




The effects of blended teaching approach on university students' English learning outcomes: A three-level meta-analytic review

Gang Yang¹ · Qiongqiong Kuang² · Renfeng Jiang¹ · Renfeng Jiang¹ 

Received: 22 May 2024 / Accepted: 3 September 2024

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

Abstract

Recognized as an important teaching approach, blended teaching has been widely applied in English language instruction within university settings. Nonetheless, discrepancies exist in the reported results regarding the efficacy of blended teaching approach on university students' English learning outcomes. This study employs a comprehensive three-level meta-analysis method, analyzing 373 distinct effect sizes from 207 empirical studies. The three levels correspond to the sampling variance, the variance among effect sizes within individual studies, and the variance between different studies, allowing for a nuanced understanding of the factors influencing learning outcomes. Our findings reveal a significant overall positive impact of blended teaching on English learning outcomes. Specifically, we identify that the relationship between blended teaching and language achievement is nuanced, moderated by variables such as duration of implementation and interaction type, while remaining unaffected by teaching method, form of learning and blended mode. Additionally, we observe that the relationship between blended teaching and learners' personal characteristics is moderated by duration of implementation, teaching method and blended mode, yet remains independent of form of learning and interaction type. These findings are poised to inform and enhance English language pedagogy within the framework of blended teaching method.

Keywords Blended teaching approach · English learning outcomes · University students · Three-level meta-analysis

Extended author information available on the last page of the article

Published online: 18 September 2024

1 Introduction

The widespread application of “student-centered classrooms” in English language education has underscored educators’ focus on student learning outcomes (Starkey, 2019). Against this backdrop, blended teaching approach, which integrates traditional classroom instruction with online learning, has emerged as a prevailing paradigm in university education in recent years (Yu et al., 2022). Departing from conventional teacher-centric approaches, it offers enhanced flexibility and individualization in the learning process, prompting educators to reassess and adapt their pedagogical practices (Müller & Mildenerger, 2021). In the domain of English as a second language (ESL) education, blended teaching represents a significant trend in higher education, with numerous studies reporting its positive impact on learning outcomes (Ramalingam et al., 2022). However, variations in research methodologies and inconsistent findings have led to uncertainties regarding the effectiveness of the blended teaching approach in promoting university students’ English language acquisition. This ambiguity impedes educators and administrators from fully grasping its role and hinders its broader adoption.

To this end, this study employs a three-level meta-analysis to systematically synthesize the diverse quantitative findings on blended teaching and ESL learning outcomes. Traditional meta-analysis methods, which typically select only one effect size per study, are insufficient for capturing the complexity of studies that report multiple effect sizes. The three-level meta-analysis allows for the extraction of all relevant effect sizes and accounts for variance at the levels of sampling, within-study effect sizes, and between-study differences (Ran et al., 2022). This approach provides a comprehensive view of the data, enabling the exploration of intricate relationships and moderating variables that might otherwise be overlooked. Compared to conventional random effects models, the three-level meta-analysis offers greater flexibility and robustness, accommodating variations at different levels of analysis and enhancing the overall interpretation of the effectiveness of blended teaching approaches.

1.1 The effects of blended teaching approach on university students’ English learning outcomes

Blended teaching, is an approach to education that integrates digital educational resources and online interaction opportunities with conventional classroom-based instruction. Its characteristic is believed to leverage the advantages of both traditional classroom instruction and online teaching while mitigating their respective shortcomings (Bernard, 2014). Blended teaching has become increasingly prevalent in ESL education, particularly for university students aiming to attain both academic and professional qualifications, as it is well-suited to meet the diverse needs of ESL students who often require flexible learning schedules while managing various personal and professional commitments. The flexibility of blended teaching allows for a more personalized learning experience, enabling students to engage with materials at their own pace and time (Zhang & Zhu, 2018). Meanwhile, it facilitates a richer interaction with the content through various digital tools and platforms that enhance language learning and engagement, which not only supports the acquisition

of language skills but also integrates digital literacy as a critical component of modern education (Banditvilai, 2016). Studies therefore indicate that such integrative approaches can lead to higher motivation, better management of learning time, and improved academic outcomes (Min et al., 2019; Zainuddin & Perera, 2017; Zhou, 2018).

Learning outcomes essentially represent what learners consciously or unconsciously acquire after engaging in a certain form of participation. Although the measurement of learning outcomes has historically lacked uniformity (Prøitz, 2010), scholarly consensus acknowledges the necessity to assess both cognitive factors (e.g. knowledge, skills, abilities) and non-cognitive factors (e.g. attitudes, values) (Bauer, 2003). Building upon this framework, Peng and Fu (2021) delineate English learning outcomes as comprising language proficiency and psychological facets of language acquisition, such as self-confidence, perseverance, and interest. Drawing from the above connotations and dimensions, this study posits a two-fold categorization of indicators for English learning outcomes which include language achievement (encompassing cognitive factors such as assessment outcomes, language proficiency, and skills) and learners' personal characteristics. In this study, learners' personal characteristics refers to the non-cognitive attributes that influence and are influenced by the educational experiences of university students. These characteristics include, but are not limited to, aspects such as self-confidence, perseverance, motivation, and satisfaction in learning.

In addressing the impact of blended teaching models on university students' English learning outcomes, extensive empirical research within academia has yielded three distinct categories of conclusions: (1) Blended teaching significantly enhances university students' English learning outcomes. For instance, Zhou (2018) conducted a comparative analysis between experimental and control groups, revealing superior performance in language skills among students exposed to blended learning environments compared to those in traditional classroom settings. Additionally, Min et al. (2019) and Zainuddin and Perera (2017) observed a notable positive effect of blended learning on non-cognitive dimensions, including motivation and emotional states, within the English learning process. (2) There is no significant difference between blended teaching approach and traditional instruction. Suranakkharin (2017) and Al-Harbi and Alshumaimeri (2016) conducted empirical studies examining the effectiveness of flipped classrooms in improving students' English grades. The results suggested a positive effect size for flipped classrooms, albeit the effect was not statistically significant. (3) There is a negative association between blended teaching and learning outcomes. For instance, Berga et al. (2021) reported that the impact of blended teaching approaches on student performance is inferior to that of traditional instruction. The inconsistency in findings calls for a meta-analysis to systematically synthesize and statistically analyze these disparate results.

Given the ongoing debate surrounding the impact of blended teaching approach on university students' English learning outcomes, a limited number of synthesized analyses have been conducted. Chen et al. (2020) and Li (2022) utilized meta-analysis to ascertain that the overall effectiveness of blended teaching on language learning is moderately high. Wang and Hu (2018) highlighted the moderately small positive influence of flipped classrooms on English major students' academic performance.

