Effect of the Blended Learning Model in English Course on Students' English Proficiency

Weiting Liu[1] and LeeHsing Lu[2]

[1,2] Graduate School of Business and Advanced Technology Management, Assumption University, Thailand
[1]E-mail: davidweitingliu@yahoo.com, ORCID ID: https://orcid.org/0009-0002-7984-1496
[2]E-mail: leelu@mail.com, ORCID ID: https://orcid.org/0000-0002-4818-1440

Abstract

Background: Traditional college English instruction in China has been largely teacher-centered, which often limits student engagement and language skill development. The growing adoption of blended learning—integrating face-to-face and digital instruction—has presented a promising solution. The Superstar mobile learning platform has emerged as a tool to support this innovative approach.

Objective: This study aimed to evaluate the effectiveness of Superstar-based blended learning in enhancing English language proficiency among college students. It also sought to assess students' engagement and their perceptions of the blended learning environment.

Methodology: A quasi-experimental design was employed with two groups of students enrolled in a college English course at Shenyang University. The control group (n=60) received traditional instruction, while the treatment group (n=62) engaged in Superstar-based blended learning over eight weeks. Pre-tests and post-tests assessed listening, reading, and writing skills using the CET-4 framework. A 5-point Likert-scale questionnaire measured student engagement and perception. Data were analyzed using SPSS 26 through paired and independent samples t-tests.

Results: Students in the treatment group showed significantly greater improvements in listening, reading, and writing scores compared to the control group (p < .05). Engagement levels were significantly higher in the treatment group across behavioral, emotional, social, and cognitive dimensions. Perception data indicated strong agreement with the pedagogical, social, and technical design of the Superstar-based learning model.

Conclusion: The Superstar-based blended learning model significantly enhances college students' English proficiency and engagement. Students positively perceived the platform's instructional design and flexibility. These findings support the integration of mobile-based blended learning tools in college English instruction to improve educational outcomes.

Keywords: Blended Learning; College English; English Proficiency

Introduction

The rapid development of information technology has brought significant changes to people's work and lives, gradually driving new educational reforms (AbuSa'aleek, 2014). With the extensive advancement of network information technology and the ongoing evolution of educational concepts, individuals increasingly recognize the potential of new technologies to enhance education (Selwyn, 2011). In recent years, according to Fukuda and Yoshida (2013), the emergence of many online courses and the widespread use of information-based teaching tools have led to a growing interest in blended online and offline teaching. This approach combines the benefits of face-to-face and online learning and yields better teaching outcomes. It is particularly effective in achieving student-centered learning and fostering a revolution in both teaching and learning (Moskal et al., 2013).

Implementing blended learning in educational activities relies heavily on the integration of technology. This includes utilizing learning management systems, video conferencing tools, digital content, and various online resources. Technology facilitates interactions, tracks progress, and provides access to various educational materials (Picciano, 2009).

In traditional English classes, a teacher-centered approach is typically employed (Hamad, 2017). For many years, teachers, particularly in Asia (Littlewood, 1999), have been the focal point of the classroom, serving as the primary source of knowledge and imparting information deemed suitable by themselves and the curriculum planners. According to Ogata and Yano (2005), a significant problem with traditional classroom teaching is its teacher-centered approach. Another issue with traditional

teaching is content overload. Modern educational technologies offer numerous tools and resources to enhance learning, but are often underutilized in traditional settings. Addressing these issues is crucial for improving the quality and effectiveness of education.

Based on the results above, as for the integration of the blended learning based on Superstar in College English classes, the following research questions were going to be solved. This research aims to explore whether the college students' English proficiency would be affected after using blended learning, how the students' engagement in college English would be influenced, and what the students' perception of the college English course using blended teaching is based on.

Objectives

Based on the research questions, the research objectives are outlined as follows:

1. To assess the effectiveness of the Superstar-based blended learning model in enhancing students' academic performance in college English.

2. To investigate how the Superstar-based blended learning model affects student engagement in college English.

3. To evaluate students' perceptions of college English courses delivered through the Superstar blended teaching approach.

Literature review

Introduction of College English in China

College English teaching (CE) refers to providing English instruction to students who are not majoring in English at Chinese colleges and universities, as defined by Xiao (2015). It is divided into two primary levels: one targets undergraduate students who are not English majors, while the other serves non-English major college students.

Evaluating college English teaching and its impact is mainly conducted through standardized testing. Both the Syllabus for Science and Engineering programs and the Syllabus for Arts and Science courses have established specific guidelines for assessing college English courses through testing. According to the stipulations of these syllabuses (Xiu, 2001), college English tests are classified into three distinct levels: department examination, semester-end college examination, and the national unified College English Test (CET). The CET is administered after two years of foundational English instruction.

According to the Chinese National Education Examinations Authority (NEEA, 2023), the College English Test (CET) is a comprehensive national standardized exam overseen by the Ministry of Education in China. The primary goals of CET-4 and CET-6 are to support the implementation of the college English curriculum guidelines, evaluate college students' English proficiency objectively and accurately, and provide a basis for assessing and improving the quality of college English instruction (SCET, 2016). College English teaching is mutually reinforcing and promotional, suggesting that the quality and effectiveness of the teaching process are interdependent and influence each other reciprocally (SCET, 2016). As teaching methods and approaches improve, so does the student's ability to learn and master the language, prompting further advancements in teaching strategies. Conversely, any limitations or challenges in the teaching process can impede student learning and vice versa.

