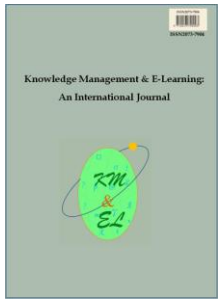
Knowledge Management & E-Learning: An International Journal, Vol.5, No.1. Mar 2013

Knowledge Management & E-Learning: An International Journal



ISSN 2073-7904

An effective approach using blended learning to assist the average students to catch up with the talented ones

Jiyou Jia
Peking University, China
Dongfang Xiang
Huiwen Middle School, Beijing, China
Zhuhui Ding, Yuhao Chen, Ying Wang, Yin Bai, Baijie Yang
Peking University, China

Recommended citation:
Jia, J., Xiang, D., Ding, Z., Chen, Y., Wang, Y., Bai, Y., & Yang, B. (2013). An effective approach using blended learning to assist the average students to catch up with the talented ones. *Knowledge Management & E-Learning, 5*(1), 25–41.

An effective approach using blended learning to assist the average students to catch up with the talented ones

Jiyou Jia*

Department of Educational Technology
Graduate School of Education
Peking University, Beijing, China
E-mail: jjy@pku.edu.cn

Dongfang Xiang

Huiwen Middle School
Dongcheng District, Beijing, China
E-mail: xdf200512@163.com

Zhuhui Ding

Department of Educational Technology
Graduate School of Education
Peking University, Beijing, China
E-mail: dvh@pku.edu.cn

Yuhao Chen

Department of Educational Technology
Graduate School of Education
Peking University, Beijing, China
E-mail: st911@gse.pku.edu.cn

Ying Wang

Department of Educational Technology
Graduate School of Education
Peking University, Beijing, China
E-mail: st10865m@gse.pku.edu.cn

Yin Bai

Department of Educational Technology
Graduate School of Education
Peking University, Beijing, China
E-mail: baiyin@pku.edu.cn

Baijie Yang

Department of Educational Technology
Graduate School of Education
Peking University, Beijing, China
E-mail: yangbaijie@pku.edu.cn

*Corresponding author

Abstract: Because the average students are the prevailing part of the student population, it is important but difficult for the educators to help average students by improving their learning efficiency and learning outcome in school tests. We conducted a quasi-experiment with two English classes taught by one teacher in the second term of the first year of a junior high school. The experimental class was composed of average students (N=37), while the control class comprised talented students (N=34). Therefore the two classes performed differently in English subject with mean difference of 13.48 that is statistically significant based on the independent sample T-Test analysis. We tailored the web-based intelligent English instruction system, called Computer Simulation in Educational Communication (CSIEC) and featured with instant feedback, to the learning content in the experiment term, and the experimental class used it one school hour per week throughout the term. This blended learning setting with the focus on vocabulary and dialogue acquisition helped the students in the experimental class improve their learning performance gradually. The mean difference of the final test between the two classes was decreased to 3.78, while the mean difference of the test designed for the specially drilled vocabulary knowledge was decreased to 2.38 and was statistically not significant. The student interview and survey also demonstrated the students' favor to the blended learning system. We conclude that the long-term integration of this content oriented blended learning system featured with instant feedback into ordinary class is an effective approach to assist the average students to catch up with the talented ones.

Keywords: Blended learning; Computer simulation in educational communication (CSIEC); Average students; Talented students; English instruction in a middle school

Biographical notes: Dr. Jiyou Jia is an associate professor from the Department of Educational Technology, Graduate School of Education, and director of the International Research Center for Education and Information, Peking University, China. His research interests include educational technology and artificial intelligence in education. Dr. Jia authored one book in German (2004), one book in Chinese (2009), and edited another one in English (2012). Both of the Chinese and English works have been registered in the Library of Congress, USA. He has published more than 60 articles in international (SCI/SSCI) and national peer-reviewed journals and conferences. He has been responsible for more than ten national key projects and international projects, and his research has been recognized by the international academic community. He serves as a member of editorial board of five international journals and as a program chairman or committee member of a dozen of international conferences.

Dongfang Xiang is an English teacher in the Huiwen Middle School, Dongcheng District, Beijing, China. Her interest is teaching with technology.

Zhuhui Ding, Yuhao Chen, Yin Bai and Baijie Yang are master students in the Department of Educational Technology, Graduate School of Education, Peking University.

Ying Wang is a doctoral candidate of the Graduate School of Education, Peking University. Her research interests include technology-enhanced learning and teacher professional development. She has been working on the development of a teacher training program named Intel ® Learn, a survey on development of higher education informationization, a UNICEF project named SkillsMotivation and Imagination for Learning Excellence and other projects of Ministry of Education (MOE).

1. Introduction

Average students in the primary and secondary education are referred to as the students whose performance in the classroom is normal, while the gifted or talented students outperform the average ones in the classroom tests. Because the average students are the prevailing part of the student population, it is an important and difficult task for the educators to help the average students by stimulating their learning interests, improving their learning efficacy and the learning outcome that can be achieved in school exams and tests.

