RALfie activities. The case study approach should help readers to understand the reasons for positive or negative change of PSTs’ confidence and attitudes. Using multiple data sources is argued as a major strength of case study because it is more likely to generate accurate and persuasive findings based on a variety of evidence (Yin, 2009).

RAL experiences

The RALfie project offered PSTs maker events, which were face-to-face, and user activities, which were online. All PSTs in the course had access to the user activities but only the 25 students enrolled on campus could participate in the maker activities.

In the maker activities, participants used LEGO Mindstorms to assemble hands-on experiments and connect them to the RALfie system for remote access. Engineering academics assisted with setting up the maker activities and interacted with on-campus participants as they worked on the activities. Two maker events, each lasting two hours, were offered. In the first, PSTs were engaged in programming using Snap!, which is a free,
graphical, drag and drop programming language (Garcia, Segars, & Paley, 2012). They learned how to use the Snap! language to program the LEGO EV3. Ready-made robots were connected to the server computers by the engineering academics. Students then used client computers to control the robots. In the second maker event, PSTs used the LEGO Mindstorms Kit to build robots, program their robot using Snap! and play robot soccer. Figure 1 shows the PSTs working with their robots and the completed robots playing soccer.

![Figure 1. Robot soccer](image1)

In the user activities, participants, who included both on-campus and online students, were able to remotely operate experiments mounted in the Engineering laboratory. Two of the experiments being tested as part of the RALfie project development were selected as the basis for learning activities that could be undertaken by pre-service teachers. Both offered experience with remote operation of the equipment and had broader relevance to the technologies curriculum. Each was presented with step-by-step instructions with illustrations in a webpage within the course materials that included background information, links to relevant resources, and questions for reflection.

Figure 2 shows the Pendulum activity. It presents an apparatus in which a ball bearing could be raised or lowered to a selected distance from the pivot point and then pulled to one side and released using a mechanism constructed with LEGO. Recording the time required for a swing at different lengths enables exploration of the relationship between length and period of a pendulum and estimation of the gravitational constant. In practice PSTs did not have time for repeated trials so they were asked to raise the ball bearing to a suitable height, set it in motion, record the time for 20 swings, and enter that time and the length of their pendulum in a Google form where the data entered by all users were aggregated and displayed on a graph driven by a Google sheet. The intention was to use the pooled data to estimate the gravitational constant which users were also invited to calculate directly for comparison with the pooled result.

![Figure 2. RALfie Pendulum activity](image2)

Figure 3. RALfie Gearbox activity
The Gearbox activity in Figure 3 presented users with a gearbox constructed using LEGO and the challenge to determine the ratios among the 4 gears, A to D. The setup included a graphical user interface similar to that shown in Figure 2 but omitted in Figure 3 to afford a clearer view of the gears. Users were able to remotely control the motor to rotate Gear C through a selected angle (in degrees) and observe and record the rotation of the other gears to determine the ratios.

**Data collection and analysis**

An online questionnaire using LimeSurvey (limesurvey.org) captured quantitative data about affective states, using PANAS (Watson, Clark, & Tellegen, 1988) before and after the RALfie intervention. The second questionnaire included open ended questions to gauge reactions to the experience. Quantitative data were extracted from LimeSurvey for analysis using SPSS. Scores for PA and NA were calculated for each participant on both the pre-test and post-test applications of PANAS. The differences between those scores for each participant were calculated to examine the changes in their PA and NA. From 168 enrolled students invited to respond, there were 122 completed questionnaires from the first administration and 47 from the second. The pre-post survey data (N = 40) were inconclusive because the numbers of respondents in different conditions were insufficient for statistical analysis. Subsequently the pre-post survey data were used to guide selection of interviewed participants for closer investigation as cases.

Qualitative data from the second questionnaire and transcribed interviews were imported into NVivo and analysed thematically to develop case themes. Semi-structured interviews were conducted with six PSTs using prompts based on the research questions. Participants were asked to recount their experience of working with hands-on and online RALfie activities and comment on which activities worked well for them and which did not and why. Participants were asked to comment on aspects of the experience that made them feel more or less confident to teach the Technologies Curriculum, and the support they need in the future.

Three overarching themes, each with sub-themes, were generated from the qualitative data. The overarching themes related to advantages of the RALfie activities experienced by participants, issues experienced while accessing RALfie activities, and effects of personal background on the experience.

This paper reports data for two PSTs for whom there were notable changes in affective states as measured by the PANAS. Each case starts with description of PANAS results, which is followed by a case description including elements of biography and reactions to the RALfie experiences (Meyers & Bagnall, 2015).

**Case studies**

**Case 1. No. 28 Sally**

Figure 4 presents a scatter plot of pre-test score (X-axis) against post-test score (Y-axis) for PA with the four groups of respondents identified as shown in the legend. Points above the diagonal line represent respondents who scored higher on the post-test than pre-test. Points below the diagonal line represent respondents who reported decreases in PA.

For the pre-post PA score, Sally was an outlier whose pre-post PA score difference was greatest among the six interviewees, indicating that Sally had a notably positive experience with RALfie. Her PA score increased from 1.50 to 3.20, and her NA score decreased from 1.70 to 1.20. The implication is that she was engaged during the RALfie experience and enjoyed it.

Sally was in her early 20s and was in the final year of her preparation to be a primary school teacher. She was an online student and participated only in the RALfie user activities, spending a total of 1.5 hours on them. The Pendulum activity did not work for her but the Gearbox activity did. She was willing to use RALfie to teach in the future and thought she could host an experiment herself for a class if she was provided with a RALfie kit. Five themes emerged during analysis of her data.
Theme 1: Science learning experience in primary school

Sally had a very positive learning experience with science when she was a child. She commented that “when I moved to Queensland the teacher I had in Queensland he was very positive toward, especially science and things and trying to make it interesting and very hands-on.”

Sally’s positive science learning experience in primary school led her to continue to learn science at university and she stated that “I know a couple of the science. I’ve actually used those experiments at Uni and I think it was in 1st year the science course. So that was another good one as well.”

Sally believed that her positive science learning experience during childhood helped her to become confident to teach science. She stated that “I’m sure that really does just being confident in that area or having a good experience myself in that area especially with science in the classroom, quite positive to teach that one.”

Theme 2: Prior experience of LEGO

Sally had seen LEGO kits when she was in her teaching practice placements and commented that “I’d actually seen that before on a previous prac that they used the little LEGO kits in the classroom and the kids…it could’ve been 2 to 3 times I’ve seen LEGO being used … I’ve had a positive use with it in the classroom.”

Theme 3: Hands-on experiences

Sally preferred participating hands-on with RALfie because the hands-on activities were concrete and playful. She explained that “just being able to touch it and feel it and play with it would be a bit easier. Sometimes you’ve got to watch something a couple of times online to fully grasp what is being done or what is being said.”

Theme 4: The ease of use of online RALfie and resources

Sally had a positive experience going through RALfie brochures, RALfie websites, and YouTube Videos. She commented that the materials “advertised it at the start on the brochure…I found it quite easy to get around and click on different things and have a look and the instructions… I found getting through the different activities and quests wasn’t too bad and your website as well”

Theme 5: Scaffolding by professional engineers and educators

Sally commented on the cooperation between the course lecturer and the RALfie team, stating that “I think that was done really well especially with the quests and it was really easy to access from the course.”
Case 2: No. 20 Anissa

Figure 5 shows the pre-post comparisons of NA scores. PSTs who are located below the diagonal line experienced a decrease in NA between first and second administration of the PANAS, demonstrating a generally positive response to their experiences. PSTs located above the diagonal line recorded an increase in anxiety. For Anissa (No. 20), her pre-post NA increase was greatest among 6 interviewees which was a negative indicator. Her NA score increased from 1.40 to 2.00. The implication is that she experienced increased negative affect possibly including heightened anxiety during the RALfie experiences. However, her PA scores increased from 2.4 to 2.9 which suggests that RALfie experiences offered something enjoyable to her.

Anissa was in her early 20s and was in the final year of her preparation to be a primary school teacher. She participated in both maker events and the user activities. She had a positive experience with the maker events, spending a total of 3 hours across the two maker events. She experienced some difficulties with operation of the online user activities at both university and home and spent just 20 minutes attempting the online user activities.

Theme 1: Science learning experience in primary school

Anissa did not enjoy learning science and technology when she was in primary school. When Anissa was asked why she was scared of learning science and technology, she commented that “I just never have liked it at school. Like the perception was ruined for me I think – like the way they taught it and what was expected and stuff.”

Anissa did not learn much science at primary school which made it hard for her to continue to learn science at high school. “Well for primary school we hardly did any science so when I got to high school it was like you should have had all this knowledge which I didn’t have because at my school science wasn’t a big deal.”

Anissa further commented on the negative preconception of science learning, stating that “I think there is a perception at school. Like science is like the smart subjects and like you can’t do them at school – you’re not going to be able to do them at university.”

Anissa commented on her increased confidence as a science and technology teacher and stated that “it’s developing as I go further through my degree. If you’d asked me at the beginning I would have been scared to teach science and technology but now I’m getting more confident like knowing things.” It was consistent with her low pre-post PA and high pre-post NA result.

Theme 2: Prior experience of LEGO

Anissa commented on her successful learning experience with Robot Soccer activities thus, “I can do it; ... I shouldn’t be afraid of technology as much as I am. Like the making of it wasn’t hard and once you got Scratch down – like it was quite easy.”
After the RALfie experience, she was more likely to join in Robogals and teach robotics. Robogals was a robotics organization with a branch at University of Southern Queensland. Robogals use LEGO Mindstorms to program robotics and teach them at local primary and secondary schools. Anissa stated that “I have been asked to join Robogals a couple of times but I didn’t think I would be confident enough to do it...but probably now I would. Because from the RALfie like seeing it’s not as hard as I thought it would be I think was my preconceptions that it was going to be really difficult but it wasn’t so.”

**Theme 3: Hands on experiences**

Anissa liked the maker event “Because I was engaged – like I had stuff to play with...the making one was more engaging – like seeing the whole process from start to finish.”

Anissa was more positive and willing to try new technologies after the experience of the RALfie maker event. She stated that “Probably anything is possible. Like you don’t have to be scared of it because it is like doable like it is attainable if you set your mind to it.”

**Theme 4: Resources**

For the user activities, Anissa enjoyed the Gearbox activities “because you had to think – like the maths side of it comes into it as well. Like thinking about the degrees and if I turn that one then that one’s going to turn that far and then it’s going to go opposite and thinking about all the different aspects that come together just to turn one little thing was really interesting.”

**Theme 5: The ease of use of online RALfie**

Anissa gave up when the RALfie system did not work, stating that “it wasn’t working and there was a glitch with the computer so it wasn’t working so I gave up pretty easily.”

**Theme 6: Scaffolding provided by professional engineers and educators**

The instructor’s encouragement and scaffolding was important for engaging PSTs to try RALfie. Anissa commented that “he really wanted us to learn Scratch and do the activities. Like there was always time set out to do it so we had the opportunity. If we couldn’t do it at home we could always do it in class. Like he was always very encouraging that we at least try to do it.”

The course instructor taught PSTs how to use Scratch which helped them to use Snap!. “I think if you didn’t have a knowledge of Scratch and you were told to program something you would fail. Like just the little knowledge of Scratch that I had at the first Maker event – it helped immensely.”

**Discussion**

Based on the data presented above, it is suggested that the following factors make a difference in PSTs’ experience of RAL: science learning experience in primary school, prior experience of LEGO, scaffolding by engineers and educators, access to resources, hands-on experiments, and the ease of use of the online RALfie system. The following sections discuss the results from the cross-case analysis with references to the related literature (Ebrahimí, Faghíh, & Marandi, 2016).

**Science learning experience in primary school**

Sally and Anissa had contrasting science learning experiences in primary school. It is evident that Sally had a positive science learning experience in primary school. She commented that “the teacher I had in Queensland, he was very positive toward science...and very hands-on.” On the contrary, Anissa did not learn much of science in primary school, which contributed to her insufficient background knowledge for science learning at tertiary level. She stated that “for primary school we hardly did any science so when I got to high school it was like you
should have had all this knowledge which I didn’t have.” This was consistent with the literature suggesting that the decrease in study of science at high school develops from lack of science learning in primary education (Westerlund et al., 2011).

Sally and Anissa demonstrated different attitudes towards science teaching because of their different science learning experiences in primary school. Sally stated that “I’m sure that really does. Having a good experience myself in science in the classroom make it quite positive to teach.” Sally was very confident to teach science because she has a positive science learning experience. Anissa was scared of science due to her lack of science learning in her background. She commented that “I just never have liked it at school.” Prior experience and background knowledge are important to construct new knowledge (Vygotsky, 1978). Sally’s positive learning experience of science in primary school led to her positive attitudes and confidence to teach science. Whereas for Anissa, lack of prior knowledge and background learning resulted in her negative attitudes and anxiety to teach science.

Prior experience of LEGO

Sally had a very positive experience with RALfie. Her prior successful experience of using LEGO to teach in her teaching practicum gave her a positive outlook to use the online RALfie activities. She knew that children will be excited about using LEGO to learn in the classroom. She was very confident to use the Gearbox activity. Anissa’s lack of prior experience of using LEGO contributed to her anxiety. It is consistent with the literature that learners construct knowledge based on their prior experience and background (Vygotsky, 1978).

Scaffolding by engineers and educators

Scaffolding provided by the RALfie team supported PSTs to use the online RALfie activities. The course instructor taught them how to use the Scratch programming language before the maker events. Anissa adapted her background knowledge of programming to use the Snap! language, demonstrating it was easy for her to build on her prior knowledge and internalize new knowledge (Vygotsky, 1978). There were engineering academics who provided the RALfie system and also offered face-to-face interaction with PSTs in the maker events. Social interactions between PSTs and the engineers were important to create a supportive learning environment. The course instructor’s encouragement also helped PSTs to alleviate their anxiety levels and to try new technology.

Resources

Resources provided by the RALfie project were important to engage PSTs. Brochures, websites and videos were used for demonstration which was very helpful to engage PSTs to use RALfie. Sally enjoyed watching the videos and navigating the RALfie website. Providing high quality resources was important to build up PSTs’ confidence and capacity to teach (Albion & Spence, 2013). Resources are also important to PSTs’ professional learning.

Hands-on experiences

Hands-on experiences provide a sense of playfulness which was helpful to alleviate the sense of anxiety and frustration with using robotics. Anissa stated “I was engaged – like I had stuff to play with,” which was in line with her increased PA score. Hands-on activities are engaging as PSTs can tinker, play, and build things. Playing and tinkering with hands-on equipment was important for PSTs who were at the beginning level of using robotics. Moving from concrete maker activities to abstract user activities was in line with Piaget’s learning stages theory (Piaget, 1974). It was also in line with the RALfie project’s concept that “f stands for fun.” Before the engagement with maker activities, Anissa was reluctant to join in the Robogals. However, after engagement with the maker activities, she was more willing to participate in the Robogals. Hands-on experiences were important to build up Anissa’s confidence and capacity to use robotics. Moreover, hands-on experiments were powerful to engage learners to learn science for future career pathways (Westerlund et al., 2011).
The ease of use of online RALfie

The ease of use of online RALfie activities contributed to Sally’s enjoyment and engagement during her interactions with RALfie experiences. Overall, Sally had a very positive experience with RALfie which is evident from her increased PA score and decreased NA score. However, Anissa experienced difficulties with using the online activities. Anissa stated that “It wasn’t working and there was a glitch with the computer so it wasn’t working so I gave up pretty easily.” Her increased NA score after her RALfie experience was consistent with experiencing anxiety while working with RALfie. The difficulties of use and unreliability of the online activities contributed to Annisa’s anxiety and disengagement. Therefore, the ease of use of the online system affects PSTs’ emotional status when using RALfie. It is important to make the online system user-friendly.

