Implementing online question generation to foster reading comprehension

Hui-Chin Yeh and Pei-Yi Lai
National Yunlin University of Science and Technology

Many studies have concluded that question generation has a positive effect on students' reading comprehension. However, few studies have delineated how students generate questions from a text and what processes are involved in question generation. This study aims to investigate how the question generation processes improve students' reading comprehension, using an online question generation system including the organisation, composition and peer assessment modules. 19 out of 106 non-English major college students were recruited as participants. They were required to complete question generation tasks in the organisation, composition and peer assessment modules. Students' scores on the pre- and post-tests, action logs in the online question generation system, and interview transcripts were collected and analysed. In a micro view, results of this study indicated that college students who showed more progress in reading comprehension demonstrated similar question generation patterns. In the organisation module, those who made more progress had a higher frequency of adding new vocabulary, sentences, and main ideas and editing their previously organised information. In the composition module, they had a higher frequency in reviewing the previously organised information from a text to generate questions and in editing the organised information. In the peer assessment module, those who showed more progress were much more active in viewing peers' questions, providing comments on peers' questions, reading and responding to peers' comments on the questions. In a macro view, the intensive engagement and the actions of editing to retrieve the organised information to compose the online questions and reviewing peers' questions online were found to be critical factors for enhancing students' reading comprehension.

Introduction

Reading academic textbooks has been recognised as a crucial skill for English as a Foreign Language (EFL) college students to acquire content knowledge in their domains (Chien, 2000; Day & Bamford, 1996; Dlugosz, 2000; Liu, Chen & Chang, 2010; Salinger, 2003). However, a lot of EFL students fail to prepare themselves for the reading demands of higher education (Dreyer & Nel, 2003). Yang (2010) highlighted that 71% of Taiwanese college students were underprepared for the course reading in technological and vocational education system (TVES). In general, many college students in Taiwan have difficulty in their reading comprehension, and have a lack of cognitive strategies to remedy their comprehension breakdown.

Some of the drawbacks in English reading instruction are identified in the literature. First, students often act as passive learners in reading classes, simply following the teachers' instruction and answering teacher-generated questions (Gillespie, 1990; Miciano, 2002; Singer & Donlan, 1982). As a result, students seldom actively engaged

themselves in the reading comprehension process, which has resulted in poor reading comprehension (Bristow, 1985). Second, the majority of reading instruction is test-oriented. Teachers often devote a great amount of time to testing students' reading comprehension rather than teaching reading strategies (Liang & Dole, 2006; van Keer, 2005). Consequently, most college students struggle or fail to comprehend their English textbooks. Third, most teachers evaluate students' reading comprehension based merely on students' performance on reading comprehension tests (Dreyer & Nel, 2003). However, "performance on a standardized reading comprehension test reflects the number of correct answers readers select from a list of alternate choices, but fails to provide information about how readers cope with the various cognitive demands of the task" (Tal, Siegel & Maraun, 1994, p. 387). Hence, teachers have few clues to know the difficulties their students encounter in their reading processes, and thereby fail to provide effective scaffolding to help students solve their reading problems. It should be noted that students' passive comprehension process and the lack of teachers' effective scaffolding contribute to students' poor reading comprehension.

One of the effective approaches to improve students' reading comprehension is through self-generated questions (Singer, 1978) that engage students in comprehension fostering and comprehension monitoring (Palincsar & Brown, 1984). Question generation can be specifically defined as an activity in which students generate exam questions based on the reading content, using multiple choice, matching, short answer, true-false, and fill-in-the-blank formats or word puzzles (e.g., Belanich, Wisher & Orvis, 2005; Rafaeli, Barak, Dan-Gur & Toch, 2004; Wilson, 2004; Yu, Liu & Chan, 2005). For example, when generating multiple choice questions, students need to concentrate on the important information in their texts, and then provide correct answers and distracters (Yu, Liu & Chan, 2002). In the process of designing distracters for a question, students need to know the reason why the distracters are incorrect. Belanich, Wisher and Orvis (2005) maintained that questions which are authorised by a student can reveal what information a student knows and does not know. Fellenz (2004) further claimed, students' abilities to explain why the answer options they create are correct or incorrect, reveal whether students really understand the reading materials or not. In the process of generating a question, students need to have both micro and macro understandings of the reading materials (Horgen, 2007), and connect new information in the materials to their prior knowledge (Wong, 1985; Draaijer & Boter, 2005). By undertaking the process of question generation, students can more actively comprehend a text, and monitor their comprehension (Graesser & McMahen, 1993; King, 1990; King, 1995; Spires & Donley, 1998). As a result, through question generation, teachers could identify students' reading problems and thus provide adaptive instruction (Lan & Lin, 2011).

A number of online learning systems related to question generation have been developed to support this activity, such as the *TEAMThink* program (Belanich, Wisher & Orvis, 2005), *QuestionBank* (Draaijer & Boter, 2005), *PeerWise* (Denny, Hamer, Luxton-Reilly & Purchase, 2008), *Concerto III* (Hirai, Hazeyama & Inoue, 2010), *ExamNet* (Wilson, 2004), *QPPA* (Yu, Liu & Chan, 2005), *QuARKS* (Yu, 2009), *QPIS* (Lan & Lin, 2011). Yu, Liu and Chan (2002) reported that through online question generation, students paid more attention to course reading materials, clarified their understanding of specific phrases in the texts, and utilised different avenues to generate questions. Yu, Liu and Chan (2004) collected data from questionnaires and also indicated the effectiveness of online exam question generation in enhancing students' reading comprehension. Students who undertook the multiple choice question generation activity reported that exam-question generation helped them better understand

reading materials. Yu and Liu (2005) conducted a study to examine the potential value of online multiple-choice question generation in supporting students' reading. Most of the students agreed that generating questions enabled them to master the content of the reading materials. Some students further reported that through exam-question generation, they reflected on the learning materials to evaluate their understanding. The students also reported that question generation activity helped them focus on the mistakes they made. In sum, the students indicated that their awareness, comprehension, and retention of the reading materials were fostered after going through the question generation process (Yu & Liu, 2005). Denny, Hamer, Luxton-Reilly and Purchase (2008) used a web-based question generation system named *PeerWise*, for students to generate questions and answer the peer-generated questions. They concluded that the students who engaged intensively in question generation had deeper levels of comprehension of reading materials.

Peer assessment incorporated in question generation has been proposed to ensure the quality of the questions. Peer assessment refers to the process that students serve as evaluators to provide feedback to peer-generated questions (Palincsar & Brown, 1984; Van Zundert, Sluijsmans & Van Merrienboer, 2010). Students need to spend more time reviewing and comprehending the text before generating a question. Such competitive and collaborative learning environments would enhance students' ability (Cozens, 1997; Wilson, 2004). Peer assessment fosters the interactive conversations and information exchanges in the online question generation process between question evaluators and question generators (Yu, Liu & Chan, 2003). Yu et al. (2005) found that when students discussed the generated question items, the answers to the questions, and the distracters, they could identify more detailed information, clarify their misunderstandings, and develop their deeper understanding of the reading materials. That is, when evaluating peers' questions, students were motivated to read the texts to examine if they could answer their peers' questions (Denny, Luxton-Reilly & Hamer, 2008). All in all, enhancing reading comprehension through the peer assessment method of question generation empowers the students to take more initiative for their learning through the collective effort.

Despite the potential for using online exam question generation to foster reading comprehension, several problems remain unresolved. Many scholars have argued that question generation could improve students' reading comprehension (Denny, Hamer, Luxton-Reilly & Purchase, 2008; Hirai, Hazeyama & Inoue, 2010; Wilson, 2004; Yu & Liu, 2005), but these studies have failed to disclose what reading comprehension processes are involved in question generation. Specifically, how the students comprehend reading materials to generate questions has not been explicitly investigated. Most previous studies were inclined to use students' performance on final examinations to prove the effectiveness of online question generation in enhancing reading comprehension (Wilson, 2004). Nevertheless, the lack of data regarding students' question generation processes to illustrate reading comprehension improvement makes these studies less convincing. Another problem found in previous studies is that some studies merely used students' perceptions in questionnaires (Yu, Liu & Chan, 2005) to claim that it is effective to use online question generation to enhance their reading comprehension. Unfortunately, there is a discrepancy between what students reported in questionnaires and the actual extent of their learning, as demonstrated on subsequent evaluations (Yeh & Yang, 2011). Hence, in previous studies, there are some limitations to showing through self-reported questionnaires that online question generation has a positive impact on reading comprehension.