Since the modern computer was born in the 1940s, one of its important application fields is education at every educational stage from kindergarten to higher education. Computer Assisted Instruction (CAI) is one early definition that describes instruction assisted by computer technology. Though the computer hardware and software have evolved through several generations from 1940s up to date, this definition still can designate the nature of computer application in instruction with all kinds of forms, no matter what it is called, such as Computer Based Education (CBE), Computer Based Instruction (CBI), electronic learning (e-learning), Web Based Learning (WBE), etc. In the new millennium, a new term, called blended learning or blending learning, has been adopted and widely used to replace the old-fashioned notation CAI and to describe the instructional design that blends the traditional classroom and Information and Communication Technology (ICT). Thus in order to ensure the consistency in this paper, we just use the term CAI to represent all kinds of computer's application in instruction.

2. Related work

Can CAI help all kinds of students including disabled, average and talented ones improve their learning outcome and to what extent? This question has drawn great attention since 1950s. A number of meta-analysis studies of CAI analyzed dozens, hundreds or even thousands of studies dealing with thousands of subjects, and found that CAI generally can have a more positive effect on learning performance than traditional instructional approaches (Burns, 1981; Hartley, 1978; Kulik & Kulik, 1991; Liao, 2007; Tamim, Bernard, Borokhovski, Abrami, & Schmid, 2011; Yueh, Lin, Huang, & Sheen, 2012; Wakefield, Warren, Rankin, Mills, & Gratch, 2012). For the average or low-performance students, much research has shown that CAI can have a positive impact on their learning outcome (Lynch, Fawcett & Nicholson, 2000; O'Byrne, Securro, Jones & Cadle, 2006; Huang, Yang, & Hwang, 2010).

How can CAI be used to help the low-performance students improve their learning performance? The answer to this question depends on the learning content, the learner's age and other learner characteristics. Despite of disciplinary content and learner difference, the learning time plays a key role in general. Mann, Shakeshaft, Becker, and Kottkamp (1999) conducted a study of West Virginia's Basic Skills/Computer Education (BS/CE) by analyzing results from a representative sample of 950 fifth-grade students from 18 elementary schools across the state. The study showed that the longer students participated in the BS/CE, the higher their test scores on Stanford Achievement Test (SAT): SAT-9. Ligas (2002) conducted a five-year longitudinal study to examine the impact of CAI on reading achievement of 'at-risk' elementary and middle school students in Florida. The study found that the students group who used the software for 12 hours or more outperformed the students group who did not use the software, or used it less than 5 hours, by 7.74 points on the SAT-8 Reading Comprehension average normal curve equivalent (NCE) scores. Liao (2007) revealed that for the duration variable, the largest mean ES (1.182) was associated with studies lasting 4–18 hours.

Summarizing the aforementioned literature review, it is inferable that the CAI can have positive effects on average students' learning performance, and that the longer usage of CAI can produce a better performance improvement.

For English instruction as a second language in middle schools, which is often listed as a core subject, most research has shown the positive effect of CAI on learning performance within a short duration, for example, several hours within several weeks (Tsou, Wang, & Li, 2002; Liu, 2009; Liu & Chu, 2010; Chen, Ho, & Yen, 2010;Fujishiro & Miyaji, 2010). The average or low-performance students' learning improvement, compared with the excellent or high-performance ones, varied case by case.

However, we have found few research papers on the long-term integration of CAI into English instruction in middle schools, for example for a school term, and its effect on the school test performance of existing classes comprised of average students. Our previous study (Jia, Chen, Ding, & Ruan, 2012) indicated that the blended learning setting can facilitate the vocabulary acquisition and improve the students' examination performance. The experimental class, starting with a higher pre-test score mean and participating in the blended learning with one weekly school hour in the computer pool throughout the experimental term, enlarged its examination score mean difference to the control class. What will be the result if the experimental class with average students performs much worse in the pre-test than the control class with talented students? Can the blended learning help the average students catch up with the talented ones? This is the key problem reflected in this research paper.

In the instruction of English as a second language, vocabulary acquisition is the most important foundation, because it is the fundamental prerequisite to the four skills of a language: listening, speaking, reading and writing. Linguistic experts believe that vocabulary knowledge and the ability to comprehend text are inextricably linked, and the breadth and depth of a student's vocabulary is a key forecaster of his/her ability to understand a wide range of texts (Anderson & Freebody, 1981; Thorndike, 1973). This is true for both native speakers of English and second language learners (Coady, 1993; Stoller & Grabe, 1993).

A large amount of researches investigate the vocabulary instruction supported by emerging technologies in university and college, such as (Chen, Hsieh, & Kinshuk, 2008; Chen & Chung, 2008; Jones, 2004; Huang & Liou, 2007; Lu, 2008; Peters, 2007), etc. However, we can only find very few literature studies on computer assisted vocabulary teaching and learning for formal secondary school students throughout a long time such

as a whole school term. This is just the gap between the theoretical research and the pedagogical practice which we would like to reduce.

3. System architecture

Research has shown that children can be taught new word meanings through rote methods involving synonyms and definitions (McKeown, Beck, Omanson, & Pople, 1985; Stahl, 1983). Moreover, in the case of L2 learners, there is great value in the repetition and immediate access to definitions for unknown words, especially when those words are rarely used in English (Stoller & Grabe, 1993). Because technology generally improves performance if the application directly supports the curriculum content, specifically the vocabulary learning, we still use the same system, namely CSIEC, as the one described in (Jia, Chen, Ding, & Ruan, 2012) to support the blended learning for an English class. This web-based system comprises exercises for every module in a textbook. The exercise for every module is logically composed of two parts, vocabulary and dialogue, as shown in Fig. 1.