While these studies further enrich the discourse on the effectiveness of blended teaching, certain limitations persist: (1) Present meta-analyses assessing the effectiveness of blended teaching primarily concentrate on academic grades. Although certain studies categorize learning outcomes into cognitive and non-cognitive dimensions, few have explicitly investigated the correlation between blended teaching and various dimensions of learning outcomes. (2) Studies assessing the effects of blended teaching on student learning outcomes often overlook the influence of factors such as teaching method, form of learning, blended mode and type of interactions. (3) The prevailing approach in existing studies involves the utilization of traditional meta-analysis methods. In contrast to traditional meta-analysis, three-level meta-analysis is more conducive to extracting and explaining correlations within studies, maximizing the utilization of effect sizes from original literature (Assink & Wibbelink, 2016).

1.2 Moderator variables

1.2.1 Duration of implementation

One of the important factors contributing to the positive impact of the blended teaching approach on student learning is the allocation of time, as students engage in learning activities for a longer duration compared to traditional classroom settings (Han, 2022). Yet, consensus is lacking regarding whether extended periods of blended teaching are more beneficial and what constitutes the optimal duration. Li et al. (2022) asserted that an excessively prolonged duration of blended teaching may be detrimental to learning outcomes. In their experimental report, they indicate that the optimal duration for blended teaching ranges from one to three months, yielding moderately high improvements with significant intergroup differences. However, Vo et al. (2017) argued, based on quasi-experimental research, that the duration of experimental interventions did not moderate learning outcomes. This study adopts Vo et al.'s (2017) classification method for implementation duration in blended teaching, categorizing the duration into two groups: one semester (short-term) and longer than one semester (long-term).

1.2.2 Teaching method

Blended teaching can be seen as an amalgamation of diverse instructional tools and methods. As elucidated by Chen et al. (2020), this pedagogical approach has evolved from teacher-centered, behaviorist, and content-focused methods to more student-centered, constructivist, and collaborative approaches, which are deemed to be more effective than traditional teaching methods. However, it is worth noting that some other studies have suggested that teaching method may not be the sole determinant of instructional heterogeneity, underlying the importance of cautiously integrating different teaching approaches (Alammary et al., 2014). This study divides common teaching methods employed in blended teaching into three main categories: traditional methods, characterized by lecture-based instruction; student-centered methods, which include interactive and task-based approaches; and mixed methods that blend elements from both traditional and student-centered methods.

Among the methods, task-based approach, which requires learners to actively use language for language acquisition purposes, is based on the concept that tasks serve as frameworks for linguistic activity. It provides an alternative to traditional language teaching or present-practice-produce pedagogies by emphasizing interaction during authentic tasks (Bryfonski & McKay, 2019). Interactive teaching, on the other hand, serves to address the limitations of traditional lecture-based approaches by facilitating systematic and purposeful information exchange between instructors and learners within specific contexts, emphasizing the critical role of teacher-student interaction. An interactive teaching style is characterized by learning activities in which students participate in the process of learning and reflect on their knowledge, thoughts, and beliefs (Veeraiyan et al., 2022).

1.2.3 Form of learning

Constructivism posits that knowledge is not simply transmitted by educators; instead, learners actively construct meaning within specific social and cultural contexts, often in collaboration with teachers and learning partners (Zhang, 2019). They engage with relevant learning materials to construct their understanding. Accordingly, this study categorizes form of learning into self-directed and collaborative learning. Self-regulated learning is an educational approach that encourages students to assume responsibility for their individual learning journey (Zimmerman, 1990). Within this framework, students receive guidance in effectively planning, monitoring, and reflecting on their academic endeavors (Zimmerman, 2015). Collaborative learning involves small groups, encouraging cooperation and mutual assistance among students to optimize learning outcomes (Wang & Liao, 2017). Compared to self-directed learning, collaborative learning enhances students' mastery of knowledge, problem-solving abilities, and overall satisfaction with the learning process. Moreover, students' engagement in collaborative learning stimulates their interest in self-directed learning (Kang & Kim, 2021). This argument is also supported by cognitive learning theory, which suggests that collaborative learning yields superior results compared to individual efforts.

1.2.4 Blended mode

The integration of online and face-to-face components in blended teaching exhibits significant variation. For example, Allen and Seaman (2010) proposed that online instruction should constitute 30–79% of blended teaching. Nonetheless, the impact of this proportion, along with others, on learning outcomes remains a subject of debate (Margulieux, 2016), necessitating further investigation into its impact. This study adopts and adapts Li et al.'s (2022) classification of blended modes into the English learning context, delineating the categories as “online+offline”, “online+offline+online”, and “alternating between online and offline regularly” to suit diverse educational strategies. Additionally, a category labeled “technology-integrated classroom” is introduced, where blended learning transcends the distinction between online and offline modalities by incorporating technological tools into face-to-face classroom instruction (Poláková, 2022). While blended teaching involves a

mixture of online and face-to-face learning, its deeper essence lies in the integration of any form of instructional technology with face-to-face pedagogy.

1.2.5 Interaction type

Teacher-student interaction plays an important role in student's cognitive development and academic performance (Çiğdem & Kıymet, 2016). Categorically, teacher-student interaction can be delineated into three primary types based on temporal dimensions: synchronous interaction, synchronous combined with asynchronous interaction, and asynchronous interaction (Li et al., 2022). Previous research suggests that distinct forms of teacher-student interaction may lead to varying instructional outcomes within specific educational contexts. For example, Wu et al. (2011) suggested that learners' academic performance naturally declines in the absence of synchronous interaction opportunities. Therefore, it is evident that types of teacher-student interaction may moderate the effects of blended teaching approach on English learning outcomes.

1.3 Research questions

Based on the preceding discussion, existing literature presents conflicting findings regarding the extent to which blended teaching enhances university students' English learning outcomes, with limited exploration of its influence on diverse dimensions of learning outcomes. Moreover, previous systematic reviews utilizing conventional meta-analysis methods failed to elucidate whether factors such as duration of implementation, teaching method, form of learning, blended mode, and interaction type could moderate this relationship. To address these gaps, this study employs a three-level meta-analysis approach to comprehensively and accurately assess the impact of blended teaching on various facets of university students' English learning outcomes. By clarifying existing research controversies and overcoming the limitations of prior meta-analyses, this research aims to advance knowledge within this domain.

Specifically, this study addresses the following two questions:

- 1) What is the magnitude of the general relationship between blended teaching and university student's English learning outcomes?
- 2) Does the relationship between blended teaching and university student's English learning outcomes vary as a function of duration of implementations, teaching method, form of learning, blended mode, or interaction type?