Conclusion

The cross-case analysis described above yielded useful insights into the areas probed by the research questions. RAlfie is a variant of RAL. Hence the user experiences and feedback on the technical problems and teacher support PSTs needed for RALfie are relevant to any consideration of more extensive use of RAL to support professional development of teachers or for direct use in primary school classrooms.

The most influential factor influencing PSTs’ experience of RAL was their background knowledge and prior STEM experience in their own primary and secondary education. PSTs who had a positive experience with STEM learning in primary school were more likely to be positive and confident to use RAL. PSTs who had a negative experience with STEM learning in primary school were more inclined to hold a negative attitude and be anxious about using RAL. This highlights the importance of primary STEM education for prospective teachers’ confidence and capacity to engage with STEM activities.

A second influential factor was direct access to the activities. Hands-on experiments were powerful to motivate and engage PSTs to learn robotics. For PSTs, especially for those who did not learn much science and technology in their own schooling, hands-on activities were playful and enjoyable which was helpful to alleviate their STEM anxiety. Compared to conventional RAL systems, the distributed feature of RALfie offers potential for users to take full advantage of RALfie as remote makers and users regardless of location.

The third factor that affected PSTs’ responses to working with RAL was their user experience, which was influenced by the reliability of the systems and the levels of support available either directly or indirectly. Scaffolding was of great importance, both instructional scaffolding provided by professional engineers and educators in the laboratory or in the presentation of online activities and technical scaffolding to make the RALfie system reliable and user-friendly. The technical scaffolding was particularly important to alleviate PSTs’ anxiety and increase their engagement.

Each of these factors affecting PSTs’ experience of RAL points to lessons for the future application of RAL in primary education. PSTs’ own experience of STEM influenced their willingness to engage in the activities. Hence, it seems logical that teachers with positive experiences of STEM will be more likely to engage with and offer STEM activities in their classrooms. Thus it is important to provide more STEM activities for primary school teachers and teacher candidates to boost their prior experience with STEM. RAL offers advantages for overcoming the challenges of distance in Australian education systems by making available a broader range of experiences than would otherwise be possible. Those experiences could be valuable for teacher professional development and for direct use with students in school classrooms.

Direct access to hands-on activities was especially motivating for the PSTs and the same is likely to apply for teachers and students in school classrooms. The distributed feature of RALfie offers particular advantages because it would allow schools to share experiments and to take full advantage of RALfie as remote makers and users. RALfie has the potential to provide hands-on and online opportunities for teachers to develop capability and confidence for implementing the Australian Curriculum: Technologies and to offer a wider range of learning opportunities for students in their classrooms.

Finally, user experience is an important factor to be considered in any system, especially one which is intended for use by children and inexperienced users. If RAL is to be used successfully to enhance the learning opportunities available to teachers and students in isolated schools, the system needs to be designed for ease of use and supported by adequate documentation and other supports.
This study, though small in scope, identified factors influencing PSTs’ use of RAL and some implications of those findings for broader application of RAL for use by teachers and students in isolated schools. If those findings are heeded RAL, especially in the distributed format facilitated by RALfie, can make a valuable contribution to the development of teachers’ capacities for implementing the Australian Curriculum: Technologies and to the resources available to support students’ learning about technologies. By doing so it will enable Australians to create their preferred futures with greater opportunities for shared prosperity through application of STEM knowledge.

Acknowledgements

This research was supported through the Australian Government's Collaborative Research Networks (CRN) program.

Reference


Fostering EFL teachers’ CALL Competencies Through Project-based Learning

Sheng-Shiang Tseng1 and Hui-Chin Yeh2*

1Graduate Institute of Curriculum and Instruction, Tamkang University, New Taipei City, Taiwan // 2Department of Applied Foreign Languages, National Yunlin University of Science and Technology, Yunlin, Taiwan // u9241346@gmail.com // hyeh@gemail.yuntech.edu.tw

*Corresponding author

(Submitted October 22, 2016; Revised March 28, 2017; Accepted May 12, 2017)

ABSTRACT

Project-based learning (PBL), a learning-by-doing practice, has been used for enhancing English as a Foreign Language (EFL) students’ language skills. However, the extent to which and how EFL teachers develop or improve Computer-Assisted Language Learning (CALL) competencies while experiencing PBL remain unexplored. For this study an 18-week PBL project was designed to improve EFL teachers’ CALL competencies. A total of 12 EFL prospective teachers were recruited to participate in a sequence of activities: class observations, group discussions, and the design of lesson plans. Pre- and post- TPACK (technological pedagogical content knowledge) surveys were administered to measure participants’ improvement of CALL competencies. Qualitative data, including class observation notes, lesson plans, group discussion records, and reflective essays, were collected to triangulate and complement survey results. The survey results showed that the prospective teachers demonstrated higher levels of CALL competencies after the PBL project. Using the qualitative data, this study explicitly documented the benefits which prospective teachers may obtain and the problems they may face when participating in a PBL project. The findings can help future teacher educators understand how to design and implement effective training for CALL competency development.

Keywords

Project-based learning, CALL competencies, Teacher training

Introduction

Many technological tools have been integrated into language teaching to enhance English as a Foreign Language (EFL) students’ language skills. For example, information and communication technologies (ICTs), such as Twitter and Facebook, help overcome the scarcity of exposure to English in EFL contexts by enabling students to communicate with native speakers (Kim, 2011; Sun, 2010). Web-based tools, such as Google Docs, Wiki, and Blogs, have also been used by EFL teachers to create a collaborative learning environment for writing and reading (Arslan & Şahin-Kizil, 2010; Lin & Yang, 2013; Lin & Yang, 2011; Marzban, 2011; Sato, Matsunuma & Suzuki, 2013; Tozeu & Coady, 2004; Tseng & Yeh, 2018; Vanderplank, 2010; Yaghoobi & Razmjoo, 2016; Zaini & Mazdayasna, 2014; Yeh, Tseng, & Chen, 2019). To effectively integrate these tools into language teaching, EFL teachers should possess Computer-Assisted Language Learning (CALL) competencies, which combine instructional and content knowledge with technological knowledge (Golonka, Bowles, Frank, Richardson, & Freynik, 2014; Hong, 2010; Jones, 2001; Liu & Kleinsasser, 2015; Yang & Wu, 2012). The three types of knowledge align with Mishra and Koehler’s (2006) and Koehler and Mishra’s (2008) technological pedagogical content knowledge (TPACK) framework, comprising the three fundamental knowledge bases of content knowledge (CK), pedagogical knowledge (PK), and technological knowledge (TK), which provided a theoretical framework for this study. CK refers to the knowledge of the subject matter to be taught, including concepts, organizational frameworks and evidence for claims (Shulman, 1986). PK includes knowledge of learning theories, general classroom management skills, lesson planning, and student assessment. PK enables teachers to understand how students construct knowledge and use this understanding to implement instructional practices that scaffold students’ learning. TK, which is not limited to understanding how to operate various technologies, includes knowledge of how to choose the right technologies to accomplish tasks in specific situations. CK, PK, and TK intersect with each other to form technological content knowledge (TCK), and technological pedagogical knowledge (TPK) (see Figure 1). Therefore, the TPACK framework has been used by researchers (e.g., Liu & Kleinsasser, 2015) to understand teachers’ technology integration competencies.

Although many teacher training venues such as conferences and workshops have been used to foster EFL teachers’ Computer-Assisted Language Learning (CALL) competencies, most EFL teachers still have difficulty adapting technologies to their own teaching contexts (Lin & Lu, 2010; Teo, 2009; Wang, Ertmer, & Newby,
2004). Even in-service teachers equipped with high levels of technical knowledge find it difficult to develop appropriate technology supported teaching materials and activities for their classrooms (Jang, 2010), often because previous teacher training programs focused on delivering content, demonstrating skills, or employing instructional strategies without providing opportunities for hands-on experience with various technologies (e.g., Chao, 2015; Hong, 2010; Kim, 2011). An effective teacher training program should provide teachers with experiences in which they practice combining teaching content, teaching activities, and using technology tools for authentic teaching purposes (Debski, 2006; McNeil, 2013).

One way to provide such experiences is to combine CALL with Project-based learning (PBL), a learning-by-doing practice in which teachers acquire knowledge through the design of usable artifacts such as lesson plans. This approach enables teachers to connect new knowledge to their own teaching contexts (Krajcik & Blumenfeld, 2006) and thus build connections among content, activities, and technology tools. However, only a few studies have investigated using PBL to develop teachers’ CALL competencies and some caveats concerning the data in the investigation of PBL were identified. For example, Liu and Kleinsasser (2015) used self-reported data such as surveys and interviews to investigate the effects of PBL upon the development of teachers’ CALL competencies, but the self-reported data could not capture how teachers develop CALL competencies and might not reveal teachers’ actual CALL competencies. Therefore, the extent to which and how teachers develop or improve CALL competencies while experiencing PBL remains unexplored. To address this gap and seek ways to use PBL to support EFL teachers’ CALL competency development, this study examined the effects of a PBL project featuring a sequence of activities including class observations, group discussions, lesson plan design, guided by the following research questions:

- What is the impact of a PBL project on the EFL teachers’ CALL competencies?
- How do EFL teachers develop CALL competencies through PBL?

Theoretical foundation of PBL

PBL aligns with a constructivist view of learning, which advocates learning by collaboratively investigating and solving real-world problems. Three elements of PBL identified by Krajcik and Blumenfeld (2006), (1) generating driving questions, (2) developing artifacts, and (3) collaborating with peers, were adapted in designing the PBL project in this study. First, generating driving questions means that learners produced the questions to help them organize meaningful project goals and activities. Good driving questions are open-ended but feasible for learners to solve, relevant and important to their realities, and therefore interesting. Also in the case of inquiry, they provide worthwhile data (Krajcik & Blumenfeld, 2006). Poor driving questions have predictable outcomes, leaving no space for learners to explore alternatives, exercise judgment, and gain new insights and skills. One way to generate effective driving questions for teacher training is to engage teachers in class observations and stimulate their critical reflection on the problems of real-world teaching practices (Bell & Mladenovic, 2015).

Second, constructivist theory suggests that knowledge is not acquired through transmission but is developed through involvement in collaboratively developing artifacts to address the issues raised by driving questions. Developing artifacts involves learners in a series of activities including planning, searching for information, analyzing the information, and making products while sharing ideas with others (Blumenfeld, Soloway, Marx, Krajcik, Guzdial, & Palincsar, 1991). Such activities help learners build information literacy, problem-solving, and critical thinking skills (Garrison, 2007; Howard, 2002). Accordingly, to develop EFL teachers’ CALL competencies and the ability to integrate technology into their teaching, participants in the present study were asked to develop a CALL lesson plan to address an observed teaching problem. The resulting artifacts were

---

Figure 1. The TPACK framework (Mishra & Koehler, 2006)
expected to represent the teachers’ emergent state of CALL competencies and their understanding of the driving question, which could inform further efforts to promote teachers’ CALL competencies.

Finally, constructivist theory contends that collaboration plays a key role in knowledge construction because knowledge is developed and co-constructed with others rather than individually (Lave & Wenger, 1991; Wenger, McDermott & Snyder, 2002; Vygotsky, 1978). PBL provides the opportunity for teachers with different CALL competencies and teaching experiences to work together to solve teaching problems, sharing information and discussing and challenging ideas (Lin & Hsieh, 2001; Ward & Tiessen, 1997). Several studies have shown that collaboration enhances teachers’ problem-solving abilities and willingness to undertake challenges and adopt new approaches in their classrooms (Cochran-Smith & Lytle, 1999; Hopkins, 2000; Van Horn, 2006). Collaboration can thus enable teachers to go beyond their current knowledge and build a shared understanding of effective CALL instruction.

A few empirical studies have examined whether PBL can be used effectively in EFL teacher education and some caveats concerning the data in the investigation of the PBL were identified. For example, Liu and Kleinsasser (2015) investigated how six vocational high school EFL teachers developed CALL competencies in a year-long technology-enriched professional development (TEPD) program, in which the teachers designed two WebQuest projects. The teachers learned about the theory and design of project-based instruction, discussed its implementation, and learned how to use a WebQuest procedure through a Moodle platform. A pre- and post-TPACK survey was administered to measure the improvement of their CALL competencies. The results showed that teachers achieved higher levels of CALL competencies through the analysis of pre- and post-TPACK surveys. However, the self-reported TPACK surveys might not accurately reveal teachers’ CALL competencies learned through the PBL project, as there is usually some discrepancy between what individuals have reported on surveys and what they have actually done (Egbert, Paulus, & Nakamichi, 2002). In addition, the TPACK survey data itself could not capture how teachers developed CALL competencies throughout the PBL project. Other kinds of data should be collected to complement and validate survey data.

Methodology

Participants

The participants were a group of 12 prospective teachers majoring in an applied foreign language program and enrolled in an elective three-credit course, “Teaching children English,” at a University of Science and Technology in central Taiwan. The objective of the course was to engage these prospective teachers in using technology to design a lesson plan for primary school students. Eight of the participants had no teaching experience, two had tutored English for one year, and two had tutored for two years, and all were in training to become certified English teachers. Participants were informed of the research procedures and data to be collected. All were given pseudonyms to ensure their anonymity.

Research design of this study: A PBL project

An 18-week teacher training course was designed to align with the three elements of PBL: driving questions, collaboration, and artifact development (Krajcik & Blumenfeld, 2006) (see Figure 2 and Table 1). First, to generate driving questions, the participants were divided into three groups to observe two English classes. Each group used a language classroom observation guide to document the strengths and weaknesses of the class and summarize two or three teaching problems that interested them. These teaching problems were then converted into the driving question of the group’s PBL project.

Second, each group developed an artifact, a CALL lesson plan, to address the driving questions. As shown in Figure 2, the design of the CALL lesson plan involved the procedures of setting teaching objectives based on the driving questions, exploring a variety of online English teaching tools, selecting at least two online tools, and designing teaching activities based on the selected tools. These procedures were intended to help the participants to improve their CALL competencies. The CALL lesson plans designed by the participants were the artifacts that represented the emergent state of their CALL competencies and their understanding of the driving question.

Third, collaboration was implemented in this study through group discussions. Two rounds of online group discussions were conducted using Google Docs. In the first group discussion, which occurred right after the class observation, the members of each group reviewed their observation notes and identified lesson plan topics. The
second group discussion was held after the prospective teachers had received instruction in how teaching theories inform CALL pedagogical design for EFL learners. The goal of the second group discussion was for the prospective teachers to design teaching activities using technology tools. To facilitate each group discussions, a list of guiding questions was given to the participants, such as “What problems of English learning and teaching can you identify from your observations?” “What do you want students to learn from your lesson plans?” and “What are the instructional activities and tools that you will use in your project?” The participants discussed those questions with group members and kept records of their conversations in written form using Google Docs.