This research would concentrate on the foundational level of college English teaching, mainly targeting undergraduate students not majoring in English across various universities.

Introduction to Superstar Application

Superstar is a professional mobile learning platform for smartphones, tablets, and other mobile devices (Zhang & She, 2021). According to the Ministry of Education of the People's Republic of China (2020), Superstar applications can be utilized for teaching and learning activities in education. It offers many learning resources and supports mobile reading, video browsing, live streaming, social sharing, and more anytime and anywhere. Its goal is to create a private library. Superstar features

multiple sections, including lesson plans, chapters, materials, notifications, assignments, exams, discussions, and statistics, which provide excellent hardware conditions for developing blended learning methods.

Teaching activities based on Superstar can be divided into those for teachers and students. Teacher activities include establishing courses, recording, integrating, and uploading learning resources, publishing and notifying learning tasks, online communication and interaction, monitoring teaching, preparing lessons, and other related tasks. Student activities encompass entering the classroom, receiving learning tasks and notifications, self-directed learning, online communication, asking questions, submitting learning tasks, and other related tasks.

Superstar offers robust teaching support for educators and efficient services for students to learn independently. Its characteristics include ease of use, abundant resources, powerful functions, wide application, and rapid classroom evaluation. Therefore, this study demonstrates that Superstar provides a viable solution for integrating blended learning in online and offline environments.

Introduction to Blended Learning

Until recently, many definitions of blended learning have emerged. According to Valiathan (2002), the term "blended learning" refers to an educational approach that incorporates diverse delivery methods, including collaboration software, web-based courses, electronic performance support systems (EPSS), and knowledge management practices. This learning methodology encompasses event-based activities such as traditional face-to-face classrooms, live e-learning sessions, and self-paced independent study.

Bruff et al. (2013) advocate using blended teaching methodologies that align with contemporary educational trends. Blended learning emerges as an innovative approach that integrates traditional classroom methods with new techniques, enabling educators to identify and address issues in conventional teaching practices to enhance effectiveness. Halverson et al. (2014) found that the design of blended learning environments and the evaluation of their effectiveness are currently at the forefront of academic discussion and inquiry.

Bonk et al. (2009) highlighted the process of blended teaching design, outlining five key steps. Initially, teachers should provide effective classroom instruction. Second, building a positive relationship with students through interactive communication is crucial. Third, promoting group collaboration and exploration among students is essential. Fourth, teachers administer classroom practice tests tailored to real-world teaching scenarios. Finally, after the class, teachers assign homework that aligns with the learning outcomes achieved.

According to a survey conducted by Porter and Graham (2014), the blended learning process encompasses three key stages. First, teachers must devise an overarching plan for blended learning. Secondly, they must effectively manage the organizational and execution aspects during the teaching process. Lastly, they need to ensure the smooth progression of the blended teaching approach.

In addition to the significant research focus on blended learning models, scholars have also devoted attention to evaluating their teaching effectiveness. McCarthy and Murphy (2010) argue that blended learning (BL) contributes to the redesign of course curricula, leading to improved student learning outcomes, such as higher grades, more excellent content knowledge, and a deeper understanding of course materials. In his study, Sembiring (2018) suggested that blended learning can effectively enhance students' overall learning experience by employing various delivery methods, ultimately resulting in improved academic achievement.

Geçer and Dağ (2012) utilized a case study approach to highlight students' positive perceptions of distance education and satisfaction with the blended learning environment. Conversely, Medina (2018) found that blended learning enabled the delivery of online resources for students to prepare before class, ultimately reducing their cognitive load and enhancing their academic performance. Blended learning methods are more effective than traditional teaching approaches (Bath & Bourke, 2010). Supported by evidence from numerous scholars' teaching practices, many educators adopt

blended learning in their classrooms, making it a preferred teaching method among contemporary instructors.

Blended Learning in Language Learning Activities Based on Technology

In today's rapidly evolving educational landscape, technology plays an increasingly vital role in language learning (Puentedura, 2009). One approach that has gained significant attention is blended learning, which combines traditional classroom instruction with digital tools and resources (Hass & Joseph, 2018). This method leverages the advantages of face-to-face and online learning environments, providing learners with a more dynamic and engaging educational experience (Selwyn, 2011).

Blended learning can take many forms in language learning. For instance, learners may participate in synchronous online sessions with a teacher to discuss grammar rules or vocabulary while engaging in asynchronous activities such as reading digital texts or completing online exercises (Sharma & Barrett, 2007). Combining online and offline elements creates a personalized learning path that caters to learners' needs and preferences (Hass & Joseph, 2018). Technology is essential for implementing blended learning in language activities (Sharma & Barrett, 2007). Here are some key technologies that enable and enhance this approach:

Online Learning Platforms: MOOCs (Massive Open Online Courses) and LMS (Learning Management Systems) provide learners access to various language courses and resources. These platforms often include interactive tools such as forums, quizzes, and simulations, facilitating active learning (Boyd & Kasraie, 2013).

Mobile Applications: Mobile devices have transformed language learning by allowing learners to study on the go. Language learning apps offer various features, including vocabulary practice, speaking exercises, and language immersion through authentic content (Hass & Joseph, 2018).

Artificial Intelligence (AI) and Machine Learning: AI-powered tools can assist learners in numerous ways, from personalized learning plans to automated grammar and pronunciation correction. These technologies provide learners instant feedback and help them progress at their own pace (Garg, 2020).