2 Methods

2.1 Selection of studies

We conducted our study following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Page et al., 2021), a widely referenced and validated method that ensures transparency and methodological rigor in systematic reviews and meta-analyses. A comprehensive literature search was

conducted using several online databases, including China National Knowledge Infrastructure, Wanfang Data, CQVIP Database, Web of Science, ProQuest, and ScienceDirect. Keywords such as English, blended learning, blended instruction, hybrid learning, mixed mode learning, mixed teaching, blending learning, blended teaching approach, blended teaching model, blended teaching, learning outcomes, learning effect, learning achievement, learning gains, learning performance, academic achievement, learning effectiveness, and study performance were utilized to perform advanced searches on the titles, abstracts, and keywords of each database. The search was restricted to studies published between January 2000 and December 2022.

Several inclusion criteria were established to determine the suitability of studies for inclusion in this research. Firstly, studies had to be published in peer-reviewed journals and available in either English or Chinese. Secondly, they needed to be empirical studies examining the impact of blended teaching models on university students' English learning outcomes. Thirdly, the study design should either be cross-sectional or longitudinal, involving experimental and control groups or pre-tests and post-tests of blended teaching. Fourthly, the studies had to provide comprehensive effect size data, including sample size, mean, standard deviation, t-value, or correlation coefficient. Lastly, the research subject had to be university students.

The selection process for our systematic review began with an initial database search yielding 9,194 studies, from which 3,633 duplicates were removed. The remaining 5,561 studies were then screened based on titles and abstracts, reducing the pool to 4,686 studies suitable for full-text review. Further scrutiny for compliance with our strict inclusion criteria excluded 3,554 studies, leaving 1,132 studies for detailed eligibility assessment. This phase removed an additional 777 studies, primarily due to methodological inadequacies or insufficient data. The final rigorous assessment of the remaining 355 studies led to the exclusion of 148 studies, culminating in a robust set of 207 studies included in the meta-analysis. These studies met all specified research quality and relevance criteria, providing a comprehensive view of the impacts of blended teaching on university students' English learning outcomes. This process is visually summarized in the flow diagram presented in Fig. 1.

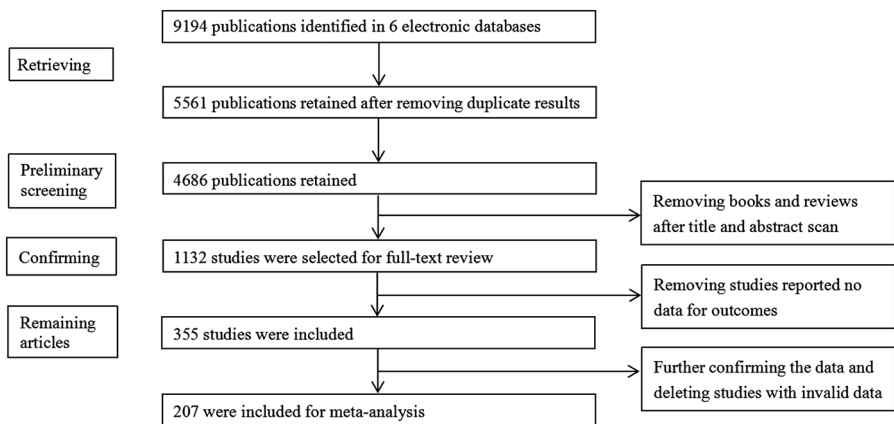


Fig. 1 Flow diagram for search and inclusion procedure

The final sample included 207 studies, comprising 119 articles in Chinese and 88 in English. These studies were diverse in terms of design, sample size, and geographical focus, but all examined the effects of blended teaching on English learning outcomes at the university level. An overview of the characteristics of these studies is presented in Table 1.

2.2 Coding of study features

Each study was coded in a standardized Microsoft Excel sheet, with the extraction and coding of the following characteristics: (1) authors and year of publication; (2) number of effect size; (3) sample size; (4) duration of implementation (short term and long term); (5) teaching method (lecture-based, interactive, task-based, and mixed); (6) form of learning (self-directed, or self-directed+collaborative learning); (7) blended mode (online+offline, online+offline+online, technology intervention during class, and alternating between online and offline regularly); (8) interaction type (asynchronous, synchronous+asynchronous, and non-interactive); (9) learning outcomes (learners' personal characteristics and language achievements).

Several criteria were formulated during coding to ensure consistency and accuracy in data extraction: (1) For studies presenting multiple samples, effect sizes were included corresponding to each distinct sample wherever feasible; (2) In cases where specific subgroup sample sizes were not disclosed, the approach recommended by Quarmley et al. (2022) was used, which entails dividing the overall sample size by the number of subgroups to estimate the size of each independent group; (3) For instances of studies reporting on identical samples, preference was given to the source that offered more information. This coding procedure was carried out independently by two researchers, who meticulously recorded all pertinent details from the studies included in the meta-analysis. Any disagreement or inconsistencies at this stage were

Table 1 The characteristics of the studies included in the meta-analysis

Duration of implementation	Short term	171
	Long term	41
Teaching method	Lecture-based	20
	Mixed	15
	Interactive	138
	Task-based	43
Form of learning	Self-directed	58
	Self-directed+collaborative learning	156
Blended mode	Online+offline	34
	Technology intervention during class	13
	Online+offline+online	67
	Alternating between online and offline regularly	92
Interaction type	Asynchronous	172
	Synchronous+asynchronous	15
	Synchronous	2
	Non-interactive	12

Note. There is an overlap in the studies considering the same moderator, indicating that some studies may have been counted multiple times under different classifications

discussed and fully resolved between the coders. Following the completion of the coding, a total of 373 independent effect sizes were obtained from 207 articles, forming the empirical basis for this analysis.

2.3 Statistical analyses

2.3.1 Calculation of effect sizes

This study employs the standardized difference between the mean of two different groups (Cohen's d) as the effect size and utilizes the professional software CMA 3.7 (Comprehensive Meta-Analysis 3.7) to calculate the effect sizes. A positive d value indicates better learning outcomes with blended teaching in comparison to alternative methods, whereas a negative value indicates the contrary. According to Cohen's (2016) criteria, d values of 0.2, 0.5, and 0.8 can be interpreted as small, moderate, and large effects, respectively.

Traditional meta-analyses typically assume independence among effect sizes, often extracting only one effect size per study (Assink & Wibbelink, 2016). However, this study incorporates multiple effect sizes from selected literature. The rationale behind this lies in the fact that when literature was included, it often reported multiple learning outcome indicators. In cases where the same study reported multiple effect sizes, assuming independence among these effect sizes may not be appropriate (Cheung, 2014). Utilizing effect sizes from the same study can potentially inflate correlations between variables, violating the assumption of independence in traditional meta-analysis (Lipsey & Wilson, 2001). To address this concern, a three-level meta-analysis approach is employed in this study.