*Figure 2. Design of the PBL project*

*Table 1. Timeline of the PBL project*

<table>
<thead>
<tr>
<th>Weeks</th>
<th>Activities</th>
<th>Collected data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 1</td>
<td>Taking pre-TPACK survey</td>
<td>Pre-TPACK survey</td>
</tr>
<tr>
<td>Week 2</td>
<td>Introducing the PBL project</td>
<td>None</td>
</tr>
<tr>
<td>Week 3</td>
<td>Observing an English class</td>
<td>Class observation notes</td>
</tr>
<tr>
<td>Week 4</td>
<td>Discussing class observations to generate driving questions</td>
<td>Group discussion records</td>
</tr>
<tr>
<td>Week 5</td>
<td>Observing an English class</td>
<td>Class observation notes</td>
</tr>
<tr>
<td>Week 6</td>
<td>Discussing class observations to generate driving questions</td>
<td>Group discussion records</td>
</tr>
<tr>
<td>Week 7-8</td>
<td>Designing a CALL lesson plan: Setting teaching objectives based on the driving questions through group discussions</td>
<td>CALL lesson plans</td>
</tr>
<tr>
<td>Week 9-11</td>
<td>Designing a CALL lesson plan: Searching and selecting online English teaching tools through group discussions</td>
<td>CALL lesson plans</td>
</tr>
<tr>
<td>Week 12-14</td>
<td>Designing a CALL lesson plan: Designing teaching activities based on the selected English teaching tools</td>
<td>CALL lesson plans</td>
</tr>
<tr>
<td>Week 15-16</td>
<td>Presenting the CALL lesson plan</td>
<td>CALL lesson plans</td>
</tr>
<tr>
<td>Week 17</td>
<td>Writing the reflective essays about the PBL project</td>
<td>Reflective essays</td>
</tr>
<tr>
<td>Week 18</td>
<td>Taking post-TPACK survey</td>
<td>Post-TPACK survey</td>
</tr>
</tbody>
</table>

In the final stage of the PBL project, the prospective teachers were asked to compose a 500- to 750-word reflective essay to describe what they learned from class observations, group discussions, and the design of lessons. The reflective essay was guided by seven open-ended questions, e.g., “How did the class observations improve your CALL competencies?” “How do you think your ability to use technology tools in teaching has improved after completing the PBL project?” and “Which of the PBL activities did you feel benefitted your CALL competencies the most? Please explain in detail.”
Data collection and analysis

A mixed-methods design was employed to investigate how the PBL project improved prospective teachers’ CALL competencies and which aspects had the most influence. The quantitative data were pre- and post-TPACK surveys. The qualitative data included reflective essays, class observation notes, group discussions, and lesson plans.

TPACK survey

The TPACK survey, administered before and after the PBL project, was used to investigate the impact of the PBL project upon prospective EFL teachers’ CALL competencies (Research question 1). The survey was adapted from Schmidt, Baran, Thompson, Mishra, Koehler, and Shin’s (2009) instrument, designed to measure “teachers’ understanding of and self-reported ability to apply the domains of TPACK throughout their teacher preparation programs and in classrooms during practicum experiences” (p. 128). The survey comprised four items for collecting demographic information and 33 Likert scale items related to TK, CK, PK, PCK, TCK, and TPK. The 33 scale items together represented teachers’ CALL competencies. To ensure the reliability of the 33 Likert scale items, Cronbach’s (1951) coefficient alpha was calculated, and a value of .96 was obtained, indicating a highly reliable survey (Schmidt et al., 2009). The pre- and post-surveys were analyzed through the paired sample t-test.

Reflective essays, Class observation notes, Group discussions, and Lesson plans

Reflective essays, class observation logs, group discussion logs, and lesson plans complemented the TPACK survey data by providing qualitative data showing how the participants developed CALL competencies throughout PBL (Research question 2). An inductive approach (Creswell, 2013) was used to analyze the data, which included (1) organizing and reading through data, (2) coding the data, (3) generating themes, (4) interrelating the themes, and (5) interpreting the themes. The researchers started the analysis by reading through the reflective essays to discover emerging themes about the participants’ class observations, group discussions, and the design of lessons. While reading the reflective essays, the researchers tagged interesting passages with one of four themes relevant to the research question 2: (1) teaching problems in English classes, (2) teaching topics in CALL lesson plans, (3) online English teaching tools, and (4) English teaching activities. The researchers then used the themes that emerged from the reflective essays to code the class observation notes, group discussions, and lesson plans and generated revised themes, which were broad enough to cover the ideas of the coded passages. The researchers arranged the themes in a network to visually represent their interrelationships. The researchers then interpreted the themes by drawing inferences, constructing explanations, and suggesting conclusions, which are presented in the sections of results and discussion.

Results

Teachers’ progress in their CALL competencies after the PBL project

A paired-samples t-test was performed to examine participants’ CALL competencies before and after the PBL project. The independent variable was the PBL project. The means and standard deviations for the prospective teachers’ CALL competencies are shown in Table 2, which indicates that there were significant pre and post differences in TK (10) = -3.35, CK (10) = -2.86, PCK (10) = -2.58, TCK (10) = -3.19, TPK (10) = -4.56, and overall CALL competencies (10) = -3.61. However, there was no significant difference in PK (10) = -1.32, η² = .40, although the participants scored higher in PK in the post survey. The effect size (Eta Square, η²) was calculated to examine the magnitude of difference when a significant difference was observed. Based on Cohen (1992), the effect size of .20, .50, and .80 denote small, medium, and large effect sizes respectively. The results showed a large effect size in the TK, = 1.01, CK, η² = .86, PCK, η² = .78, TCK, η² = .96, TPK, η² = 1.37, and overall CALL competencies, η² = 1.09. Based on the descriptive and paired-sample t-test results, the researchers concluded that the prospective teachers significantly improved their CALL competencies overall and in TK, CK, PCK, TCK, and TPK after the PBL project.
### Table 2

<table>
<thead>
<tr>
<th>CALL competencies</th>
<th>Pre Mean</th>
<th>Pre SD</th>
<th>Post Mean</th>
<th>Post SD</th>
<th>t</th>
<th>df</th>
<th>Sig.</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>TK</td>
<td>3.95</td>
<td>1.12</td>
<td>5.01</td>
<td>.89</td>
<td>-3.35</td>
<td>10</td>
<td>.007</td>
<td>1.01</td>
</tr>
<tr>
<td>CK</td>
<td>4.35</td>
<td>1.11</td>
<td>5.56</td>
<td>.77</td>
<td>-2.86</td>
<td>10</td>
<td>.017</td>
<td>.86</td>
</tr>
<tr>
<td>PK</td>
<td>4.66</td>
<td>.85</td>
<td>5.07</td>
<td>.76</td>
<td>-1.32</td>
<td>10</td>
<td>.215</td>
<td>.40</td>
</tr>
<tr>
<td>PCK</td>
<td>4.67</td>
<td>.98</td>
<td>5.34</td>
<td>.69</td>
<td>-2.58</td>
<td>10</td>
<td>.027</td>
<td>.78</td>
</tr>
<tr>
<td>TCK</td>
<td>4.31</td>
<td>.87</td>
<td>5.30</td>
<td>.77</td>
<td>-3.19</td>
<td>10</td>
<td>.010</td>
<td>.96</td>
</tr>
<tr>
<td>TPK</td>
<td>4.16</td>
<td>.97</td>
<td>5.35</td>
<td>.80</td>
<td>-4.56</td>
<td>10</td>
<td>.001</td>
<td>1.37</td>
</tr>
<tr>
<td>CALL</td>
<td>4.34</td>
<td>.71</td>
<td>5.10</td>
<td>.65</td>
<td>-3.61</td>
<td>10</td>
<td>.005</td>
<td>1.09</td>
</tr>
</tbody>
</table>

### Developing CALL competencies through PBL

An 18-week PBL project was implemented to improve 12 prospective English teachers’ CALL competencies based on the three features of PBL: driving questions, collaboration, and artifact development. This section describes how the prospective teachers developed CALL competencies through the PBL project.

#### Generating driving questions from class observations

The class observations were implemented to enable the prospective teachers to generate focused driving questions by witnessing actual teaching problems that arose in English classes. The class observation notes show that two essential teaching problems were observed by the prospective teachers. The first was that students of different levels of language proficiency were in the same class, making it difficult to meet the learning needs of each individual student. Several prospective teachers observed that nearly one third of the students were low-achievers, who shared that they sometimes could not understand what teachers were saying in English or their English textbooks. The second problem was that many students had low levels of motivation and participation. The prospective teachers observed that some students were completely disengaged and talking among themselves when they were supposed to be participating in activities because these activities consisted largely of mechanical drills that did not allow students to use their creativity to practice what they learned. Based on the two teaching problems, the prospective teachers proposed two driving questions: “How can teachers make teaching materials comprehensible for students of different language proficiency levels?” and “How can teachers enhance students’ learning motivation?”

#### Developing artifacts through CALL lesson plan design

The prospective teachers worked in groups to design CALL lesson plans as final projects to address the two driving questions mentioned above. The design of CALL lesson plans comprised the procedures of (a) identifying topics of CALL lesson plans, (b) selecting online English teaching tools, and (c) designing teaching activities based on the selected online teaching tools. Three CALL lesson plans, one by each group, were developed (see Table 3). The three CALL lesson plans and prospective teachers’ group discussion records were analyzed to determine how the prospective teachers went through the procedures of CALL lesson plan design.

### Table 3. Three CALL lesson plans

<table>
<thead>
<tr>
<th>Topics</th>
<th>Teaching objectives</th>
<th>Activities</th>
<th>Online teaching tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classroom</td>
<td>Reading</td>
<td>Memory games</td>
<td>Popplet</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Crossword puzzles</td>
<td>YouTube</td>
</tr>
<tr>
<td>Storymaking</td>
<td>Writing</td>
<td>Story writing</td>
<td>BBC-School Radio</td>
</tr>
<tr>
<td></td>
<td>Speaking</td>
<td>Lyrics writing</td>
<td>Pinky Dinky Doo</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nursery rhymes singing</td>
<td>Story Starter</td>
</tr>
<tr>
<td>SuperHero</td>
<td>Reading</td>
<td>Vocabulary recognition</td>
<td>PBSKids</td>
</tr>
<tr>
<td></td>
<td>Listening</td>
<td>Story listening</td>
<td></td>
</tr>
</tbody>
</table>
Identifying topics of CALL lesson plans based on the driving questions

The prospective teachers determined the topics of their CALL lesson plans based on the driving question of how to enhance students’ learning motivations. To solve the driving question, the prospective teachers proposed that teaching topics of CALL lesson plans should be designed in relation to students’ life experiences. For example, the prospective teachers in Group A suggested that “classroom” was an interesting topic for students, because the classroom was a space in which students spent a good part of their lives and their classroom experiences gave them vocabulary to practice in daily life. Similarly, Group B’s discussion records show that to make lesson plan topics relevant to students’ life experiences, the prospective teachers intended to allow students to choose topics of their own. The two discussion records demonstrated that the driving questions derived from observing English classes were latter used by the prospective teachers as references to identify the topics of CALL lesson plans.

Lisa: I think the topic about classroom is ok.
Sean: I think we can talk about supermarkets.
Cindy: Why do you want to teach the topic of classroom?
Lisa: Because a classroom is a place where the students spend time almost every day.
Sean: And they can use the everyday vocabulary they learn about classrooms.
Cindy: So the topic is close to their daily life! I love the idea. So we all agree with the topic?
Sean: Yes (Group A discussion records)

Nina: Any ideas about lesson plan design?
Jack: How about story making? We can ask students to create stories.
Nina: Good idea. But what are the possible topics?
Alice: I think we can allow students to choose their own topics? Students may feel more motivated in writing stories on their own.
Jack: I agree. We can also list three to five topics for less proficient students to choose if they have no ideas. So those students will not feel overwhelmed (Group B discussion records).

Selecting online English teaching tools

After identifying CALL lesson plan topics, the prospective teachers proceeded to the selection of online tools to use in CALL lesson plans. A total of six online tools, including Popplet, YouTube, BBC-School Radio, Pinky Dinky Doo, Story Starter, and PBSkids, were incorporated by the prospective teachers into CALL lesson plans. These teaching tools were purposefully selected to address the driving question of how to make learning materials comprehensible for students at different language proficiency levels. Therefore, the selection criteria for the online tools called for a multimedia environment in which students could comprehend texts through contextual information including visual and audio cues. In the selection process, the prospective teachers found that many open educational websites and tools with kid-friendly visual and audio components were available online to support the learning of English and other subjects. However, they also pointed out that selecting appropriate educational tools from the Internet was not a simple task but rather required them to evaluate the pros and cons of the educational tools based on the value of the activities for their teaching objectives and the tools’ appropriateness for a particular group of students. Group A’s discussion records, for example, demonstrated how the prospective teachers evaluated Pinky Dinky Doo and Story Starters for their CALL lesson plan. Nina asked Tim to evaluate which of the two storytelling websites would be better for their lesson plan. Tim preferred Pinky Dinky Doo, because its content and activities were easier than Story Starters for students whose native language was not English. However, Nina argued that Pinky Dinky Doo was not a good choice for it did not allow students to practice oral skills, which was one of the teaching objectives in their lesson plan. She suggested that they could have students practice both writing and speaking skills by asking them to write lyrics for a song based on the stories they created on Pinky Dinky Doo. As another example, Group C discussed the advantages of the use of Popplet and YouTube in English class. These examples revealed that the selection of online tools engaged the prospective teachers in analyzing and evaluating online educational resources, which improved their technological knowledge. The next section will describe how the prospective teachers integrated these online tools into their teaching activities.

Tim: The first one, Pinky Dinky Doo.
Nina: Yes, why do you want to choose Pinky Dinky Doo?
Tim: I think story-starters might be a little bit too hard for the kids. Because English is not our native language, Pink Dinky is easier for students. Pinky Doo is an exciting game. There are lots of fun stories to choose like horror stories, silly stories, fairy stories, etc. Students can complete the story by choosing different characters, background, etc. Another interesting thing about it is that you can learn new vocabularies. Pinky Dinky would give students the meaning of vocabulary, and read the text for students.
Nina: Do you think Pinky Dinky meets our objectives? I think we need to think about our objectives first. Basically, we want to motivate students to speak more in class. I think we can still use Pinky Dinky. But we need to change it a little. Maybe we can create a song without lyrics. Students need to write the lyrics for the song based on their stories. (Group B discussion records)

Cindy: I found a concept map tool, http://www.popplet.com/ What do you think?
Sean: This website is good. It can create a map that can help us make an outline. So students can know what they are going to learn in this class.
Lisa: We can describe our outline before we teach vocabulary. In addition, I also think that we can look for some songs that related to our teaching theme on YouTube. Music can draw students’ attention and establish a pleasant learning environment.
Sean: That’s right. We can find lots of songs or videos on YouTube.
Lisa: I think learning by songs can be more interesting. And we can sing with students (Group A discussion records)

Designing teaching activates based on the selected online tools

Based on the online tools they selected, the prospective teachers designed such activities as crossword puzzles, story writing, and nursery rhyme singing to help students improve English reading, listening, speaking, or writing skills. They used a concept map tool and YouTube links to combine vocabulary learning with a game. They created a map of themes, each of which was linked to a related YouTube video that gave clues to the words needed to fill in a crossword puzzle. For example, as shown in Figure 3, when students clicked on the theme “clothes,” they would be linked to a YouTube video featuring a clothes song by Peter Weatherall, to which they would listen for clues to the adjacent crossword puzzle. After completing the activity, students would be expected to have learned the meanings, pronunciations, and spelling of the target vocabulary.

