Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 13. Number3. September 2022 Pp.236-253
DOI: https://dx.doi.org/10.24093/awej/vol13no3.15

A Case Study on the Effectiveness of Applying Content and Language Integrated Learning in an Artificial Intelligence English Reading Course

Jingwei Tang
Graduate School of Human Sciences
Assumption University. Bangkok, Thailand
Corresponding Author: 50745497@qq.com

Joseph Foley
Graduate School of Human Sciences
Assumption University. Bangkok, Thailand

Received: 07/06/2022 Accepted: 08/27/2022 Published:09/24/2022

Abstract
The present study was a case study on the application of Content and Language Integrated Learning (CLIL) for an Artificial Intelligence (AI) English reading course in the Chinese general English education (GEE) context. It aimed to investigate the effectiveness of CLIL in improving EFL learners' AI general knowledge. It was significant for the practice of applying CLIL in China and provided some recommendations for the teaching of AI English. This study proposed two research questions: 1. What are the effects of CLIL in improving students' AI general knowledge?; 2. How do students perceive CLIL? An Artificial Intelligence English reading course was conducted as a teaching experiment in a GEE class of 45 students at Chongqing Technology and Business University, China. Both quantitative and qualitative methods were applied in data collection and data analysis. Research results indicated that Content and Language Integrated Learning effectively improved EFL learners' AI general knowledge and was a teaching method applicable to most students taking part in the teaching experiment. CLIL was worthy of being further implemented and studied in the Chinese GEE context. Besides, the research results also suggested that AI-related content was feasible in Chinese GEE. Students demonstrated the need to learn both general and professional AI knowledge. AI English was a subject that needs to be further explored.

Keywords: Artificial Intelligence, Artificial Intelligence English reading course, Artificial Intelligence general knowledge, Chinese general English education, Content and Language Integrated Learning

Cite as: Tang, J., & Foley. J. (2022). A Case Study on the Effectiveness of Applying Content and Language Integrated Learning in an Artificial Intelligence English Reading Course . *Arab World English Journal, 13* (3) 236-253. DOI: https://dx.doi.org/10.24093/awej/vol13no3.15

Introduction

Chinese college English education has been undergoing a profound educational reform. General English education (GEE), as a component of general education and college English education in China, is the orientation of College English Teaching Reform. The emphasis of college English teaching is turning from the practice and improvement of English language skills such as listening, speaking, reading, writing, and translation to GEE, English for Specific Purposes (ESP), and English for Academic Purpose (EAP). This new educational idea focuses more on cultivating students' social and cultural background in a foreign language, English language used in a specific industry, and English for academic writing and communication purposes.

Among all the GEE courses colleges and universities have provided to undergraduates, English concerning Artificial Intelligence (AI) is nearly a total blank. AI is a cross-edge discipline of natural science and social science. Its fast development and highly-internationalized cooperation ask for talents both with AI background or AI professional knowledge and skills and with English ability that can satisfy their needs for international communication. In this situation, English education should also probe into the possibility and feasibility of involving AI-related content in English language teaching. But up to now, few language textbooks were related to AI, nor has any exploration of English for AI purposes made in China, whether in the field of GEE, ESP, or EAP. The gap between English education and the development of AI economy and AI science and technology is becoming increasingly apparent. It is possible and necessary to explore the integration of AI-related knowledge and English learning, especially in the context of Chinese GEE. For the lack of AI-related courses in the Chinese GEE context, the present study was led by the researcher as an empirical and exploratory study into English for AI purposes under the guidance of CLIL. It hoped to fill the gap of lacking AI-related English courses in the Chinese GEE and further explored the application of CLIL in China. The researcher proposed the notion of 'AI English' as opposed to Business English, Legal English, Tourism English, Agriculture English, etc. To explore the effectiveness of applying CLIL in teaching an AI English Reading course in the domain of GEE, this study involved the implementation of an AI English Reading course and proposed two researcher questions which could be formulated as 1. What are the effects of CLIL in improving students' AI general knowledge?; 2. How do students perceive CLIL?

Literature Review

Content and Language Integrated Learning (CLIL) is a language teaching ideology that originated in Europe in the 1990s. The Council of Europe comes up with the term 'CLIL' firstly in its Division of Language Policies, where Professor David Marsh proposes that CLIL refers to situations where subjects, or parts of subjects, are taught through a foreign language with dual-focused aims, namely, the learning of content and the simultaneous learning of a foreign language. It is an alternative to Content-Based Instruction (CBI), which has gained more popularity in the United States and Canada, since they are similar (Lorenzo, 2007; Mi, 2015). CLIL is a dual-

focused teaching method incorporating language into disciplinary content (Coyle, 2007). Thus, it has been thought optimal for second language learning (Harley, Allen, Cummins, & Swain, 1990). The best approach to learning a foreign language is not to study the language itself but to learn the language through the contents conveyed by a specific language (Wolff, 2003). In ELT, CLIL integrates a specific disciplinary or interdisciplinary content into the teaching of English as a foreign language. It represents a significant departure from traditional foreign language teaching methods in that language proficiency is achieved by shifting the focus of instruction from the learning of language to the learning of language through the study of subject matter (Stryker & Leaver, 1997).