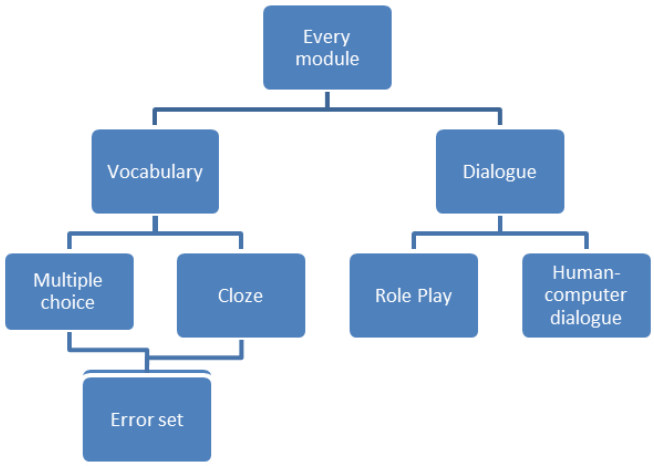


Fig. 1. The architecture of every module with vocabulary and dialogue assessment functions

The first part is the course management system that mainly supplies question banks and quizzes about the vocabulary required in a certain module. The questions and the quizzes have four features.

The first feature is the multiple choice question and cloze in which a sound file can be played so that pronunciation and listening based questions can be embedded. For example, a multiple choice question or a cloze about the spelling and meaning of an English word or phrase is raised to the students after its pronunciation is played back.

The second feature refers to the randomized items of the quiz based on a question bank, as well as the randomized sequence of the choice items to a multiple choice question. This intelligent feature challenges all the students sitting in front of the computers in a computer pool and doing the same quiz simultaneously.

The third feature is the instant feedback including score, comments and correct answers after the student submits his or her answers to a quiz. Nevertheless, the scores of all students in the class can not only be read by the students themselves, but also be

browsed by the teacher. Both the individual feedback and the collective scores can inform the students and the teacher about the learning outcome, and motivate the students to compete with each other in the blended learning setting.

The fourth feature is the individualized error set that includes the words and phrases, with which one student has made mistakes in the multiple choice question and cloze. A new individualized cloze quiz can be generated based on the words and phrases within the error set. The student can review the words and phrases with which the mistakes have been made by doing this cloze quiz.

The second part besides the vocabulary exercises is the dialogue simulation for specific topics or scenarios defined in the teaching module in the textbook. Two or more than two roles participate in this kind of dialogue about a specific topic. Two types of simulation with multiple agent technology have been designed. The first addresses the talk show of multiple agent characters to role play the dialogue, in which the main content is semantically the same as the one given in the textbook, but the expressions are randomly generated according to the predefined script. The second represents the interactive dialogue with the student participating as a role in it. During the interactive dialogue the student should input the semantically same or similar expressions as the textbook in order to ensure the dialogue process. Both in the talk show and the interactive dialogue, the user can select one of the twelve avatars to represent one role in the dialogue according to his/her preference. The avatar is in fact a Microsoft agent character that can speak the text with synthesized voice and carry out some actions. The dialogue simulation can stimulate the students to participate in the dialogue and learn its content, and strengthen the listening comprehension.

The two parts, vocabulary and dialogue, are not separated. Some words and phrases drilled in the first part occur in the corresponding dialogue in the same module. Through the dialogue simulation the students can understand how the words or phrases learned are used in the practical dialogue. Wilkins (1972) argued: ‘‘without grammar very little can be conveyed, without vocabulary nothing can be conveyed.’’ The two parts are intended to help the average students master the vocabulary and its usage in the dialogue.

4. Methodology

4.1. Research hypothesis

The blended learning setting in the computer pool with the specific web-based vocabulary and dialogue system throughout a school term can decrease the mean difference of the learning performance in an ordinary English test between the ordinary class and the talented class.

4.2. Participants

The participants in this research came from two existing classes of Grade one of a junior school in Capital Beijing, one was an ordinary class and another was an excellent class. The 34 students in the excellent class were selected from the primary schools in the entire city with their excellent performance in tests of three main subjects, specifically mathematics, Chinese language and English language, while the 37 students in the ordinary class were randomly picked out with a lower performance in tests of the three main subjects. Therefore in the final exam of the first term of Grade one, which we used

as the pre-test in this research, the excellent class achieved much better scores in the English subject than the ordinary class. The mean difference between the two classes in the test, 13.5%, was statistically significant based on the independent sample T-Test analysis with statistical software SPSS (V16.0), as shown in Table 1.

The two classes' teacher X was interested in and experienced with computer assisted language learning. The school managers agreed to our blended learning experiment for Teacher X's two classes in the second term of Grade one, and arranged one school hour in the school class schedule for the ordinary class to be held in one multimedia computer pool of this school. The initial hope was that our blended-learning could help the normal students improve their learning outcome, and decrease the difference between the two classes. We defined the ordinary class as an experimental class and the school hour in the computer pool as an experimental hour, while the excellent class as a control class still held its class in a traditional classroom.