Interactive Media: Incorporating videos, audio files, and animations in language learning materials makes the content more engaging and easily comprehended. These interactive media elements enhance learners' ability to retain information and improve their understanding skills (Peake & Reynolds, 2020).

Conceptual Framework

This research was to explore students' English proficiency after using blended teaching based on Superstar, students' engagement, and their perception. Before and after an 8-week pilot of blended learning using Superstar in College English classes, the two selected students took pre- and post-tests. The assessment tool was the College English Test Band-4, which evaluated listening, reading, and writing skills. To examine academic achievement, the study analyzed scores in these areas. Additionally, students from the two groups completed a questionnaire after the eighth week to share their views on the engagement; students from the experimental group took part in the questionnaire to respond to some questions about the perception of this teaching mode. Consequently, the research focused on four variables: listening comprehension, reading comprehension, writing skills, and students' engagement and perception. The conceptual framework for this study is outlined below:
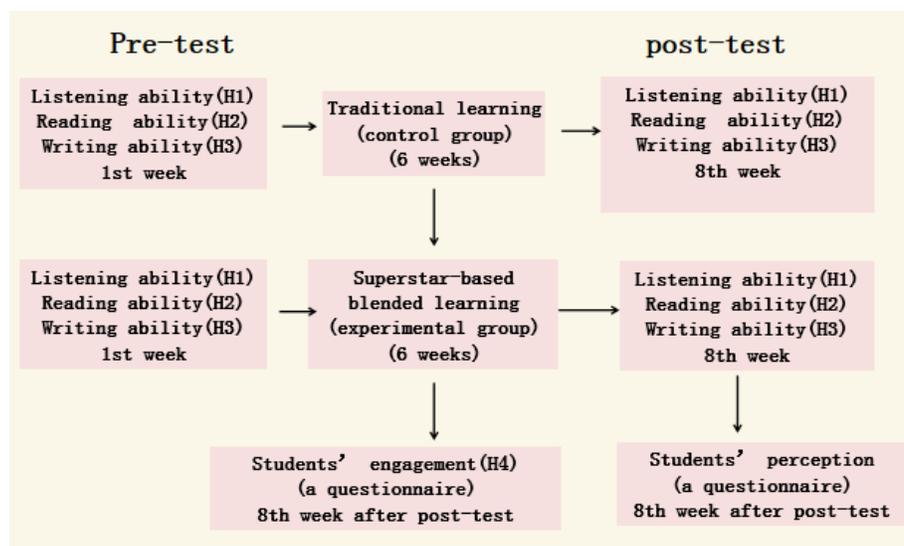
Figure 1 Conceptual Framework of Superstar-based Blended Learning
Source: Constructed by the Author

According to the conceptual framework, there were such hypotheses as follows:

$H_0 1$ There is no difference in listening comprehension ability between pretest and post-test in the control group.

$H_a 1$ There is a difference in listening comprehension ability between the pretest and post-test in the control group.

$H_0 2$ There is no difference in listening comprehension ability between the pretest and post-test in the experimental group.

$H_a 2$ There is a difference in listening comprehension ability between the pretest and post-test in the experimental group.

$H_0 3$ There is no difference in listening comprehension ability between the control and experimental groups.

$H_a 3$ There is a difference in listening comprehension ability between the control and experimental groups

$H_0 4$ There is no difference in reading comprehension ability between pretest and post-test in the control group.

$H_a 4$ There is a difference in reading comprehension ability between the pretest and post-test in the control group.

$H_0 5$ There is no difference in reading comprehension ability between pretest and post-test in the experimental group.

$H_a 5$ There is a difference in reading comprehension ability between the pretest and post-test in the experimental group.

$H_0 6$ There is no difference in reading comprehension ability between the control and experimental groups.

$H_a 6$ There is a difference in reading comprehension ability between the control and experimental groups.

$H_0 7$ There is no difference in writing ability between pretest and post-test in the control group.

$H_a 7$ There is a difference in writing ability between the pretest and post-test in the control group.

$H_0 8$ There is no difference in writing ability between pretest and post-test in the experimental group.

H_a8 There is a difference in writing ability between the pretest and post-test in the experimental group.

H_09 There is no difference in writing ability between the control and experimental groups.

H_a9 There is a difference in writing ability between the control and experimental groups.

H_010 There is no difference in students' engagement in College English learning activities between the control and experimental groups.

H_a10 There is a difference in students' engagement in College English learning activities between the control and experimental groups.

Methodology

This research adopted a mixed method, incorporating a quasi-experiment and a structured questionnaire. A quasi-experimental design would be employed to evaluate the effectiveness of blended learning using Superstar in college English. Additionally, a questionnaire would be administered to gather insights into students' engagement and perceptions of this blended learning environment.

In the quasi-experiment, two distinct instructional methods would be implemented in parallel: the experimental group would participate in College English classes utilizing blended learning based on Superstar, while the control group would follow traditional teaching methods for College English. Subsequently, data collected through pre- and post-tests would undergo rigorous analysis to determine if the blended learning approach leveraging Superstar enhances students' academic performance and whether any statistically significant differences emerge from the traditional teaching method.

The research also aims to investigate whether there are noticeable differences in students' listening comprehension, reading comprehension, and writing abilities after experiencing the studied technical intervention or treatment. The objective is to assess the potential impact of this treatment on their English proficiency.