Scholars have made studies on CLIL from the angles of the prominence of language objective and content objective (Snow et al., 1989; Short, 2002; Bigelow, 2010; Baecher, Farnsworth, & Ediger, 2014), classroom discourse analysis, and teacher-student interactions and instructions (Gibbons, 2003; Liebscher & Dailey-O'Cain, 2005; Hoare, 2010; Evnitskaya & Morton, 2011; Huang, 2011; Jakonen & Morton, 2015), teaching strategies to promote CLIL inherent patterns (Lingley, Nunn, & Otlowski, 2006; Lorenzo, 2007; Lyster, 2007; Schleppegrell, Achugar, & Orteíza, 2012; Short, 2012), and CLIL teacher development (Arkoudis, 2005; Echevarria, 2009; Cammarate & Tedick, 2012; Short et al., 2012). Chinese scholars have mainly studied the effectiveness of CLIL and the development of learners' language skills (Chang, Liu, & Deng, 2009; Yuan, 2012).

In the Chinese foreign language teaching context, CLIL has been gaining more and more attention from educators. It has been applied by many teachers in teaching college English. CLIL teaching model can provide instruction in College English Teaching Reform in China, and it is necessary and feasible to implement CLIL in college English education. This study adopted CLIL for the teaching of the AI English Reading course. It was a localized application of CLIL in teaching Chinese GEE courses.

Method

This study used both the quantitative data analysis method and the qualitative data analysis method. All data were collected from two English proficiency tests, a 5-point Likert Scale questionnaire, and two one-to-one semi-structured interviews by the researcher with the assistance of two trained and qualified assistants. The AI general knowledge performance was measured by the scores a student got on the English proficiency tests before and after the course. Data from the 5-point Likert Scale questionnaire and the one-to-one semi-structured interviews showed students' perceptions of their learning of AI general knowledge. Data analysis methods used in this study were percentage, mean, standard deviation, paired-samples t-test, Pearson correlation coefficients, one-sample K-S test, Mann-Whitney U test, and the thematic analysis.

Participants

The participants of the present study were 45 undergraduate students at Chongqing Technology and Business University (CTBU). This study used convenience sampling. Students chose the AI English Reading course at their will as a compulsory optional course according to the English teaching program of CTBU.

Research Instruments

English Proficiency Tests

Before the teaching experiment, the researcher developed two English proficiency tests, EPT1 and EPT2, to assess students' AI general knowledge before and after taking the AI English Reading course. Both tests had 50 marks in total and followed the same format: Multiple choice questions (full 10 marks of 10 questions), Noun explanations (full 20 marks of 4 terms), and Essay-type questions (full 20 marks of 1 essay writing). Both tests were piloted by another ten students and reviewed and adapted by three ELT experts.

Questionnaire

One 5-point Likert Scale questionnaire was developed by the researcher to investigate students' perceptions of their AI general knowledge and CLIL. The questionnaire was piloted by another ten students and examined by the same three ELT experts, and finally adopted by the researcher according to the pilot study feedback. The questionnaire used a 5-point Likert Scale including 10 items, with each item involving five choices of strongly disagree (1 point), disagree (2 points), neither agree nor disagree (3 points), agree (4 points), and strongly agree (5points). Items in the questionnaire are illustrated in Table one as follows.

Table 1. *Format of the questionnaire*

Items
1. I think AI is interesting.
2. I like the AI English Reading textbook.
3. I think the AI English Reading textbook is helpful.
4. I think the AI English Reading textbook is engaging.
5. I like the AI English Reading course.
6. I think the AI English Reading course is helpful.
7. I think the AI English Reading course is engaging.
8. I like the teaching method.
9. My AI general knowledge is improved.
10. My vocabulary of AI is improved.
11. I gained more insight into AI applications.

One-to-one Semi-structured Interviews

The one-to-one semi-structured interview was conducted after the questionnaire by the researcher with five randomly chosen students of their perceptions of their acquisition of AI general knowledge and CLIL. The interview was mainly about the following three questions: 1.

How do you feel about your AI general knowledge after taking the course?; 2. How do you feel about the teaching method applied in the course? 3. How do you feel about the AI English reading textbook and course? The second one-to-one semi-structured interview was conducted on specific students about the reason why they chose totally negative answers to some of the items in the questionnaire.

Research Procedures

The whole study included four phases that can be seen in figure one as follows.

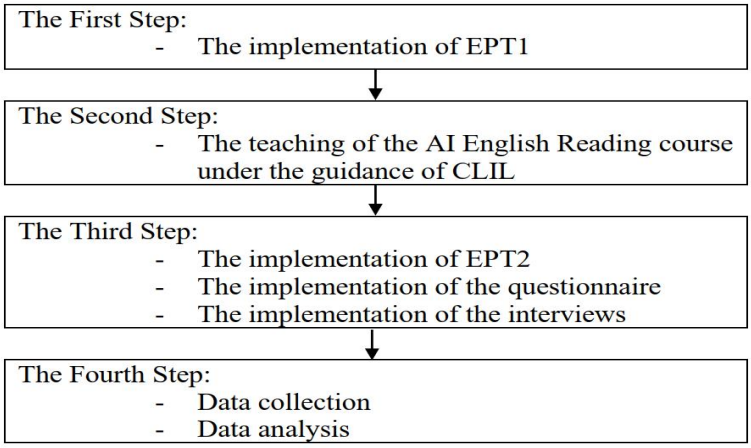

Figure 1 Research design

Teaching Procedure

The teaching experiment lasted for one academic semester, which was 16 weeks, according to the college English education syllabus in CTBU. One course of 90 minutes would be taught for one unit at a time in one week. Each class would follow the same five-step teaching procedure, which can be seen in figure two as follows.

