Evaluating the Effectiveness of Computer Applications in Developing English Learning

James Todd Whitaker
International College or Suan Sunandha Rajabhat University, Nakhon Pathom, THAILAND.

ABSTRACT

I examined the effectiveness of self-directed learning and English learning with computer applications on college students in Bangkok, Thailand, in a control-group experimental-group pretest-posttest design. The hypothesis was tested using a t test: two-sample assuming unequal variances to establish the significance of mean scores between the two groups for the pretest and the posttest. Both groups studied under normal classroom conditions; however, the experimental group supplemented their learning by using English language learning computer applications outside normal class hours in a self-directed manner. In addition, I reviewed usage and performance metrics with the students at the start of each class to activate and maintain momentum in the learners. At the conclusion of the experiment and after the posttest, the experimental group was surveyed using a semantic differential rating scale to better understand their computer learning experience. The findings were significant; the experimental group learning with the software outperformed the control group in terms of posttest mean scores. The students felt the software tool was “effective” and rated it highly; however, finding time to use the software was scored as “easy” but with a bias toward “hard.”

KEYWORDS
Computer language learning, self-directed learning, English software learning applications

ARTICLE HISTORY
Received 3 September 2016
Revised 30 November 2016
Accepted 22 December 2016

Introduction

As a prerequisite to graduating from International College, Suan Sunandha Rajabhat University (SSRUIC), students must pass an English exit exam. The curriculum is conducted in English. Students from Thailand make up 99% of the student body and are therefore working towards obtaining a bachelor’s degree in their non-native language. Most students’ English skills, upon entering SSRUIC, are limited and would therefore be problematic in passing the English exit exam for graduation. Academic achievement in terms of a degree for the students is twofold: satisfactorily completing the required coursework and passing the English exit exam. Supposing effort is required for said academic achievement, this study focuses on students’ demonstrated effort in using technology to learn English outside the classroom in the form of self-directed learning.

Self-directed learning is a key factor in this research. Other important factors considered in this paper include effort, discipline, motivation, self-
regulated learning, conditioning, attitude, and self-efficacy beliefs. Implementing “various difficulties” for the student results in an effortful learning process essential for maximizing test performance (Schmidt & Bjork, 1992). The “various difficulties” students faced in this research was the challenge to use computer technology through self-directed learning outside the classroom to acquire and improve their English skills. In essence, students were engaged in a conditioning process of frequent, if not daily, use of the software to enhance and improve their English proficiencies balanced with regular performance reviews. This task, or “various difficulty,” implies that a certain conditioning process must take place in order to get through the “various difficulties” and achieve success. Conditioning implies repetition. Achievement or goal attainment implies a certain amount of expertise has been reached through concentrated effort (conditioning and repetition or frequency).

Given the demands by industry and government, especially in light of the impending Asean Economic Community (AEC), the International College is committed to yielding graduates proficient in English. This paper looks at the effectiveness of using a computer application in developing students’ English skills while simultaneously compelling effort on the part of the learners. “Compelling effort,” meaning challenging students in self-directed and outside of normal class hours English learning. Recognizing the need to improve the English skills of not only International College students, but also to students located in provinces outside of Bangkok, SSRUIIC anticipates adding a computer instruction tool to enhance students’ development of English skills. So, evaluating the effectiveness of the computer instruction tool must be accomplished. I intend to evaluate the effectiveness of such tool. This study aimed to evaluate the effectiveness of using computer applications to learn English and whether or not this method of instruction produced results significantly different between SSRUIIC students using the application and those not using the application.

**Literature Review**

Since this study was conducted in Thailand where English is not the first language, the review of literature relies mainly on research done in countries where English is not the first or official language.

Klentien and Kamnungwut (2015) studied the impact of English learning with electronic media for elementary and high school students in Thailand. They held that students need to learn English both inside and outside the classroom and that learning outside the classroom assumed the use of electronic media (computer software). They tested the relationship between time usage of the software and achievement on lessons completed during a 30-day data collection period. They concluded that time spent using the software was positively correlated with higher lesson scores, increased vocabulary learning, and therefore enhanced English skills and proficiency levels.