Purpose of the study

This study aimed to investigate students' reading comprehension improvement and processes after implementing online question generation. Based on the research purposes, two research questions were addressed: 1. What was students' reading comprehension after implementing online question generation?; 2. What processes were involved in students' online question generation that affected their reading comprehension?

The instructional framework of online question generation in the present study was based on Chang, Tung and Chan's (2005) question generation model, including the organisation, composition and peer assessment modules. Students in the organisation module needed to organise vocabulary, select important sentences, identify main ideas of each paragraph, and construct a summary of texts. The composition module supports students in their generation of exam questions. In this module, students could review their previously organised vocabulary, important sentences, main ideas and summaries of texts to search for the needed information, and combine them while generating questions (Lai, 2011). In the peer assessment module, a list of guidelines and examples of how to evaluate exam questions were provided and discussed in class. Students were allowed to evaluate peers' questions, respond to peers' comments and evaluations, and revise their questions according to peers' comments.

The online question-generating system

An online *ICan* system was developed to support students undertaking question generation processes, through the organisation, composition, and peer assessment modules.

Organisation module

In the *Vocabulary* component, students could add new vocabulary through clicking *Vocabulary Helper* which automatically helped them to filter out their vocabulary, based on their pre-chosen TOEIC reading levels. With a list of the vocabulary that fits into their reading level, they could choose to use the online dictionary built into the system (see Figure 1). All of the vocabulary organised by students could be saved to form a personal vocabulary database. Students also could at any time edit the vocabulary information they had organised previously in their vocabulary databases.

In the *Sentence* component, students could select and edit the key sentences from texts (see Figure 2). The key sentences were used to help students in comprehending the text. *References* on the lower left side was built into the system to provide students a quick and convenient access to review their previously organised information. For example, students could click on *Vocabulary* from the menu to use their organised vocabulary to help them select or edit sentences.

In the *Paragraph* component, students needed to construct the main ideas of each paragraph in the text (see Figure 3). *References* on the lower left side is built into the system to allow students to review the previously organised vocabulary and sentences when writing or editing main ideas of each paragraph.
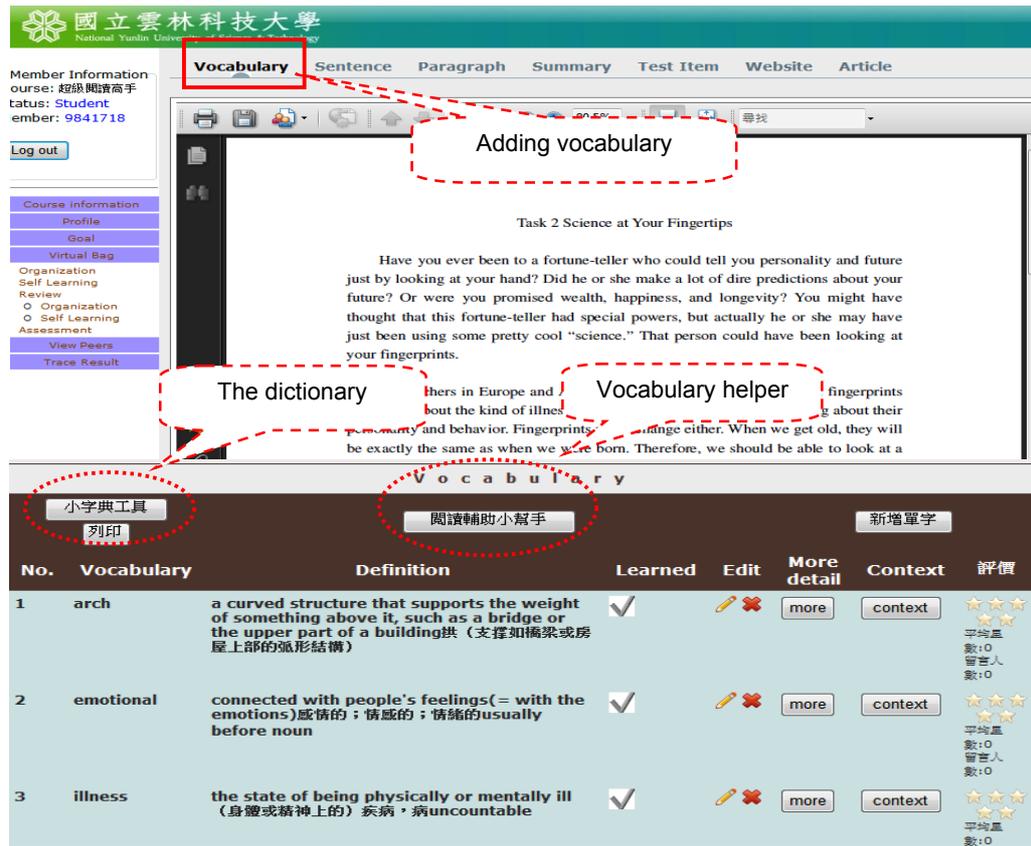
Figure 1: A student's organisation of vocabulary
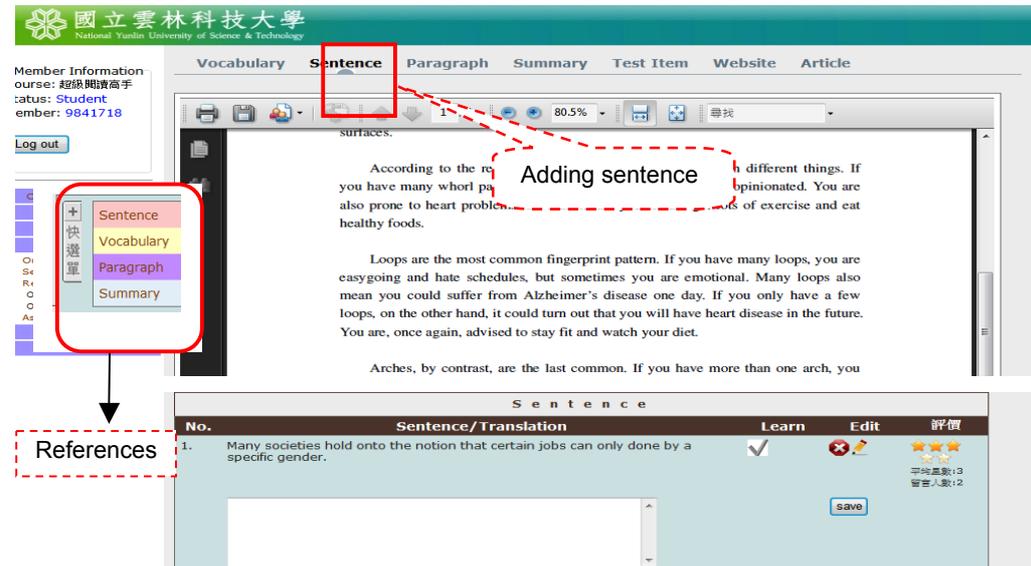

Figure 2: A students' selection of sentences
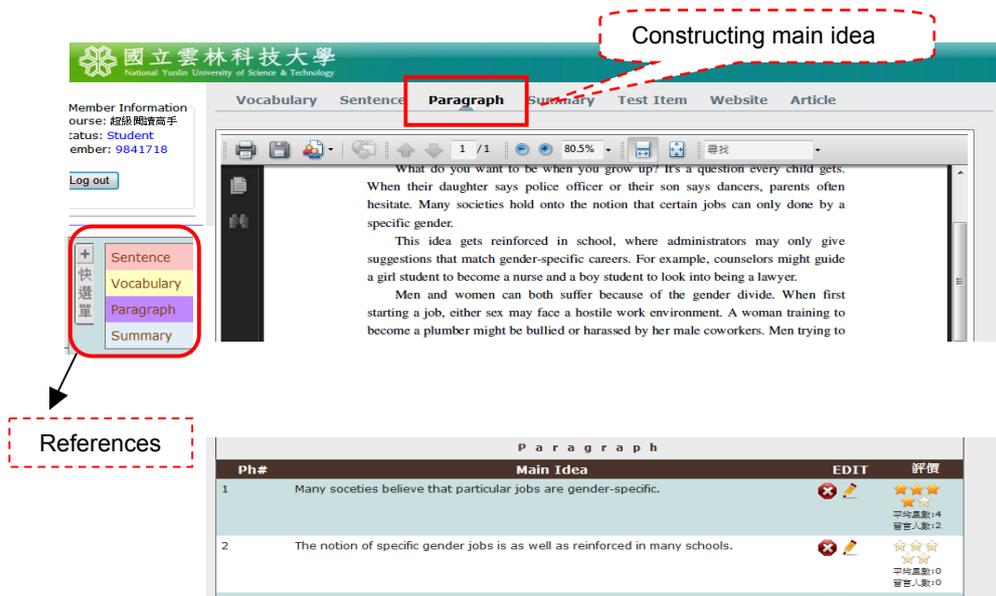
Figure 3: A student's identification of main ideas