A three-level model is executed in R statistical program (version 4.2.1), with each level serving distinct purposes: level 1 represents sampling variance, level 2 denotes variance among effect sizes extracted from the same study, and level 3 represents variance between studies. The "rma.mv" function from the metafor package (Viechtbauer, 2010) is employed for modeling and computing the overall effect. This three-level meta-analysis is implemented in R program following the tutorial by Assink and Wibbelink (2016). Sampling variance is computed using CMA 3.7, and the results are imported into R program for one-tailed log likelihood ratio tests on levels 2 and 3 to confirm their significance. In instances where both levels 2 and 3 are significant, further moderation effect tests are conducted to ascertain the sources of heterogeneity, with all moderating variables recoded as dummy variables.

2.3.2 Model selection

Current meta-analyses predominantly employ either fixed-effects model or random-effects model. The fixed-effects model assumes a consistent true effect size across all studies, with observed variation attributed solely to random error or sampling variability. The random-effects model acknowledges the potential for different true effect sizes across studies due to factors beyond random error, such as variations in populations or methodologies (Borenstein et al., 2009). Considering the likelihood of diverse moderating variables among the literature included in this study, which

indicates the presence of heterogeneous populations, the random-effects model was chosen to conduct the meta-analysis as it accounts for variability both within samples and across populations, allowing for more generalizable results.

2.3.3 Publication bias

Publication bias refers to the discrepancy where the array of articles published does not thoroughly and systematically represent the complete scope of research conducted within a specific domain. In this study, publication bias is assessed by funnel plots, Egger's regression test, Rosenthal's fail-safe N (indicative of the number of studies required to negate the observed effect), and updated p -curve techniques. A funnel plot is essentially a scatter plot with effect size on the x-axis and the sample size on the y-axis. Support for the absence of publication bias is indicated if the funnel plot exhibited a symmetrical distribution (Borenstein et al., 2009). The fail-safe N denotes the minimum number of insignificant findings necessary to invalidate the current conclusions as non-significant. A higher fail-safe N suggests a reduced probability of bias, with bias being presumed when the fail-safe N falls below the threshold of $5k + 10$ (where k represents the count of original studies) (Rosenthal's, 1979). For p -curve test, should the effect size of a study genuinely exist, the distribution of p -values is expected to be right-skewed, meaning the interval of p -values from 0 to 0.025 should manifest a greater frequency compared to the interval from 0.025 to 0.5. In contrast, a left-skewed distribution indicates the presence of publication bias (Lai et al., 2018).

3 Results

3.1 Overall effect sizes

The funnel plot and Egger's test results ($t = 15.037$, $p < 0.05$, intercept = 3.97) suggested the presence of potential publication bias in the current study. According to Rosenthal's N value, a substantial number (> 1552) of additional relevant research would be needed to render the overall effect size non-significant. Further investigation using Trim and Fill method (Duval & Tweedie, 2000) to examine the impact of publication bias on the meta-analysis results revealed that, after adjustment, the overall effects obtained by the Random-Effects model remained statistically significant. Moreover, Additionally, the p -curve analysis exhibited a significant right-skewed distribution (Binomial test: $p < 0.0001$, Continuous test: $z = -50.384$, $p < 0.0001$), with 292 out of 373 effect sizes presenting p -values below 0.05, and 253 below 0.025. Taken together, these findings suggested that while there was a slight presence of publication bias in this meta-analysis, the results remained robust and valid.

Subsequently, a Random-Effects model was employed to explore the relationship between blended teaching and learning outcomes. Results from the main effects tests indicated a substantial effect size ($d = 0.916$, 95% CI [0.614, 1.218]) for differences in learner characteristics between the blended learning group and the control group. In terms of total variance, the sampling variance (level 1) accounted for 2.26%, within-

study variance (level 2) accounted for 16.74%, and variance between-study (level 3) accounted for 81%. The one-tailed log likelihood ratio test revealed significant differences in both level 2 ($p < 0.001$) and level 3 ($p < 0.001$). Similarly, the effect size for language achievement differences between the experimental and control groups was substantial ($d = 0.953$, 95% CI [0.84, 1.067]), with sampling variance (level 1) representing 5.71%, within-study variance (level 2) 43.68%, and between-study variance (level 3) 50.61%. The one-tailed log likelihood ratio test indicated significant differences in level 2 ($p < 0.001$) and level 3 ($p < 0.001$). Therefore, these findings clearly demonstrate the significant impact of blended teaching approaches on both learner characteristics and language achievement, with the latter showing a slightly higher effect size. The observed significant heterogeneity between studies signals the influence of moderator variables, warranting further investigation to elucidate their roles (Assink & Wibbelink, 2016).

3.2 Moderator analyses

The use of a three-level meta-analysis allowed for the identification of nuanced moderating effects that might have been missed with traditional meta-analytic approaches. This method enabled us to capture the intricate relationships between blended teaching and learning outcomes, considering the diversity of study designs and contexts. By incorporating multiple levels of variance, the analysis provides a more comprehensive understanding of how blended teaching affects English learning outcomes. The study respectively examined how the variables of duration of implementation, teaching method, form of learning, blended mode, and interaction type moderate the overall average effect size. Figure 2 presents an overview of these moderating variables on the relationship between blended teaching and English learning outcome.

The results of the moderator analyses on the relationship between blended teaching and learners' personal characteristics (see Table 2) indicated significant moderating effects for duration of implementation, teaching method, and blended mode, while the moderating effects of learning style and teacher-student interaction type are not significant. 1) The moderating effect of experiment period was significant, with the effect size for short term ($d = 0.837$) being larger than that for long term ($d = 0.46$). This suggested that blended teaching with a shorter time span is more effective in enhancing students' English learning outcomes in terms of personal characteristics. 2) The moderating effect of teaching method was significant, with the largest effect size for the lecture-based method ($d = 1.562$), followed by the mixed teaching approach

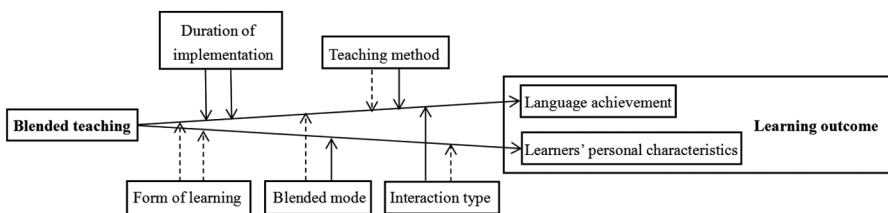


Fig. 2 Conceptual framework illustrating the influence of blended teaching on English learning outcome

Table 2 The moderating effect of the relationship between blended teaching and Learners' personal characteristics