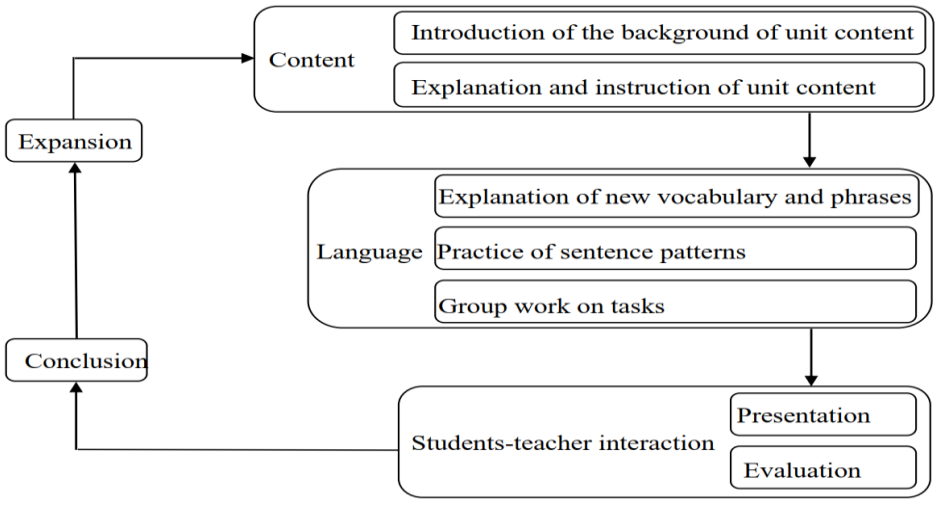

Figure 2. Teaching procedure

Such a teaching procedure was a progressive circle among which each step strengthened the former and introduced the next. Finally, it formed a systematic and integral teaching program for the whole teaching experiment.

Research Findings

After the teaching experiment, 39 students provided effective data for the present study, including one set of quantitative data for students' performance of AI general knowledge in both EPT1 and EPT2, one set of quantitative data for questionnaire, and one set of qualitative data for interviews. All the quantitative data would be tested by authorized computer programs, including percentage, mean, standard deviation, paired-samples t-test, Pearson correlation coefficient, one-sample K-S test, and Mann-Whitney U test. The qualitative data would be analyzed manually by the researcher with thematic analysis. All research results are illustrated as follows.

Statistical Results for Students' Performance in EPT1 and EPT2

The descriptive statistical results of students' performance of AI general knowledge in EPT1 in pretest are shown in Table two as follows, with 'Toral 1' referring to the total scores of EPT1.

Table 2. *Descriptive statistics for 'AI general knowledge' in EPT1*

	N	Minimum	Maximum	Mean	SD
Multiple choice questions	39	1	7	4.10	1.65
Noun explanations	39	0	7	2.49	1.73
Essay-type question	39	0	7	3.26	2.11
Total 1	39	2	17	9.85	3.78

The descriptive statistical results of students' performance of AI general knowledge in EPT2 in posttest are shown in Table three as follows, with 'Toral 2' referring to the total scores of EPT1.

Table 3. *Descriptive statistics for 'AI general knowledge' in EPT2*

	N	Minimum	Maximum	Mean	SD
Multiple choice questions	39	0	8	5.46	1.46
Noun explanations	39	0	17	6.41	4.96
Essay-type question	39	2	9	5.21	2.07
Total 2	39	6	29	17.08	6.17

The descriptive statistical results of EPT2 showed an improvement in students' command of AI general knowledge in comparison with that of EPT1. But it should be further measured whether this improvement was significant.

Results of Pearson Correlation Coefficients and One-sample K-S Test

Before students' improvement in EPT2 after the teaching experiment was measured, the Pearson correlation coefficient was applied to test the correlation between students' scores for

'Multiple choice questions, 'Noun explanations', and 'Essay-type questions' in both EPT1 and EPT2 respectively, and between their total scores for both tests. The results of Pearson correlation coefficients for students' scores for each question type and total scores (referring to Total 3) for both tests can be seen in Table four as follows.

Table 4. *Pearson correlation coefficients for EPT1 and EPT2*

	N	Correlation	Sig. (2-tailed)
Multiple choice questions	39	.556**	.000
Noun explanations	39	.047	.778
Essay-type question	39	.536**	.000
Total 3	39	.325*	.043

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

Pearson correlation coefficients in table four showed that students' performance on 'Multiple choice question' in EPT1 was significantly correlated with 'Multiple choice question' in EPT2 ($r=0.556$, $p=0.000<0.01$). The scores for 'Essay-type question' in EPT1 were also found to be positively correlated with that in EPT2 ($r=0.536$, $p=0.000<0.01$). Students' total performance in EPT1 was also positively correlated with EPT2 ($r=0.325$, $p=0.043<0.05$). No correlation was found between students' performance on 'Noun explanations' ($r=0.047$, sig.=0.778).