In the *Summary* component, students were required to construct a summary from the text (see Figure 4). This activity was designed to help students learn to use previously organised vocabulary, sentences and main ideas through *References,* to construct summaries and understand the gist of the text.
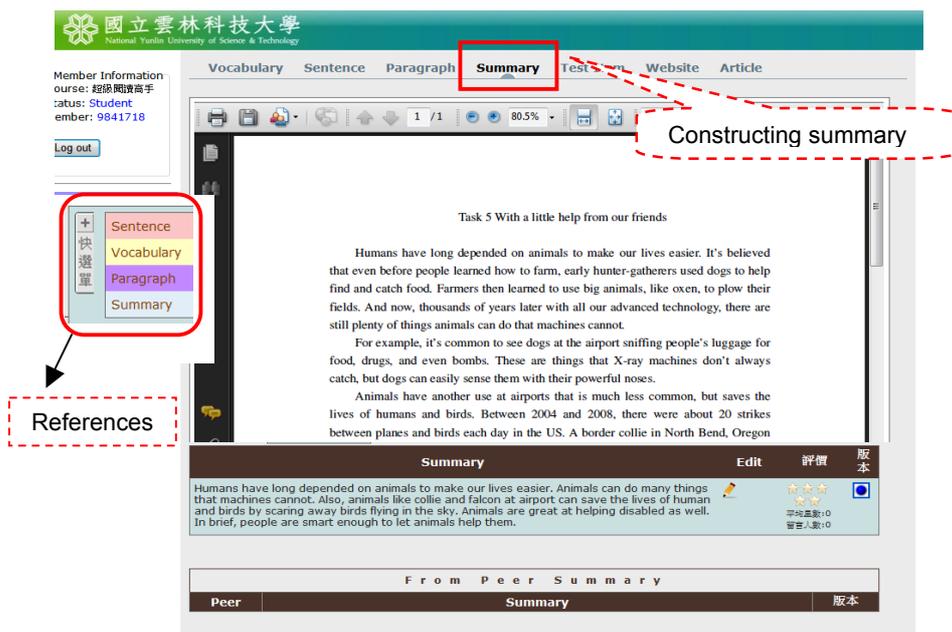

Figure 4: A student's construction of a summary

Composition module

When students click *Test Item*, they are allowed to generate questions from the text. The system provides several question formats such as multiple choice, matching, short answer, true-false and fill-in-the-blank for students to select from (Figure 5). Students can retrieve *Vocabulary, Sentence, Paragraph and Summary* from the *References* on the right side to facilitate the generating of questions. Then, students can click *View Peers* on the left side to rate peers' questions on the scales of 1-5 (Figure 5).
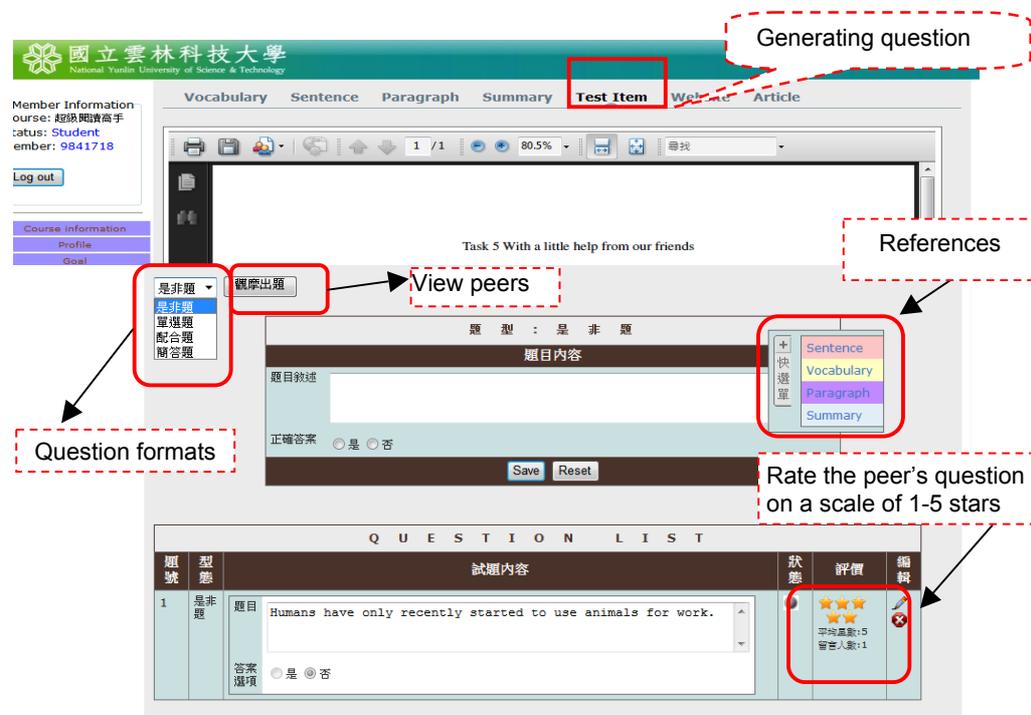

Figure 5: Functions for question generation

Peer assessment module

After generating exam questions, students could view peers' comments on their generated questions (see Figure 6) by providing evaluation on a scale of 1-5 (see Figure 5). If exam question generators disagreed with peers' ratings or comments, they could use the *Message* built into the system to ask further questions or provide their comments to their peers (see Figure 7). Then, students could revise their questions based on peers' comments or retrieve previous organised vocabulary, sentences, main ideas and summaries to rewrite the questions and answers.

In order to investigate students' question generation processes in the question generation system, their action logs in the *Trace Result* were recorded and analysed (see Figure 8). Through analysing students' action logs and behaviours, what difficulties students encountered when reading a text to generate questions, and what strategies they took to overcome their difficulties can be analysed for further instruction.
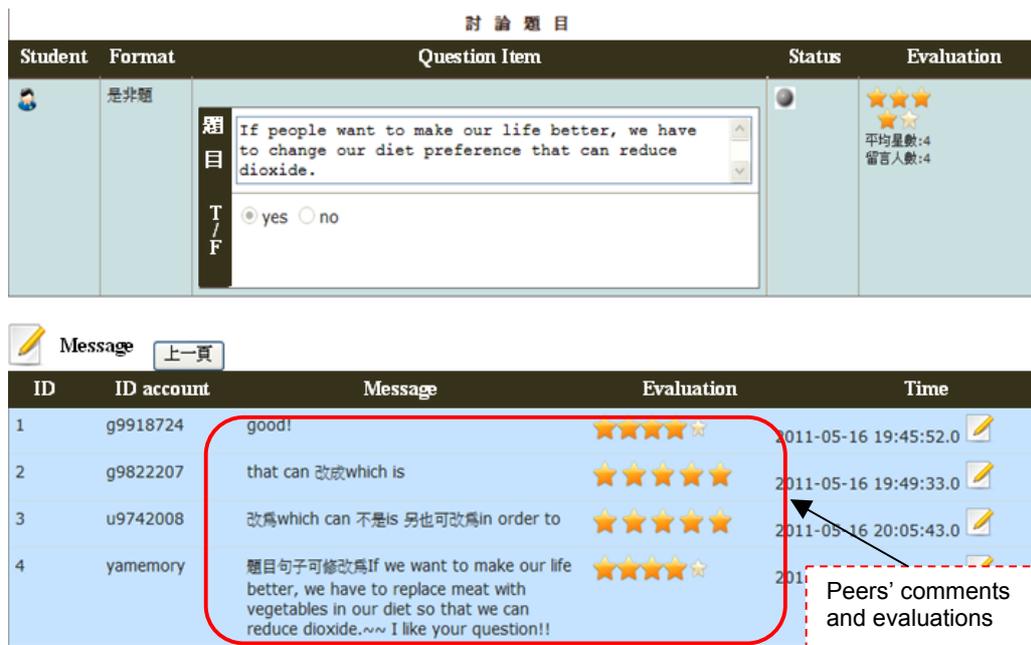
Figure 6: A student's comments on a question

Figure 7: A student's responses to a peer's comments

Method

Participants

The *English Reading Instruction Program* has been running for fifteen years to help non-English major students enhance their English ability so that they can acquire their domain knowledge from English textbooks. The participants were all non-English major students from a National University of Science and Technology in central

Taiwan. Between 80 and 100 students sign up for this program voluntarily every year. A total of 106 students registered for the year of this study and were divided into 5 classes based on their *Testing of English for International Communication* (TOEIC) reading scores. One of the classes was randomly selected for this study. A total of 19 participants in this class were recruited as participants for this study and received 11 weeks of English instruction. Six students were from the Engineering School, six from the School of Design, five from the Business School, and two were from the School of Humanities.