Another group used Pinky Dinky Doo (see http://www.pinkydinkydoo.com/storybox.html) to improve students’ writing skills. As Figure 4 shows, Pinky Dinky Doo, for example, provides students with six story genres. When they click on one, they find a blank story frame with five pictures on the left side and a short text below in which some vocabulary and phrases are left out. Students are asked to fill in each blank in the story by clicking on the appropriate picture on the left. As students fill out the story, they click on “read aloud” to listen to the story and watch the animations in order to move to the next part. After all the students have finished the Pinky Dinky Doo activities, they are given worksheets to write lyrics for a song based on the stories they made in Pinky Dinky Doo.
Discussion

The purpose of this study was to investigate how prospective teachers’ CALL competencies could be improved through participating in a PBL project. The results support previous research (Liu & Kleinsasser, 2015) which showed that PBL projects can help teachers develop CALL competencies. In addition, this study extended the existing literature by applying Mishra and Koehler’s (2006) TPACK framework to explore whether participation in a PBL project improved CALL competencies in the specific areas of TK, CK, PK, PCK, TCK, and TPK. The survey results indicated that the prospective teachers demonstrated significant improvement in all areas except PK, although post-survey scores were higher for this component too. PK refers to knowledge about learning theories and the design of teaching practices based on specific learning theories (Mishra & Koehler, 2006). The less significant improvement in PK might be explained by prospective teachers’ attention to particular theories suitable to their lesson plans rather than to a broad range of learning theories. Thus they showed some PK gains, but these reflected their knowledge of only the particular theories with which they adopted in this project.

Apart from the pre- and post-surveys, the qualitative data, including class observation notes, group discussion transcripts, lesson plans, and reflective essays, which provided rich description of prospective teachers’ actual processes of developing CALL competencies, were valuable complements to the self-reported TPACK survey results. Class observations were found to provide a crucial foundation for the development of CALL competencies through involvement in PBL projects. Being exposed to a situated learning environment stimulated the prospective teachers to generate driving questions and critically reflect on authentic teaching problems. The driving questions derived from observing real classrooms gave the prospective teachers a starting point for developing creative and student-centered ideas for integrating technology into language teaching. Liu and Kleinsasser (2015) found that in previous CALL training programs, teachers often confined themselves to the traditional use of technology by asking students to recite the content from online websites. The driving questions concerning low English learning motivation and multilevel classrooms stimulated the prospective teachers in this study to use the audio and visual features of online websites to keep students motivated to learn, and make learning contents personalized and comprehensible for students of different language proficient levels.

As the culminating activity of the project, the design of CALL lesson plans required the prospective teachers to explore online tools and integrate selected tools into teaching activities, which provided hands-on experiences that involved combining pedagogical, content, and technical knowledge required for CALL competencies. For example, echoing Vanderplank (2010), who studied applications of technology in language teaching, this study showed that selecting the right Internet resources to integrate into teaching was the major challenge facing the prospective teachers. Conflicts arose regarding individual preferences for specific technologies to use in teaching, and these forced the prospective teachers to debate the extent to which particular technologies met teaching objectives and could be adapted into particular teaching contexts. This process led the prospective teachers to increase their technological and instructional knowledge so they could articulate their reasons for claiming the effectiveness and viability of particular technologies in language teaching.

Although this study demonstrated the extent to which and how teachers developed or improved CALL competencies through the PBL pre-teacher training, there are some limitations. First, the prospective teachers did not implement the lesson plans they developed in real classrooms. Future studies might explore how pre-service or in-service teachers use their lesson plans in real teaching contexts, and ask them to reflect upon the effectiveness of their lesson plans. Second, this study did not require the prospective teachers to compose a CALL lesson plan before the PBL project for the purpose of pre- and post-project comparison, because the participants were the prospective teachers who did not have sufficient knowledge to design lesson plans before
the PBL project. Future studies may compare the lessons developed by the in-service teachers before and after the PBL to investigate changes in CALL competencies.

Conclusion

This study contributes to the CALL teacher education knowledge base by proposing a PBL project based on Krajcik and Blumenfeld’s (2006) elements of PBL to improve EFL teachers’ CALL competencies. The PBL project involved class observations, group discussions, and the design of lesson plans. The results showed that the prospective teachers demonstrated higher levels of CALL competencies after the PBL project developed for this study, which can provide a model for developing both pre-service and in-service EFL teachers’ CALL competencies.

In addition, previous research relied solely on self-reported survey data to investigate the impact of PBL upon teachers’ CALL competencies. This study extended the existing literature by qualitatively exploring additional data sources to investigate the processes by which teachers developed CALL competencies, including class observation notes, group discussion records, and lesson plans. Using these concrete artifacts, this study explicitly illustrated the benefits which prospective teachers may obtain and the problems they may face when participating in a PBL project. The findings can inform teacher educators’ efforts to design and implement effective teaching training for CALL competency development.

Acknowledgments

This research was supported by the Ministry of Science and Technology of Taiwan (MOST 104-2511-S-224-002-MY3). This support makes the implementation of the project possible.

Reference


Using a Bilingual Concordancer for Text Revisions in EFL Writing

Yu-Fen Yang¹, Ruey-Fen Harn¹ and Gwo-Haur Hwang²

¹School of Applied Foreign Languages, National Yunlin University of Science & Technology, Taiwan // ²Program in Interdisciplinary Studies, National Yunlin University of Science & Technology, Taiwan // yangy@yuntech.edu.tw // g9841705@yuntech.edu.tw // ghhwang@yuntech.edu.tw

*Corresponding author

(Submitted January 28, 2018; Revised May 27, 2018; Accepted June 19, 2018)

ABSTRACT

Many students who study English as a foreign language (EFL) have difficulty transferring linguistic features between languages while writing. This study aims to investigate how college students improve their EFL writing with text revisions by using a bilingual concordancer (Chinese and English). A sample of 32 college students consented to participate in a writing program. According to pre-test scores, they were grouped into an experimental (N = 15) group, which used the bilingual concordancer, and a control (N = 17) group, which used Yahoo’s Chinese-English online dictionary. The results of the study indicate that the experimental group made greater improvement in text revision than those in the control group, as they raised metalinguistic awareness to evaluate word choice and sentence construction, by viewing authentic language examples retrieved from the bilingual concordancer. The bilingual concordancer also provided the experimental group with rich linguistic contexts to compare the differences between two languages and induce the rules of the target language (English), which lead to a reduction in errors due to the interference of the native language in EFL writing. In contrast, the control group relied mostly on limited sample sentences listed in the online dictionary, and lacking contexts resulted in little writing improvement.

Keywords

Authentic language examples, Bilingual concordancer, Language interference, Metalinguistic awareness, Text revision

Introduction

Students’ using their native language in organizing an essay is viewed as one effective writing strategy (Kim & Yoon, 2014; Manchón, Roca de Larios, & Murphy, 2009; Manchón & Roca de Larios, 2007) as students often find it difficult to comprehend a target language when the language principles are different from those of their native language. When students encounter difficulties, some may resort to their native language to apply rules and structures to their second or foreign language learning (Harun, Massari & Puteh Behak, 2014; Storch & Wigglesworth, 2003; Payant & Kim, 2015). As a result, some interference may occur from the application of their native language in various areas of linguistic components including phonology, morphology, grammar, syntax, lexis and semantics (Hohenstein, Eisenberg & Naigles, 2006), leading to errors like misuse of word choice, incorrect sentence construction, and logical order.

To overcome learning difficulties, college students who are studying English as a foreign language (EFL) often use bilingual dictionaries to confirm word meanings, choice, or use. But language examples provided by bilingual dictionaries “may not be sufficient for students to learn word meaning and usage, and may even lead to inappropriate word use” (Lai & Chen, 2015, p. 341). As a consequence, students often write incorrect sentences and select inappropriate words when engaged in EFL writing. This is usually the case, because bilingual dictionaries do not include many “collocations, sentence constructions, or metaphorical interpretations” (Gao, 2011, p. 256).

To further support EFL writing, many studies have proposed that a bilingual concordancer may improve learning at both word and sentence levels or even aid in writing proficiency (Gao, 2011; Lai & Chen, 2015) as the bilingual examples from collections of articles with mutual translations provided by the bilingual concordancer may facilitate students to apply prior knowledge to examine linguistic resources and revise writing errors to construct new language knowledge, and therefore help EFL students with their writing. Yeh, Liou, and Li (2007), however, indicate that students may “spend so much time searching for appropriate lexical items but still have difficulty expressing themselves precisely” (p. 132). Even with the assistance of a bilingual concordancer, teachers still have to guide their EFL students on how to use a bilingual concordancer with the engagement of metalinguistic awareness to express their thoughts appropriately in written communication (Yang, Wong & Yeh, 2013).
A bilingual concordancer with metalinguistic awareness to improve EFL writing

A bilingual concordancer provides a large number of language examples, with keywords displayed in contexts, and allows students to query in both the source and the foreign language. While making a query, students can benefit from the teacher’s guidance in fostering students to integrate their previous and emerging knowledge to monitor and evaluate unfamiliar expressions in a foreign language. For example, a teacher can guide his/her students to query unfamiliar words or phrases encountered in the process of reading or writing a foreign language text, and help his/her students to understand the meanings by evaluating authentic language examples and translations, so as to enhance the transfer of word knowledge (Kaur & Hegeleimer, 2007), self-correction of expression (Gaskell & Cobb, 2004; O’Sullivan & Chambers, 2006), and the transfer of language knowledge to academic writing tasks. A bilingual concordancer may help to improve students’ EFL writing, but most students need to be trained to orchestrate their metalinguistic awareness when using the concordancer to get the maximum benefit.

When students use a bilingual concordancer to examine their word choice or sentence construction, they are encouraged to monitor and evaluate the related language examples for revising errors (Gao, 2011). In writing processes, writing and revision are regarded as rational negotiations between writers’ thoughts and their emerging texts. To analyse linguistic rules, structures, and functions for writing meaningful sentences, it is important for students to raise their metalinguistic awareness. Jessner (2008) defined metalinguistic awareness as “the ability to focus on linguistic form, to categorize words into parts of speech, to switch focus between form, function, and meaning, explain why a word has a particular function” (p. 277). Roehr (2008) and Szerencsi (2010) also proposed that a student’s language proficiency, in their target language, was highly correlated with metalinguistic knowledge that could be manipulated to evaluate his/her language production and learning aptitude. In EFL writing, metalinguistic awareness plays an important role in helping students examine their writing process. When they encounter difficulties in terms of different linguistic features and structures, the metalinguistic process is conductive to thinking about what one already knows, monitoring emerging understanding, applying strategies, evaluating the effectiveness of these strategies, and revising strategies (Baker, Gersten, & Scanlon, 2002). A bilingual concordancer functions to raise EFL students’ metalinguistic awareness as students are able to reflect on the language features by examining the connections between words or sentence structures in two languages to respond to a particular query posed either by the teacher or students themselves (Gao, 2011). Many researchers have employed bilingual corpora to help students study grammatical patterns from authentic texts (e.g., Chujo, Anthony, Oghigian & Uchibori, 2012; Francis, 2012; Gao, 2011; Liou & Chan, 2005; Wu, Chen, Chang & Chang, 2017).

Background of this study

This study aimed to investigate how the students improved their EFL writing using a bilingual concordancer or an online bilingual dictionary with their engagement in metalinguistic awareness in the process of text revision. Text revision can “increase student engagement and attention to detect and correct writing problems” (Ferris, 2003, p. 52), and then elevates writing proficiency (Ferris & Roberts, 2001). When using a bilingual concordancer, an online Chinese-English dictionary, or the teacher’s guidance, students may improve EFL writing which features word choice, sentence patterns, and logical structures in text organization, from translating native (Chinese) to a foreign (English) language. A bilingual concordancer that retrieves authentic language examples from Taiwan Panorama concerning Taiwan’s political, economic, social, and cultural evolution with insightful views was authorized and developed in this study based on the data-driven learning (DDL) approach (Johns, 1991), emphasizing the consultation of authentic corpus examples and the discovery of rules by induction (Gao, 2011). The concordances, retrieved from the corpus that contains 8,594,904 English words and 5,757,930 Chinese words, are aligned at a paragraph level manually (Yang, Wong & Yeh, 2013).

The bilingual concordancer “enables users to search for language patterns in a corpus” (Sun, 2007, p. 324) and to evaluate whether the language examples containing the keywords used are helpful to their text revision and then decide what the revising strategies are. In this study, metalinguistic awareness in the use of a bilingual concordancer refers to students’ abilities to plan strategies, monitor learning processes, evaluate the effectiveness of the strategy, and revise strategies. EFL students are encouraged to self-correct and explain their errors in both their first language (Chinese) and foreign language (English). In searching for Chinese language examples and their English translation, entering Chinese keywords is required. Students are encouraged to look for the keywords in the context (Chinese) that helps express his/her thoughts in the target language (English).
In Figure 1, two Chinese keywords are submitted, for example, “看到” (see/saw) and “追” (chase/chase after), as an entry to search for English examples on how the keywords relate to their adjacent words in different contexts. In this case, there are 20 examples retrieved from the corpus. The bilingual concordancer allows a multiple-word entry in both Chinese and English interface in order to retrieve more examples with contexts for the students to check how the keyword or phrases are used in various contexts as well as compare the difference between two languages in the same context. As shown in Figure 1, the keywords are marked in the context of the Chinese language example in the first row. Following this, it is the English translation for the Chinese example, in which the student can look for the English equivalent for the keywords, also marked, in the context of the English language example. All useful language examples can be saved for later reference.

To investigate how college students improve EFL writing with text revision by using a bilingual concordancer or an online bilingual dictionary, three research questions were addressed: (1) What writing progress is made in text revision between the two groups?; and (2) How do students operate their metalinguistic processes in text revision for writing improvement?; and (3) What are students’ perceptions toward using a bilingual concordancer and an online bilingual dictionary for text revision in EFL writing?

Method

Participants

Thirty-two participants studying English as a foreign language (EFL) from different departments and colleges of a university in central Taiwan volunteered to participate in a writing program, an elective course. Before entering the writing program, the students were asked to take a text-structured pre-test to evaluate their ability to comprehend texts and develop a logical flow of main and supporting ideas in a text as well as to master content words such as phrases and transition words. Selected from the Testing Center for Technological and Vocational Education in Taiwan, the text-structured pre-test consisted of two parts, a topic sentence part and a supporting part (see Appendix A). There were 10 test items in each part and the maximum score on the test was 100. The reliability of the test was .94 (see http://www.tcte.edu.tw). According to the text-structured pre-test, the 32 students were grouped into experimental and control groups with similar mean scores. There were 15 students in the experimental group and 17 students in the control group. The mean score and standard deviation on the pre-test for the experimental group was 36.32 and 15.54, while the mean score and standard deviation for the control group was 39.17 and 13.20. A paired-sample t-test was employed and represented no significant difference between the two groups’ English language proficiency (t(31) = 3.75, p > .05). The Levene’s test for equality of
variances was also calculated (Levene Statistic= 4.01; p = .32). The result indicated that there was no difference between the variances in the experimental and the control groups.