The of Pearson correlation coefficients for both tests showed that paired-samples t-tests could be applied to compare students' performance on 'Multiple choice questions, 'Essay-type questions', and their total scores. No correlation was found for 'Noun explanations' between pretest and posttest, which meant that it was not applicable to use a paired-sample t-test in comparing students' performance on this question type. Before further processing of these two sets of data of 'Noun explanations', their normal distributions were tested respectively. One-sample K-S test was applied to test the normal distributions of 'Noun explanations' in EPT1 and EPT2 because whether the data distributions were normal might determine the statistical analysis method for comparing these two sets of data. Table five presents the results of the one-sample K-S test for students' performance on 'Noun explanations'. N1 stood for 'Noun explanations' of EPT1 and N2 stood for 'Noun explanations' of EPT2.

Table 5 *One-sample K-S test for 'Noun explanations'*

		N1	N2
N		39	39
Normal parameters[a,b]	Mean	2.4872	6.4103
	Std. Deviation	1.73010	4.95622
Most extreme differences	Absolute	.201	.150
	Positive	.201	.150
	Negative	-.107	-.098

Test statistic	.201	.150
Sig. (2-tailed)	.000[c]	.026[c]

a. Test distribution is Normal
b. Calculated from data
c. Lilliefors Significance Correction

The results of the one-sample K-S test for students' performance on 'Noun Explanation'were not normally distributed (Sig.=0.000<0.01 [EPT1] and Sig.=0.026<0.05 [EPT2]), which meant that independent-samples t-test was not suitable either to be applied to measure these two sets of data. In that case, the researcher adopted the Mann-Whitney U test to analyze studentslied to measure these two sets of da.

Results of Paired-samples T-tests

Paired-samples t-tests were applied to show the differences between students' performance on 'Multiple choice questions', 'Essay-type question', and the total 'AI general knowledge' in EPT1 and EPT2. Table six presents the results of paired-samples t-tests for these three parts, with 'Total 4' referring to the total scores of both tests.

Table 6 *Paired-samples t-tests for 'AI general knowledge'*

	Mean	Std. Deviation	Std. Error Mean	*t*	df	Sig. (2-tailed)
Multiple choice questions	-1.35897	1.47768	.23662	-5.743	38	.000
Essay-type question	-1.94872	2.01244	.32225	-6.047	38	.000
Total 4	-7.23077	6.10203	.97711	-7.400	38	.000

$p<0.01$

As demonstrated in Table six, there were significant differences between mean scores for 'Multiple choice questions', 'Essay-type question' and the total performance in EPT1 and EPT2: 'Multiple choice questions' (t=-5.743, p=.000<.05); 'Essay-type question' (t=-6.047, p=.000<.05); and total scores (t=-7.400, p=.000<.05). All the t values were negative, which meant that data of either question type and total performance collected from EPT2 were larger than EPT1. The results showed that the differences in students' performance on EPT1 and EPT2 were statistically significant. After the teaching of the AI English Reading course, students' performance on 'Multiple choice questions' and 'Essay-type question', and students' total performance on 'AI general knowledge' varied to a great extent when compared with before the class. The comparison of mean scores for either question type (mean ['Multiple choice questions' in pretest]=4.10; mean ['Multiple choice questions' in posttest]=5.46; mean ['Essay-type question' in pretest]=3.26; mean ['Essay-type question' in posttest]=5.21), and for total 'AI general knowledge' (mean [pretest]=9.85; mean [posttest]=17.08) also showed that students' acquisition of AI general knowledge had been improved a lot.

In 'Multiple choice question' of EPT2, more students got scores in questions about Strong AI (question 1 with 27 correct answers), AI movie works (question 2 with 30 correct answers), machine learning (question 3 with 34 correct answers), natural language processing (question 5 with 27 correct answers) and AI products – drones (question 9 with 25 correct answers). When compared with their performance on 'Multiple choice questions' in EPT1, where students got lower scores for AI conceptual problems such as speech recognition (question 3 with 13 correct answers), AI definition (question 6 with 8 correct answers), AI schools (question 8 with 5 correct answers) and natural language processing (question 10 with 9 correct answers), students did better in similar AI knowledge categories. But in 'Multiple choice questions' of EPT2, students also showed a deficiency in knowledge of AI applications, namely, the functions AI technology could realize, especially in specific industries. The analysis of students' answers to 'Multiple choice questions' in both tests found that after taking the AI English Reading course, students had a deeper understanding and better command of AI general knowledge. They acquired a better understanding of AI-related concepts though their knowledge of AI applications and specific technological functions of AI still fell short.

In 'Essay-type question' of EPT2, more students could at least express something about their reflections on AI, with five students getting the lowest 2 scores. In contrast to their performance on this question type in EPT1, where six students submitted a total blank paper, after attending the course, no blank paper was submitted, and all students shared their insights into AI more or less, though with different levels. Besides, more students provided longer writing lengths in this question type in EPT2. Eighteen students provided more than 120 words, among whom two students wrote between 200-250 words, and six students wrote between 150-200 words. Fifteen students provided between 100-120 words, and the rest wrote no more than 100 words. While in EPT1, only five students wrote more than 100 words but less than 120 words, five students wrote between 80-100 words, and the rest wrote no more than 80 words. The ability to write more also suggested that students might have acquired more knowledge of AI, in which case they could express more of their reflections on AI. But a further detailed analysis of students' writing also found that though students could provide more viewpoints on AI, they still lacked deep enough thinking about AI.