This research employed a questionnaire to gather students' feedback on their engagement and perceptions within the Superstar-based blended learning model for college English. The goal is to analyze and evaluate these responses to comprehensively understand students' experiences with this mixed learning approach in the Superstar-facilitated college English course. Ultimately, the objective is to offer constructive suggestions that could enhance the overall quality of college English education.

Population

The population of this study consists of college students who are required to enroll in a college English course at Shenyang University. These students may come from various grades, majors, and backgrounds. However, they share the common trait of learning English at the college level, incorporating online and offline teaching resources and learning activities in their educational process.

Sample

In this study, the sample would consist of two classes of students majoring in Mathematics Education from the Mathematics Department at Shenyang University of Liaoning in China, who are scheduled to participate in the college English course together: the same teacher, textbook, and time. In June 2023, they took part in the University Entrance Examination, and their performance in English is expected to reflect similar proficiency.

Research Instruments and Tools

The research instruments in this study are performance tests and questionnaires that the researcher needs to use to collect data.

In the experiment, the CET-4 examination papers would serve as assessment tools for students. Xiao (2015) summarized that the objectives of both CET-4 and CET-6 are to refine the design of college English teaching curricula, evaluate students' English proficiency, establish a foundation for enhancing language skills, and foster a mutually reinforcing and restrictive dynamic within college English instruction. Two standardized CET-4 examination papers would be utilized for the pre-and post-testing

phases. The treatment and control groups would undergo a pre-test and a post-test, consisting of three distinct sections: listening comprehension, reading comprehension, and writing.

A questionnaire survey would serve as an additional research instrument for data collection. Following the conclusion of the 8th-week post-test for both groups, students in both classes would be administered a questionnaire utilizing a five-point Likert scale. In the data collection process, questionnaires are used to measure students' engagement and satisfaction. A 5-level Likert is used to collect the data. CVI results of the questionnaire for students' engagement in college English learning activities and for students' perception towards Superstar-based blended learning in college English are proven to be accepted.

Data Collection and Analysis

The experiment lasted for eight weeks. In the first week, a pretest was administered for both groups. The traditional teaching method was implemented in the control group from the second to the seventh week. Simultaneously, the experimental group adopted the blended learning approach based on Superstar. In the eighth week, a posttest was conducted for both groups. At the end of June 2024, a content validity index (CVI) and a pilot test were performed to assess the validity and reliability of the questionnaire. During the eighth week, questionnaires were distributed among both groups.

The questionnaire comprises three parts: demographic information, engagement, and perception. Participants in the control group completed parts one and two of the questionnaire, while those in the experimental group responded to the entire questionnaire. To ensure that the process is efficient, effective, and respectful of the participants' time and privacy, the first step in distributing the questionnaire is to determine the timing for distribution and collection to allow ample time for completion. The second step involves obtaining approval from school administrators, relevant authorities, and students. The third step is to distribute the questionnaires directly to students during class using the online survey tool Questionnaire Star.

In this research, score improvement was calculated between two groups through pre-tests and post-tests to identify the impact on students' language proficiency in Superstar-based blended learning for college English. The statistical analysis tool SPSS (Statistical Package for the Social Sciences) (version 26) was used for data analysis, independently. The paired samples T-tests were used to compare the score improvement between the two groups.

Questionnaires were employed to collect data on students' engagement and perception. The internal consistency of the items in the questionnaire was confirmed by calculating Cronbach's Alpha. The data collected from the questionnaire were analyzed using SPSS 26 with the independent samples T-test.

Results

Results for Students' Language Proficiency of Listening, Reading, and Writing

This study assessed the improvement in listening, reading, and writing scores between the pre-test and post-test within and between the two groups. To determine if there was a significant difference in these scores within the control group (using traditional teaching methods) and the treatment group (using Superstar-based blended learning), paired and independent samples t-tests were conducted.

The first group of hypothesis testing for the control group was outlined as follows:

H_01: There is no difference in listening comprehension ability between the pretest and post-test in the control group.

H_a1: There is a difference in listening comprehension ability between the pretest and post-test in the control group.

H_04: There is no difference in reading comprehension ability between the pretest and post-test in the control group.

H_a4: There is a difference in reading comprehension ability between the pretest and post-test in the control group.

H$_0$7: There is no difference in writing ability between the pretest and post-test in the control group.

H$_a$7: There is a difference in writing ability between the pretest and post-test in the control group.

Table 1 Paired Samples T-test Results for Listening, Reading, and Writing Scores between Pre-test and Post-test in Control Group (n=60)

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
air 1	Listening scores Pre-test & Post-test	-10.767	9.468	1.222	-13.212	-8.321	-8.809	59	.000
air 2	Reading score Pre-test & Post-test	-13.767	26.154	3.377	-20.523	-7.010	-4.077	59	.000
air 3	Writing score Pre-test & Post-test	-7.333	18.520	2.391	-12.118	-2.549	-3.067	59	.003

Table 1 displays the results of the paired-sample t-test for the control group's listening, reading, and writing scores. A total of 60 participants took part in both the pre-test and post-test.

The mean of paired differences for listening scores is 10.767, indicating a slight increase in the post-test. The t-value is 8.809, with degrees of freedom (df) of 59 and a two-tailed significance level (Sig. (2-tailed)) of 0.000. Since 0.000 < 0.05, there is a significant difference between the pre-test and post-test listening scores.