Moderators	K	Intercept/ d(95% CI)	β (95% CI)	F(df1,df2)	p	Level 2 variance	Level 3 variance
(1) Duration of implementation							
Short term	89	0.837[0.619,1.055]***		F(1,100)=4.327	0.04	0.183***	0.363***
Long term	13	0.461[0.081,0.840]*	-0.377[-0.736,-0.017]*				
(2) Teaching method							
Lecture-based	8	1.562[0.892,2.232]***		F(3, 97)=3.275	0.024	0.167***	0.889***
Mixed	3	1.297[-0.128,2.723]	-0.264[-1.840,1.311]				
Interactive	52	0.673[0.310,1.035]***	-0.889[-1.544,-0.234]**				
Task-based	38	1.029[0.653,1.404]***	-0.533[-1.243,-0.176]				
(3) Form of learning							
Self-directed	32	0.721[0.286,1.156]**		F(1,103)=1.826	0.18	0.168***	1.034***
Self-directed+collaborative learning	73	1.001[0.661,1.342]***	0.280[-0.131,0.692]				
(4) Blended mode							
Online+offline	15	1.531[0.874,2.188]***		F(3,103)=3.014	0.033	0.147***	1.092***
Technology intervention during class	6	-0.076[-1.628,1.475]	-1.608[-3.293,0.078]				
Online+offline+online	37	0.771[0.302,1.240]**	-0.76[-1.325,-0.195]**				
Alternating between online and offline regularly	49	0.996[0.599,1.392]***	-0.536[-1.257,0.186]				
(5) Interaction type							
Asynchronous	56	0.945[0.606,1.283]***		F(3,93)=1.059	0.37	0.192***	0.583***
Synchronous+asynchronous	26	0.461[-0.12,1.043]	-0.483[-1.156,0.190]				
Synchronous	2	0.124[-1.191,1.444]	-0.82[-2.179,0.538]				
Non-interactive	13	0.903[0.198,1.607]*	-0.042[-0.824,0.740]				

Note k=number of effect sizes; CI=confidence interval; β =estimated regression coefficient; Level 2 variance=variance between effect sizes extracted from the same study; Level 3 variance=variance between studies. * $p<0.05$, ** $p<0.01$, *** $p<0.001$

($d=1.297$), task-based teaching method ($d=1.029$), and the smallest effect size for the interactive method ($d=0.673$), the latter reaching only a moderate effect level. Therefore, the lecture-based blended teaching method was more effective in promoting students' learning outcomes in terms of personal characteristics. 3) Form of learning, however, did not present a significant moderating effect ($p>0.05$), and Bayesian variance analysis ($BF_{10}=0.448$) provided weak evidence against the hypothesis that form of learning influenced the relationship between blended teaching and personal characteristics, suggesting a negligible impact of form of learning on this relationship (Wagenmakers et al., 2017). 4) Regarding blended mode, the combination of "online+offline" mode yielded the largest effect size ($d=1.531$), followed by alternating between online and offline regularly" ($d=0.996$), "online+offline+online" ($d=0.771$). The strategy of "technology intervention during class" did not demonstrate a moderate effect within this study's context. 5) The type of teacher-student interaction did not show a significant moderating effect, with Bayesian variance analysis ($BF_{10}=0.203$) indicating moderate evidence that this variable did not significantly influence the relationship between blended teaching and the personal characteristics of learners.

The results of the moderator analyses on the relationship between blended teaching and language achievements (see Table 3) indicated significant moderating effects for duration of implementation and interaction type, while teaching method, form of learning, and blended mode do not show significant effects. (1) The duration of the blended teaching implementation showed a significant moderating effect, with short-term implementations ($d=1.017$) resulting in greater improvements in language achievements than long-term implementations ($d=0.655$). This suggested that blended teaching strategies applied over shorter periods are more efficacious in enhancing students' language achievements than those extended over longer periods. (2) Teaching method fails to significantly moderate the relationship between blended teaching and language achievements ($p>0.05$). Bayesian variance analysis ($BF_{10}=0.113$) provided moderate evidence indicating that the efficacy of blended teaching on language achievements is not dependent on the teaching method employed. Notably, the mixed teaching approach yielded the highest effect size ($d=1.005$), followed by interactive teaching ($d=0.958$), task-based teaching ($d=0.923$), and lecture-based methods ($d=0.721$), showing a contrast with the findings related to learners' personal characteristics. (3) The moderating effect of form of learning on the relationship between blended teaching and language achievements was not significant ($p>0.05$), with Bayesian analysis ($BF_{10}=0.168$) offering moderate evidence of no influence from form of learning on this relationship. The effect size for self-directed+collaborative learning ($d=0.951$) was slightly larger than that for self-directed learning ($d=0.935$), consistent with the moderator analysis results under learners' personal characteristics. (4) Blended mode also did not significantly moderate its relationship with language achievements ($p>0.05$), with Bayesian analysis ($BF_{10}=0.182$) suggesting moderate evidence against any significant influence. The effect size for "technology intervention during class" was the largest ($d=1.367$), followed by "alternating between online and offline regularly" ($d=0.97$), "online+offline+online" ($d=0.922$), and the smallest effect size for "online+offline" ($d=0.775$). (5) Interaction types exhibited a substantial effect, where "synchronous+asynchronous" interaction presented the highest

Table 3 The moderating effect of the Relationship between Blended Teaching and Language achievements

Moderators	K	Intercept/ d(95% CI)	β (95% CI)	F(df1,df2)	p	Level 2 variance Level 3 variance
(1) Duration of implementation						
Short term	208	1.017[0.890, 1.145] ***		F(1,243)=5.946	0.015	0.277*** 0.332***
Long term	37	0.655[0.383, 0.928] ***	-0.362[-0.654, -0.069]*			
(2) Teaching method						
Lecture-based	18	0.721[0.321, 1.121] ***		F(3,247)=0.45	0.717	0.261*** 0.318***
Mixed	17	1.005[0.609, 1.401] ***	0.284[-0.278, 0.847]			
Interactive	171	0.958[0.822, 1.095] ***	0.238[-0.185, 0.660]			
Task-based	45	0.923[0.651, 1.195] ***	0.202[-0.282, 0.686]			
(3) Form of learning						
Self-directed	67	0.935[0.716, 1.154] ***		F(1,262)=0.016	0.901	0.252*** 0.351***
Self-directed+collaborative learning	197	0.951[0.818, 1.085] ***	0.016[-0.238, 0.271]			
(4) Blended mode						
Online+offline	54	0.775[0.534, 1.016] ***		F(3,255)=1.853	0.138	0.271*** 0.287***
Technology intervention during class	15	1.367[0.924, 1.809] ***	0.591[0.087, 1.095]*			
Online+offline+online	78	0.922[0.727, 1.116] ***	0.147[-0.162, 0.455]			
Alternating between online and offline regularly	112	0.97[0.796, 1.144] ***	0.195[-0.103, 0.492]			
(5) Interaction type						
Asynchronous	233	0.895[0.779, 1.012] ***		F(2,244)=6.325	0.002	0.280*** 0.278***
Synchronous + asynchronous	6	2.042[1.407, 2.676] ***	1.146 [0.502, 1.790] ***			
Synchronous	8	1.213[0.604, 1.822] ***	0.318[-0.302, 0.937]			

Note k=number of effect sizes; CI=confidence interval; β =estimated regression coefficient; Level 2 variance=variance between effect sizes extracted from the same study; Level 3 variance=variance between studies. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

effect size ($d=2.042$), surpassing “non-interactive” ($d=1.213$) and “asynchronous” types ($d=0.895$), all achieving large effect sizes. This indicated that teacher-student interaction type played a critical role in enhancing language achievements through blended teaching, with synchronous and asynchronous methods proving particularly effective.