The results of paired-samples t-tests for 'Multiple choice questions', 'Essay-type question', and total scores for 'AI general knowledge' of EPT1 and EPT2 indicated that statistically, after taking the AI English Reading course, students' performance in EPT2 showed an improvement when compared with their performance in EPT1, which partly provided positive evidence for the research question of the present study. The results partially supported that the AI English Reading course effectively enhanced students' AI general knowledge.

Results of Mann-Whitney U test

Students' scores for 'Noun explanations' in EPT1 and EPT2 were found to have no correlation, which meant that it was not suitable to measure these two sets of data with a paired-samples t-test. These two sets of data were not normally distributed either, which meant that the independent-samples t-test was neither applicable to be applied to measure the difference between students' performance on 'Noun explanations' in both tests. A nonparametric test, the Mann-Whitney U test, was adopted to analyze these two sets of nonnormal distribution data. Table seven presents the results of the Mann-Whitney U test for students' performance on 'Noun explanations' in EPT1 and EPT2.

Table 7. *Mann-Whitney U test for 'noun explanations'*

Ranks

	N	Mean Rank	Sum of Ranks
EPT1	39	29.91	1166.50
EPT2	39	49.09	1914.50
Total	78		

Test Statistics[a]

	Noun explanations
Mann-Whitney U	386.500
Wilcoxon W	1166.500
Z	-3.766
Sig. (2-tailed)	.000

a. Grouping Variable: Pretest and Posttest

The results of the Mann-Whitney U test for 'Noun explanations' in table seven showed that the difference between students' performance on 'Noun explanations' in EPT1 and EPT2 was statistically significant ($Z=-3.766$, $p=0.000<0.01$), which meant that students' performance on 'Noun explanations' in EPT2 was significantly improved when compared with their performance on this question type in EPT1. But a further detailed analysis of students' answers also indicated that students still displayed different levels of understanding of various categories of AI general knowledge. Students got higher scores for terms concerning AI concepts such as 'Machine learning' (totally 76 marks) and 'Expert system' (a totally 69 marks), which were basic technologies or concepts of AI. Students got lower scores for terms concerning AI applications such as 'Traveling salesman problem' (totally 50 scores) and 'Monte Carlo tree search (totally 44 scores). This finding corresponded to the result reflected in students' performance on 'Multiple choice questions' in EPT2.

Generally speaking, the results of the Mann-Whitney U test for 'Noun explanations' also partly answered the research question, and partially supported that the AI English Reading course effectively enhanced students' AI general knowledge.

Results of Questionnaire

After students had finished EPT2, the researcher conducted the questionnaire to collect data on students' perceptions of the AI English Reading textbook and course and students' self-cognition of their AI general knowledge. The descriptive statistics of the questionnaire can be seen in Appendix A.

As could be seen in Appendix A, the mean values of all items were larger than the median 3, which indicated that after taking the AI English Reading course, students showed general positive attitudes to items in the questionnaire. But apart from their integral acknowledgement of the interestingness of AI, there were still a few students who had negative attitudes to their preference for the AI English Reading textbook and the course, the usefulness and interestingness of the textbook and the course, the teaching method, and their gains in taking the course. After checking the questionnaire, the researcher purposedly invited students who showed negative attitudes in the questionnaire to take one more one-to-one semi-structured interview to investigate their reasons.

Results of Interviews

With thematic analysis, students' transcriptions of the interviews were analyzed by the researcher manually. The main findings of question 1, 'How do you feel about your AI general knowledge after taking the course?', were illustrated as follows.

a. A generally positive attitude to the improvement of AI general knowledge was reported, especially in AI concepts, AI-related jargon, and AI products such as drones and games;
b. AI general knowledge taught in the course was generally labeled as useful, learnable, and systematic;
c. Some students reported that they had gained inspiration and insights into the integration of AI and their majors;
d. Some students reported that their accumulation of English vocabulary was improved;
e. One student reported uncertainty about the improvement of AI general knowledge because he lacked inner motivation and showed indifference and disinterest in the course;

The main findings of question 2, 'How do you feel about the teaching method applied in the course?', were illustrated as follows.

f. Most students showed a preference for the teaching method, but two students also mentioned that they did not like the teaching method because teaching all in English was difficult for them to follow the course and it greatly did harm their interests in taking the course;
g. More high-tech teaching methods and teaching appliances should be used in the course, such as MOOCs, online courses, video clips, electronic books, learning software, apps, etc.;
h. An intelligent teaching environment composed of high-tech teaching appliances could facilitate students' learning and provide a vivid learning experience;
i. It would be better if instructions and exercises of language knowledge points could be added to the teaching process;

The main findings of question 3, 'How do you feel about the AI English reading textbook

and course?', were illustrated as follows.

j. One student reported the timeliness of AI general knowledge involved in the course needed to be improved, and the course content should keep updated with the fast development of AI.
k. Practice courses might help teach students the application of AI in a more vivid way, and they should be added to the curriculum;
l. The length of text passages should be shorted to make them more readable;
m. After-class exercises in the textbook were not necessary.