The mean of paired differences for reading scores is 13.767, suggesting a significant improvement following the test. The t-value is 4.077, df = 59, and the significance is 0.000 < 0.05, indicating a significant change between the two testing periods.

The mean of paired differences for writing scores is 7.333, reflecting an increase in the post-test. The t-value corresponding to a significance level of 0.003 < 0.05 indicates a significant difference in writing scores between the two testing periods for the control group.

Consequently, the results indicate that all three null hypotheses were rejected. This signifies that the control group experienced significant changes in scores across these three areas when traditional learning methods were applied.

The second group of hypothesis testing for the treatment group is as follows:

H$_0$2: There is no difference in listening comprehension ability between the pre-test and post-test in the experimental group.

H$_a$2: There is a difference in listening comprehension ability between the pre-test and post-test in the experimental group.

H$_0$5: There is no difference in reading comprehension ability between the pre-test and post-test in the experimental group.

H$_a$5: There is a difference in reading comprehension ability between the pre-test and post-test in the experimental group.

$H_0$8: There is no difference in writing ability between the pre-test and post-test in the experimental group.

H_a8: There is a difference in writing ability between the pre-test and post-test in the experimental group.

Table 2 Paired Samples T-test Results for Listening, Reading, and Writing Scores between Pre-test and Post-test in Treatment Group (n=62)

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
air 1	Listening scores Post-test & Pre-test	18.403	37.569	4.771	27.944	8.862	3.857	61	.000
air 2	Reading score Post-test & Pre-test	23.565	17.073	2.168	27.900	19.229	10.868	61	.000
air 3	Writing score Post-test & Pre-test	15.290	10.775	1.368	18.027	12.554	11.173	61	.000

Table 2 displays the paired-sample t-test results for the listening, reading, and writing scores of a treatment group consisting of 62 participants in the pre-test and post-test.

For listening scores, the mean of paired differences is 18.403, indicating that, on average, the post-test scores are 18.403 points higher than the pre-test scores. With a t-test value and a two-tailed significance level (Sig. (2 2-tailed)) of .000 (less than the standard threshold of 0.05), there is a highly significant difference between the pre-and post-test listening scores.

Regarding reading scores, the mean of paired differences is 23.565, suggesting a substantial increase after the test. The significance level is also .000, indicating a significant improvement in reading scores from the pre-test to the post-test.

The mean of paired differences for writing scores is 15.290, demonstrating an upward trend. The significance level of .000 implies a significant difference between the pre-test and post-test writing scores.

Therefore, the results indicate that all three null hypotheses were rejected. Overall, the treatment had a significant positive impact on the participants' performance in these three areas.

The third group of hypothesis testing between the control group and the treatment group is as follows:

$H_0$3: There is no difference in listening comprehension ability between the control and experimental groups.

H_a3: There is a difference in listening comprehension ability between the control and experimental groups.

$H_0$6: There is no difference in reading comprehension ability between the control and experimental groups.

H_a6: There is a difference in reading comprehension ability between the control and experimental groups.

$H_0 9$: There is no difference in writing ability between the control and experimental groups.

$H_a 9$: There is a difference in writing ability between the control and experimental groups.

By calculating and comparing the score improvements in listening, reading, and writing, this study measured the increase in these three areas from pre-test to post-test between both groups and applied an independent samples t-test to investigate whether there was any difference in these score improvements between the control group (using traditional teaching) and the treatment group (using Superstar-based blended learning) after the experiment to explore the effect of Superstar-based blended learning.

Table 3 Mean Summary for Students' Improvement of Listening, Reading, and Writing Scores between the Control Group and Treatment Group

	Group	N	Mean	Std. Deviation
Listening improvement	Control	60	1.77	9.468
	Treatment	62	18.40	37.569
Reading improvement	Control	60	13.77	26.154
	Treatment	62	23.56	17.073
Writing improvement	Control	60	2.33	18.520
	Treatment	62	15.29	10.775

Table 4 Independent Samples T-test Results for Students' Improvement of Listening, Reading, and Writing Scores between the Control Group and Treatment Group

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
Listening improvement	Equal variances assumed	41.525	.000	-3.329	120	.001	-16.637	4.998	-26.531	-6.742
	Equal variances are not assumed.			-3.378	68.962	.001	-16.637	4.925	-26.463	-6.811
Reading improvement	Equal variances assumed	7.329	.008	-2.458	120	.015	-9.798	3.986	-17.690	-1.906

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
	Equal variances are not assumed.			-2.442	101.071	.016	-9.798	4.013	-17.758	-1.838
Writing improvement	Equal variances assumed	9.137	.003	-4.742	120	.000	-12.957	2.732	-18.367	-7.547
	Equal variances are not assumed.			-4.703	94.209	.000	-12.957	2.755	-18.427	-7.487

Table 4 displays the independent samples t-test results regarding the improvement of students' listening, reading, and writing scores between the control group and the treatment group.

Levene's Test for Equality of Variances: The F-values for listening, reading, and writing are 41.525, 7.329, and 9.137, respectively, with all Sig. Values are less than 0.05 (0.000 for listening, 0.008 for reading, and 0.003 for writing). This indicates that the variances between the two groups are not equal in these three aspects, violating the assumption of equal variances for a standard t-test.