**Procedures of data collection**

Writing instruction from the same teacher was conducted over a period of 12 weeks (2 hours per week) from March 11th, 2017 to June 10th, 2017. First, the students were categorized into the experimental and the control groups, according to their scores on the text-structured pre-test in week 1. After that, both groups were required to complete three writing tasks during the following weeks, entitled “One good deed” (week 2-4), “I had a dream” (week 5-7), and “Listen to me” (week 8-10). The length of each essay was from 150 to 200 words. Before the end of the writing instruction, the students were asked to take a text-structured post-test (week 11) with different version from the pre-test to examine their learning progress after using the bilingual concordancer or the online bilingual dictionary for text revisions during the writing instruction. Moreover, the students’ perspectives on using the bilingual concordancer and Yahoo’s Chinese-English online dictionary to improve text revisions were also reported in the two versions of the open-ended questionnaires (week 12), adapted from Yoon & Hirvela (2004) and administered to each of the two classes during a regular class period. Instructions on how the bilingual concordancer and online bilingual dictionary work for text revisions were given throughout the writing program. The keywords used to search the language examples from the bilingual concordancer as well as the online dictionary were documented.

For the experimental group’s first writing task, the teacher provided scaffolding prompts (Appendix B) and modeled the students on how to use the bilingual concordancer by entering keywords for example sentences in contexts. Next, the teacher illustrated the students’ errors on their first draft “One good deed,” and instructed them on how to enter keywords to improve their writing. The students were also encouraged to practice entering keywords in the bilingual concordancer to clarify any unfamiliar words and grammar use when revising their first drafts into their final ones. In task 2, the students would learn to search for appropriate words or expressions in different contexts by consulting the bilingual concordancer, they were asked to write their first draft of “I had a dream.” After they completed their first drafts, the teacher would then point out the writing errors for them to revise. They then individually practiced using the bilingual concordancer for text revisions, and were reminded again to write useful phrases from the sentence examples. The teacher only provided necessary instructional support on searching for keywords when needed. In task 3, the students wrote their first draft of “Listen to me.” The students revised their first drafts into final ones, using the bilingual concordancer to revise writing errors.

On the other hand, to accomplish the three writing tasks, the control group used Yahoo’s Chinese-English online dictionary which is one of the most frequently used online bilingual dictionaries in Taiwan and is free of access and functional as users look up a single word or a string of keywords in L1 or L2 for an equivalent to check its usage and sentence structure, confirm meaning, spelling, parts of speech, inflection and count ability or even find an alternative (Lai & Chen, 2015). In Task 1 and 2, based on the teacher’s feedback, the students revised their first drafts into final ones. They were required to take notes on each error correction and explain their decisions for accepting or rejecting their teacher’s suggestions. In Task 3, they were asked to correct their writing errors by themselves and revise their first drafts into final ones.

While the Chinese-English online dictionary offers definitions, and sometimes with sentences examples to show how a word can be used, the authentic sentence examples provided by the bilingual concordancer come in with paragraphs.

**Procedures of data analysis**

The data collected in this study included the text-structured pre- and post-tests, descriptive statistics of text revisions, log files of the students’ interaction with the bilingual concordancer and Yahoo’s Chinese-English dictionary, and two open-ended questionnaires. First, two paired-sample t-tests were administered to compare the students’ writing progress in both the experimental and control groups. Second, in the three writing tasks for text revisions, based on O’Sullivan and Chambers’ (2006) classification, the students’ writing errors were classified into three types, including (1) grammatical errors (incorrect use of nouns/adjectives/adverbs, verb tense/verb voice, and subject-verb agreement); (2) lexical errors (misuse of word choice), and (3) syntactic errors (incorrect sentence construction and logical order). For example, “I do many things for the social.” “social” would be identified as a grammatical errors in this context due to an incorrect use of adjectives. “They thought of many ways to care for the stray dog, like feeding them or playing with them.” “Care for” in this particular context
would be seen as a lexical error. “I talk with you was very loud” is an example of a syntactic error, showing an incorrect sentence construction.

The students’ error corrections from first to final drafts in tasks 1 to 3, were divided into correct and incorrect revisions, and were then calculated by descriptive statistics. When the errors identified in the first drafts were correctly revised in the final ones, they were regarded as correct revision; otherwise, they were regarded as incorrect revision. Third, to investigate any writing improvement between the experimental and the control groups after using the bilingual concordance and Yahoo’s Chinese-English online dictionary, two sample cases were randomly selected from the two groups as they underwent different metalinguistic processes. Finally, two trained raters, who had more than 10 years experiences in EFL teaching followed the steps of content analysis (Patton, 2002) to evaluate the students’ text revisions in their final drafts. They also examined the students’ perceptions about using the bilingual concordancer and Yahoo’s Chinese-English online dictionary in two versions of an open-ended questionnaire.

Four steps of content analysis were included in this study: coding, categorization, description, and interpretation. First, two raters coded the statements from the students’ text revisions recorded in the log files, as well as, their perspectives on using the bilingual concordancer and Yahoo’s Chinese-English online dictionary. Next, the students’ text revisions were categorized into three types, including grammatical, lexical and syntactic errors (O’Sullivan & Chambers, 2006). The students’ statements from the open-ended questionnaires were also categorized. The researchers described the statements by summarizing the main points. Finally, the researchers interpreted the main ideas by offering explanations, drawing conclusions, and making inferences. The inter-rater reliability for the students’ text revisions and open-ended questionnaires were .83 and .86 in this study. Disagreements between the raters and researchers were resolved through discussion.

Results

The results of this study are presented in three aspects. First, the students’ text revisions for EFL writing progress is shown by the paired-sample *t*-tests between the two groups from the pre- and post-tests, as well as, the descriptive statistics of the text revisions. Second, the log files of two randomly selected students, S1 from the control group and S10 from the experimental group, were examined in order to understand how the students used the bilingual concordancer and Yahoo’s Chinese-English online dictionary to reduce writing errors through the metalinguistic awareness. Finally, the students’ perspectives toward the bilingual concordancer and the online bilingual dictionary are reported.

Students’ text revision for writing improvement between the two groups

The answer to the first research question, “What writing progress is made in text revision between the two groups?” is illustrated as follow.

The two paired-sample *t*-tests were conducted to evaluate the two groups’ text revisions. For the experimental group, the mean scores and standard deviation increased from 36.32 and 15.54 on the pre-test to 53.33 and 27.31 on the post-test. A significant difference between the pre- and post-test scores are shown (*t* = -4.09, *p* < .01). A medium effect size was found (ES = 0.61). In contrast, the control group made a slight improvement in their writing, as seen in the mean scores and standard deviation which increased from 39.17 and 13.20 on the pre-test to 45.94 and 22.89 on the post-test. There was a significant difference between the pre- and post-tests (*t* = -2.42, *p* < .03). There was a small effect size (ES = 0.44) in text revisions for writing in the control group. The results reveal that the students in both the experimental and control groups improved their writing with text revisions, while the experimental group made greater writing improvement than those in the control group. Regarding the descriptive statistics between the two groups (see Table 1), grammatical errors (e.g., incorrect use of nouns/adjectives/adverbs, verb tense/verb voice, and subject-verb agreement), lexical errors (misuse of word choice) and syntactic errors (e.g., inappropriate word expression and incorrect sentence structure) were examined. The two groups made more grammatical errors than other types of errors; however, the students of the experimental group were able to make more lexical and syntactic revisions than the students in the control group. In particular, the students in the experimental group diagnosed lexical errors by themselves in corrective revisions which increased from 6 to 8 in tasks 1 to 3. Some of them also corrected syntactic errors in task 3 without the teacher’s instructional assistant. In contrast, the students in the control group relied mostly on the teacher’s scaffolding to detect grammatical errors. They revised 17 grammatical errors in corrective revisions.
with the teacher’s help in task 1; however, most of them were unable to revise grammatical errors in task 3 without the teacher’s instructional support, and they only diagnosed three errors in corrective revisions.

<table>
<thead>
<tr>
<th>Writing errors</th>
<th>Correct revisions</th>
<th>Incorrect revisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical errors</td>
<td>Task1</td>
<td>Task2</td>
</tr>
<tr>
<td>12</td>
<td>18</td>
<td>21</td>
</tr>
<tr>
<td>Lexical errors</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Syntactic errors</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1. Text revisions between the two groups

<table>
<thead>
<tr>
<th>Writing errors</th>
<th>Correct revisions</th>
<th>Incorrect revisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical errors</td>
<td>Task1</td>
<td>Task2</td>
</tr>
<tr>
<td>17</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>Lexical errors</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Syntactic errors</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Students’ text revision to improve EFL writing with metalinguistic awareness

To answer the second research question, “How do students operate their metalinguistic processes in text revision for writing improvement?” two sample cases, S10 from the experimental group and S1 from the control group, were selected by simple random sampling. Each student was given a random number with the same probability of being chosen from the experimental and control groups to compare their text revisions through metalinguistic process. S10 reduced his writing errors, improving his writing quality when he actively examined his revising process. S1, however, displayed little error correction ability, and mostly relied on the teacher’s instructional support.

S10’s text revision to improve writing in metalinguistic process

In Task 1, S10 received instruction from the teacher on using the bilingual concordancer to get to know how text revision in EFL writing can be done by observation and investigation on the language examples. And in Task 2, he started to revise his first drafts to final ones based on the teacher’s suggestions. He planned and carried out the search strategies for text revisions. His revision process revealed in Table 2, in the first draft, syntactic errors such as “I do many things for the social, like public service activities,” S10 typed in the search strings Chinese phrases, such as “社會”(society), “公益活動”(public service activities) or “慈善事業”(charities) in the bilingual concordancer. For the second sentence, “I accomplish my dream of a childhood,” S10 made a search on “兒時的”(of childhood). For the third sentence, lexical errors such as “After traveling, I come to home, and I am so tired,” he searched for sentence examples to revise the phrase, “come to home” by typing in the Chinese search string, “回家”(return home).

S10’s final draft appears to be more coherent and logical, indicating his ability to check incorrect word expressions and sentence structures and to choose appropriate words and phrases for expressing his thoughts in writing. In view of metalinguistic awareness, S10 monitored his own learning processes to reduce writing errors. Table 2 shows S10’s writing process for task 2, demonstrating how the students in the experimental group underwent metalinguistic processes by monitoring language examples with the help of the bilingual concordancer and by evaluating the teacher’s suggestions to improve text revisions. It may seem challenging for him to completely understand the English language examples retrieved by the concordancer. Nevertheless, with careful reading the bilingual language examples and the assistance from the teacher, S10’s final draft shows both fluency and accuracy after his undergoing metalinguistic processes.

Table 2. S10’s revising process in task 2

<table>
<thead>
<tr>
<th>Phase</th>
<th>Revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>First draft</td>
<td>I do many things for the social, like public service activities. Also I accomplish my dream of a childhood. ......After traveling, I come to home, and I am so tired.</td>
</tr>
<tr>
<td>Suggestions from the teacher</td>
<td>For revising “I do many things for the social,” please make searches in on “社會”(society), “慈善事業”(charities).</td>
</tr>
</tbody>
</table>
For revising, "my dream of a childhood," please make a search on "兒時的" (of a childhood).
For revising the grammatical error, "come to home," please make a search on "回家" (go home). Please pay attention to the tenses for some of the sentences.

1. "the social" couldn’t be found in any of the sentence examples, but I found "the society."
2. "engaged in charitable activities" seems to be a better way to express what I want to say.
3. I couldn’t find "go to home" in the concordances; however, I can say "went home" or "returned home."

I engaged in charitable activities for the society, like public service activities. Also I accomplished a childhood dream of mine……After traveling, I returned home, and I was so tired.

Table 3 shows S10’s text revisions in task 3, using the bilingual concordancer without the teacher’s instructional support. In the first draft, it is obvious that S10’s English sentence structure is under the influence of her native language, Chinese. A few errors are identified as followed: (1) a syntactic error, "I talk with you was very loud"; (2) a syntactic error and lexical errors, "Because last week I force important subject’s midterm"; (3) a syntactic error and two grammatical errors, "So when I return home. I wanted to prepare for study this subject immediately." S10 searched the bilingual concordancer for words and phrases in contexts to revise his writing errors. The remarks he made during the revising process provides evidence that he detected the differences between the Chinese and English language examples, deduced appropriate expressions, and then finally reconstructed his sentences in accordance with English sentence patterns.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>First draft</td>
<td>I talk with you was very loud. Because last week I forced important subject’s midterm. The test content was very hard. So when I return home. I wanted to prepare for study this subject immediately.</td>
</tr>
<tr>
<td>Errors pointed out by the teacher</td>
<td>1. I talk with you was very loud.</td>
</tr>
<tr>
<td>S10’s remarks on the error revision</td>
<td>2. Because last week I force important subject’s midterm.</td>
</tr>
<tr>
<td></td>
<td>3. So when I return home. I wanted to prepare for study this subject immediately.</td>
</tr>
<tr>
<td>Final draft</td>
<td>I talked too loudly to you. Last week I faced important subject midterm. The test content was very hard. When I returned home, I wanted to prepare for the midterm immediately.</td>
</tr>
</tbody>
</table>

Table 4 shows how the students in the experimental group would query the words in the bilingual concordancer, a strategy leading to language examples, to improve his writing in task 3. First, S10 queried a keyword “大聲說” meaning raising voice or speaking out loud as the first sentence in the first draft lacks the clear predicate relationship in sentence construction, an syntactical error pointed out by the teacher. The notes taken during his revising process and saved in the online notebook indicate S10 was able to correct the sentences by observing correct English sentence construction in paragraphs. That is, the keyword in contexts in the language examples could be advantageous for his induction and deduction of grammatical rules and word usage. In this case, he observed the contexts containing the keywords and discovered that “speaking out loud” or “announced loudly” could not exactly express what he meant. As he continued to examine the rest of the sentence examples, he found a phrase “大聲說,” meaning “too loudly,” S10 identified and extracted the examples that were appropriate to express his thoughts in revising his final draft. Although there are still some errors in S10’s revisions, the final draft was markedly different from the first draft, the metalinguistic process which S10 was engaged in while using the bilingual concordancer prompted him to monitor and evaluate the word choice and sentence patterns, allowing him to recognize his writing problems and revise his sentences into a logical flow.
Table 4. Student 10’s records of searching in Task 3

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Sentences</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>大聲說</td>
<td>2017/6/7 Hou hates to bother people (to the point that he gets embarrassed if his stereo is turned up too loudly). But an aimless year in Taiwan and a depressing 24th birthday gave the impetus to try busking in Ximending to break out of his rut.</td>
<td></td>
</tr>
</tbody>
</table>

Table 5. S1’s revising process in task 2

<table>
<thead>
<tr>
<th>Phase</th>
<th>Revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>First draft</td>
<td>I found myself on corridor of elementary school. As soon as I saw they chase me, I scream and run as fast as possible. Amazingly, their speed was very fast to me to run.</td>
</tr>
</tbody>
</table>
| Suggestions from the teacher | 1. the corridor  
2. an elementary school  
3. As soon as I saw them chasing after me  
4. screamed and ran  
5. Amazingly, they were too fast. I could not escape. |
| Keywords for searches in Yahoo online bilingual dictionary | 1. corridor  
2. elementary  
3. chase |
| Remarks on Teacher’s suggestions | 1. For “corridor,” S10 noted down “Article + Noun.”  
2. For “elementary,” S10 noted down “a/an (an indefinite article),” used before a word that begins with a vowel.  
3. No remark on “chase” |
| Final draft        | I found myself on the corridor of an elementary school. As soon as I saw them chasing after me, I screamed and ran as fast as possible. Amazingly, they were too fast. I could not escape. |

In task 3, S1 revised the errors identified by her teacher in the first draft (Table 6). She managed to revise most grammatical and lexical errors. As for the syntactical errors, she seemed to have trouble finding answers by only...
consulting the online bilingual dictionary. One of the possible reasons could be the inadequacy of the language examples provided by the dictionary. Another may result from the number of words in a search string allowed to be entered as an entry in the dictionary. To be specific, she cannot use multiple words or phrases as a search string to get language examples from the online dictionary. Consequently, she was not able to find the language examples that meet her needs. Syntactical errors usually involve multiple words or phrases. Under the circumstance, it is understandable that she revised “Two weeks ago, you were asking me” into “Two weeks ago, you were asking me” instead of “Two weeks ago, you asked me.” Another example, the phrase, “Originally, I said, “O.K.,” but I regretted when last Wednesday,” was translated into English from Chinese word for word. First, language influenced S1’s metalinguistic process in constructing the clause. Without a context, she could not observe how words were related to one another, nor the formation of sentences in a specific context between Chinese and English. These findings indicate that the online bilingual dictionary alone could not help EFL students find English equivalents for their Chinese concepts when their EFL writing is under the influence of their native language since they did not receive a large number of language examples and translations to help monitor their writing problems in word choice and sentence construction.