Students who were purposedly invited to take one more one-to-one semi-structured interview after the researcher's check on the questionnaire mainly provided feedback as follows.

n. Four students held a totally negative attitude toward some aspects of the AI English Reading textbook and course for too much load;
o. Two students proposed that though AI was influential and popular at present, everyone did not need to learn;
p. One student lost interest in taking the course for too much time to be spent after class to accomplish homework;
q. One student mentioned that the reason why he took the course was purely to gain the two credits necessary for graduation.
r. One student mentioned that the teaching content could be more professional.

Generally speaking, students held a positive attitude toward their improvement of AI general knowledge and showed a preference for the teaching method. They still provided many productive suggestions for the course.

Discussion

The quantitative findings of EPT1 and EPT2 showed from the above statistical analysis results that students' scores in EPT2 were higher than their scores in EPT1 with a statistical significance at a 0.01 level, whether each question type or the total scores. The findings indicated that students improved their AI general knowledge after taking the AI English Reading course. Students were more skilful in answering questions concerning AI development, AI products, and AI basic concepts especially related to language, but least in more professional AI knowledge such as AI algorithms and AI application in specific industries. Besides, students also showed a deficiency in their ability to acquire deep insights into AI.

Further detailed analysis of students' test answers also showed that students had a better command of knowledge concerning AI concepts and AI products after attending the AI English Reading course, which could be seen in students' performance on 'Multiple choice question' and 'Noun explanations' in EPT2. In 'Multiple choice question', more students got scores in questions about Strong AI (question 1 with 27 correct answers), AI movie works (question 2 with 30 correct answers), machine learning (question 3 with 34 correct answers), natural language processing (question 5 with 27 correct answers) and AI products – drones (question 9 with 25 correct answers). When compared with their performance on 'Multiple choice questions' in EPT1, where students

got lower scores for AI conceptual problems such as speech recognition (question 3 with 13 correct answers), AI definition (question 6 with 8 correct answers), AI schools (question 8 with 5 correct answers) and natural language processing (question 10 with 9 correct answers), students did better in similar AI knowledge categories. But in 'Multiple choice questions' of EPT2, students also showed a deficiency in knowledge of AI applications, namely, the functions AI technology could realize, especially in specific industries. The analysis of students' answers to 'Multiple choice questions' in EPT1 and EPT2 found that after taking the AI English Reading course, students had a deeper understanding and a better command of AI general knowledge. They acquired a better understanding of AI-related concepts though their knowledge of AI applications and specific technological functions still fell short.

Further detailed analysis of students' test answers to 'Noun explanation' found that though students did much better for this question type in EPT2 than in EPT1 ($t=-6.05$, $p=.000<.05$), which showed an overall improvement in students' performance, they still displayed different levels of understanding of various categories of AI general knowledge. Students got higher scores for terms concerning AI concepts such as 'Machine learning' (totally 76 marks) and 'Expert system' (totally 69 marks), which were basic technologies or concepts of AI. Students got lower scores for terms concerning AI applications such as 'Traveling salesman problem' (totally 50 scores) and 'Monte Carlo tree search' (totally 44 scores). This finding corresponded to the result reflected in students' performance on 'Multiple choice questions' in EPT2.

Further detailed analysis of students' answers to 'Essay-type question' in EPT2 suggested that students could at least express something about their reflections on AI, with five students getting the lowest 2 scores. In contrast to their performance on this question type in EPT1, where six students submitted a total blank paper, after attending the course, no blank paper was submitted, and all students shared their insights into AI more or less, though at different levels. Besides, more students provided longer writing length in this question type of EPT2. Eighteen students provided more than 120 words, and among whom two students wrote between 200-250 words, and six students wrote between 150-200 words. Fifteen students provided between 100 and 120 words, and the rest wrote no more than 100 words for this question type of EPT2. While in EPT1, only five students wrote more than 100 words but less than 120 words, five students wrote between 80-100 words, and the rest wrote no more than 80 words. The ability to write more also suggested that students might have acquired more knowledge of AI, in which case they could express more of their reflections on AI. But a further detailed analysis of students' writing also found that though students could provide more viewpoints on AI, they still lacked deep enough thinking about AI.

The results of paired-samples t-tests for 'Multiple choice questions', 'Essay-type question', and total scores for 'AI general knowledge , and the results of the Mann-Whitney U test for 'Noun explanations' in both tests indicated that statistically, after taking the AI English Reading course

students' performance on EPT2 was improved when compared with their performance on EPT1, which provided positive evidence for research question 1: How effective is the AI English Reading course in enhancing students' AI general knowledge?. The results supported that the AI English Reading course effectively enhanced students' AI general knowledge.