Al-Mansour and Al-Shorman (2012) researched the influence of computer-assisted English learning on Saudi Arabian students. Their objective was to determine if learning English by computer instruction together with the traditional method was significantly different from non-computer learning (traditional method only). The study was carried out using a control-group experimental-group pretest-posttest design. The findings showed student achievement to be statistically different, favoring the experimental group using
computer-aided instruction together with the traditional method over the control group using the traditional method alone.

Mahmoudi, Samad, and Razak (2012) studied Iranian postgraduate students to evaluate the impact Computer Assisted Language Learning (CALL) had on learning vocabulary and to what extent having a positive attitude affects learning outcome. The students were given a questionnaire to measure attitude and were also given a vocabulary pretest and posttest. Following the pretest, a separate group of students studied vocabulary using websites. The results demonstrated that students were positive about using computers to learn English vocabulary and that attitude measurement showed a positive association with CALL performance.

Tang, Wong, and Wong (2015) interviewed Chinese nursing students at Tung Wah College in Hong Kong, via focus group interviews, to determine if their completion of an online English language learning course had improved their English skills, especially the use of clinical English associated with their profession. Of the 100 participants in the study, only 20 completed the online course. The outcomes suggested that the online English learning course did improve the nursing students English with the most positive aspects about the experience being interactive software, schedule flexibility, and pertinence to their career field.

Studying the effect of images used in dynamic graphical interfaces on computer-aided English vocabulary learning, Chang, Lin, and Lee (2005) found that the use dynamic graphical images enhanced the subject’s learning experience by increasing interest, motivation, and achievement in learning English.

Many studies in this review consider the importance of self-efficacy and its impact on improved or positive learning. Zimmerman (2000) claims that self-efficacy is a key incentive to learning. Bandura (1977) labels supposed self-efficacy as an individual belief structure where the force of one’s certainty in their own capability will influence whether or not effort is spent and that “cognitive processes play a prominent role in the acquisition and retention of new behavior patterns” (p. 192). Bandura (1977) states “the stronger the perceived self-efficacy, the more active the efforts” (p. 194) and “Given appropriate skills and adequate incentives, however, efficacy expectations are a major determinant of people’s choice of activities, how much effort they will expend, and of how long they will sustain effort in dealing with stressful situations” (p. 194).

In a study of web-based English learning (WBEL), Chen (2014) examined which factors most affected Taiwan college students’ proactive “stickiness” or stick-to-iteniveness within a framework of Social Cognitive Theory (SCT) and Uses and Gratifications (U&G) theory. Results of the study revealed the most significant elements affecting students’ stick-to-iteniveness to WBEL were learning gratifications, learning outcome expectations, followed by learning climate. Chen’s research proved consistent with the Uses and Gratifications (U&G) theory study performed by Guo, Tan, and Cheung (2010), referenced by Chen (2014):

Based upon the U&G perspective, they concluded that the anytime and anywhere features of web-based learning technologies were seen to play a crucial role in gratifying students by reinforcing their learning expectancy value and creating a comfortable learning climate with open-minded social interactions. (p. 168)
Lai (2015) explored the use of self-directed language learning (outside the classroom) by university students in Hong Kong using computer technology in demonstrating what impact, if any, teachers had on self-directed learning. Lai (2015) interviewed 15 students and also conducted an online survey of 160 participants to collect data for her research. Results showed that students were positive towards using technology for language learning, were convinced they could succeed in using technology to learn a language, and were using technology outside the classroom on average 1–3 hours per week. Teacher recommendation, teacher guidance of use, and teacher use of technology in the classroom (setting the example) proved to be the key positive influences on student use of self-directed technology learning.

Continuing with self-directed inquiries, Kim, Olfman, Ryan, and Eryilmaz (2014) researched the use of online education in terms of self-directed learning. They argued that although learning online is growing and popular, there are many difficulties blocking students' achievement in using electronic media. Factors mentioned are lack of motivation, lack of discipline, and lack of satisfactory time-management skills. To address these difficulties, they developed a self-directed learning system (SDLS) that would support the students' self-directed learning and include attributes such as being able to record learning goals, keeping a list of resources, maintaining learning schedules, and generating milestone completions. Their study consisted of a control group and an experimental group using a pretest and posttest design. The experimental group used the SDSL. The results were positive regarding the use of the SDSL: comparing the means between the two groups in accomplishing self-directed learning activities, the experimental group's mean was statistically different and superior to that of the control group. Kim et al. (2014) emphasized the need for designing a system that is both "personalized" and "collaborative" in order to help the students achieve success in self-directed learning.