Table 1
The English test scores (with the full score 100) of the treatment (average class) and the control (excellent class)

		Pre test	Midterm test	Final test	Vocabulary test
Month		January	April	July	July
Treatment: average class (N=37)	Mean	67.73	77.38	91.21	91.99
	Std. Dev.	14.649	10.364	5.702	7.737
Control: talented class (N=34)	Mean	81.20	87.68	95.00	94.38
	Std. Dev.	7.892	4.946	2.256	4.199
Absolute mean difference between two classes		13.48	10.30	3.78	2.38
Relative mean difference compared with control class		16.60%	11.74%	3.98%	2.52%
Difference between the Std. Dev. of the two classes		6.756	5.417	3.445	3.537
Significance of the independent samples T-test between two classes (2-tailed and equal variances assumed)		0.000	0.000	0.000	0.1894

4.3. Syllabus design

Twelve modules were required to be taught and learned in this school term. Certain amount of English words or phrases was required to be mastered in every module. So we designed both multiple choice quizzes and cloze quizzes in every module for vocabulary acquisition and assessment. Totally there were 436 required English words and phrases in this school term, therefore we produced 436 cloze questions and 436 multi-choice questions.

Because in every module there was also a multiple roles dialogue, we authored twelve dialogue scripts for role play and other twelve scripts for human-computer dialogue based on the dialogue contents, and embedded them on the blended learning website so that the teacher and students in the experimental class can access it.

There were 19 weeks in the experimental school term starting from February, 2011 and ending in July, 2011. Every week the experimental class held one school hour in the multimedia computer pool, and the other school hours still in the normal classroom

as usually. On the contrary, the control class still had all its English class in the normal classroom. While the students in the experimental class reviewed and assessed their vocabulary and dialogue by using the blended learning system, the students in the control class did it via traditional approaches without computer support, such as paper-based or with peers. Except this experimental hour, all the other syllabus design and implementation of the experimental class and the control class remained the same.

In the computer pool, all the multimedia computers are connected via the Internet, so that every student can use one computer individually. A computer and a projector can also be used by the teacher for instructional purposes. In the experimental hour, the students browsed the website of the CSIEC system via Internet and logged into it with their own account and password. Then they did the quizzes by themselves. After submitting the answers the student can read the mark he/she achieved and find the mistakes and feedback. If the student encountered difficulties by finding the answers, he/she can look for them in the textbook or get help from the teacher. This search action strengthened the student's memory of English words or phrases.

By the blended learning in the computer pool, the teacher was still the leader of the instructional process. He/she can encourage or affect the students through his/her speech and body language, and can use the computer projector to show the marks of all students after they have submitted their answers. This instant feedback motivated the students to finish the exam more focused and carefully. Because the computers were connected to the Internet, the teacher's presence prevented the students from browsing games or other websites not related to the class.

Fig. 2. One school hour scenario in the computer pool

Fig. 2 shows the school hour scenario in the computer pool, while the students were doing their quiz.

5. School test results and findings

We collected the ordinary paper-based test scores of two classes throughout the experiment and paid attention to the difference between them. The average scores of the two classes and their standard deviation (Std. Dev.) are listed in Table 1. Their changes along the time line are illustrated by the diagrams in Fig. 3.

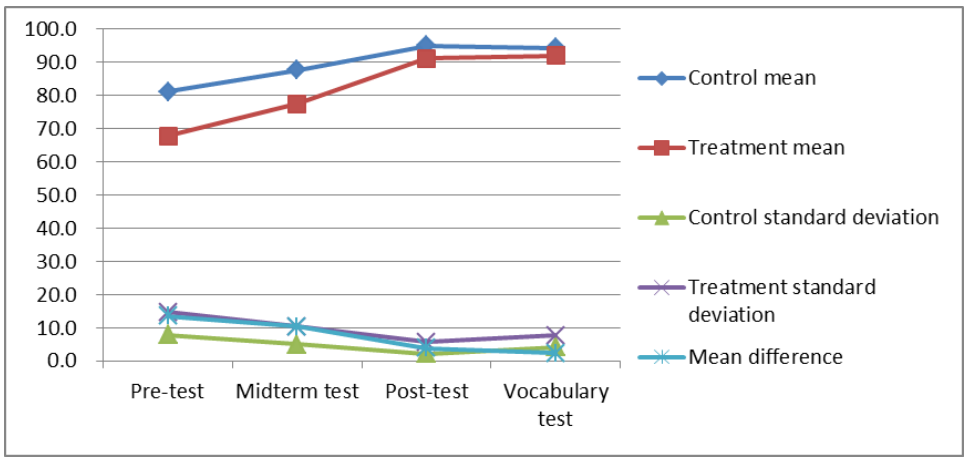


Fig. 3. The comparison of the test performance of the treatment and the control class through the school term

In the pre-test, the excellent class performed much better than the ordinary class with the mean difference 13.5. In April and July there was the midterm exam and the final test, respectively. All the tests assessed the listening, reading and writing skills of the examinees. In all the exams, the excellent class still achieved a better performance than the ordinary class, and the independent sample T-Test analysis with SPSS always showed statistically significant difference between the means of two classes. However, the absolute mean difference was decreased gradually from 13.5 to 3.8. The mean difference of the two classes was decreased by 71.9% throughout the term.

Especially, in order to test the vocabulary acquisition of the students, a vocabulary test was held in July. Though the excellent class performed better than the ordinary class, the mean difference was just 2.4, and was statistically not significant based on the independent sample T-Test analysis with SPSS ($p=0.189>0.05$).