As for the test for equality of means: When equal variances are assumed for listening, the t-value is 3.329; df is 120, and Sig. (2-tailed) is 0.001. When not assuming equal variances, the t-value is 3.378, df is 68.962, and Sig. is also 0.001. The mean difference of 16.637 indicates that, on average, the treatment group improves listening scores more than the control group. The 95% confidence interval (6.742 - 26.531 for equal variances assumed; 6.811 - 26.463 for not assumed) does not include zero, suggesting a significant difference. Regarding reading, the equal variances assumed and not assumed cases have a Sig. (2-tailed) value of 0.015 and 0.016, respectively. The mean difference of 9.798 implies that the treatment group has a significantly greater improvement in reading scores. The 95% confidence intervals (1.906 – 17.690; 1.838 – 17.758) further confirm the substantial gap between the two groups. For writing, the Sig. (2−tailed) Values are 0.03 in both variance-assuming scenarios. The mean difference of 12.957 indicates that the treatment group has a higher average improvement in writing scores. The 95% confidence intervals (7.547 - 18.367; 7.487 - 18.427) also support the existence of a significant difference.

Therefore, the results show that all three pairs of null hypotheses were rejected. In conclusion, in all three areas of listening, reading, and writing, there are significant differences in score improvement between the control and treatment groups, with the treatment group generally demonstrating more substantial improvements. In brief, the outcomes of both paired and independent sample t-tests determined the acceptance or rejection of nine pairs of hypotheses during the testing process. Table 5 outlines the results of the hypothesis testing conducted in this study.

Table 5 Summary of Hypothesis Testing and Results

Hypotheses	Statements	Results
H_01	There is no difference in listening comprehension ability between pretest and post-test in the control group.	Rejected
H_02	There is no difference in listening comprehension ability between the pretest and post-test in the experimental group.	Rejected
H_03	There is no difference in listening comprehension ability between the control and experimental groups.	Rejected
H_04	There is no difference in reading comprehension ability between the pretest and post-test in the control group.	Rejected
H_05	There is no difference in reading comprehension ability between pretest and post-test in the experimental group.	Rejected
H_06	There is no difference in reading comprehension ability between the control and experimental groups.	Rejected
H_07	There is no difference in writing ability between the pretest and post-test in the control group.	Rejected
H_08	There is no difference in writing ability between the pretest and post-test in the experimental group.	Rejected
H_09	There is no difference in writing ability between the control and experimental groups.	Rejected

Results for Students' Engagement

For this study, SPSS 26 was used to analyze student engagement quantitatively using the questionnaire. An independent samples t-test was performed to determine whether a significant difference exists in student engagement between two groups: the experimental group, which utilized Superstar-based blended learning in College English, and the control group, which followed traditional teaching methods.

The questionnaire survey explored students' engagement in four aspects: behavioral, social, emotional, and cognitive. This research utilized a 5-point Likert scale to collect participants' attitudes towards these variables. To analyze the collected data, predefined levels were employed to interpret the average value of each variable.

Table 6 Arbitrary Level for Interpretation of Questionnaire Data

Likert scale	Range	Interpretation
5	4.51 - 5.00	Strongly Agree
4	3.51 - 4.50	Agree
3	2.51 - 3.50	Neutral
2	1.51 - 2.50	Disagree
1	1.00 - 1.50	Strongly Disagree

Source: Norman, G. (2010). Likert scales, levels of measurement, and the "laws" of statistics. Advances in Health Sciences Education, 15(5), 625-632.

Table 7 Results of Independent Samples Test for Students' Engagement (n-62 treatment group; n-60 control group)

| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
| | | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
									Lower	Upper
Behavioral Engagement	Equal variances assumed	5.216	.024	36.173	120	.000	3.125	.086	2.954	3.296
	Equal variances are not assumed.			36.150	119.410	.000	3.125	.086	2.954	3.296
Social Engagement	Equal variances assumed	7.610	.007	28.009	120	.000	2.941	.105	2.733	3.149
	Equal variances are not assumed.			27.846	104.649	.000	2.941	.106	2.732	3.151
Emotional Engagement	Equal variances assumed	9.656	.002	31.781	120	.000	3.005	.095	2.818	3.193
	Equal variances are not assumed.			31.583	102.722	.000	3.005	.095	2.817	3.194
Cognitive Engagement	Equal variances assumed	19.628	.000	30.291	120	.000	3.158	.104	2.952	3.364
	Equal variances are not assumed.			30.120	105.309	.000	3.158	.105	2.950	3.366

Table 7 shows that the assumption of variance homogeneity is invalid (with a significance level of F less than 0.05), indicating a significant variance difference between the experimental and control groups. When examining the t-statistic, its significance level (two-tailed) is 0.000, which is less than 0.05, resulting in the rejection of the null hypothesis. This indicates a significant difference in participation between the two teaching methods. In summary, the results of the independent sample t-test during the hypothesis testing process led to the rejection of the original hypothesis. Table 8 summarizes the hypothesis testing results for students' engagement in this study.

Table 8 Summary of Hypotheses for Students' Engagement

Hypotheses	Statement	Result after Analysis
$H_0 10$	There is no difference in students' engagement in College English learning activities between the control and experimental groups.	Rejected

Results for Students' Perception in the Experimental Group

In this study, SPSS 26 was used to analyze students' perception questionnaire data quantitatively. The questionnaire survey includes three variables: perception of pedagogical design, perception of social design, and perception of technical design. This study utilized a 5-point Likert scale (Agreement)

to assess the sample's attitudes toward these variables. To interpret the collected data, any score below the mean value of each variable is used to explain the overall perception.