Table 6. S1’s revising process in task 3

<table>
<thead>
<tr>
<th>Phase</th>
<th>Revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>First draft</td>
<td>Dear Judy,</td>
</tr>
<tr>
<td></td>
<td>I apologize for having a quarrel with you last week.</td>
</tr>
<tr>
<td></td>
<td>Two weeks ago, you were asking me, “Do you want to go to Tainan with me next week?”</td>
</tr>
<tr>
<td></td>
<td>Originally, I said, “O. K.,” but I regretted when last Wednesday because I was very tired and I wanted to go home. About this thing I felt guilt and sorry for you.</td>
</tr>
<tr>
<td>Errors pointed out by the teacher</td>
<td>1. Two weeks ago, you were asking me, “Do you want to go to Tainan with me next week?”</td>
</tr>
<tr>
<td></td>
<td>2. …, but I regretted when last Wednesday because I was…</td>
</tr>
<tr>
<td>Keywords for searches in Yahoo online bilingual dictionary</td>
<td>3. About this thing I felt guilt and sorry for you.</td>
</tr>
<tr>
<td>1.</td>
<td>when last Wednesday</td>
</tr>
<tr>
<td>2.</td>
<td>about</td>
</tr>
<tr>
<td>3.</td>
<td>with</td>
</tr>
<tr>
<td>4.</td>
<td>guilty</td>
</tr>
<tr>
<td>5.</td>
<td>guilty</td>
</tr>
<tr>
<td>6.</td>
<td>sorry for</td>
</tr>
<tr>
<td>Final draft</td>
<td>Dear Judy,</td>
</tr>
<tr>
<td></td>
<td>I apologize for having a quarrel with you last week.</td>
</tr>
<tr>
<td></td>
<td>Two weeks ago, you were asking me, “Do you want to go to Tainan with me next week?”</td>
</tr>
<tr>
<td></td>
<td>Originally, I said, “O. K.,” but I regretted when last Wednesday because I was very tired and I wanted to go home. About this thing I felt guilty and sorry.</td>
</tr>
</tbody>
</table>

The students’ perceptions toward using the bilingual concordancer and Yahoo’s bilingual dictionary

To answer the third research question, “What are students’ perceptions toward using a bilingual concordancer and an online bilingual dictionary for text revision in EFL writing?” two open-ended questionnaires were conducted to understand the students’ perceptions when using the bilingual concordancer and Yahoo’s Chinese-English dictionary in the metalinguistic process on improving writing. As shown in Table 7, most of the students in the experimental group (93%) expressed a positive response towards the bilingual concordancer. They indicated that the bilingual concordancer could help them to clarify word use and sentence patterns in writing as they reviewed many authentic language examples. Eleven participants (73%) figured out grammatical rules and sentence structures between Chinese and English, such as lexical features. Vocabulary in different English contexts may have different meanings, but this linguistic feature may not exist in Chinese. Nine participants (60%) remarked that the teacher’s scaffolding prompts for searching keywords and associated strings were very helpful to them when revising their texts. However, three participants (20%) stated that they still had trouble using the bilingual concordancer because the language examples were too long and difficult to understand. Simple language examples would be better for them. The results indicate that the students in the experimental group raised their metalinguistic awareness, as they learned to distinguish different linguistic features between English and Chinese. They also reduced their errors in the two languages, as they could observe both grammatical rules and sentence structures from authentic language examples in the bilingual concordancer.

Regarding the students in the control group, as shown in Table 8, all of them (100%) agreed that using the Yahoo’s Chinese-English online dictionary is beneficial, and could help them find single word equivalence in writing. Ten participants (59%) described how they used the online dictionary to check synonyms, but they could
not ensure the correctness of their use. By using the online dictionary, they also discovered that the differences between Chinese and English are the subject and tense. In English, obvious subjects and tense are necessary, but in Chinese they may be ignored. However, seven participants (40%) claimed that they were unable to detect syntactical errors, such as sentence structures when using the online dictionary, among which may have to rely on teacher feedback to revise their essays. The results indicate that the control group could revise their grammatical errors and sometimes lexical errors with the help of Yahoo Chinese-English online dictionary, but when it comes to syntactical errors, they needed to seek other type of assistance.

Table 7. Students’ use of the bilingual concordance (N = 15)

<table>
<thead>
<tr>
<th>Statements</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The bilingual concordancer provides language examples in different contexts to help me learn word expression and sentence structure.</td>
<td>14</td>
</tr>
<tr>
<td>2. I could discover different grammatical rules and sentence structures between Chinese and English. For example, vocabulary in different contexts might stand for different meanings. The lexical feature often exists in English but does not exist in Chinese.</td>
<td>11</td>
</tr>
<tr>
<td>3. I could select the appropriate words or phrases and place them in the right context in term of my teacher’s scaffolding prompts that introduced me how to enter keywords and associated strings.</td>
<td>9</td>
</tr>
<tr>
<td>4. Entering Chinese keywords to find equivalent English words and understand its word usage in multiple contexts facilitated me to avoid errors because I could observe and imitate correct word expression and sentence structure.</td>
<td>7</td>
</tr>
<tr>
<td>5. Pronouns and passive voice are often used in English sentences, but in Chinese they are not used very often. The bilingual concordancer helps me to discover how to use pronouns and passive voice correctly.</td>
<td>5</td>
</tr>
<tr>
<td>6. The language examples in English are difficult for me. I hope I can read more simple sentences after entering keywords.</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 8. Students’ use of Yahoo’s Chinese-English online dictionary (N = 17)

<table>
<thead>
<tr>
<th>Statements</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I usually used an online dictionary to find English equivalents for a single word.</td>
<td>17</td>
</tr>
<tr>
<td>2. It is helpful for me to check word meanings and the spelling of unfamiliar words.</td>
<td>13</td>
</tr>
<tr>
<td>3. I think consulting an online dictionary is helpful for me to check synonyms, but I cannot ensure whether using synonyms are correct or not when writing essays.</td>
<td>10</td>
</tr>
<tr>
<td>4. In English, the subject usually exists in the beginning of a sentence, followed by obvious grammatical tense. In Chinese, the subject and tense are often ignored.</td>
<td>8</td>
</tr>
<tr>
<td>5. Consulting an online dictionary may help me detect peers’ grammatical errors at word-level, but I have problems detecting sentence-level errors such as sentence structures.</td>
<td>7</td>
</tr>
<tr>
<td>6. I still cannot distinguish sentence structures between Chinese and English when using the online dictionary. I rely on teacher feedback to revise my sentences.</td>
<td>5</td>
</tr>
</tbody>
</table>

Discussion and conclusion

This study is a response to previous studies, confirming that a bilingual concordancer can help EFL students find word expression and sentence structure to improve their writing through text revision, in terms of providing authentic language examples and translations (Lai & Chen, 2015; Gao, 2011; Yang, Wong & Yeh, 2013). The results also correspond to the findings of Gao’s study (2011), indicating that metalinguistic awareness raised the students’ awareness of how to understand and apply language knowledge to text revisions with the use of a bilingual concordancer.

The findings show that viewing a large number of language examples and translations is essential for EFL students to diagnose their writing problems as well as check on their own learning process, and especially solve the problem of misuse of word choice, incorrect sentence construction due to the interference from the application of students’ native language. The teacher’s useful guidance facilitated the students to engage in the metalinguistic process to diagnose more lexical and syntactical errors, since they were able to apply new language knowledge of word choice, sentence structures, and constructions to text revisions. On the contrary, the students in the control group mostly looked up the meanings of words by consulting an online dictionary and relied on the teacher’s suggestions to revise their drafts. It is difficult for EFL students to induce or deduce language rules and sentence structures based on inadequate language examples. As a result, they often produce grammatically incorrect sentences and inappropriate word choices. It is suggested that EFL students refer to
various language tools in addition to online bilingual dictionaries. Finally, learning how to use a bilingual concordancer is of a great advantage for those who adopt self-correction approach to EFL writing improvement.

Some limitations and suggestions for future research were also elaborated in this study. First, the group size and number of participants was a major limitation of the study. The results of this study may not be representative enough to fully interpret all the problems that EFL students encounter, and the solutions that could be improved upon by using a bilingual concordancer. Second, the bilingual language examples from Taiwan Panorama may be challenging for lower-level EFL students. Additionally, future studies can investigate the effects of peer feedback in metalinguistic awareness to improve EFL students’ language transfer in text revisions. Finally, a longitudinal study can be carried out to investigate the students’ application of concordance learning to different fields with more participants and increasingly complicated writing tasks.

Acknowledgements

Heartfelt thanks to the grant support of Ministry of Science and Technology in the Republic of China, Taiwan (MOST 104-2511-S-224-001-MY3).

Reference


Francis, N. (2012). Bilingual development and literacy learning: East Asian and international perspectives. Hong Kong: City University of Hong Kong Press.


Appendix A. Sample Items of Text-structured Pre- and post-test

Part I: Each paragraph lacks one topic sentence. Please select an appropriate topic sentence from the following test items.

1. ____________ One enduring aspect of the cultures of Kenya is the family. Traditionally, Kenyan people received all their education from their parents and grandparents. Today there are schools to educate children, but the family is still very important as a social unit. Another enduring aspect of Kenyan cultures is respect for the old. Traditional society was organized around not only family life but also relationships with a group of people of the same age. These same-age groups went through all the stages of life together until they became the much-respected elders who made decisions for the community. Today old people are still respected, but they don’t have the responsibility of leadership that they once had.
   (A) Although Kenya has many different ethnic groups, the cultures across the nation share some common characteristics.
   (B) Kenya is a country that embraces modernity with new characteristics.
   (C) If you want to travel to Kenya, here are a few things you need to know.
   (D) To know a country, you have to know its cultures.

Part II: Each paragraph lacks one sentence or clause. Please select an appropriate answer from the following test items.

2. When visiting a home in Japan, presenting a housewarming gift can either make or break your relationship with the recipient. As with most countries around the world, Japan has very ancient traditions concerning gift giving and detailed rules for everything from the color of the wrapping paper to the time of the gift presentation.
   ____________
   (A) To be appreciated in a culture, it is important to honor those customs as not to offend the receiving party.
   (B) In Japan, gifts are avoided wrapped with brightly covered papers or bows.
   (C) In Japanese culture, the giver hands the gift to the recipient at the end of the evening. This is the opposite of the traditions in the United States, where a housewarming gift is usually given as soon as you enter the home.
   (D) It is customary to make a humble comment that the gift you are presenting, such as “This isn’t a big gift,” or “This is just a box of tea.”

Appendix B. Bilingual concordancer training

The following steps are helpful for you to understand how to query keywords and associate strings in the bilingual concordancer.

A Chinese example: 澎湖每年都花大筆預算在海上放煙火吸引觀光人潮。

Step 1: Underline words/phrases you want use as keywords.

Step 2: For example, when you enter “放煙火” in the bilingual concordancer, you can find the following results.

<table>
<thead>
<tr>
<th>Chinese sentence</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>在海上放煙火的確是與眾不同，再加上澎湖的空氣明淨，煙火特別耀眼。</td>
<td>Setting off fireworks over the ocean is, indeed, out of the ordinary.</td>
</tr>
<tr>
<td>「浪漫耶誕夜」、「跨年」等慶祝活動，永遠都是放煙火、找明星跨刀那一套，搞得既無創意，甚至演變成行銷首長形象的樣板。</td>
<td>“Romantic Christmas Eve,” New Year’s Eve shows, and things like that, which never got to be more than formulaic fireworks displays and appearances by pop stars, completely lacking in creativity, and in some ways they are just image-building exercises for politicians.</td>
</tr>
</tbody>
</table>

Step 3: Read the language examples, especially pay attention to the keywords highlighted in red color, checking its word use in sentences.

Step 4: Look back at the keyword in the Chinese example and its surrounding words.
**Step 5:** Compare and analyze the different results of the keyword and figure out which one would be better to translate this Chinese example.

**Step 6:** Induce the grammar use and sentence patterns to the keywords and surrounding words based on the examples.
Effects of Gamified Comparison on Sixth Graders’ Algebra Word Problem Solving and Learning Attitude

Hong-Zheng Sun-Lin* and Guey-Fa Chio
Graduate Institute of Information and Computer Education, National Taiwan Normal University, Taipei City, Taiwan // hzsunlin@gmail.com // gueyfa@ntnu.edu.tw

*Corresponding author

(Submitted September 15, 2016; Revised June 18, 2017; Accepted August 8, 2017)

ABSTRACT
This study examined effects of gamified comparison on sixth graders’ performance of algebra word problem solving and attitude toward algebra learning. Seventy-two sixth graders were invited to participate in a four-week experimental instruction and assigned to three groups: gamified comparison, comparison, and control. The results showed (1) a significant effect on solving similar problems: the gamified comparison group performed significantly better than the comparison group and the control group respectively, and the comparison group performed significantly better than the control group; (2) a significant effect on solving transfer problems: the gamified comparison group gained significantly higher scores than the comparison group and the control group respectively, and the comparison group gained significantly higher scores than the control group; (3) a significant effect on students’ learning attitude: while no significant differences found on students’ confidence, the gamified comparison group made significantly more positive responses than the comparison group and the control group in terms of enjoyment, motivation, and perceived value. This study proposed a feasible combination of game rule and comparison strategy, as well as exploring implications for teachers’ teaching design and students’ gamified learning activities.