Historical comparison shows that the final test mean (91.21) of the ordinary class was 34.7% greater than that of the last term (67.73), while the final-test mean of the excellent class (95.00) was just 17.0% greater than that of the last term (81.20). Though both classes reported a statistically significant increase at the 0.000 level, the mean difference of the absolute score gain (post test score minus pre-test score) between the treatment and the control is statistically significant at the 0.000 level, as shown in Table 2. The longitudinal improvement is often used to demonstrate the students' performance advancement in educational research. Therefore from the historical view of the student performance evolution, we can conclude that the ordinary class increased its exam performance during the experimental term much more significantly than the excellent class.

As the two classes were in Grade One, their previous test scores were comparable only in one term, specifically in the first term of Grade One in which both of them existed.

The score mean difference at the beginning of the first term was 12.4. Through the first term without blended learning, the mean difference was slightly increased to 13.5.

Table 2
The independent sample T-Test analysis with SPSS software for absolute score gain

Class	N	Mean	Standard deviation	T-test for equality of means (Equal variances assumed)
Average	37	23.49	11.30	t = -4.214, Sig. (2-tailed) =.000
Talented	34	13.79	7.53	

Because the average score stands for the collective performance of an observed class as a whole, we reveal the following findings about the treatment and the control.

1) Before the experiment, the examination mean of the excellent class was statistically significant much more than that of the ordinary class, and this difference already remained for one term without blended learning.
2) Throughout the experiment both classes improved their test performance significantly.
3) The ordinary class's improvement was greater than that of the excellent class so that the mean difference between the two classes was decreased gradually.
4) Though the mean difference between the two classes in the post test remained statistically significant, the difference was just 3.8. Compared with the greater difference 13.5 before the experiment, the difference was decreased by 71.9%.
5) Though the vocabulary test performance of the ordinary class was worse than that of the excellent class, the difference was statistically not significant more, proved by the independent sample SPSS T-Test result.

From those findings, and the fact that the only instructional difference between the two classes was that the ordinary class adopted one school hour blended learning every week with our CSIEC vocabulary and dialogue system, while the control class did not, we can come to the conclusion that the blended learning with our CSIEC system, or the integration of the vocabulary and dialogue assessment system into the ordinary English instruction, improves the students' test performance, and especially the vocabulary acquisition.

6. Student interview results

On May 5th 2011, i.e. two months after launching the experiment, we conducted an interview with five randomly selected students in the average class. We asked them to give free suggestions and comments on this blended learning. The main content of their feedback is summarized in the following citations.

"I have not taken part in any other blended learning class."

"Sometimes it is very slow to start the system' homepage. But it is faster to browse the webpage at home."

"In the cloze question the system requires the exact input of the Chinese meaning of the English word and punctuations, as defined in the answers, what is too inflexible." Other students also complained this problem.

"The interface of the web-based system is very clear and is very easy to use."

"This blended learning in the computer pool is very helpful for vocabulary learning. We have to learn the words by heart every week and do not need to review all the words in a hurry just before the exam."

"Though the cloze question is harder for me than the single choice question, which I like to do in fact, the cloze question can help me better remember the new words."

"This kind of blended learning in a computer pool is more effective to help me learn the vocabulary and other content than the learning in a traditional classroom, because we must concentrate on it, otherwise we cannot get higher scores from the computer."

"Just the words learning function is somewhat boring. However, it improves my learning outcome."

"I wish to continue to use this system in the next term."

"Sometimes it is difficult for me to hear the voice in English."

From these comments and suggestions we are informed that instant scoring and feedback function can motivate the students to master the vocabulary, and the problems such as Chinese meaning input and different pronunciation were challenging the students.

7. Student survey results and findings

To investigate the attitude and feeling of the students toward the blended learning setting, we designed a web-based survey and implemented it in the last experiment school hour in July 2011. The English teacher in the computer pool asked the students in the ordinary class to fill in the survey by clicking a link, which was apparently seen in the course homepage, to the survey website. All thirty-seven students submitted their complete answers to the survey. The survey questionnaire is composed of three parts: basic data, feeling and attitude, and comments and suggestions. We introduce these parts and analyze the results in the following subsections.

7.1. Basic data

The first two questions refer to the students' age and gender. The survey answers indicate that the average age is 13 years, 16 (43.2%) students are male, and 21 (56.8%) are female.

Then two questions address the students' experience in participating in a blended learning setting: "How many English classes with blended learning have you participated in?" and "How many non-English classes with blended learning have you participated in?" The answers indicate that only five students (13.5%) have taken part in one English class with blended learning, and five (13.5%) in more than one classes, while the other 27 (73.0%) have not participated in any class. Only six students (16.2%) have taken part in one non-English class with blended learning, and three (8.1%) in more than one classes, while the other 28 (75.7%) have not participated in any similar class.

7.2. Feeling and attitude

The second part of the survey contains 18 items dealing with the subjective feeling and attitude toward the blended-learning instruction. The answers to those items are measured by a continuous five-point Likert scale with 1 as strong disagreement, 2 as disagreement, 3 as neutral, 4 as agreement and 5 as strong agreement. We use the statistical software SPSS V16.0 to analyze the reliability of the 18 items. For the 18 items the Cronbach's Alpha is 0.903. So the reliability of this survey is very good. The students' answers to the 18 items about their feeling and attitude toward the blended learning setting are listed in Table 3.