Sandberg, Maris, and de Geus (2011) investigated mobile English learning with primary school students to determine if students who used the mobile learning application would outperform classroom-only learning students and if those students using the mobile application were sufficiently motivated for self-directed learning away from the classroom. The study was structured using an English vocabulary pretest posttest. The findings showed the group of students using the mobile application in class and at home outperformed the groups not using mobile learning and not using self-directed study. The higher performing group demonstrated sufficient motivation for self-directed learning by using the application during non-directed or unscheduled uses (using their own initiative).

Kim, Wang, Ahn, and Bong (2015) investigated Korean students' self-efficacy beliefs and how they use self-regulation tactics while studying English. By means of convenience sampling, they asked 167 Korean undergraduate students about their self-efficacy beliefs and self-regulated learning strategies. In addition, students were required to submit their Test of English for International Communication (TOEIC) as a measurement of their English proficiency for the conducted research. The findings showed three self-efficacy profiles: low, medium, and high. The medium and high efficacy profiles were dominated by females and represented students with more years of English learning than those in the low self-efficacy profile; additionally, the low self-efficacy profile was
significantly different from the medium and high profile in terms of self-regulated learning strategies.

Razak, Ismail, Aziz, and Babikkoi (2012) assessed the English language learning strategies employed by secondary students in Malaysia and whether there was a difference between female and male learning strategies. They used a modified version of the Oxford Strategy Inventory for Language Learning (SILL) to evaluate six language learning strategies: memory, cognitive, compensation, metacognitive, affective, and the social domains. The findings showed that all six strategies were used by the students with the affective strategy favored most and that females applied more strategies than males and females were more proficient at learning language as well.

Abdolahzadeh and Nia (2014) sought to understand Iranian students’ language learning beliefs, and what factors affected those beliefs and whether or not their self-efficacy was positively correlated to their English language proficiency level. They administered a paper-based type of Key English Test (KET) to measure student proficiency levels as well as a Beliefs about Language Learning Inventory (BALLI) questionnaire. The results indicated that students’ language learning beliefs about motivation and expectations were most important and that Iranian learners believed they would learn to speak English very well.

Lee, Thang, Tan, Ng, Yoon, Chua, and Shirlenna (2014) explored, using a descriptive survey, Malaysian university learners’ sensitivities and plans for using technology for English teaching and learning. The current assumption is that technology can improve upon traditional classroom English teaching and learning instruction. The findings of the survey revealed that although students maintained that technology helped improve their language learning, they actually used technology more in their daily activities rather than for academic purposes.

Methodology

This research study was based on a pretest and posttest design comprising two states, control and experimental. The population I studied was first year undergraduate students enrolled in the Bachelor of Arts Airline Business Program at International College, Suan Sunandha Rajabhat University in Bangkok, Thailand. The instruments used in this study were a pretest, posttest, English learning computer software, and a paper-based questionnaire completed by the experimental group surveying their experience using the software. A semantic differential rating scale was used in the questionnaire. The control group and the experimental group were randomly selected among four groups of students (5, 6, 7, & 8) enrolled in English for Airline Operations and Management summer class 2015. One thousand iterations of random numbers were generated by assigning a probability of 25% to each group number. The least frequent number generated (5) was selected as the control group and the most frequent number generated (8) was selected as the experimental group. The null hypothesis is there is no significant difference between the observed mean of the control group and that of the experimental group, $H_0$: $\mu_{control} = \mu_{experimental}$. Both the control and experimental groups were instructed by the same teacher (the author of this research). The control group consisted of 21 students; 14 (67%) were female and 7 (33%) were male. The experimental group consisted of 24 students; 14 (58%) were female and 10 (42%) were male.
The control and experimental groups were pretested prior to experiment launch and post tested with the same test at experiment completion. The 39-question test, designed by the researcher, tested reading comprehension, vocabulary, sentence and word completion, and conversation. The experimental group completed a questionnaire based on their experience using the computer application to learn English. Students in the control group attended and participated in class normally. In addition to attending and participating in class normally, each student in the experimental group received a CD containing a computer application for learning English. Once loaded, this tool functioned on the internet and allowed the students to study and learn English outside of class in a self-directed learning manner. I encouraged the students to use the software outside of class to increase their English skills. The duration of the study took place during the June – July 2015 timeframe for a period of 30 days. The duration between the pretest and posttest was approximately 60 days. To test the hypothesis, a t test (two-sample assuming unequal variances, heteroscedastic) was calculated to determine the significance of mean scores between the control and experimental groups for the pretest and the posttest. The questionnaire utilized a 7-step semantic differential rating scale measuring 15 responses to assess the students’ true feelings about the self-directed computer application English learning experience in the following dimensions: evaluative, strength, and activity (Osgood, 1952). There was a two-question comments section at the end of the survey, which some students completed. The data in this study were analyzed fully in Microsoft Excel 2013.