The quantitative analysis of the questionnaire and the thematic analysis of the interviews also showed that students had a generally positive attitude to the acquisition of their AI general knowledge after taking the course. In general, students interviewed mentioned that they had improved their knowledge of AI development and praised highly of the systematicity and learnability of the AI general knowledge due to the explanation of AI schemas. But to some students, the improvement did not meet their expectations. Some students gained inspiration and insights into AI, and constructed their own AI schemas, which were significant meanings of the course. Most of them also admitted the value of the AI English reading course (usefulness). Besides, some students' English language ability (vocabulary) was also improved. Some students explained the lack of inner motivations (indifference and disinterest) leading to their underperformance in their acquisition of AI general knowledge. And, the timeliness of AI general knowledge involved in the course also needed to be improved. The results of interview 2 also demonstrated students' positive attitudes to researcher question 1 and provided productive suggestions for research question 2. Besides, students also provided meaningful and constructive suggestions for the improvement of the localization of the teaching method – CLIL, and for the development of AI-related English courses under the GEE context in China, especially the decrease of the difficulty, the timely update of AI general knowledge and the trade-off of Chinese-English in-class teaching language.

In summary, the results of both quantitative and qualitative analyses of the data collected indicated the effectiveness of the AI English Reading course. It was reasonable to conclude that the implementation of the AI English Reading course was effective in enhancing students' AI general knowledge. But there still existed problems in students' acquisition of AI general knowledge and their insights into AI. Firstly, students' AI general knowledge was unbalanced in their different levels of command and different domains of AI. Students' personal interests, their different majors, and the time consumed all influenced their command of AI general knowledge. Secondly, students' understanding of AI general knowledge was still comprehensive. Because AI was a domain covering a mass of knowledge and with rich contents, it seemed explicable that students lacked some AI general knowledge such as conceptual problems of AI technologies and algorithms, and their applications. Thirdly, students' insights into AI were not as profound as expected. Though all students showed deeper insights into AI when compared with before the course, their perceptions of AI and its development in the future still remained at a superficial level, which in fact could be attributed to their incomprehensive understanding of AI general knowledge and lack of in-depth thinking. Fourthly, some students still lacked inner motivation of

learning AI general knowledge. They showed quite a lot of indifference and disinterest in AI. This study only probed into the effectiveness of CLIL on improving students' AI general knowledge in a very fundamental way. The improvement of students' English language ability was left without being investigated. Further studies can be made on the integration of AI knowledge and English learning. Besides, more applications of CLIL in the Chinese GEE context should also be explored.

Conclusion

The present study aimed to investigate the implementation of CLIL in Chinese GEE courses. Compared with the initial level of their command of AI general knowledge before the course, the overall performance of students' AI general knowledge after the course was improved. The findings of this study indicated that it was possible and practical to implement CLIL in Chinese GEE courses. The theme of AI was valuable content that could be included in GEE. Besides, students also demonstrated a need for professional AI-related English learning, and ESP or EAP for only AI purposes remained a total blank up to now. To benefit from the cultivation of AI talents, whether professional AI English should be constructed to help cultivate well-rounded AI talents was still a theme that needed to be further studied.

About the Authors:

Jingwei Tang is an ELT doctoral candidate at Graduate School of Human Scoences of Assumption University, Thailand. She is also a lecturer in the School of Foreign Languages at Chongqing Technology and Business University, China. Her professional interest include English Language Teaching and English for Specific Purposes.ORCID: https://orcid.org/0000-0002-9668-4501.

Prof. Dr. Joseph Foley is one of the Senior Lecturers in the Graduate School of Human Sciences, English Language Teaching Department at Assumption University, Thailand. Previously, he was the Head of English Language in the Department of English Language and Literature at the National University of Singapore. He was for many years the editor of the RELC Journal. His Publications are mainly in the area of Systemic Linguistics, Language Development, World Englishes, and the Common European Framework of Reference in Asian Context. ORCID: https://orcid.org/0000-0002-7293-0560.

References

Arkoudis, S. (2005). Fusing pedagogic horizons: language and content teaching in the mainstream. *Linguistics and Education, 16*(2), 173-187. https://doi.org/10.1016/j.linged.2006.01.006.

Baecher, L., Farnsworth, T., & Ediger, A. (2014). The challenges of planning language objectives in content-based ESL instruction. *Language Teaching Research, 18*(1), 118-136. https://doi.org/10.1177/1362168813505381.

Bigelow, M. H. (2010). Learning to plan for a focus on form in CBI: The role of teacher knowledge and teaching context. In J. Davis (Ed.), *World Language Teacher Education: Transitions and Challenges in the Twenty-first Century* (pp.35-56). Charlotte, NC: Information Age Publishing.

Cammarata, L., & Tedick, D. J. (2012). Balancing content and language in instruction: The experience of immersion teachers. *The Modern Language Journal, 96*(2), 251-269. https://doi.org/10.1111/j.1540-4781.2012.01330.x.

Coyle, D. (2007). Content and language integrated learning: Towards a connected research agenda for CLIL pedagogies. *International Journal of Bilingual Education & Bilingualism, 10*(5), 543-562. https://doi.org/10.2167/beb459.0.

Echevarria, J., Vogt, M., & Short, D. (2009). *The SIOP Model for Teaching Mathematics to English Learners*. Boston: Allyn & Bacon.

Evnitskaya, N., & Morton, T. (2011). Knowledge construction, meaning-making and interaction in CLIL science classroom communities of practice. *Language and Education*, *25*(2), 109-127. https://doi.org/10.1080/09500782.2010.547199.