In addition, in this meta-analysis, a multiple regression analysis was performed on identified significant moderating variables, following the method proposed by Assink and Wibbelink (2016), to address and mitigate potential collinearity among these variables. The results pertaining to the moderating variables influencing learners’ personal characteristics are presented in Table 4, with duration of implementation (short-term), teaching method (lecture-based), and blended mode (online+offline) as reference variables. The findings from this regression analysis revealed that at least one of the moderating variable’s regression coefficients significantly deviates from zero. Similarly, the regression analysis results for moderating variables affecting language achievements are documented in Table 5, where duration of implementation (short-term) and interaction type (asynchronous) served as reference variables. The result also indicated that at least one regression coefficient for the moderating variables significantly differs from zero. Therefore, there was no significant collinearity present among the moderating variables across the two dimensions of learning outcomes.

Table 4 Multiple regression analysis of moderating variables of Learners’ personal characteristics

Moderators		K	B(95% CI)	
	Intercept		1.84[0.986,2.694]***	
Duration of implementation	Long term	13	-0.446[-0.796, -0.096]*	
Teaching method	Mixed	3	-0.080[-1.379 ,1.219]	
	Interactive	52	-0.669[-1.318,-0.020]*	
	Task-based	38	-0.481[-1.187 ,0.226]	
Blended mode	Technology intervention during class	6	-1.273[-2.527,-0.020]*	
	Online + offline + online	37	-0.493[-1.111 ,0.125]	
	Alternating between online and offline regularly	49	-0.454[-1.128 ,0.220]	
Multiple regression model	$k=97$ $F(7, 89)=2.491$	$p=0.022$	Level 2 variance	Level 3 variance
			0.146***	0.455***

Table 5 Multiple regression analysis of moderating variables of Language Achievements

Moderators		K	B(95% CI)	
	Intercept		0.95[0.817, 1.082]***	
Duration of implementation	Long	37	-0.308[-0.6 , -0.016]*	
Interaction type	Synchronous + asynchronous	6	1.084[0.430, 1.737]**	
	Non-interactive	8	0.447[-0.249, 1.143]	
Multiple regression model	$k=226$ $F(3, 222)=5.745$	$p<0.001$	Level 2 variance	Level 3 variance
			0.283***	0.284***

4 Discussion

4.1 The overall impact of blended teaching on university students' English learning outcomes

The results of meta-analysis indicate a significant positive impact of blended teaching on both dimensions of university students' English learning outcomes - language achievement and learners' personal characteristics. This finding corroborates with other meta-analytical findings, substantiating the superiority of blended teaching approaches over conventional face-to-face instruction in promoting English learning outcomes. Specifically, the magnitude of impact on language achievement surpasses that reported in preceding research (Chen et al., 2020; Li, 2022). Prior studies encompassed a broad spectrum of educational levels, ranging from primary and secondary education to higher education and adult learning contexts. The findings indicate that effect sizes related to primary, secondary, and adult education are relatively lower compared to those pertaining to university settings, generally falling within the moderate to low range. This difference can be ascribed to the fact that university student, compared to primary and secondary group, tends to possess advanced digital literacy skills and a more solid foundation in English, which collectively facilitate a smoother transition to blended learning environments; Adult learners, on the other hand, may encounter difficulties in juggling professional responsibilities with their educational pursuits. Another possible reason is related to the varying levels of learner autonomy, which emerges as a critical element in the context of blended learning environments. Autonomy in learning refers to the ability of students to take charge of their own learning process, which includes setting goals, selecting strategies, and self-assessing progress (Little, 1991). Students with higher levels of autonomy tend to perform better academically because they are more capable of leveraging the flexible nature of blended learning to suit their individual learning preferences and schedules (Benson, 2011). In primary and secondary education, where students are still developing self-regulatory skills, their level of autonomy may not be as advanced as that of university students, and they may require additional support to fully benefit from such environments. Meanwhile, the observed improvement in personal characteristics attributable to blended learning exceeds findings from earlier studies (Yu et al., 2022), possibly reflecting disciplinary and educational level discrepancies. Therefore, the beneficial effects of blended teaching on the English learning outcomes of university students demonstrate multidimensional characteristics, which is congruent with the modern educational focus on fostering competencies, including critical thinking, communication, collaboration, and creativity. In the contemporary era, mere knowledge acquisition is deemed inadequate for personal development in real-world contexts (Laar et al., 2017).

4.2 Moderating effects

The findings of the present meta-analysis indicate the significant moderating role of duration of implementation on the two dimensions of English learning outcomes, with evidence suggesting a more pronounced impact in short term as opposed to long

term. This observation corroborates the insights provided by Du et al. (2022), indicating that abbreviated time frames lead to more favorable improvements in learning outcomes. However, this stance contrasts with the findings posited by Vo et al. (2017), where a disparity might be discerned, potentially due to the smaller effect size reported in their study ($k=51$) and the inclusion of varied academic disciplines (encompassing both STEM and non-STEM fields). The theoretical framework of Schramm's Media Selection Law, articulated as $\text{Expected Selection Probability} = \text{Possible Rewards} / \text{Costs Incurred}$, offers a pertinent lens for understanding this phenomenon within the context of blended teaching. "Possible Rewards" herein encapsulate the achievement of pedagogical objectives, and "Costs Incurred" represent the cumulative expenses and effort associated with media production, incorporating factors such as complexity, time investment, and additional considerations (Schramm, 1954). This theory elucidates the rationale behind the optimal confinement of duration of implementation to one academic semester for augmenting learning outcomes through blended teaching. In addition, extended periods of engagement in online learning environments have been identified as a substantial barrier for students to sustain concentration on their academic pursuits (Hwu, 2023), which requires the educators to intensify their supervision and evaluation of students' online learning activities.