**Results**

SSRUIC students in the experimental group demonstrated significantly better English language skills through the self-directed use of an English learning computer application, as measured by mean posttest scores (M = 28.54, SD = 4.55, N = 24), than did those in the control group who did not learn with such application (M = 24.19, SD = 5.56, N = 21), t(39) = -2.85, p = 0.007. Therefore, this computer application was effective in developing English learning for students at SSRUIC in Bangkok, Thailand, and so we reject the null hypothesis and conclude the difference in means is significant. Furthermore, the students in the experimental group improved their mean scores, measured from pretest to posttest, by 8.90% compared to 0.00% in the control group.

The expected result from the pretest should provide evidence that both groups were essentially equivalent in demonstrated English skills prior to experiment start. The expected result was realized. There was no significant difference in the mean scores between the control group (M = 24.19, SD = 5.07, N = 21) and the experimental group (M = 26.21, SD = 4.38, N = 24), t(40) = -1.42, p = 0.164.

In analyzing computer application usage and performance during the 30 day experimental period, weekly hourly usage was (M = 3.21, SD = 1.49, Min = 1.10, Max = 7.57). Students learned 968.21 new vocabulary on average (SD = 498.76 Min = 95, Max = 2502). I used new vocabulary learned per hour to measure student productivity (M = 77.99, SD = 30.69, Min = 13.34, Max = 135.69) (see Summary and Discussion and Figure 1). The software operates with artificial intelligence algorithms specific to each student’s level and proficiency and assigns a performance rating score, which includes level of difficulty, time usage, unit completion, unit score, and new vocabulary learned. Students learned at a level
of difficulty of their choosing, from least difficult (level 1) to most difficult (level 20). At the start of the experiment with initial usage data, student levels were (M = 11.36, SD = 4.85, Min = 1, Max = 20) and at experiment conclusion (M = 13.13, SD = 3.77, Min = 2, Max = 20). Student overall performance ratings were (M = 2380.70, SD = 897.76, Min = 528.73, Max = 4486.76) and total points accumulated were (M = 5616.21, SD = 3447.45, Min = 425, Max = 18772). In terms of gender, female weekly hourly usage (M = 3.39, SD = 1.54, Min = 1.10, Max = 7.57) was greater than male usage (M = 2.96, SD = 1.37, Min = 1.25, Max = 6.34). Female students learned more new vocabulary (M = 1071.21, SD = 568.87, Min = 95, Max = 2502) than male students (M = 824.00, SD = 329.14, Min = 248, Max = 1402). However, male students outperformed female students on overall performance rating (M = 2444.26, SD = 1018.51, Min = 769.71, Max = 3881.63), (M = 2335.31, SD = 797.34, Min = 528.73, Max = 4486.76), respectively. They also outperformed on the posttest (M = 30.20, SD = 4.92, Min = 24, Max = 39), (M = 27.36, SD = 3.67, Min = 19, Max = 33), respectively. Yet, females did better on the pretest (M = 26.29, SD = 3.71, Min = 20, Max = 33) than did males (M = 26.10, SD = 4.99, Min = 18, Max = 38).