Gibbons, P. (2003). Mediating language learning: Teacher interactions with ESL students in a content-based classroom. *TESOL Quarterly, 37*(2), 247-265. https://doi.org/10.2307/3588504.

Harley, B., Allen, P., Cummins, J., & Swain, M. (Eds). (1990). *The Development of Second Language Proficiency*. Cambridge: Cambridge University Press.

Hoare, P. (2010). Content-based language teaching in China: Contextual influences on implementation. *Journal of Multilingual and Multicultural Development, 31*(1), 69-86. https://doi.org/10.1080/01434630903367207.

Huang, K.-M. (2011). Motivating lessons: A classroom**-**oriented investigation of the effects of content-based instruction on EFL young learners' motivated behaviours and classroom verbal interaction. *System, 39*(2), 186**-**201. https://doi.org/10.1016/j.system.2011.02.002.

Chang, J., Liu, X., & Deng, Y. (2009). Impacts of Content-based Instruction Reform on Students' Reading Comprehension. *Chinese Foreign Languages, 29*(3), 40-53. https://doi.org/10.3969/j.issn.1672-9382.2009.03.008.

Jakonen, T., & Morton, T. (2015). Epistemic search sequences in peer interaction in a content-based language classroom. *Applied Linguistics, 36*(1), 73-94. https://doi.org/10.1093/applin/amt031.

Liebscher, G. & J. Dailey-O'Cain. (2005). Learner code-switching in the content-based foreign language classroom. *The Modern Language Journal, 89*(2), 234-247. https://doi.org/10.1111/j.1540-4781.2005.00277.x.

Lingley, D., Nunn, R., & Otlowski, M. (2006). ELT curriculum reform project: collaborative design of units of instruction. *Kochi University Reports of Educational Research & Activity*, 10.

Lorenzo, F. (2007). An analytical framework of language integration in L2 content-based courses: The European dimension. *Language and Education, 21*(6), 502-514. https://doi.org/10.2167/le708.0.

Lyster, R. (2007). *Learning and Teaching Languages Through Content: A Counter-balanced Approach*. Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.18.

Mi, B. (2015). New development in Content and Language Integrated Learning research abroad. *Modern Foreign Languages (Bimonthly), 38*(5), 715-725. https://doi.org/CNKI:SUN:XDWY.0.2015-05-014.

Schleppegrell, M., Achugar, M. & Orteíza, T. (2012). The grammar of history: Enhancing content-based instruction through a functional focus on language. *TESOL Quarterly*, *38*(1), 67-93. https://doi.org/10.2307/3588259.

Short, D. (2002). Newcomer programs: an educational alternative for secondary immigrant students. *Education & Urban Society, 34*(2), 173-198. https://doi.org/10.1177/0013124502034002004.

Short, D. (2012). Language learning in sheltered social studies classes. *TESOL Journal, 11*(1), 18-24. https://doi.org/10.1002/j.1949-3533.2002.tb00062.x.

Snow, M. A., Met, M., & Genesee, E. (1989). A conceptual framework for the integration of language and content in second/foreign language instruction. *TESOL Quarterly*, *23*(2), 201-217. https://doi.org/10.2307/3587333.

Stryker, S. B., & Leaver, B. L. (eds.). (1997). Content-based instruction foreign language education: models and methods. *NUML Journal of Critical Inquiry*, *13*(1), 1-22. https://doi.org/10.2307/346085.

Wolff, D. (2003). Content and language integrated learning: A framework for the development of learner autonomy. In D. Little, J. Ridley, & E. Ushioda, (eds.), *Learner Autonomy in the Foreign Language Classroom: Teacher, Learner, Curriculum and Assessment* (pp.211-222). Dublin: Authentik.

Yuan, P. (2012). On the Impact of CBI on EFL Learners' Motivation and Anxiety in College English Context. *Journal of PLA University of Foreign Languages*, *35*(3), 41-45. https://CNKI:SUN:JFJW.0.2012-03-010.

Appendices

Descriptive statistics for questionnaire

Score / Number	1	2	3	4	5	Total	Mean
Item 1	0	0	6	14	19	169	4.33
Percentage	0	0	15%	36%	49%	100%	
Item 2	0	3	18	13	5	137	3.51
Percentage	0	8%	46%	33%	13%	100%	

Item 3	1	2	12	22	2	139	3.56
Percentage	3%	5%	31%	56%	5%	100%	
Item 4	0	9	12	17	1	147	3.26
Percentage	0	23%	31%	44%	3%	100%	
Item 5	0	3	13	19	3	138	3.54
Percentage	0	8%	33%	49%	8%	100%	
Item 6	1	3	13	19	3	137	3.51
Percentage	3%	8%	33%	49%	8%	100%	
Item 7	1	2	15	18	3	137	3.51
Percentage	3	5%	38%	46%	8%	100%	
Item 8	0	3	15	16	5	140	3.59
Percentage	0	8%	38%	41%	13%	100%	
Item 9	0	2	6	26	5	151	3.87
Percentage	0	5%	15%	67%	13%	100%	
Item 10	1	2	7	25	4	146	3.74
Percentage	3%	5%	18%	64%	10%	100%	
Item 11	0	2	14	22	1	139	3.56
Percentage	0	5%	36%	56%	3%	100%	