The 19 students were required to take pre- and post-TOEIC reading tests to investigate the extent to which they made progress in their reading. Based on the gain scores of the pre- and post TOEIC scores, students were classified into two groups: students who made more progress (the MP Group) and students who made less progress (the LP Group). The MP Group was comprised 5 students whose gain scores were in the top 25% of the participants, while the LP group included 4 students whose gain scores were in the bottom 25% of the class. Specifically, the mean gain score that the 5 students in the MP Group obtained was 44, whereas the mean gain score the 4 students obtained in the LP Group was 17. One case was selected from the MP and LP group respectively to represent how most of the students from the two groups underwent the three different modules to generate online questions in similar patterns.

Figure 8: A student's action logs

Data collection

In order to explore how online question generation had an impact on students' reading comprehension, we collected data from the pre- and post-test scores, students' action logs in the online system, and semi-structured interviews. Firstly, students were required to take a pre-test simulation of a TOEIC exam as a placement test. Secondly, in-class instruction lasted for eleven weeks, two hours per week, beginning in March 2011 and ending in June 2011. At the beginning of the course, students were instructed

to become familiar with how to use each function in the online system, to generate questions, answers, and distracters. After students became familiar with the system, and they were required to read a text while undergoing the three instructional modules for online question generation. The students needed to complete the tasks in the organisation module, composition module, and peer assessment module sequentially, in and after class. Thirdly, after eleven weeks of instruction a simulated TOEIC exam was conducted as a post test to evaluate students' progress in reading comprehension. Action logs in the trace results recorded in the system were collected to understand the students' question generating processes in the three modules. A final semi-structured interview with the focal students (the selected students from both the MP and LP groups) was conducted to explore the reasons behind the actions and to understand their reactions toward the online question generating process, thereby, obtaining more information to explain their reading processes through online question generation.

Data analysis

The data was analysed in terms of the students' pre-test and the post-test TOEIC scores, students' action logs in the trace result, students' task performance in the three modules, and the semi-structured interviews. Firstly, a t-test was adopted to examine students' reading comprehension improvement in the pre-test and the post-test simulated TOEIC exam through *Statistical Package for Social Sciences* version 12.0 (SPSS). From the students' action logs, we explored how students generated exam questions, how they interacted with peers, and how they took actions to revise their questions. The students' action logs along with students' task performance in each module were also analysed, and thus we obtained a better understanding of the students' question generation processes. The information pertaining to why the students took some actions in the question generation processes would be further analysed in the semi-structured interviews.

Students' task performance (including students' organised vocabulary, sentences, main ideas and summaries) and the interview transcripts were examined in content analysis. The four procedures of content analysis include (1) coding, (2) categorisation, (3) description, (4) interpretation (Patton, 1990). In the phase of coding, we identified the meaningful units from the students' task performance in undertaking question generation and the transcripts of interviews. In the phase of categorisation, the meaningful units were classified into different categories, such as the error types which appeared in students' task performance and the revised versions of students' tasks. In description, we proposed the main ideas of the categorised units by narrowing down and summarising the information the participants presented. In interpretation, we explained, drew conclusions, made inferences, and built linkages between the main ideas. Meetings were held weekly to discuss and negotiate on the codings. The inter-rater reliabilities with the codings of the students' task performance and transcripts of interviews were 0.87 and 0.83 respectively.

Results

Reading comprehension improvement through online question generation

A paired-sample t-test was conducted to investigate whether online question generation is an effective approach to enhance students' reading comprehension. As

shown in Tables 1 and 2, a significant difference (t=6.58, p<.05) between the pre-test (m=276, sd=10.32) and post-test (m=320, sd=7.59) was identified for the MP Group who made more progress in the post TOEIC (N=5). There was no significant difference (t=0.17, p>.05) between the pre-test (m=267, sd =15.33) and the post-test (m=281, sd=2.07) for the LP Group who made less progress in the post TOEIC (N=4). The results of the Pearson product-moment correlation coefficient between the frequency of the students' actions in the three modules of the question generation and reading comprehension are shown in Tables 3 and 4. The results highlighted that the MP Group's and LP Group's frequency of actions in online question generation had a positive relationship with their reading comprehension.

Table 1: Results of the t-test on the pre- and post-test for the MP Group

	N	Min	Max	Mean	SD	t	Sig.
Pre-test	5	253	284	276	46.52	2.585	.019
Post-test	5	297	346	320	42.39		

*p < .05

Table 2: Results of the t-test on the pre- and post-test for the LP Group

	N	Min	Max	Mean	SD	t	Sig.
Pre-test	4	244	282	267	15.33	0.17	.031
Post-test	4	268	291	281	2.07		

p > .05

Table 3: Pearson's correlation between the MP Group's frequency of actions in the three modules and reading comprehension

Number of participants	Organisation module	Composition module	Assessment module
5	.82	.80	.81

Table 4: Pearson's correlation between the LP Group's frequency of actions in the three modules and reading comprehension

Number of participants	Organisation module	Composition module	Assessment module
4	.76	.80	.75

Question generation processes in the organisation, composition, and peer assessment modules for the MP and LP students

To investigate the processes which were involved in the organisation module, composition module, and peer assessment module of the online question generation, the students' gain scores were used to classify the class into the MP and LP groups. Specifically, the MP Group included 5 students who benefited from the online question generation in the top 25% of the class, and made significant progress on average from 276 to 320 in the TOEIC tests. The LP Group included the 4 students who made less progress in the TOEIC tests, namely in the bottom 25% of the class. They only made slight progress from 267 to 281 in the TOEIC tests.

Figure 9 shows that the MP group shared a similar tendency with the LP group (see Figure 9). Compared to the LP Group, the MP Group demonstrated almost twice the frequency of actions in each module in the online system.
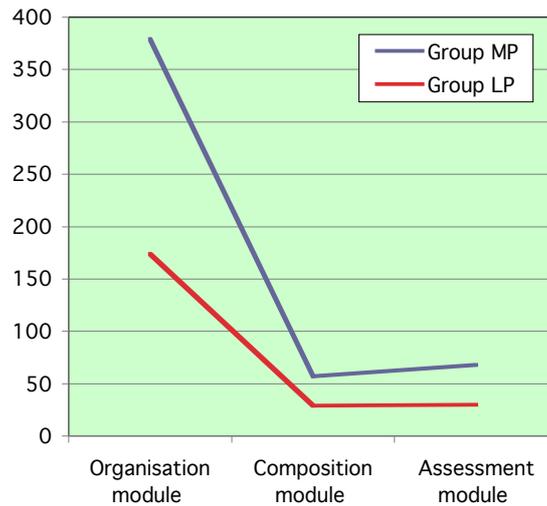
Figure 9: The MP and LP students' frequency of actions

The more progress (MP) students versus less progress (LP) students

The results in Table 5 reveal that the MP Group was more actively engaged in the online question generation processes than the LP Group. The frequency of the MP Group's actions in online question generation (m=504, sd=21.54) exceeded double the frequency of actions of the LP Group's actions (m=233, sd=27.72) by 271 instances. Particularly, the MP Group was found to have put in much more effort (m=315, sd=11.07) than the LP Group (m=174, sd=12.96) in the organisation module, which included the tasks of organising vocabulary, sentences, main ideas, and summaries.