Table 3
Questionnaire items and the answer scores

No.	Question	Mean	Std. Dev.	Disagree	Neutral	Agree
Q1	The navigation is so noticeable that I can use all the functions easily.	4.3	0.9	5.4%	5.4%	89.2%
Q2	All links are so reliable that no link error happens.	3.2	1.2	24.3%	35.1%	40.5%
Q3	All the texts in the web pages are understandable.	4.5	1.0	5.4%	5.4%	89.2%
Q4	There is not any grammar, spelling, format or layout error.	3.9	1.2	18.9%	5.4%	75.7%
Q5	This system is adaptive to my learning habit and cognition level.	3.8	1.1	13.5%	18.9%	67.6%
Q6	Some data is lost by using the system and some errors happen.	2.7	1.1	45.9%	29.7%	24.3%
Q7	The content is related to daily life and can be applied to normal English study.	3.9	1.2	10.8%	18.9%	70.3%
Q8	The content can guide me to recall the knowledge learned in the classroom.	3.9	1.0	8.1%	24.3%	67.6%
Q9	The content difficulty is appropriate.	4.1	1.0	5.4%	18.9%	75.7%
Q10	The exams are oriented to the teaching objectives.	4.1	1.1	8.1%	18.9%	73.0%
Q11	The vocabulary and dialogues are strongly associated with the textbook.	4.0	1.0	10.8%	13.5%	75.7%
Q12	This system intrigues and maintains my attention and interest to English learning.	3.8	1.0	10.8%	24.3%	64.9%
Q13	The instant feedback and score can understand and correct the errors.	4.1	1.1	5.4%	18.9%	75.7%
Q14	The individual error set provides me with reflection and retry chance.	4.1	0.9	2.7%	16.2%	81.1%
Q15	The vocabulary pronunciation and dialogue voice can improve my listening and speaking ability.	4.0	1.0	8.1%	16.2%	75.7%
Q16	The system can improve my learning efficiency and test performance.	3.9	0.9	5.4%	29.7%	64.9%
Q17	There are too many quizzes in each module for me to finish them on time.	3.0	1.3	40.5%	21.6%	37.8%
Q18	I would like to continue to use this system in the future study.	4.2	1.2	10.8%	2.7%	86.5%

According to Table 3, we summarize the following findings.

1) The only item with a lower score mean (2.7<3.0) is about the system's defect (Q6). Data lost or other system errors happened to some students (24.3%), but nearly half (45.9%) of them did not experience those problems.

2) The only item with just the neutral score mean (3.0) is about the system's cognitive load (Q17). This neutrality indicates that the system's average cognitive load is appropriate for the whole class.

3) All the other items were answered with a higher score mean than the neutral (3.0). Both the usability, content and pedagogical effect of the blended learning system were recognized by most of the students. As for the usability aspect, the system is easy to use (Q1), reliable (Q2), understandable (Q3), error free (Q4) and adaptive (Q5). The content is practical (Q7), at the point of difficulty (Q9), oriented to the teaching objectives (Q10), associated with the textbook (Q11), and helpful to recall the knowledge (Q8). As for the learning effect, the system is intriguing and interesting (Q12) and can improve the learning performance (Q16), the instant feedback and individual error set are helpful (Q13, Q14), and the vocabulary pronunciation and dialogue voice can improve the listening and speaking ability (Q15). Most students (86.5%) hoped to continue to use this system in the future (Q18).

7.3. Comments and suggestions

The last question is an open text area for the students to write their comments, suggestions and critics on the blended-learning English class. Two students left this text area empty. One made a negative comment:" it is too difficult to use". Six students pointed out the challenges they faced: "Sometimes the word pronunciation cannot be heard", "There should be other types of exercises other than just pure vocabulary quizzes." The other 29 students mainly made positive comments but also gave some suggestions, such as: "It is very useful for vocabulary acquisition and listening comprehension, but it would be better if the Chinese meaning answers could be more flexible", "The network response is too slow."

Summarizing the students' responses in the four parts of the survey, we come to the following conclusion:

1) The CSIEC vocabulary assessment system is easy to use for most students due to its simple interface and clear navigation, although it is the first time for most of them to participate in such a blended learning class.

2) The integration of the vocabulary assessment system into the ordinary English teaching can facilitate the students' vocabulary acquisition including English pronunciation comprehension, spelling writing and Chinese meaning's mastering. It can improve their learning efficacy and test performance.

3) The most students (86.5%) hope to continue to use this system in the future English class.

8. Conclusion and analysis

The collected school test scores of average and talented students confirmed the effectiveness of blended learning on the average student learning outcome, especially on the vocabulary test. The students' survey and interview result also showed their favor to this kind of blended learning with the CSIEC system featured with vocabulary and dialogue exercises. Our hypothesis is established.

The reason for this establishment lies on the immediate, relevant and individualized feedback function both of vocabulary tests and dialogue simulation. Assessment is the most powerful lever with which teachers influence the way students respond to courses and behave as learners (Gibbs, 1999). In the vocabulary assessment part, either in the multiple choice question or cloze, the student must be concentrated on the pronunciation of the word or phrase, exactly type its alphabetic letters one by one in the sequence, and even insert its Chinese meaning as defined in the textbook. As soon as the students submit their answers to the questions, the instant feedback notifies the correctness of the answers, and the students can try again with the right answers. In the human-computer dialogue simulation part, the students must insert the required words or phrases learned in the unit; otherwise the dialogue cannot be continued. This drill and practice function can strengthen the students' memory and retention of the required vocabulary and lead to the statistically not significant mean difference of the both classes in the final vocabulary test. Because the vocabulary is the foundation of foreign language learning, the improved vocabulary retention and usage in practical dialogues can facilitate the students' listening, reading and writing skills to some extent, which are all examined in the academic tests.