The results from the semantic differential rating scale survey, in terms of highest mean score (+3.00 to -3.00), indicated the students’ beliefs: (1) (good-bad) the computer application was “good” (M = 2.58, SD = .76), (2) (like-dislike) “like” the computer application (M = 2.46, SD = .58), (3) (weaker-stronger) English became “stronger” by learning new vocabulary (M = 2.25, SD = .72), (4) (effective-ineffective) using the computer application was “effective” (M = 2.21, SD = .87), (5) (negative-positive) higher performance ratings caused “positive” feelings (M = 2.08, SD = 1.61), (6) (weaker-stronger) English is “stronger” after using the tool (M = 2.04, SD = 0.93), (7) (better-worse) English skills are “better” after using the computer application (M = 2.04, SD = 0.84), (8) (clear-confusing) the computer application is “clear” (M = 1.96, SD = 0.89), (9) (worse-better) vocabulary is “better” after using the application (M = 1.92, SD = 1.32), (10) (awful-fantastic) self-directed learning is “fantastic” (M = 1.83, SD = 1.11), (11) (fantastic-awful) learning on my own outside of class is “fantastic” (M = 1.79, SD = 1.00), (12) (fun-boring) the software application is “fun” (M = 1.75, SD = 1.30), (13) (happy-sad) when I am using the application I feel “happy” (M = 1.71, SD = 1.14), (14) (easy-

![Figure 1](image-url)
hard) finding time to use the computer application is “easy” with a bias toward “hard” (M = 0.67, SD = 1.82), and (15) (hard-easy) doing the exercises is “easy” with a bias toward “hard” (M = 0.58, SD = 1.11) (see Figures 2 and 3).

Osgood (1952) submitted three factors for measuring semantic differential meaning: evaluative, strength, and activity. Grouping the survey questions in terms of dimension produced evaluative (M = 2.03, SD = 1.16), strength (M = 2.04, SD = 0.97), and activity (M = 1.14, SD = 1.52) results. Activity measured the lowest mean result indicating the students were challenged by finding time to use the software and by the difficulty of the exercises. Evaluative and strength dimensions indicated students’ overall positive attitude toward using the software and the belief that their English skills were strengthened.

<table>
<thead>
<tr>
<th>Good</th>
<th></th>
<th>Bad</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective</td>
<td></td>
<td>Ineffective</td>
</tr>
<tr>
<td>Clear</td>
<td></td>
<td>Confusing</td>
</tr>
<tr>
<td>Fun</td>
<td></td>
<td>Boring</td>
</tr>
<tr>
<td>Worse</td>
<td></td>
<td>Better</td>
</tr>
<tr>
<td>Better</td>
<td></td>
<td>Worse</td>
</tr>
<tr>
<td>Hard</td>
<td></td>
<td>Easy</td>
</tr>
<tr>
<td>Happy</td>
<td></td>
<td>Sad</td>
</tr>
<tr>
<td>Negative</td>
<td></td>
<td>Positive</td>
</tr>
<tr>
<td>Easy</td>
<td></td>
<td>Hard</td>
</tr>
<tr>
<td>Like</td>
<td></td>
<td>Dislike</td>
</tr>
<tr>
<td>Weaker</td>
<td></td>
<td>Stronger</td>
</tr>
<tr>
<td>Awful</td>
<td></td>
<td>Fantastic</td>
</tr>
<tr>
<td>Stronger</td>
<td></td>
<td>Weaker</td>
</tr>
<tr>
<td>Fantastic</td>
<td></td>
<td>Awful</td>
</tr>
</tbody>
</table>

Figure 2. Bullet points positioned on a 7-step semantic differential rating scale (Osgood, 1952) reflect student’s attitude toward using the computer application in developing their English learning.

Figure 3. Results of 15 question semantic differential rating scale survey from 24 student respondents. Each bullet point depicts the mean result from a question of semantic measurement of students’ feelings towards using the application in developing English learning. Upper and lower limits measuring the 7-step rating scale were scored from +3 to -3.
Summary and Discussion

The goal of this study was to determine whether or not using computer applications was an effective tool for learning English. Results indicated that the computer application used in this experiment was an effective tool for learning English. These findings are consistent with other studies (e.g., Klentien & Kamnungwut, 2015; Al-Mansour & Al-Shorman, 2012; Lai, 2015; Razak, Ismail, Aziz, & Babikkoi, 2012; Sandberg, Maris, & de Geus, 2011; Kim, Wang, Ahn, & Bong, 2015). Comparing mean scores on the posttest, the experimental group outperformed the control group by 18% and improved their mean score by 8.9%.