Table 5: The MP and LP students' actions in the three modules

Participants	Instructional framework	Means of the frequency of actions	SD
MP Group	Organisation module	379	11.07
	Composition module	57	12.93
	Assessment module	68	12.68
	Total	504	21.54
LP Group	Organisation module	174	12.56
	Composition module	29	13.29
	Assessment module	30	10.34
	Total	233	27.72

The MP Group's online question generation action pattern is presented in Figure 10, and the LP Group's online question generation action pattern is shown in Figure 13. A significant difference between the MP and LP Group in undertaking online question generation was found in the action of "editing." The MP Group took more actions in retrieving the information to edit their previously organised information (as shown in action [1], [2] and [3] in Figure 10). For example, when the MP Group generated questions, they reviewed the important sentences (see Figure 12) by clicking *References* to retrieve the related information. They continued to edit the sentence that was previously organised (see Figure 13). The MP Group demonstrated their retrieval

actions of reviewing and editing in the organisation and composition modules in order to generate online questions.
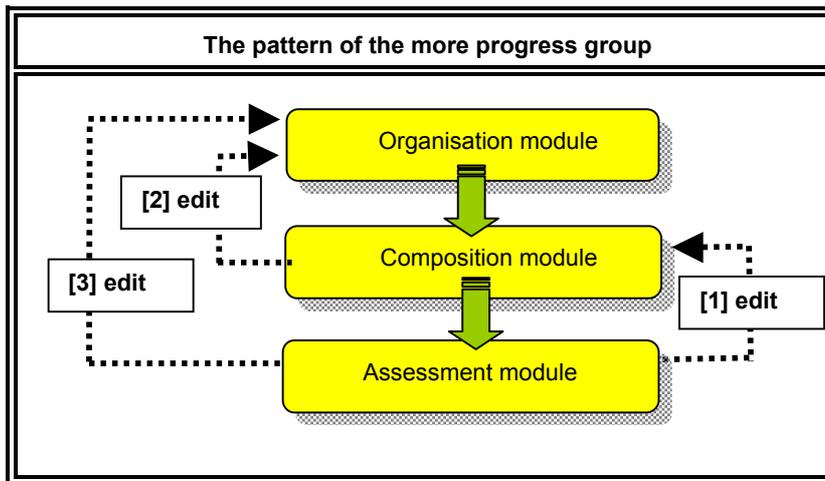
Figure 10: The MP's Group online question generation pattern

Sentence	
No.	Content
1	Humans have long depended on animals to make our lives easier.
2	Thousands of years later with all our advanced technology, there are still plenty of things animals can do that machines cannot.
3	Because humans are clever enough, they try to train animals to help them.

Figure 11: An example of the sentences reviewed by the MP Group

Previously organised sentences

Module	Activity	Action	Review a sentence	Time
Composition	Question generation	reviews	SI reviews [sentence] through tool in Task 5.	2011-05-02 19:26:45
Composition	Question generation	posts	SI posts a question [Humans have only recently started to use animals for work.] in Task 5.	2011-05-04 14:27:52
Organization	Sentence	edits	SI edits the sentence by changing [Because humans are clever enough, they try to train animals to help them.] into [**Humans are clever enough to train animals to help them.**] in Task 5.	2011-05-05 14:34:18

Edit a sentence

Figure 12: An example of the MP Group's action logs in reviewing and editing sentences

The LP Group did not have as high frequency of taking these actions as the MP group. In Figure 13, the editing action was seldom found in the LP group compared to the MP group. The LP group mainly completed the tasks in each module linearly. In other words, most of the students in the LP Group seldom went back to edit their questions or retrieve the information by clicking *References* after receiving peers' comments on their self-generated questions. It was found they seldom took further actions to either review or edit their previously organised information in each module.

The pattern of the less progress group

Organisation module

↓

Composition module

↓

Assessment module
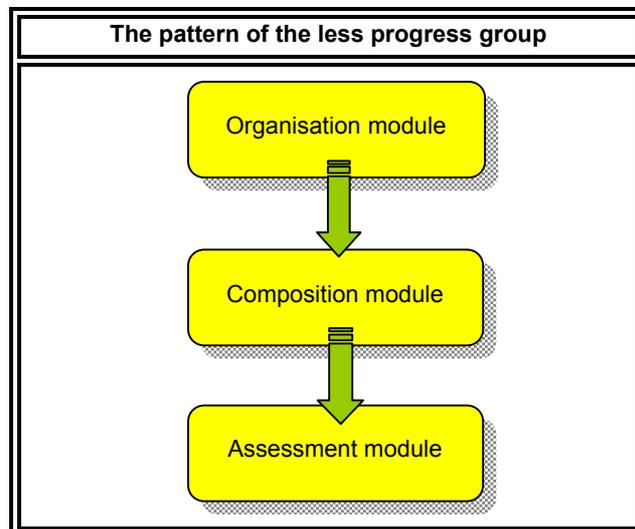
Figure 13: The LP Group's online question generation pattern

Selected cases from the MP group ('Amy') and the LP group ('Sandy')

Two students were selected as representative cases to illustrate the processes the MP Group and LP Group undertook in the organisation, composition, and assessment modules of the question generation. 'Amy' was selected to represent the MP Group, and 'Sandy' was selected for the LP Group.

In the Organisation module - vocabulary organisation
In organising vocabulary, Amy's and Sandy's actions included adding new vocabulary, querying online dictionaries, editing vocabulary, and clicking the button "more," to obtain more information on the vocabulary, such as parts of speech, synonyms, antonyms, and example sentences. The major difference between Amy and Sandy in organising vocabulary was that Amy added more vocabulary, used a dictionary, edited the vocabulary, and viewed more details. According to Table 6, Amy added 151 words while Sandy added only 75 words; Amy used the online dictionary 31 times, while Sandy used it 5 times.

In the Organisation module - sentence selection
In selecting sentences, Amy's and Sandy's actions mainly included reviewing vocabulary, posting important sentences and editing the posted sentences. The most significant difference between Amy and Sandy in selecting sentences was that Amy reviewed vocabulary 4 times and Sandy did not review the vocabulary. Amy edited the sentences 9 times, while Sandy edited the posted sentences only once (see Table 7).

Table 6: Amy's and Sandy's online actions in vocabulary

Participant	Instructional framework	Learning activity	Learning actions	Frequency
Amy	Organisation module	Vocabulary	add	151
			query	31
			edit	25
			more details	9
Sandy	Organisation module	Vocabulary	add	75
			query	5
			edit	5
			more details	3

Table 7: Amy's and Sandy's online actions in sentence selection

Participant	Instructional framework	Learning activity	Learning actions	Frequency
Amy	Organisation module	Sentence	review (vocabulary)	4
			post	39
			edit	9
Sandy	Organisation module	Sentence	review (vocabulary)	0
			post	36
			edit	1

In the Organisation module - main idea identification
In identifying main ideas, Amy and Sandy learned to retrieve previously organised vocabulary and sentences to write the main ideas, post the main ideas of each paragraph, and then edit the posted main ideas. The major difference between Amy and Sandy was the number of postings. Table 8 demonstrates that Sandy reviewed the vocabulary 8 times and Sandy did not review to identify the main ideas. In addition, Amy posted 17 main ideas, while Amy posted 7.

Table 8: Amy's and Sandy's online learning actions in main idea identification

Participant	Instructional framework	Learning activity	Learning actions	Frequency
Amy	Organisation module	Paragraph	review (vocabulary)	8
			review (sentence)	4
			post	17
			edit	9
Sandy	Organisation module	Paragraph	review (vocabulary)	0
			review (sentence)	0
			post	7
			edit	2

In the Organisation module - summary construction
In constructing a summary, the actions of Amy and Sandy included reviewing previously organised vocabulary and sentences, as well as recording main ideas to write the summary. Also, Amy and Sandy posted a summary and edited the summary. Based on Table 9, while constructing her summary, Amy actively reviewed the vocabulary, sentences, and paragraphs to construct her summary, whereas Sandy took fewer actions in reviewing the previously organised information. Amy edited her summary 6 times, whereas Sandy edited her summary only once.

Table 9: Amy's and Sandy's online learning actions in the summary

Participant	Instructional framework	Learning activity	Learning actions	Frequency
Amy	Organisation module	Summary	review (vocabulary)	5
			review (sentence)	9
			review (paragraph)	3
			post	4
			edit	6
Sandy	Organisation module	Summary	review (vocabulary)	2
			review (sentence)	3
			review (paragraph)	1
			post	4
			edit	1

In the Composition module - question generation

In generating questions, Amy and Sandy's actions included reviewing previously organised vocabulary, sentences, main ideas, and the summary. Amy and Sandy also posted and edited questions with different formats (multiple choice, matching, short answer, true-false and fill-in-the-blank). As shown in Table 10, a notable difference was found between Amy and Sandy's actions in generating questions. Amy used previously organised vocabulary, sentences, main ideas, and the summary, in total 45 times, while Sandy used the information to generate questions a total 13 times. Additionally, Amy used previously organised vocabulary, sentences, main ideas, and summary to generate questions, whereas, Sandy merely retrieved summary to generate questions.