However, the mean difference between the two classes in the final exam (posttest) was still statistically significant. This result can be explained by the trivial relevancy of the test content to the learned vocabulary. After scrutinizing the test papers both in the mid-term and the final exam assisted by text mining technology such as word frequency statistics, we find that the test papers and their answers contained only aproximately 20% of the required vocabulary in the experimental term. In other words, 80% of the test contents were not related to the words and phrases learned in this textbook, which were extensively exercised by this blended learning system. The school tests and their results could not reflect the learning content of the students in this term.

From the view point of pedagogical effectiveness, our findings coincide with some findings in Liao (2007), as shown in Table 4. Thus our research represents a new example of the positive effect of computer assisted instruction or blended learning on the students' learning outcome, and especially on the average students' performance. Our findings add credibility to the premise that long-term blended learning or computer assisted instruction can help average students catch up with the talented ones. While our research design and implementation are oriented to the existing average and excellent classes within authentic school context, it can reduce the gap between the theoretical research of educational technology and the pedagogical practice.

This finding is promising because the average students are the very target group with the most potential for further improvement. The CSIEC system and the instructional design of its integration into normal classroom can serve the education policy that attempts to promote the educational equivalency for all students. Therefore the average students should be paid special attention to because they are the most part of the student population. The CSIEC system as a good example of an ICT application can be widely used by the average students to enhance their learning interest and improve their learning

outcome. It can also decrease the teachers' daily heavy burden with the students' assignment distributing and checking.

Table 4
The coinciding findings both in our research and in Liao (2007)

Variable influencing ES	Finding in Liao (2007)	Our case
Duration	The largest mean ES (1.182) was associated with studies lasting 4–18 hours.	One term, including 16 school hours for the blended learning in the computer pool.
Grade level	The ES associated with junior school subjects (7th–9th graders) was the highest (0.847).	First year junior school students, i.e. seventh grade students.
Subject	The mean ES of language/social study (0.664) was the second highest.	English subject as a second language.
Sample size	The small sample (1-50) had the highest ES 0.690.	37
System type	The intelligent CAI software reached the highest ES 1.591.	Web-based intelligent vocabulary and dialogue assessment system.

9. Limitation and future work

Some existing problems were addressed in the student interview and survey. First of all, some students suggested that the input of Chinese meaning for a given English word or phrase should not be limited to the accurate one defined in the textbook. This suggestion is reasonable because various Chinese expressions can have the same semantic meaning and an English word can have different meanings, too. However, as the textbook selected by this school is written by both Chinese and foreign experts in English instruction, and has been widely used in Chinese schools, we believe that the Chinese meaning specified in the textbook for a given English word or phrase is the most exact one in the teaching module or unit, and should be learned by heart by the average students in the first year of the junior school. This content oriented exercise and assessment is designed to reinforce the students' mastery of fundamental vocabulary knowledge. Of course, in the future we should supply more semantically equivalent Chinese expressions to a given meaning of one English word or phrase as the answers to the cloze in the vocabulary assessment.

The same issue appeared in the dialogue simulation function. A great number of expressions, even unlimited, can express the same semantic meaning. But the human computer dialogue function required that the students must insert the expressions containing the specific word or phrase required in the current module. This design was also aimed at enhancing the students' understanding and mastery of the required word or phrase. In the future, we should include more expressions that contain the required word or phrase, on the one hand, and can be understood by the students at their cognitive level, on the other hand.

In our quasi-experiment, the treatment and control group were not equivalent, though both the teacher and the instruction method was identical except the experimental school hour. While the control group had a much higher mean score than the treatment in the pre-test, it is almost impossible for the control to make the same or greater longitudinal jump as the treatment, no matter with traditional pedagogy or blended learning approach. Therefore we should try to find two classes of average students with

the equivalent academic performance in the pre-test, or to divide one class into two equivalent groups. However, in a practical school setting, it is difficult to find such two equivalent classes. The idea of class division would cause prejudicial feeling of the students in the same class, and more importantly, the syllabus design and implementation with two groups would mean doubled work load for the teacher. Such practical school settings cannot be neglected in our academic research.

Acknowledgements

The research in this paper is a part of one ongoing project "The blended instruction research of middle school English subject supported by intelligent tutoring system CSIEC" (10BYY036), sponsored by the National Social Science Council, China. It is also sponsored by the Program for New Century Excellent Talents in University (NCET-10-0179), Ministry of Education, China. We would like to thank all the principals, teachers and pupils in the junior high school for their voluntary participation in this project. We are also grateful for anonymous reviewers' valuable critics, comments and suggestions.

References

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In Guthrie, J. T. (Ed.), *Comprehension and Teaching: Research Perspectives* (pp.71–117). Newark, DE: International Reading Association.

Burns, P. (1981). A quantitative synthesis of research findings relative to the pedagogical effectiveness of computer assisted instruction in elementary and secondary schools. *Dissertation Abstracts International, 42*, 2946A.

Chen, C., & Chung, C. (2008). Personalized mobile English vocabulary learning system based on item response theory and learning memory cycle. *Computers & Education, 51*(2), 624–645.