The demands on the students in the experimental group were to use the computer application on their own and outside of class (i.e., self-directed learning) on a daily basis. At the start of each class during the experimental period, I reviewed usage and performance statistics with the students. The data were presented in Excel for all group members to see. My purpose was to provide real-time performance metrics to the students as a sort of “scoreboard” to help the learners stay focused and motivated in using the application. I believe this process served as a positive factor for subsequent effort and use of the tool, along with developing discipline, time management skills, and self-efficacy.

On average, students used the application 3.21 hours per week, however; the top quartile of students was at four plus hours per week. My personal expectation was 3 hours per week, which was not revealed to the students. On mean lesson difficulty levels: students increased this metric by almost two points (15.6%) from 11.36 to 13.13. Five students learned at a difficulty rating of 15 or above and three of those students were learning at 20, the most difficult level.

As a teacher my initial challenge was measuring and understanding students’ usage of the application in a self-directed manner outside of class, because, in analyzing time spent using the software certain discrepancies became apparent. For example, early usage and performance reviews with the learners was dominated by actual time spent using the application, whereby, top performers were recognized for putting the most time in and lower performers were recognized for not doing so. However, following initial reviews and deeper investigation, the data revealed that the number of hours spent using the software was not the primary determinant of developing English skills. Discussions with the class on time usage revealed that in some cases students would turn on the application, do exercises, leave the application, and then return to doing exercises later in the day, resulting in a high hourly usage statistic, but in fact actual learning time was much less. Therefore, the primary determinant to actual learning, at a minimum the best assumption, was the number of new vocabulary learned, i.e., actual learning, combined with the time factor produced a valuable productivity metric: New vocabulary learned per hour (see Figure 4). This productivity measure became a key metric in measuring student output and was presented to the experimental group in all subsequent performance reviews. This enabled students to grasp their actual learning performance in terms of time and learning.

In reviewing the results of the semantic differential survey, students rated the application highly, near the maximum score of “good.” They liked using the software to learn English and felt that the application was effective and strengthened their vocabulary. Not surprisingly, students reported that finding time to use to tool was scored lower at “easy” but with a bias toward “hard” and
that doing exercises in the application was “easy,” also with a bias toward “hard.” I assume the students felt the “hard” part was in self-directed learning and in learning a second language. However, they found that the application itself was “clear” as opposed to “confusing.” As to self-directed learning, the students scored that aspect as “fantastic,” but at the low end, as opposed to “awful.”

![Figure 4. Student productivity measure compared to level of difficulty. Although the figure displays declining productivity in terms of mean new vocabulary learned per hour (left axis), students increased learning effort in terms of mean level of difficulty (right axis), which may explain the divergence.](image)

Notwithstanding the significant results found herein, this research has limitations that may be useful to future research. To supplement the survey, a one-on-one detailed “exit interview” with each student in the experimental group could provide a deeper understanding of the learner’s experience using the computer application (e.g., attitude, behavior, study habits, time management, actual time of day usage, frustrations, application deficiencies or desired improvements). How would the results differ if I did not conduct the frequent in-class reviews of usage and performance statistics? By not implementing performance reviews, more self-directed or self-motivated learners might be easily identified, thus providing potential insights into the habits of higher performing learners. Further investigation is needed into student output in terms of the new vocabulary learned per hour productivity measure (see Figure 4): Increases in level of difficulty intuitively explains the decrease in productivity, but what else could be behind the decline? Also, the speaking module in the application, requiring students to obtain a microphone, was not utilized to its fullest extent, therefore, emphasizing this feature in future studies may have a positive impact on future results as well as improving students’ English skills.
In light of this research, I recommend the self-directed use of computer applications outside of class to develop English learning for SSRUIC students, Thailand students in general, adult learners, and English learners not living in a native English speaking country. However, to help students succeed, it is important that teachers implement and lead a performance or measurement review, in class, of student output on at least a weekly basis. This process benefits the student in the form of immediate feedback, facilitates in maintaining learner focus and motivation, and benefits the teacher with valuable insights and discoveries from continuous data analysis.

Disclosure statement
No potential conflict of interest was reported by the authors.

References