Table 10: Amy's and Sandy's online learning actions in question generation

Participant	Instructional framework	Learning activity	Learning actions	Frequency
Amy	Composition module	Question generation	review (vocabulary)	12
			review (sentence)	10
			review (paragraph)	8
			review (summary)	15
			post	6
			edit	1
Sandy	Composition module	Question generation	review (vocabulary)	0
			review (sentence)	1
			review (paragraph)	0
			review (summary)	12
			post	6
			edit	1

In the Assessment module - peer assessment

In undertaking peer assessment, Amy and Sandy were required to view peers' questions and then give comments and evaluations of their peers' questions. Regarding peer assessment, Amy demonstrated greater engagement in undertaking peer assessment than Sandy. As shown in Table 11, Amy viewed peers' test items 36 times, gave peers' questions comments 5 times, read peers' comments 5 times, and responded to peers' comments 2 times. However, Sandy viewed peers' test items 28 times, gave peers' questions comments 3 times, read peers' comments once, and never responded to peers' comments. Hence, the results in Table 11 indicate that Sandy was not as actively engaged as Amy in peer assessment.

Table 11: Amy and Sandy's online actions in peer assessment

Participant	Instructional framework	Learning activity	Learning actions	Frequency
Amy	Peer assessment module	Peer assessment	view peers' questions	36
			give comments	5
			read	5
			respond	2
Sandy	Peer assessment module	Peer assessment	view peers' questions	28
			give comments	3
			read	1
			respond	0

Amy and Sandy's motivations were revealed in the interviews. Amy viewed peers' questions in order to give comments, and generate her own questions. She also noted that sometimes her questions were similar to her peers', but her correct answers to the questions were different from those of her peers'. In these cases, she reread the article and realised that she had misunderstood the text, so she edited her questions. Amy stated,

> I viewed my peers' questions to learn how I can generate questions and how to give peers comments. I often reviewed peers' questions again before the quiz. I benefited from peer review through examining my peers' questions. By doing so, I could clarify my understanding of the text, and reflected on the problems I encountered when giving peers' comments. While I view peers' questions and comments to me, sometimes I did not agree with those comments. I would constantly go back to read again to verify my understanding. So for the discrepancy for understanding, I need to read and reread several times.

When disagreement occurred, Amy would reread the text and confirm the correct questions and answers. While reviewing and evaluating peers' online questions, she read and reread the content to enhance her comprehension. The following figures show that Amy found she misunderstood the text from viewing peers' questions. As shown in Figure 14 and 15, Amy generated the question "If we want to protect our planet, trying a vegetarian diet is a good way to start," providing the answer "true" to the question. When she viewed her peer's question (Figure 16), she found it to be similar to her own. Nevertheless, her answer "trying a vegetarian diet" was different from her peer's "eating less meat." After rereading the text, she found that her peer's answer might be more accurate, and then she edited the question by changing it into "If we want to protect our planet, then trying to eat extra vegetables more than meat is a good way to start" (Figures 15 and 17).

Version	Format	Content	Time
Version 1	True/false	Question_ If we want to protect our planet, trying a vegetarian diet is a good way to start. Answer ⊙ True ○ False	2011-05-16 14:26:45

Figure 14: Amy's original question

Post a question (first version)

Module	Activity	Action	Content	Time
Assessment	Question generation	posts	[**If we want to protect our planet, trying a vegetarian diet is a good way to start**.] in Task 6.	2011-05-16 14:26:45
Assessment	Peer assessment	views	Amy views peer Ivy's question [According to the article, what is the best way to protect our earth?] in Task 6.	2011-05-16 14:27:52
Assessment	Question generation	edits	[**If we want to protect our planet, then trying to eat extra vegetables more than meat is a good way to start.**] in Task 6.	2011-05-16 14:40:01

View a peer's question

Edit a question (the revised version)
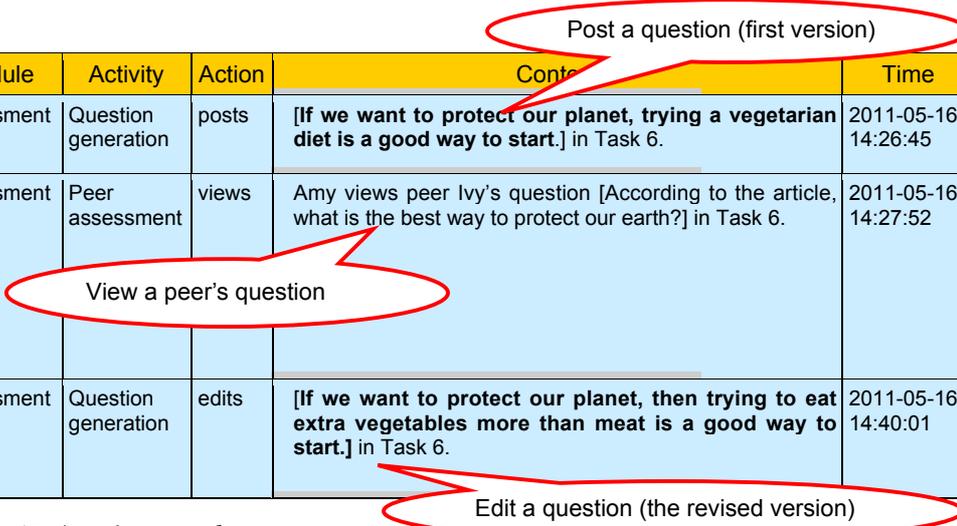
Figure 15: Amy's action logs

Version	Format	Content	Time
Version 1	Multiple choice	Question: According to the article, what is the best way to protect our earth? Answer ○ Take public transport more ◉ Reduce meat consumption ○ Become a vegetarian ○ Fight against deforestation	2011-05-16 00:49:08

Figure 16: Amy viewed a question from her peer

Version	Format	Content	Time
Version 2	True/false	Question_ If we want to protect our planet, then trying to eat extra vegetables more than meat is a good way to start. Answer ◉ True ○ False	2011-05-16 14:40:01

Figure 17: Amy's revised version of her question

Sandy, on the other hand, claimed she viewed peers' question items because she wanted to quickly understand what kinds of questions her peers had generated. Sandy explained,

> I viewed peer questions because I wanted to see what types of questions my peers came up with. I also viewed peers' questions when I need to evaluate the quality of my peers' questions. However, I seldom read and responded to peers' comments because peers' comments were usually too general to help me improve my questions. And also my vocabulary size is limited. The overwhelming vocabulary intimidated me to view and give comments to my peers. I don't feel comfortable writing in English either.

Sandy further explained that when peers' comments were not specific and helpful, she lacked the motivation to read or respond to peers' comments. Additionally, she admitted that her limited vocabulary size discouraged her from giving and responding to her peers' questions. She did not have the confidence in her writing to provide comments upon her peers' questions.

Discussion and conclusion

Major findings can be summarised as follows. First, the result of the t-test showed that there was a significant difference in students' reading comprehension improvement after implementing the online question generation system. The finding was in line with previous studies which suggest that question generation can enhance the level of understanding of the reading materials (Barak & Rafaeli, 2004; Belanich, Wisher & Orvis, 2005; Cohen, 1983; Dreher & Gambrell, 1985; Yu & Liu, 2005; Yu, Liu & Chan, 2005).

This study aimed to look into the different processes the MP and LP groups had undertaken through online question generation, in terms of the organisation, composition, and peer assessment modules. In the organisation module, the results also showed that the MP Group engaged in a higher frequency of editing their previously organised vocabulary, sentences, main ideas, and summaries. Such findings might imply that the MP Group was equipped with a higher level of metacognitive skill to monitor the question generation processes by retrieving related information, and take remedial actions to revise the previously organised vocabulary, sentences, main ideas, and summaries. In the composition module, the MP Group also had a higher frequency in reviewing the previously organised information to generate online questions. In other words, the review of the previously organised information while generating questions was helpful in enhancing students' reading comprehension. In the peer assessment module, the MP Group demonstrated greater involvement in undertaking peer assessment. They were much more active in viewing peers' questions, providing comments on peers' questions, reading and responding to peers' comments. Allowing students to read their peers' questions and answers is effective in encouraging students to reread the texts and re-examine their understanding, especially when students disagree with their peers on the questions and answers.