Chen, L., Ho, R., & Yen, Y. (2010). Marking strategies in metacognition-evaluated computer-based testing. *Educational Technology & Society, 13*(1), 246–259.

Chen, N., Hsieh, S., & Kinshuk, K. (2008). Effects of short-term memory and content representation type on mobile language learning. *Language Learning and Technology, 12*(3), 93–113.

Coady, J. (1993). Research on ESL/EFL vocabulary acquisition: Putting it in context. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second Language Reading and Vocabulary Learning* (pp. 3–23). Norwood, NJ: Ablex Publishing.

Fujishiro, N., & Miyaji, I. (2010). The effects of blended instruction on oral reading performance and their relationships to a five-factor model of personality. *Knowledge Management & E-Learning, 2*(3), 225–245.

Gibbs, G. (1999). Using assessment strategically to change the way students learn. In S. Brown & A. Glasner (Eds.), *Assessment Matters in Higher Education* (pp. 41–53). Buckingham, UK: S.R.H.E. and Open University Press.

Hartley, S. (1978). Meta-analysis of the effects of individually paced instruction in mathematics. *Dissertation Abstracts International, 38*(7-A), 4003.

Huang, H., & Liou, H. (2007). Vocabulary learning in an automated graded reading program. *Language Learning and Technology, 11*(3), 64–82.

Huang, A. F. M., Yang, S. J. H., & Hwang, G. (2010). Situational language teaching in ubiquitous learning environments. *Knowledge Management & E-Learning, 2*(3), 312–327.

Jia, J., Chen, Y., Ding, Z., & Ruan, M. (2012). Effects of a vocabulary acquisition and assessment system on students' performance in a blended learning class for English

subject. *Computers & Education, 58*(1), 63–76.
Jones, L. (2004). Testing l2 vocabulary recognition and recall using pictorial and written test items. *Language Learning and Technology, 8*(3), 122–143.
Kulik, C.-L., & Kulik, J. (1991). Effectiveness of computer-based instruction: an updated analysis. *Computers in Human Behavior, 7*, 75–94.
Liao, Y. (2007). Effects of computer-assisted instruction on students' achievement in Taiwan: A meta-analysis. *Computers & Education, 48*(2), 216–233.
Ligas, M. R. (2002). Evaluation of broward county alliance of quality schools project. *Journal of Education for Students Placed at Risk, 7*(2), 117–139.
Liu, T. (2009). A context-aware ubiquitous learning environment for language listening and speaking. *Journal of Computer Assisted Learning, 25*(6), 515–527.
Liu, T., & Chu, Y. (2010). Using ubiquitous games in an English listening and speaking course: Impact on learning outcomes and motivation. *Computers & Education, 55*(2), 630–643.
Lu, M. (2008). Effectiveness of vocabulary learning via mobile phone. *Journal of Computer Assisted Learning, 24*(6), 515–525.
Lynch, L., Fawcett, A. J., & Nicholson, R. I. (2000) Computer assisted reading instruction in a secondary school: an evaluation study. *British Journal of Educational Technology, 31*(4), 333–348.
Mann, D., Shakeshaft, C., Becker J., & Kottkamp, R. (1999). *West Virginia's basic skills/computer education program: an analysis of student achievement*. Santa Monica, CA: Milken Family Foundation.
McKeown, M. G., Beck, I. L., Omanson, R. C., & Pople, M. T. (1985). Some effects of the nature and frequency of vocabulary instruction on the knowledge and use of words. *Reading Research Quarterly, 20*(5), 522–535.
O'Byrne, B., Securro, S., Jones, J., & Cadle, C. (2006). Making the cut: the impact of an integrated learning system on low achieving middle school students. *Journal of Computer Assisted Learning, 22*(3), 218–228.
Peters, E. (2007). Manipulating l2 learners' online dictionary use and its effect on l2 word retention. *Language Learning and Technology, 11*(2), 36–58.
Stahl, S. (1983). Differential word knowledge and reading comprehension. *Journal of Reading Behavior, 15*(4), 33–50.
Stoller, F. L., & Grabe, W. (1993). Implications for L2 vocabulary acquisition and instruction from L1 vocabulary research. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second Language Reading and Vocabulary Learning* (pp. 24–45). Norwood, NJ: Ablex Publishing.
Tamim, R. M., Bernard, R. M., Borokhovski, E., Abrami, P. C., & Schmid, R. F. (2011). What forty years of research says about the impact of technology on learning: A second-order meta-analysis and validation study. *Review of Educational Research, 81*(1), 4–28.
Tsou, W., Wang, W., & Li, H. (2002) How computers facilitate English foreign language learners acquire English abstract words. *Computers & Education. 39*(4): 415–428
Thorndike, R. (1973). Reading as reasoning. *Reading Research Quarterly, 9*, 135–147.
Wakefield, J. S., Warren, S. J., Rankin, M. A., Mills, L. A., & Gratch, J. S. (2012). Learning and teaching as communicative actions: Improving historical knowledge and cognition through Second Life avatar role play. *Knowledge Management & E-Learning, 4*(3), 258–278.
Wilkins, D. A. (1972). *Linguistics in language teaching*. London: Edward Arnold.
Yueh, H.-P., Lin, W., Huang, J.-Y., & Sheen, H.-J. (2012). Effect of student engagement on multimedia-assisted instruction. *Knowledge Management & E-Learning, 4*(3), 346–357.