The results of moderating effect of teaching methods on language achievement show that the lecture-based method has the smallest effect size, at upper medium level, followed by the task-based teaching method. The mixed teaching method has the highest effect size. Nonetheless, these moderating effects across different teaching methods do not reach statistical significance, and there exists moderate evidence to suggest that the relationship between blended teaching and language achievement remains unaffected by the specific teaching methods employed. This observation aligns with the findings reported by Chen et al. (2020), who noted similar effect sizes for lecture-based and interactive methods. However, their findings diverge concerning the effect size associated with the mixed teaching method ($d=0.375$, $p>0.05$), a variance possibly due to their analysis incorporating only a single study focused on the mixed teaching approach. In a parallel vein, Alammary et al. (2014) found the moderating effects of various teaching methods to be statistically insignificant, indicating the important role of active learning strategies within blended learning environments. They advocated for the strategic utilization of a diverse array of teaching methods. These voices were reinforced by Cetin and Ozdemir (2018), who posited that the efficacy of blended learning modalities does not hinge on the teaching methods implemented, and lecture-based method can achieve outcomes comparable to those of inquiry-based approaches. On the contrary, the moderating effects of teaching methods on learners' personal characteristics are significant, with the lecture-based method having the largest effect size, followed by the task-based teaching method, and the mixed teaching approach. This result is consistent with Cheng et al.'s (2023) finding that the choice of teaching method emerges as a pivotal determinant of student satisfaction within blended learning environments in higher education settings. Moreover, the finding in relation to the high effect size of lecture-based method corroborate Zuzovsky's (2013) supposition that lecture-based methods facilitated by the teacher do not inherently preclude student engagement. Instead, through strategic teacher guidance and the incorporation of classroom questioning, such

methods can actively foster a conducive classroom climate and enhance the learning milieu. Despite an increasing recognition of teaching methods such as autonomous cooperation and guided inquiry, their effective implementation depends on a strong foundation of knowledge and basic skills, which are particularly emphasized in lecture-based method. Without the guidance of foundational knowledge, the differences in student abilities could be magnified, potentially having a negative impact on teaching outcomes (Fan & Liu, 2022). In a word, the reinforcement of learner personal characteristics can be moderated through teaching methods. Therefore, it is necessary to provide training and guidance for teachers, enhancing their ability to design teaching content and method processes.

The moderating effect of form of learning on the two dimensions of learning outcomes is not significant. This aligns with the empirical results of Bietenbeck (2014) and Li et al. (2022), wherein the intergroup effect size was not significant, indicating that the use of collaborative learning does not effectively enhance learning outcomes. This phenomenon could be ascribed to the diminished involvement of teachers within group discussions, coupled with a lack of substantive feedback from both peers and teachers, which hinders the improvement of learning effectiveness (Fan & Liu, 2022). Within the context of blended learning, particularly regarding its online instructional components, it is imperative for educators to undertake exhaustive evaluations encompassing a multitude of perspectives. This approach should encapsulate supervision, support, and instructional direction, aimed at fully activating the moderating potential of collaborative learning practices in elevating the outcomes of English language education.

The meta-analytical findings suggest that the moderating impact of blended mode on language achievement is negligible. This revelation indicates the presence of alternative, more pivotal moderating factors that influence language achievement improvements within the blended teaching paradigm. Yet, it exhibits a noteworthy moderating effect on the personal characteristics of learners. Specifically, the blended mode “online+offline” demonstrates the most substantial effect size, with “alternating online and offline regularly” and “online+offline+online” showing moderate effect sizes. Although the effects of blended mode of varying complexities differ (Kintu et al., 2017), “technology intervention during class” does not yield significant enhancements in this investigation. Through reviewing the coded literature, we found that studies related to “technology intervention during class”, particularly those with effect sizes less than 0.2, mostly engaged in asynchronous teacher-student interactions, including activities such as second language (L2) writing on wikis, vocabulary acquisition through applications, and computer utilization during class sessions. It is inferred that the absence of a positive effect from this moderating variable could be attributed to the minimal involvement of teachers in the blended teaching environment, leading to insufficient guidance and support for students. This conclusion is consistent with the preceding discussion regarding the influence of teaching method on learning outcomes.

Teacher-student interaction type has a significant moderating effect on language achievement, with the effect size of “synchronous+asynchronous” surpassing that of “no interaction,” and “asynchronous” interactions having the smallest effect size, all achieving a large effect magnitude. This finding is inconsistent with the research con-

clusions of Li et al. (2022), who discovered that the moderating effects, in descending order of effect size, were “synchronous + asynchronous” (1.189), “asynchronous” (0.521), and “no interaction” (0.13). The observed discrepancy can be attributed to the limited representation of studies within this research that focused on “synchronous + asynchronous” and “no interaction” as categories of teacher-student interaction, in comparison to those classified under “asynchronous”. Specifically, the former two categories contain only six studies, while the latter category encompassed eight studies. This imbalance in study distribution inadequately captures the moderating effects of “synchronous + asynchronous” and “no interaction” on learning outcomes. It is premature to conclusively determine the impacts of these interaction types on educational achievements, necessitating further research for more comprehensive understanding. On the other hand, the relationship between blended teaching and learners’ personal characteristics appears to be unaffected by interaction type. Partly, this is because individual learner characteristics are intrinsic and difficult to alter, being subject to other complex factors. However, similar to language achievement, the effect sizes of “asynchronous” and “no interaction” are much larger than those of “synchronous” and “synchronous + asynchronous” interactions. English communication holds a prominent position in English teaching and learning environments worldwide, as well as in personal education and career development (Helen et al., 2018; Xu et al., 2022). Despite this, learners of a second language frequently exhibit a reluctance towards engaging in English communications or participating in spoken language activities within the classroom (Xu et al., 2022). The inclination suggests that synchronous forms of interaction within blended teaching may intensify feelings of anxiety related to foreign language acquisition. In the absence of adequate instructional guidance, it could potentially lead to detrimental effects on the learning process.

5 Conclusion

This study conducted a quantitative synthesis of existing data through a three-level meta-analysis, and the results indicate that blended teaching has a significant positive effect on the English learning outcomes of university students. Specifically, it is discerned that the effectiveness of blended teaching on language achievement is subject to moderation by duration of implementation and interaction type, yet remains unaffected by teaching method, form of learning, and blended mode. The relationship between blended teaching and learners’ personal characteristics is moderated by duration of implementation, teaching method, and blended mode, but is not influenced by form of learning and interaction type. These insights advocate for universities to meticulously plan the scheduling of course hours, select appropriate pedagogical methods, and bolster instructor guidance during online components of blended English language instruction.

The application of a three-level meta-analysis in this study was crucial in uncovering the complex interactions and moderating variables influencing the effectiveness of blended teaching. By accounting for variance at multiple levels, this method provided a more detailed picture of how blended learning impacts language achievement

and personal characteristics. This methodological choice underscores the importance of using advanced analytical techniques to navigate the complexities of educational research and offers a model for future studies to explore similar educational interventions.