Amy and Sandy were selected from the MP group and LP group respectively to illustrate how students from the two groups undertook the three different modules to construct self-generated questions. Specifically, in the organisation module, Amy made a lot of effort in adding to the vocabulary and editing the previously organised vocabulary. These findings are in agreement with the results of previous studies which

indicated that breadth and depth of vocabulary knowledge is essential for students to better comprehend a text (Mehrpour & Rahimi, 2010; Qian, 2002). The breadth of vocabulary refers to the numbers of words that a student needs to acquire to comprehend a text (Qian, 1999; Nation, 2001). The depth of the vocabulary means that students need to know how to identify the appropriate word meaning to fit a context and how to use synonyms and antonyms (Qian, 1999). From Amy's action logs, she added vocabulary into her vocabulary databases and edited the previously organised vocabulary by adding synonyms or phrases to expand her vocabulary knowledge. Thus, the improvements in Amy's vocabulary knowledge might be a factor that led to her progress in reading comprehension. In order to generate questions, Amy acquired and learned extensively vocabulary from her reading.

Apart from making an effort in organising vocabulary, it was found that Amy edited the previously organised vocabulary, sentences, main ideas, and summaries more frequently. That is, the students who made progress in the post-test usually monitored their comprehension and took action to revise the information they previously organised. The results indicate that in the question generation processes, when students continuously monitored, evaluated and self-regulated their reading (Wong, 1985), their reading comprehension could be enhanced.

With respect to the composition module, Amy made more progress in her reading comprehension and had a higher frequency than Sandy in reviewing the previously organised information to generate online questions. From Amy's action logs and interview data, we learned that she not only reviewed previously organised information, but also reviewed the text and verified her understanding to generate questions. The findings suggest that generating questions from a text facilitates students to activate their prior knowledge, and connect it to the text (Miciano, 2002; Wong, 1985), and also engages them in a deeper processing of text material (Craig & Lockhart, 1972).

Regarding the peer assessment module, Amy demonstrated that she was more engaged in the processes of peer assessment than Sandy. This finding suggests that the peer assessment module plays a pivotal role in improving students' reading comprehension through online question generation. In addition, Amy was much more actively engaged in viewing peers' questions, providing comments on peers' questions, and reading and responding to peers' comments. The results showed that students gained greater understanding and retention of the text when they actively engage in peer review in their question generation processes (Wong, 1985). It can be asserted that students' engagement in peer assessment is a critical factor in improving reading comprehension through question generation. Allowing students to read the questions and the answers generated by their peers is effective in helping them verify and cross-examine their understanding. It echoed the findings in Belanich, Wisher and Orvis (2005) as well as Yu, Liu and Chan (2005), in that peer assessment in the question generation processes served as the stimulus for students to revisit the texts and evaluate their understanding, when they were not sure of the appropriateness and the correctness of the answers generated by their peers. As a result, intensive engagement in peer assessment is crucial for question generation and reading comprehension improvement.

Several implications can be drawn from this study. Teachers should encourage students to generate questions from those tasks designed in the three modules. While students are engaged in the process of question generation, they take an active role in

comprehending the texts through organising, composing, and assessing each other's questions. Different from answering teacher generated questions, students assume the roles of question designers who are engaged in a high level of cognitive function to test their peers' understanding and in using their self-regulatory cognitive strategy (Palincsar & Brown, 1984). The processes of organising, composing, and assessing each other's questions deepen students' reading comprehension and allow students to have more transactions with the texts. Question generation can be deemed as both summative and formative assessment to examine students' reading comprehension and reading processes, while their questions themselves and the processes involved reveal a lot of information about their comprehension. Teachers can tap into students' comprehension breakdown and provide adaptive instruction by evaluating their self-generated questions.

Although the findings of this study provide an insight into potential enhancements for students' reading comprehension through online question generation, the study has some limitations. The sample size is relatively small, as only nineteen English college students were recruited as participants. Therefore, findings obtained through the online question generation system might not be assuredly generalised to all EFL college students' reading comprehension. Several important processes were involved in the question generating processes so as to enhance students' reading comprehension. However, specifically, the relationship between these processes identified in the three modules was not investigated in this study. Further studies could look into the relationship between the organisation, composition, and peer assessment processes involved in question generation.

References

Barak, M. & Rafaeli, S. (2004). On-line question-posing and peer-assessment as means for web-based knowledge sharing in learning. *International Journal of Human-Computer Studies*, 61(1), 84-103. http://dx.doi.org/10.1016/j.ijhcs.2003.12.005

Belanich, J., Wisher, R. A. & Orvis, K. L. (2005). Using a question generation approach to improve web-based collaborative learning. [viewed 8 Jun 2010] http://www.uwex.edu/disted/conference/Resource_library/proceedings/03_51.pdf

Bristow, S. P. (1985). Are poor readers passive readers? Some evidence, possible explanations, and potential solutions. *The Reading Teacher,* 39(3), 318-325. http://www.jstor.org/discover/10.2307/20199080

Chang, S. B., Tung, K. J. & Chan, T. W. (2005). A learning strategy of student question generation to enhance comprehension of learning materials in digital classroom environment. *Proceedings of the Fifth IEEE International Conference on Advanced Learning Technologies Comprehension of Learning Materials in Digital Classroom Environment.* http://www.computer.org/portal/web/csdl/doi/10.1109/ICALT.2005.12

Chien, Y. (2000). *The effect of predictable stories on oral and emergent literacy development: A descriptive of EFL beginners in kindergarten.* Unpublished doctoral dissertation, Tamkang University, Taiwan [in Chinese]. http://ndltd.ncl.edu.tw/cgi-bin/gs32/gsweb.cgi/login?o=dnclcdr&s=id=%22088TKU00154002%22.&searchmode=basic

Craig, F. & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior,* 11(6), 671-684. http://dx.doi.org/10.1016/S0022-5371(72)80001-X

Cohen, R. (1983). Self-generated questions as an aid to reading comprehension. *The Reading Teacher,* 36(8), 770-775. http://www.jstor.org/stable/20198324

Cozens, P. (1997). Student created tests as motivation to learning. Presented at *The First Annual Colloquium on Current Trends in English Language Testing*, UAE University Islamic Institute, Al Ain, United Arab Emirates.

Day, R. R. & Bamford, J. (1996). *Extensive reading in the second language classroom.* New York: Cambridge University Press.

Denny, P., Hamer, J., Luxton-Reilly, A. & Purchase, H. (2008). PeerWise: Students sharing their multiple choice questions. In *ICER '08: Proceedings of the fourth international workshop on Computing education research*. ACM, New York, pp. 51-58. http://dx.doi.org/10.1145/1404520.1404526

Dlugosz, D. W. (2000). Rethinking the role of reading a foreign language to young learners. *English Language Teaching Journal,* 54(3), 284-290. http://dx.doi.org/10.1093/elt/54.3.284

Draaijer, S. & Boter, J. (2005). Questionbank: Computer supported self-questioning. In Danson, M. (Ed.), *9th CAA International Computer Assisted Assessment Conference*. (pp.235-250). Loughborough, UK. https://dspace.lboro.ac.uk/dspace-jspui/handle/2134/1993

Draper, S. W. (2009). Catalytic assessment: Understanding how MCQs and EVS can foster deep learning. *British Journal of Educational Technology,* 40(2), 285-293. http://dx.doi.org/10.1111/j.1467-8535.2008.00920.x

Dreher, M. J. & Gambrell, L. B. (1985). Teaching children to use a self-questioning strategy for studying expository prose. *Reading Improvement,* 22, 2-7.

Dreyer, C. & Nel, C. (2003). Teaching reading strategies and reading comprehension within a technology-enhanced learning environment. *System,* 31(3), 349-365. http://dx.doi.org/10.1016/S0346-251X(03)00047-2

Fellenz, M. (2004). Using assessment to support higher level learning: The multiple choice item development assignment. *Assessment and Evaluation in Higher Education,* 29(6), 703-719. http://dx.doi.org/10.1080/0260293042000227245

Gagné, R. M. (1985). *The conditions of learning and theory of instruction*. Fourth Ed. New York: Holt, Rinehart and Winston.