Keywords
Gamification, Comparison, Algebra word problem solving, Learning attitude

Introduction

Algebra word problem solving

Algebra plays an important role in mathematics learning, and the success in algebra learning is necessary to access to higher mathematics (Adelman, 2006; National Mathematics Advisory Panel, 2008). Many students, however, struggle with algebra that normally is the first unit in mathematics where they usually find the abstraction and symbolisation difficult to understand (Star et al., 2015; National Research Council, 2001). In algebra learning, solving word problems presents a challenge to middle graders (Bush & Karp, 2013; Carpraro & Joffrion, 2006; MacGregor & Stacey, 1998). It is probably because many learners, especially novice learners, are unable to recognise a similar problem type across a variety of cover stories/contexts, and apply what they have learnt to different problem types as well (Bush & Karp, 2013; Reed, 1987; Reed, Dempster, & Ettinger, 1985).

Transfer has been considered as an important ability for algebra word problem solving because it enables students to apply gained knowledge to build an understanding of a new concept. It also makes appropriate connections between various concepts within a domain based on the understanding of shared structural similarity between problems (Bush & Karp, 2013; Ngü & Yeung, 2012). In algebra learning, transfer ability prompts students to apply a learnt problem-solving technique to another context to solve novel problems (Rittle-Johnson & Star, 2007, 2009). To develop the transfer ability, previous studies suggested that teachers can encourage students to solve similar problems with different methods first (Richland, Holyoak, & Stigler, 2010; Richland & Mcdonough, 2010). For example, Rittle-Johnson and Star (2007) found that seventh graders who compared two contrasting step-by-step procedures outperformed those who studied one example with one solution at a time in problem solving. They attributed the learning effect to the benefit of allowing students to explore knowledge and flexibility in their problem-solving process. In their subsequent study, Rittle-Johnson and Star (2009) further stated that students’ knowledge and ability to solve similar algebra problems was the essential pre-requisites to ensure effects of transfer when solving novel problems. According to above-mentioned studies, practising similar problems with different stories/contexts is beneficial for students’ ability to solving transfer problems. However, research focusing on related teaching strategies to develop students’ problem solving ability is still scarce.

As to mathematics learning, topics regarding students’ learning attitude have also received considerable attention (Lim & Chapman, 2013; Zan, Brown, Evans, & Hannula, 2006). Several studies have noted that learning attitude would affect students’ learning performance, and strategies to encourage students’ learning behaviour should be
emphasised in teaching design (Bouchez & Harter, 2005; Samuelsson & Granstrom, 2007). To be specific, Zakaria and Nordin (2008) proposed four favourable affective dimensions, including enjoyment of mathematics, general motivation of learning mathematics, confidence in solving mathematics problems, and perception on value of mathematics that would affect students’ learning achievement. These aspects of learning attitude should also be measured for further exploration when exploring students’ mathematics learning achievement.

**Comparison strategy**

Comparing problems and their solutions has been considered as an effective learning strategy in mathematics (National Council of Teachers of Mathematics, 2000). Numerous studies have investigated effects of differentiating similarities and differences between problems and solutions on students’ mathematics problem solving by fostering students’ reflection (Silver, Ghousseini, Gosen, Charalambous, & Strawhun, 2005; Ziegler & Stern, 2014, 2016). Some scholars further noted that comparing side-by-side examples can effectively help students understand important features of a problem as well as solving novel problems based on what they have learnt (Gentner, Loewenstein, & Thompson, 2003; Sidney, Hattikudur, & Alibali, 2015). To enhance the learning effects, the importance of systematic guidance and coherent structure have also been highlighted (Star et al., 2015; Ziegler & Stern, 2016). Alfieri, Nokes, and Schunn (2013) mentioned that guiding prompts could play an important role in students’ comparison activities. However, the benefits of comparisons might be limited when students merely read or solve given examples (Ziegler & Stern, 2016). According to studies of Sweller and Chandler (1994) and Sweller, Merrienboer, and Paas (1998), guiding prompts could encourage students to familiarise themselves with similar examples and realise the features of problems, thereby gaining a better understanding of certain problem types and possible problem-solving techniques. This probably implies that precise and explicit guiding prompts could be necessary and essential to students’ comparison activities. For example, asking students to directly describe similarities and differences of given examples (Mason, 2004; Ziegler & Stern, 2014) or respond specific comparison questions (Gadgil, Nokes, & Chi, 2012; Ziegler & Stern, 2016).

A review of numerous empirical studies in a practice guide proposed by the US Department of Education identified comparison as one of recommendations to enhance grade 4-8 students’ mathematical problem-solving abilities (e.g., Guo & Pang, 2011; Rittle-Johnson & Star, 2007, 2009; Rittle-Johnson, Star, & Durkin, 2012; Woodward et al., 2012). Each of reviewed studies generally covered two common features of comparison strategy: the use of examples and the use of prompts. According to the conclusion, students who were shown and compared two side-by-side examples performed significantly better than those who did not receive such materials. Although positive learning effects of comparison on student’s mathematics learning (e.g., arithmetic) were reported, few studies have been done specifically on middle graders’ algebra word problem solving, especially their transfer performance.

One problem usually occurs when students compare examples: high learning effort (Ziegler & Stern, 2016). It may make learning tasks challenging and have a negative effect on students’ learning attitude, especially for those who are not fairly familiar with given concepts and the process of comparison. For example, in Ziegler and Stern’s study (2014), they provided some algebra principles and problems for sixth graders who were assigned to one of two groups: contrast or sequential group. In the contrast group, algebra problems were presented in juxtaposition; in the sequential group, the problems were presented in sequence. The results showed that not only could the contrast group distinguish between examples but also apply the algebra principles in the long run; however, the contrast group performed worse on immediate learning measures than the sequential group. The authors explained that the requirement of comparison may make the learning tasks more difficult for students to accomplish. To be specific, juxtaposition of examples could help students benefit from the process of comparison, but at the same time they probably increase students’ cognitive load that might negatively affect learning attitude.

**Gamified learning**

Gamified learning is defined as a strategy to employ game elements (e.g., points, levels, and challenges) in learning contexts to promote students’ learning (Dominguez, et al., 2013; Landers, 2014). For example, in gamified learning activities, students earn experience points after successfully accomplishing certain tasks by themselves. Villagra, Fonseca, Redondo, and Duran (2014) stated that gamification is not just about gaming, it is also about ensuring that students are motivated to complete learning tasks. To develop a gamified learning system, ways of using game elements in learning activities that make digital games attractive to students should
be carefully considered (Buckley & Doyle, 2014; Domínguez, et al., 2013). For examples, points could be regarded as rewards after completing a challenge (Attali & Arieli-Attali, 2015; Sun-Lin & Chiou, 2017b), level-ups could be used to unlock new challenges after meeting certain requirements (Buckley & Doyle, 2014; Sun-Lin & Chiou, 2017a), and records of challenges could be employed to keep students’ engagement in tasks (Domínguez, et al., 2013). Additionally, several studies showed that when fun from gamified content impregnates the learning process, students’ learning motivation would increase and the stress would be reduced (Koster, 2004; Villagrá-Arnedo, Gallego-Durán, Molina-Carmona, & Llorens-Largo, 2016). This kind of gamification design would facilitate students’ learning because game elements encourage students to engage in learning tasks, perceive what they experience, and decide what to do in learning process (Mora, Riera, González, & Arnedo-Moreno, 2015).

However, studies on gamification have not always reported positive results. Here is a common problem: students may pay too much attention to where game elements are added. For example, Attali and Arieli-Attali (2015) conducted two studies to examine effects of points on students’ learning performance. In their first study, students were rewarded by points for their performance including correctness of answers and speed of answering when in assessment. In the second study, all the participants were required to express their enjoyment of the achievement test gamified by points. The results showed that while the students gave positive feedback on the gamified assessment, they did not report an entirely positive attitude toward the learning activity. Moreover, although students’ speed of answering significantly increased, the correctness did not. This is probably because they knew they would be rewarded by quick responding and thus focused more on improving speed of responses rather than correctness of answers.

Another issue is that students’ experience of gamified learning remains on the surface. In many gamified activities, students just received a layer of standardised game elements such as points and levels; however, in a gamified context, the connection between gaming process and learning process that guides students to learn certain content has not yet been much explored (Landers, 2014; Mora et al., 2015; Sun-Lin & Chiou, 2017a). For example, Villagrá-Arnedo et al. (2016) proposed an idea that adapting game rule/procedure of a famous game Pac-Man to facilitate students’ problem solving, and in which students would experience the problem-solving process through gaming process, and vice versa. Teachers could also design and present instructional materials more effectively. To date, however, few empirical studies have been done on combination of the processes.

Overall, although algebra word problem solving has been considered an important and challenging field for students, little research has been done on related innovative strategies such as gamification. In addition, many studies reported positive effects of comparison strategy on mathematics learning achievement, and it requires higher cognitive effort that may negatively affect students’ learning attitude. Feasible applications still need to be explored further. Despite the growing awareness of the educational potential of gamification, empirical studies on its applications in algebra word problem solving are scarce. The present study, therefore, seeks to fill the gap by examining effects of gamified comparison on sixth graders’ algebra word problem solving and their learning attitude through investigating the following research questions:

- What are the effects of gamified comparison, comparison, and direct practice activities on students’ performance on solving similar algebra word problems??
- What are the effects of gamified comparison, comparison, and direct practice activities on students’ performance on solving transfer algebra word problems?
- What are the effects of gamified comparison, comparison, and direct practice activities on students’ learning attitude toward algebra word problem solving?

**Method**

This study adopted a quasi-experimental design with learning activities as the independent variable. To be specific, students who were asked to compare the correct/incorrect examples in gamified tasks were considered as the gamified comparison group (the GC group). Students who were asked to compare the correct/incorrect examples in general tasks were considered as the comparison group (the CO group). The other participants who read all the learning material and did not receive any gamification elements and comparison tasks were considered as the control group (the C group).

The dependent variables were students’ learning achievement of algebra word problem solving and learning attitude. The learning achievement referred to the learning outcomes including (1) test scores for solving similar problems, and (2) test scores for solving transfer problems. The learning attitude referred to students’ perception
of algebra learning after completing learning tasks, including four dimensions: enjoyment, motivation, confidence, and perceived value.

Participants

Seventy-two sixth graders (39 females and 33 males, 12-13 years old) participated in this instructional experiment. They were randomly and evenly assigned to one of groups, meaning there were 24 students in the GC, the CO, and the C group respectively. All the participants had learnt basic algebra concepts (e.g., know how to use symbols to present unknown values) and had basic algebra arithmetic ability (e.g., know how to simplify an algebraic expression).

To minimise the possible influence of The Hawthorne Effect, the students attended learning activities as invited participants not subjects of an experiment. All of them were protected by replacing their personal information with serial numbers. They were informed that the participation was voluntary and would not affect the grade of the course and they could withdraw from the study at any time.

The learning system

To explore combination effects of gamification and comparison strategy, the researchers developed a mini learning system by using software Construct 2. The participants were asked to accomplish four tasks by studying four examples and solving four practice problems with guidance provided by the system, as described below.

The learning content

To develop the learning content properly, one middle-school mathematics teacher, an author of K-9 mathematics reference books with more than 13-year teaching experience, and two primary-school mathematics teachers with five-year teaching experience were invited. They reviewed sixth-grade mathematics instructional materials and drafted four types of algebra word problems that commonly confuse students. Examples in each learning task were designed based on these problem types, as shown in Table 1. The guiding prompts, designed for facilitating students’ comparison, were also reviewed by the invited teachers, as listed below.

- What are the similarities between the two solutions in this example?
- What are the differences between the two solutions in this example?
- Which one is incorrect (choose one of examples)? Why (your explanation for the answer)?

<table>
<thead>
<tr>
<th>Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Addition/subtraction</td>
<td>Peter has a pen and so does his sister Amy. Amy’s pen is 4 cm longer than Peter’s. The sum of their pens’ lengths is 20 cm. What is length of Peter’s pen? What is length of Amy’s pen?</td>
</tr>
<tr>
<td>Sum of differences</td>
<td>There are 500 packages in Warehouse A and 300 packages in Warehouse B. If employees take out 15 packages from Warehouse A and 5 ones from Warehouse B per day. How many days later will the number of packages in Warehouse A be equal to the number of packages in Warehouse B?</td>
</tr>
<tr>
<td>Multiple</td>
<td>A commander is 58 years old and you are 28 years old. When will the commander’s age be two times yours?</td>
</tr>
<tr>
<td>Remainder/shortage</td>
<td>There are many apples that will be distributed to children in a classroom. If we distribute 5 apples to each child, we will need 12 more apples. If we distribute 3 apples to each child, there will be no remainder. How many children are there?</td>
</tr>
</tbody>
</table>

The GC group

Students in this group learnt to solve algebra word problems through accomplishing four gamified comparison tasks. Each task consisted of three sections: a description of four types of algebra word problems, four example problems with correct and incorrect step-by-step solutions, and four practice problems.
The design of tasks was an adaption of a famous game Photo Hunt in which in each level players would receive two side-by-side photos and be asked to find out differences between them. The students compared the correct and the incorrect examples that were juxtaposed, then responded prompts by their own explanations. They could go back to previous pages for review if necessary (see Figure 1). This juxtaposed examples were developed based on the features of effective comparison strategy, i.e., to differentiate two cases easily on a single page. The step-by-step prompts with specific statements were used to guide and encourage students to reflect on the techniques of problem solving and accomplish learning tasks more effectively instead of extremely-high cognitive effort.

In gamified tasks, students played defenders to protect the Earth from the unknown attack from the space and were required to get their aircrafts ready by solving a series of word problems. There were two solutions in each example associating with an issue of the aircraft and the contingent. For example, defenders were asked to solve a Sum of Differences problem to predict how many day(s) later there will be no aircraft fuel left in two warehouses at the same time. Before solving problems, players had to study all the examples, but they could decide whether to solve the four challenging problems or not for practice. If they practised solving the challenging problems and answered correctly, they would earn faster level-up opportunities for a better aircraft. Once all the problems in a task were solved, the levels of defenders would go up and the aircraft would be significantly upgraded.

Figure 1. A screenshot of a gamified comparison task

The CO group

Students in this group learnt how to solve algebra word problems through accomplishing four comparison tasks. They were asked to compare correct and incorrect examples, then respond the prompts by their own expressions. However, these participants did not receive game elements during the process. Similarly, the students could decide whether to solve the four challenging problems and go back to previous pages for review if necessary.

The C group

The students learnt to solve algebra word problems via studying given learning materials. They read all the four examples, decided whether to solve the four practice word problems, and reviewed previous pages if necessary. All the examples were presented page by page that could be reviewed unlimitedly during the period.

Instruments

Algebra word problem tests

The achievement tests were reviewed by three primary-school mathematics teachers with over five-year teaching experience. The review suggested the content validity. Two tests were created: a similar-problem test and a transfer-problem test. Both of them consisted of eight problems. Problems in the similar-problem test were similar to practice problems presented in the system, and problems in the transfer-problem test were slightly different from the practice problems. Here are some examples:

• A practice problem: A commander is 58 years old and you are 28 years old. When will the commander’s age be two times yours?
• A similar problem: You are 50 years old and your nephew is 10 years old. In how many years will your age be three times your nephew’s age?
• A transfer problem: John has worked for 25 years and Mary has worked for 9 years in a company. How many years ago was John’s work year three times Mary’s?