Table 9 Descriptive Statistics of Students' Perception in Treatment Group(n=62)

	Item Statement	Mean	Std. Deviation	Interpretation
Perception of pedagogical design	1. Each lesson has clear learning objectives.	4.58	.641	Strongly Agree
	2. The structure of each lesson is simple to follow.	4.55	.619	Strongly Agree
	3. The learning environment's design improves my ability to concentrate on the subject matter.	4.60	.586	Strongly Agree
	4. The requirements for assignments are clearly explained.	4.55	.645	Strongly Agree
	5. The planned activities are designed carefully.	4.56	.643	Strongly Agree
	6. The course content was effectively delivered within a blended learning environment.	4.58	.641	Strongly Agree
	7. The course content was presented in a concise and easily understandable way.	4.60	.557	Strongly Agree
	Average	4.57	.620	Strongly Agree
Perception of social design	1. I communicate with my classmates in college English by Superstar.	4.58	.641	Strongly Agree
	2, I can make clear what I can't understand with my teacher's help.	4.56	.643	Strongly Agree
	3, I can make clear what I can't understand with my classmates' help.	4.56	.668	Strongly Agree
	4. Other students respond quickly and positively to my requests for help.	4.52	.620	Strongly Agree
	5. The teacher provides feedback on my work in time.	4.60	.586	Strongly Agree
	6. My classmates and I routinely engage in peer evaluation of each other's work.	4.55	.619	Strongly Agree

	Item Statement	Mean	Std. Deviation	Interpretation
	7, I was encouraged and supported by my teacher and classmates' positive attitude.	4.52	.646	Strongly Agree
	Average	4.64	.632	Strongly Agree
Perception of technical design	1. I communicate with my classmates in college English using Superstar.	4.55	.619	Strongly Agree
	2, I can make clear what I can't understand with my teacher's help.	4.58	.588	Strongly Agree
	3, I can make clear what I can't understand with my classmates' help.	4.56	.562	Strongly Agree
	4. Other students respond quickly and positively to my requests for help.	4.63	.607	Strongly Agree
	5. The teacher provides feedback on my work promptly.	4.61	.523	Strongly Agree
	6. My classmates and I routinely engage in peer evaluation of each other's work.	4.63	.550	Strongly Agree
	Average	4.59	.575	Strongly Agree

According to the survey results, the average values for all statements concerning the three aspects—perception of pedagogical design, perception of social design, and perception of technical design—are above 4.5. The survey results show students favor learning college English through a Superstar-based blended approach. They appreciate the pedagogical strategies, social aspects, and technical design of the environment, all of which contribute to their overall learning experience and satisfaction.

Discussion

The research findings showed that the Superstar-based blended learning approach was efficient and feasible for college English education. The data revealed a positive impact on students' learning abilities. These experimental results are consistent with previous research, such as the study by Seery and Donnelly (2012), which found that blended learning effectively enhances students' academic performance.

After conducting an eight-week experiment, students in the experimental group showed significant improvements in their listening, reading, and writing skills. The research findings align with previous scholars' assertions (Reima, 2021) that blended learning positively impacts these abilities. To meet the teaching objectives, the Superstar-based learning activities engaged students and played a crucial role in enhancing their language skills.

Researchers (Reima, 2021; Seery & Donnelly, 2012) found that students expressed satisfaction with their learning experience in this mode, noting significant achievements that boosted their confidence and willingness to continue learning.

Conclusion

This paper examines the application of a Superstar-based blended learning approach in college. The research focuses on the impact of this teaching model on students' academic performance in listening, reading, and writing, as well as their engagement and satisfaction.

A series of hypothesis tests and evaluations were conducted. The experiment compared students' pre-test and post-test scores in listening, reading, and writing between the treatment and control groups. The results demonstrate significant improvements in all three areas. Students exhibit better comprehension in listening, capturing more details, and understanding complex dialogues. In reading, they display enhanced skills in analyzing texts and extracting key information. Furthermore, their written compositions become more coherent, with improvements in grammar and vocabulary usage.

Regarding student engagement, the Superstar platform offers various interactive features, such as online discussions, group projects, and instant feedback, effectively attracting student participation. Surveys on student perceptions also reveal positive outcomes. Students value the flexibility of the blended learning model, which combines in-class instruction with online self-study, and feel that it aids their learning efficiency.

In conclusion, Superstar-based blended teaching in college English courses has proven effective in promoting students' academic performance, enhancing their engagement, and increasing their satisfaction with the learning process. This study encourages educators and educational institutions to consider investing more time, effort, and resources into integrating various technologies into teaching, as this has the potential to boost student motivation and improve teaching effectiveness.

Recommendation

During experimentation, several aspects of the research offer valuable insights for educators, researchers, and educational institutions, meriting further investigation with additional time, funding, and effort. Despite participants acknowledging the effectiveness of Superstar-based blended learning in language studies, they encountered challenges such as technological issues and poor self-discipline. For teachers, implementing Superstar-based blended learning presented numerous practical hurdles, necessitating advanced technical proficiency and superior professional abilities in selecting and reviewing content for students, ultimately leading to a heavier workload compared to traditional classroom settings.

References

AbuSa'aleek, A. O. (2014). A review of emerging technologies: Mobile-assisted language learning (MALL). *Asian Journal of Education and e-Learning (ISSN: 2321–2454)*, 2(06).