Gillespie, C. (1990). Questions about students-generated questions. *Journal of Reading,* 34(4), 250-257. http://www.jstor.org/discover/10.2307/40014543

Graesser, A. C. & McMahen, C. L. (1993). Anomalous information triggers questions when adults solve quantitative problems and comprehend stories. *Journal of Educational Psychology*, 85(1), 136-151. http://psycnet.apa.org/index.cfm?fa=buy.optionToBuy&id=1993-20263-001

Hirai, Y., Hazeyama, A. & Inoue, T. (2009). Assessment of learning in Concerto III: A collaborative learning support system based on question-posing. *Proceedings of the 12th IASTED International Conference on Computers and Advanced Technology in Education* (CATE2009), 36-43. St Thomas, US Virgin Islands. http://inolab.slis.tsukuba.ac.jp/publication/paper/cate2009hirai.pdf

Horgen, S. A. (2007). Pedagogical use of multiple choice tests – Students create their own tests. *Proceedings of the Informatics Education Europe II Conference*. [verified 28 Jul 2012] http://www.aitel.hist.no/~svendah/FoU/PPT/ped-use-of-tests-FINAL.pdf

King, A. (1990). Enhancing peer interaction and learning in the classroom through reciprocal questioning. *American Educational Research Journal*, 27, 664-687. http://dx.doi.org/10.3102/00028312027004664

King, A. (1994). Autonomy and question asking: The role of personal control in guided student-generated questioning. *Learning and Individual Differences*, 6(2), 163-185. http://dx.doi.org/10.1016/1041-6080(94)90008-6

King, A. (1995). *Cognitive strategies for learning from direct teaching*. In E. Wood, V. Woloshyn & T. Willoughby (Eds.), *Cognitive strategy instruction for middle and high schools* (pp. 18-65). Cambridge, MA: Brookline.

Lai, P. Y. (2011). *Enhancing college students' English reading comprehension through online question generation*. Unpublished master's thesis, National University of Science and Technology. http://ethesys.yuntech.edu.tw/ETD-db/ETD-search/view_etd?URN=etd-0825111-114024

Liu, P. L., Chen, C. J. & Chang, Y. J. (2010). Effects of a computer-assisted concept mapping learning strategy on EFL college students' English reading comprehension. *Computers & Education,* 54(2), 436-445. http://dx.doi.org/10.1016/j.compedu.2009.08.027

Lan, Y. F. & Lin, P. C. (2011). Evaluation and improvement of student's question-posing ability in a web-based learning environment. *Australasian Journal of Educational Technology,* 27(4), 581-599. http://www.ascilite.org.au/ajet/ajet27/lan.html

Liang, L. A., & Dole, J. A. (2006). Help with teaching reading comprehension: Comprehension instructional frameworks. *The Reading Teacher*, 59(8), 742-753. http://www.jstor.org/stable/20204415

Mehrpour, S. & Rahimi, M. (2010). The impact of general and specific vocabulary knowledge on reading and listening comprehension: A case of Iranian EFL learners. *System,* 38(2), 292-300. http://dx.doi.org/10.1016/j.system.2010.01.004

Miciano, R. Z. (2002). Self-questioning and prose comprehension: A sample case of ESL reading. A*sia Pacific Education Review,* 3(2), 210-216. http://dx.doi.org/10.1007/BF03024914

Nation, I. S. P. (2001). *Learning vocabulary in another language.* Cambridge: Cambridge University Press. http://catdir.loc.gov/catdir/samples/cam031/2001269892.pdf

Palincsar, A. S. & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1(2), 117-175. http://dx.doi.org/10.1207/s1532690xci0102_1

Patton, M. Q. (1990). *Qualitative evaluation and research methods.* (2nd ed) California: Sage.

Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review,* 56(2), 282-308. http://utpjournals.metapress.com/content/38u6g3t17443j441/fulltext.pdf

Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning,* 52(3), 513-536. http://dx.doi.org/10.1111/1467-9922.00193

Rafaeli, S., Barak, M., Dan-Gur, Y. & Toch, E. (2004). QSIA – a Web-based environment for learning, assessing and knowledge sharing in communities. *Computers & Education*, 43(3), 273-289. http://dx.doi.org/10.1016/j.compedu.2003.10.008

Salinger, T. (2003). Helping older, struggling readers. *Preventing School Failure,* 47(2), 79-85. http://dx.doi.org/10.1080/10459880309604434

Singer, H. (1978). Active comprehension: From answering to asking questions. *The Reading Teacher*, 31(8), 901-908. http://www.jstor.org/discover/10.2307/20194673

Singer, H. & Donlan, D. (1982). Active comprehension: Problem-solving schema with question generation for comprehension of complex short stories. *Reading Research Quarterly,* 17(2), 166-185. http://www.jstor.org/stable/747482

Spires, H. A. & Donley, J. (1998). Prior knowledge activation: Inducing engagement with informational texts. *Journal of Educational Psychology,* 90(2), 249-260. http://psycnet.apa.org/doi/10.1037/0022-0663.90.2.249

Tal, N. F., Siegel, L. S. & Maraun, M. (1994). The role of question type and reading ability in reading comprehension. *Reading and Writing,* 6(4), 387-402. http://dx.doi.org/10.1007/BF01028850

Van Keer, H. (2005). Fostering reading comprehension in fifth grade by explicit instruction in reading strategies and peer tutoring. *British Journal of Educational Psychology*, 74(1), 37-70. http://dx.doi.org/10.1348/000709904322848815; also http://www.onderwijskunde.ugent.be/downloads/Van%20Keer.pdf

Van Zundert, M., Sluijsmans, D. & van Merrienboer, J. (2010). Effective peer assessment processes: Research findings and future directions. *Learning and Instruction*, 20(4), 270-279. http://dx.doi.org/10.1016/j.learninstruc.2009.08.004

Wilson, E. V. (2004). ExamNet asynchronous learning network: Augmenting face-to-face courses with student-developed exam questions. *Computers & Education,* 42(1), 87-107. http://dx.doi.org/10.1016/S0360-1315(03)00066-6

Wong, B. Y. L. (1985). Self-questioning instructional research: A review. *Review of Educational Research*, 55, 227-268. http://dx.doi.org/10.3102/00346543055002227

Yang, Y. F. (2010). Developing a reciprocal teaching/learning system for college remedial reading instruction. *Computers & Education*, 55, 1193-1201. http://dx.doi.org/10.1016/j.compedu.2010.05.016

Yeh, H. C. & Yang, Y. F. (2011). Metacognitive process in online text construction. *Educational Technology and Society*, 14(3), 82-101. http://www.ifets.info/journals/14_3/8.pdf

Yu, F. Y., Liu, Y. H. & Chan, T. W. (2002). The efficacy of a web-based domain independent question-posing and peer assessment learning system. In *ICCE '02 Proceedings of the International Conference on Computers in Education*. http://dl.acm.org/citation.cfm?id=838238.838932

Yu, F. Y., Liu, Y. H. & Chan, T. W. (2004). A networked question-posing and peer assessment learning system: A cognitive enhancing tool. *Journal of Educational Technology Systems*, 32(2&3), 211-226. http://dx.doi.org/10.2190/2RNG-WM2B-XEVM-4EP3; also at http://chan.lst.ncu.edu.tw/publications/2003-2004-Yu-anq.pdf

Yu, F. Y., Liu, Y. H. & Chan, T. W. (2005). A web-based learning system for question posing and peer assessment. *Innovations in Education and Teaching International,* 42(4), 337-348. http://dx.doi.org/10.1080/14703290500062557

Yu, F. Y. & Liu, Y. H. (2005). Potential values of incorporating a multiple-choice question construction in physics experimentation instruction. *International Journal of Science Education*, 27(11), 1319-1335. http://dx.doi.org/10.1080/09500690500102854

Yu, F. Y. (2009). Scaffolding student-generated questions: Design and development of a customizable online learning system. *Computers in Human Behavior*, 25(5), 1129-1138. http://dx.doi.org/10.1016/j.chb.2009.05.002

Yu, F. Y. (2011). Multiple peer-assessment modes to augment online student question-generation processes. *Computers & Education*, 56(2), 484-494. http://dx.doi.org/10.1016/j.compedu.2010.08.025

Vygotsky, L. S. (1978). *Mind in society.* Cambridge, MA: Harvard University Press.

Authors: Hui-Chin Yeh and Pei-Yi Lai
Graduate School of Applied Foreign Languages
National Yunlin University of Science and Technology
123, University Road Section 3, Douliu, Yunlin, Taiwan
Email: hyeh@yuntech.edu.tw
Web: http://www.dafl.yuntech.edu.tw/program.php?lang=en