To further ensure the quality of the measurement, 96 sixth graders who did not participate in the instructional experiment were invited to complete the tests for pilot study. The reliability of tests was estimated based on Cronbach’s α measure for the pre-test α = .864 and for the post-test α = .888, and there was a strong positive correlation between the pre-test and the post-test (r = .821, p = .000). Additionally, the tests showed significant positive correlations with a regular school mathematics examination in algebra word problem solving: between the examination and the pre-test, r = .797, p = .000; between the examination and the post-test, r = .752, p = .000. The results indicated good criterion-related validity of the tests.

Algebra learning attitude scale

The scale for students’ algebra learning attitude covering enjoyment, motivation, self-confidence, and perceived value, was modified from an inventory developed by Lim and Chapman (2013). It comprised 19 items using a four-point Likert rating schema. Here are examples of the items:
• I like solving algebra word problems.
• I am confident that I can solve more challenging algebra word problems.
• I am not stressed when solving algebra word problems.
• The ability of solving algebra word problems is important.

To ensure appropriateness of the scale, the three invited teachers reviewed all the scale items, and the 96 sixth graders were also invited to complete the survey to help the researchers examine the reliability. The Cronbach’s α value for the enjoyment was .909, for motivation was .929, for confidence was .901, for perceived value was .923, and for the entire scale was .938, showing excellent internal consistency.

Procedure

The experimental instruction took about 265 minutes in four weeks. First, the researchers introduced the system and invited the participants to experience the learning activities (5 minutes), after which students finished the pre-test of learning achievement (35 minutes) and the pre-survey of learning attitude (15 minutes). During the learning activities, all students learnt how to solve algebra word problems through accomplishing different learning tasks (i.e., gamified comparison, comparison, and read-and-practice) in 160 minutes. After that, all groups answered the post-test of learning achievement (35 minutes) and responded the post-survey of learning attitude (15 minutes).

Results

Student performance on solving similar problems

This study employed covariate of analysis (ANCOVA) to examine the sixth graders’ performance of algebra word problem solving on the similar-problem test. As presented in Table 2, the results showed a significant difference among the three groups, with $F = 29.615 (p < .05)$, meaning the post-hoc comparison was needed. The result of post-hoc comparison further indicated that the GC group (adjusted $M = 9.960$, SE = .487) gained significantly higher scores than the CO group (the adjusted $M = 7.707$, SE = .495) and the C group (the adjusted $M = 4.834$, SE = .469) respectively, and the CO group gained significantly better scores than the C group. According to the definition of effect size proposed by Cohen (1988), the partial eta squared ($\eta^2$) of the analysis results indicated a strong effect ($\eta^2 = .466 > .139$).

Table 2. ANCOVA results of student performance on the similar-problem test

<table>
<thead>
<tr>
<th>Source of variance</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>$F$</th>
<th>$\text{Sig.}$</th>
<th>Post-hoc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prior knowledge</td>
<td>244.126</td>
<td>1</td>
<td>244.126</td>
<td>46.354***</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>311.933</td>
<td>2</td>
<td>155.967</td>
<td>29.615***</td>
<td>.000</td>
<td>GC &gt; CO, GC &gt; C, CO &gt; C</td>
</tr>
<tr>
<td>Error</td>
<td>358.124</td>
<td>68</td>
<td>5.267</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. ***$p < .000$. 

125
Student performance on solving transfer problems

The ANCOVA results, as presented in Table 3, indicated a significant difference among the three groups, with $F = 47.235$ ($p < .05$). The result of the post-hoc comparison indicated that the GC group (adjusted $M = 4.735$, $SE = .504$) acquired significantly better scores than the CO group (the adjusted $M = 3.336$, $SE = .358$) and the C group (the adjusted $M = .179$, $SE = .339$) respectively, and the CO group performed significantly better than the C group. The partial eta squared ($\eta^2$) of the analysis results represented a strong effect size ($\eta^2 = .581 > .139$).

### Table 3. ANCOVA results of student performance on the transfer-problem test

<table>
<thead>
<tr>
<th>Source of variance</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>$F$</th>
<th>Sig.</th>
<th>Post-hoc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prior knowledge</td>
<td>44.382</td>
<td>1</td>
<td>44.382</td>
<td>16.107***</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>260.305</td>
<td>2</td>
<td>130.152</td>
<td>47.235***</td>
<td>.000</td>
<td>GC &gt; CO, GC &gt; C, CO &gt; C</td>
</tr>
<tr>
<td>Error</td>
<td>187.368</td>
<td>68</td>
<td>2.755</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. *$p < .01$; **$p < .001$.***

Students’ learning attitude toward algebra word problem solving

MANCOVA was used to examine the sixth graders’ learning attitude. The results showed that there was a statistically significant difference in students’ attitude toward algebra word problem solving based on learning tasks (Wilks’ Lambda = .504, $p = .000, \eta^2 = .290$). After that, ANCOVA was employed to explore the differences in each dimension. Although there was no significant difference on students’ confidence ($F = 1.670, p = .196 > .05$), significant differences were found on their enjoyment ($F = 14.425, p < .05$), motivation ($F = 24.234, p < .05$), and perceived value ($F = 11.742, p < .05$), as shown in Table 4. Therefore, the post-hoc comparisons were needed.

The result of the post-hoc comparison showed that on enjoyment, the GC group (adjusted $M = 17.235$, $SE = .469$) gave significantly more positive responses than the CO group (the adjusted $M = 15.840$, $SE = .473$) and the C group (the adjusted $M = 13.675$, $SE = .474$) respectively, and the CO group responded significantly more positively than the C group. The partial eta squared ($\eta^2$) of the analysis results represented a strong effect ($\eta^2 = .298 > .139$).

As to students’ motivation, the post-hoc comparison showed that the GC group (adjusted $M = 14.917$, $SE = .387$) made significantly more positive responses than the CO group (the adjusted $M = 12.624$, $SE = .388$) and the C group (the adjusted $M = 11.194$, $SE = .391$) respectively. The partial eta squared ($\eta^2$) of the analysis results indicated a strong effect ($\eta^2 = .416 > .139$).

As for the perceived value, the post-hoc comparison indicated that the GC group (adjusted $M = 17.945$, $SE = .433$) had significantly more positive responses than the CO group (the adjusted $M = 15.981$, $SE = .445$) and the C group (the adjusted $M = 14.824$, $SE = .483$) respectively. The partial eta squared ($\eta^2$) of the analysis results represented a strong effect ($\eta^2 = .257 > .139$).

### Table 4. ANCOVA results of learning attitude

<table>
<thead>
<tr>
<th>Source of variance</th>
<th>Dimension</th>
<th>$SS$</th>
<th>df</th>
<th>$MS$</th>
<th>$F$</th>
<th>Sig.</th>
<th>Post-hoc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prior attitude</td>
<td>Enjoyment</td>
<td>104.220</td>
<td>1</td>
<td>104.220</td>
<td>19.767***</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Motivation</td>
<td>32.630</td>
<td>1</td>
<td>32.630</td>
<td>9.131**</td>
<td>.004</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Confidence</td>
<td>409.985</td>
<td>1</td>
<td>409.985</td>
<td>76.039***</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Perceived value</td>
<td>90.429</td>
<td>1</td>
<td>90.429</td>
<td>21.108***</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>Enjoyment</td>
<td>152.111</td>
<td>2</td>
<td>76.055</td>
<td>14.425***</td>
<td>.000</td>
<td>GC &gt; CO, GC &gt; C, CO &gt; C</td>
</tr>
<tr>
<td></td>
<td>Motivation</td>
<td>173.198</td>
<td>2</td>
<td>86.599</td>
<td>24.234***</td>
<td>.000</td>
<td>GC &gt; CO, GC &gt; C</td>
</tr>
<tr>
<td></td>
<td>Confidence</td>
<td>18.008</td>
<td>2</td>
<td>9.004</td>
<td>1.670</td>
<td>.120</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Perceived value</td>
<td>100.612</td>
<td>2</td>
<td>50.306</td>
<td>11.742***</td>
<td>.000</td>
<td>GC &gt; CO, GC &gt; C</td>
</tr>
<tr>
<td>Error</td>
<td>Enjoyment</td>
<td>358.530</td>
<td>68</td>
<td>5.272</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Motivation</td>
<td>242.995</td>
<td>68</td>
<td>3.573</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Confidence</td>
<td>366.640</td>
<td>68</td>
<td>5.392</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Perceived value</td>
<td>291.321</td>
<td>68</td>
<td>4.284</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. *$p < .01$; **$p < .001$; ***$p < .000$. 

126
Discussion and conclusions

A significant effect on solving similar problems

First, the GC group acquired significantly higher scores than the CO group; second, the GC group also gained significantly higher scores than the C group. These results showed that gamified comparison tasks could effectively help students learn to solve algebra word problems. The gamification design played a role in learning effects of the comparison strategy. This is because the game context could effectively immerse students in the learning tasks, thereby encouraging them to concentrate on example studying to learn how to solve certain algebra word problems. Additionally, in gamified tasks, students would not only be engaged but be able to monitor their own learning progress because they could realise their progress while accomplishing each task with game rewards. Their learning achievement on the similar-problem test, therefore, reflected learning effects of the gamified comparison tasks. These results could address the issue proposed by Attali and Arieli-Attali’s (2015): gamified group may gain lower scores than the non-gamified group by points. When adapting a game rule from an existing game (e.g., Pac-Man or Photo Hunt) that could combine learning and gaming process, rather than just employing a layer of standardised game elements (e.g., points and levels), the effects of learning gamification would be more significant (Villagrà-Arnedo et al., 2016).

Third, the CO group acquired significantly higher scores than the C group. The results obviously showed that the comparison strategy was effective in students’ learning of algebra word problem solving. The comparison process could encourage students to reflect on what they learnt and find out their own difficulties when explaining answers by themselves. The result supports findings of numerous previous studies (Rittle-Johnson & Star, 2007, 2009; Star et al., 2015): the use of comparison strategy could improve students’ algebra learning. It also addresses the issue proposed by Ziegler and Stern (2014; 2016): the effects of comparing contrasting examples only appeared with delay. However, by using appropriate guiding prompts to facilitate students’ comparison, their learning achievement on the follow-up measures can be effectively enhanced.

A significant effect on solving transfer problems

Firstly, the GC group performed significantly better than the CO group; secondly, the GC group performed significantly better than the C group as well. The results showed that the gamified comparison had significant effects on students’ performance on solving transfer problems. The reason may lie in the game rule and the challenges in tasks. The gamified comparison tasks not only motivated students to compare examples to gain effective techniques of algebra word problem solving but also encouraged them to solve the challenging practice problems with giving game rewards and level-up settings. For example, after accomplishing a similar problem following by an example, students received a message that recognised the performance then encouraged them to solve a challenging problem for practice. Such design may help students gain a sense of achievement and motivate them to attempt to practise more related problems, thereby enhancing their transfer ability. Thus, they could score higher on the transfer-problem test because they probably had enough practice to master problem-solving techniques than others students. These results might bridge the gap between gamified meta-cognitive strategy (i.e., comparison) and students’ learning performance on solving transfer algebra word problems.

Thirdly, the CO group performed significantly better than the C group. Compared with the C group, even though students in the CO group were not encouraged to solve different algebra word problems, the process of comparison could still guide them to know problem-solving techniques. The comparison activities promoted students to think deeply, reflect on what they already understood, and find out what they still felt confused about. Such process may further encourage students to learn how to analyse and solve a problem from different angles as well as monitoring their own learning progress, thereby improving the performance on solving transfer problems. The result is in accord with research findings from Christianson et al. (2012): comparison could be promising to facilitate students’ transfer ability in algebra word problem solving. In addition, as Rittle-Johnson and Star (2007; 2009), and Rittle-Johnson et al. (2012) stated: comparison may not significantly effective without appropriate guiding prompts. Teachers should consider providing clear and understandable guiding prompts to help students develop meta-cognitive ability that usually requires high cognitive load.

A significant effect on learning attitude

As to students’ enjoyment, the GC group responded significantly more positive results than the CO group and the C group respectively, and the CO group gave significantly more positive responses than the C group.
According to the results, the gamification design made learning tasks more interesting to immerse students in learning activities. This echoes the statement from Buckley and Doyle’s (2014) and Domínguez et al. (2013) in which game elements used in learning content attracted students’ attention and motivate them to complete learning tasks. In addition, the juxtaposed examples and guiding prompts made the algebra word problem solving interesting because students might reckon that they were playing a mini game (similar to Photo Hunt) rather than just solving a problem. Such design may help students engage in the problem-solving process and lead to positive responses. This might address the issue proposed by Hanus and Fox (2015): gamified group may show less satisfaction than the non-gamified group after learning activities because of inappropriate design of learning gamification. By adapting the game rule to present process of comparison, students’ learning performance and their attitude toward algebra word problem solving can be enhanced effectively.

As for students’ motivation, the GC group reported significantly more positively than the CO group and the C group respectively. The results showed that the gamification design motivated students to learn how to solve algebra word problems. This is in line with previous research findings that game elements encourage students to learn and further increase their learning motivation (Koster, 2004; Villagrá-Amedo et al., 2016). However, there was no significant difference between the CO group and the C group. This is probably because a comparison task requires more time and effort than students thought. In other words, they were more likely to feel frustrated rather than confident when failing to accomplish assigned tasks. That may make it harder to strengthen learners’ motivation.

There was no significant effect on sixth graders’ confidence probably because that various problem types and steps of algebra word problem solving make it challenging for students to retrieve appropriate techniques when solving problems by themselves. The results are consistent with findings of Hanus and Fox’s study (2015). The difficulty of learning content and requirement of tasks might pose a challenge to students even if guiding prompts and encouragement mechanism are provided. This led them not to respond as positive as on other attitude aspects.

The GC group made significantly more positive responses than the CO group and the C group respectively in terms of perceived value. Comparison activity usually required students to pay more attention to the process and accomplish the tasks (Rittle-Johnson & Star, 2007; Rittle-Johnson & Star, 2009), and the gamification design such as levels and rules provided students with a good opportunity to realise what they had achieved and reflect on what they had learnt (Domínguez et al., 2013). Such combination design could be used to raise learners’ awareness of their own learning difficulty and progress, and further help them perceive value of learning tasks and materials.

**Implications**

Several implications can be drawn from this study. Firstly, a possible combination of game rule and meta-cognitive strategy was proposed, and it effectively facilitated students’ algebra word problem solving. This provides teachers with a better understanding of the connection between learning and gaming process and some ideas of using common game rules in learning activities. Secondly, despite the simplicity of game rule, students’ enjoyment, motivation, and perception of learning could be effectively enhanced. It encourages teachers to select and apply some appropriate game rules to classroom activities, thereby improving their students’ learning attitude. Thirdly, this study suggests positive learning effects of procedural strategies (i.e., comparison rule and steps) that can be employed in class. However, exploration of possible combined effects and relationships (e.g., between rewards, challenges, rules) is limited in this study. This may point to new possibilities for future study on gamified learning strategies.

**Limitations**

Although this study has yielded findings that have both theoretical and instructional implications, its design is not without flaws. The first limitation concerns the Hawthorne effect existed in the study. It may limit the explanations and discussion of the results. The second limitation is rooted in the relatively short time (i.e., 265 minutes in four week) allowed for the intervention. In addition, since this study involved only 24 participants in each group, the generalisation of the results can be limited. The long-term effects of learning gamification on larger and different populations could be explored further in future studies.
References