Bath, D., & Bourke, J. (2010). *Getting started with blended learning*. Griffith University Press.

Bonk, C. J., Graham, C. R., & Cross, J. (2009). The handbook of blended learning: Global perspectives, local designs. *Turkish Online Journal of Distance Education*, 10(4), 181–181.

Boyd, G., & Kasraie, N. (2013). Can MOOC fires bring light to shadow education? *International Journal of Learning and Development*, 3(4), 87–95. http://dx.doi.org/10.5296/ijld.v3i4.4111

Bruff, D. O., Fisher, D. H., McEwen, K. E., & Smith, B. E. (2013). Wrapping a MOOC: Student perceptions of an experiment in blended learning. *Journal of Online Learning and Teaching*, 9(2), 187. Chinese Ministry of Education. (2020). *The College English Teaching Guide*.

Fukuda, S. T., & Yoshida, H. (2013). Time is of the essence: Factors encouraging out-of-class study time. *ELT Journal*, 67(1), 31–40.

Garg, S. (2020). Artificial intelligence and its impact on higher education in the post-COVID era. *University News*, 58(36), 2–6.

Geçer, A., & Dağ, F. (2012). A blended learning experience. *Educational Sciences: Theory & Practice*, 12, 425–442.

Halverson, L. R., Graham, C. R., & Spring, K. J. (2014). A thematic analysis of the most highly cited scholarship in the first decade of blended learning research. *Internet & Higher Education*, 20(1), 20–34.

Hamad, M. M. (2017). Using WhatsApp to enhance students' learning of the English language: Experience to share. *Higher Education Studies*, 7(4), 74–87.

Hass, A., & Joseph, M. (2018). Investigating different options in course delivery—Traditional vs online: Is there another option? *The International Journal of Information and Learning Technology*, 35(4), 230–239.

Littlewood, W. (1999). Defining and developing autonomy in East Asian contexts. *Applied Linguistics*, 20(1), 71–94.

McCarthy, M. A., & Murphy, E. A. (2010). Blended learning: Beyond initial uses to helping to solve real-world academic problems. *Journal of College Education & Learning*, 7(6), 67–70.

Medina, L. C. (2018). Blended learning: Deficits and prospects in higher education. *Australasian Journal of Educational Technology*, 34(1).

Moskal, P., Dziuban, C., & Hartman, J. (2013). Blended learning: A dangerous idea? *The Internet and Higher Education*, 18, 15–23.

NEEA: National Education Examinations Authority. (2023). *National College English Test Committee of CET-4 and CET-6*. https://cet.neea.edu.cn/html1/folder/19081/5123-1.htm

Ogata, H., & Yano, Y. (2005). Knowledge awareness for computer-assisted language learning using handhelds. *International Journal of Learning Technology*, 5(1), 435–449.

Peake, J., & Reynolds, A. (2020). Implementing social media bridges for student-teacher chasms created during the COVID-19 pandemic. *CEA Critic, 82*(3), 274–284. https://doi.org/10.1353/cea.2020.0022

Picciano, A. (2009). Blending with purpose: The multimodal model. *Journal of the Research Center for Educational Technology, 5*(1), 4–14. https://doi.org/10.1016/j.iheduc.2012.12.001

Porter, W. W., Graham, C. R., & Spring, K. A. (2014). Blended learning in higher education: Institutional adoption and implementation. *Computers & Education, 75*(3), 185–195.

Puentedura, R. R. (2009). *As we may teach: Educational technology, from theory into practice*. https://itunes.apple.com/itunes-u/as-we-may-teach-educational/id380294705

Reima, A. J. (2021). Mobile audiobooks, listening comprehension, and EFL college students. *International Journal of Research*, 9(4), 410–423.

SCET. (2016). Syllabus for College English Test - Band Four (CET-4) (2016 Revision). (2016). https://cet.neea.edu.cn/html1/folder/19081/5123-1.htm

Seery, M. K., & Donnelly, R. (2012). Implementing pre-lecture resources to reduce in-class cognitive load: A case study for higher education chemistry. *British Journal of Educational Technology, 43*, 667–677.

Selwyn, N. (2011). *Education and technology: Key issues and debates* (1st ed.). Continuum.

Sembiring, M. G. (2018). Validating student satisfaction with a blended learning scheme in the Universitas Terbuka setting. *International Journal of Mobile Learning and Organization, 12*, 394–413.

Sharma, P., & Barrett, B. (2007). *Blended learning: Using technology in and beyond the language classroom*. Palgrave Macmillan.

Syllabus for College English Test - Band Four (CET-4) (2016 Revision). (2016). https://cet.neea.edu.cn/html1/folder/19081/5123-1.htm

Valiathan, P. (2002). *Blended learning models*. American Society for Training & Development.

Valiathan, P. (2002). *Blended learning models*. American Society for Training & Development.

Xiao, Y. M. (2015). The development of the College English test and its influence upon College English teaching. *China Examinations, 5*, 59. https://doi.org/10.19360/j.cnki.11-3303/g4.2015.08.008

Xiu, X. B. (2001). The developing course of the College English teaching and learning in China. *Journal of Science of Jiamusi University, 19*(5), 161.

Zhang, F., & She, M. (2021, January). The effectiveness of "Superstar" in College English teaching. In *2020, 3rd International Seminar on Education Research and Social Science (ISERSS 2020)* (pp. 487–490). Atlantis Press.