Limitations of this study include: first, meta-analysis requires a high level of literature retrieval completeness, but limitations such as encrypted literature and personal factors may lead to partial data gaps in literature retrieval. Second, in the moderator variable analyses, the current meta-analysis only tested the moderating effects of duration of implementation, teaching method, form of learning, blended mode, and interaction type. In the future, further exploration of other potential moderators in the relationship between blended teaching and the English learning outcomes of university students should be conducted. Last, the literature included in this study predominantly originates from Asian countries and regions, with relatively few studies from English-speaking countries. Consequently, the findings of this study may inevitably reflect a certain degree of geographical influence. It would be interesting to research this topic with a focus on cultural and political contexts in future research.

Data availability The data that support the findings of this study are available from the authors upon reasonable request.

Declarations

Competing interests The authors have no competing interests to declare that are relevant to the content of this article.

References

- Al-Harbi, S., & Alshumaimeri, A. (2016). The flipped classroom impact in grammar class on EFL Saudi secondary school students' performances and attitudes. *English Language Teaching*, 9(10), 60–80.
- Alammary, A., Sheard, J., & Carbone, A. (2014). Blended learning in higher education: Three different design approaches. *Australasian Journal of Educational Technology*, 30(4), 440–454.
- Allen, I. E., & Seaman, J. (2010). Class differences: Online education in the United States. *The Sloan Consortium*, 15–25.
- Assink, M., & Wibbelink, J. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology*, 12(3), 154–174.
- Banditvilai, C. (2016). Enhancing students' Language skills through blended learning. *Electronic Journal of e-Learning*, 14, 220–229.
- Bauer, K. W. (2003). Assessment for Institutional Research: Guidelines and resources. In W. E. Knight (Ed.), *The primer for Institutional Research* (pp. 9–13). The Association for Institutional Research.
- Benson, P. (2011). *Teaching and researching autonomy*. Pearson Education.
- Berga, A., Vadnais, E., Nelson, J., Johnston, S., Buro, K., Hu, R., & Olaiya, B. (2021). Blended learning versus face-to-face learning in an undergraduate nursing health assessment course: A quasi-experimental study. *Nurse Education Today*. <https://doi.org/10.1016/j.nedt.2020.104622>
- Bernard, M., Borokhovski, E., Schmid, R., Tamim, R., & Abrami, P. (2014). A Meta-analysis of blended learning and technology use in Higher Education: From the General to the Applied. *Journal of Computing in Higher Education*, 26(1), 87–122.
- Bietenbeck, J. (2014). Teaching practices and cognitive skills. *Labour Economics*, 30, 143–153.
- Borenstein, M., Hedges, V., Higgins, P., & Rothstein, R. (2009). *Introduction to meta-analysis: West Sussex*. Wiley & Sons.

- Bryfonski, L., & McKay, H. (2019). TBLT implementation and evaluation: A meta-analysis. *Language Teaching Research*, 23(5), 603–632.
- Cetin, A., & Ozdemir, F. (2018). Mode-method interaction: The role of teaching methods on the effect of instructional modes on achievements, science process skills, and attitudes towards physics. *Eurasia Journal of Mathematics Science and Technology Education*, 14(5), 1815–1826.
- Chen, Z., Chen, W., Jia, J., & Hui, A. (2020). The effects of using mobile devices on language learning: A meta-analysis. *Educational Technology Research and Development*, 68, 1769–1789.
- Cheng, X., Mo, W., & Duan, Y. (2023). Factors contributing to learning satisfaction with blended learning teaching mode among higher education students in China. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2023.1193675>
- Cheung, L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods*, 19(2), 211.
- Çiğdem, S. Ç., & Kıymet, S. (2016). Increasing teacher candidates' ways of interaction and levels of learning through action research in a blended course. *Computers in Human Behavior*, 61, 488–506.
- Cohen, J. (2016). A power primer. *Psychological Bulletin*, 112(1), 155–159.
- Du, L., Zhao, L., Xu, T., Wang, Y., Zu, W., Huang, X., Nie, W., & Wang, L. (2022). Blended learning vs traditional teaching: The potential of a novel teaching strategy in nursing education—A systematic review and meta-analysis. *Nurse Education in Practice*. <https://doi.org/10.1016/j.nepr.2022.103354>
- Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56(2), 455–463.
- Fan, Y., & Liu, Y. (2022). Which instructional strategies are more effective: Empirical evidence from China education panel study. *Renmin University of China Education Journal*, (3): 151–165.
- Han, S. (2022). Flipped classroom: Challenges and benefits of using social media in English language teaching and learning. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2022.996294>
- Helen, B., Sharynne, M., Sarah, V., & Gail, F. (2018). The relationship between spoken English proficiency and participation in higher education, employment and income from two Australian censuses. *International Journal of Speech-Language Pathology*, 20(2), 202–215.
- Hwu, S.-L. (2023). Developing SAMM: A model for measuring sustained attention in Asynchronous Online Learning. *Sustainability*. <https://doi.org/10.3390/su15129337>
- Kang, Y., & Kim, R. (2021). Impact of blended learning on learning outcomes in the public healthcare education course: A review of flipped classroom with team-based learning. *BMC Medical Education*, 21, 78.
- Kintu, M. J., Zhu, C., & Kagambe, E. (2017). Blended learning effectiveness: The relationship between student characteristics, design features and outcomes. *International Journal of Educational Technology in Higher Education*, 14(1), 7.
- Laar, E., Alexander, D., Jan, D., & Jos, H. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in Human Behavior*, 72, 577–588.
- Lai, L., Tao, R., Ren, Z., & Jiang, G. (2018). The current practice and ethical issues of online counseling in Chinese mainland: Evidence comes from the big data and ethical assessment for online counseling websites. *Journal of Psychological Science*, (5): 1214–1220.
- Li, R. (2022). Effects of blended language learning on EFL learners' language performance: An activity theory approach. *Journal of Computer Assisted Learning*, 38, 1273–1285.
- Li, B., Yu, Q., & Yang, F. (2022). The impact of blended instruction on students' learning performance: A meta-analysis is based on 106 empirical studies at home and abroad. *Open Education Research*, 1, 75–84.
- Lipsey, W., & Wilson, B. (2001). *Practical meta-analysis*. SAGE publications Inc.
- Little, D. (1991). *Learner autonomy 1: Definitions, issues, and problems*. Authentik.
- Margulieux, E., McCracken, M., & Catrambone, R. (2016). A taxonomy to define courses that mix face-to-face and online learning. *Educational Research Review*, 19, 104–118.
- Min, Q., Wang, Z., & Liu, N. (2019). Integrating a cloud learning environment into English-medium instruction to enhance non-native English-speaking students' learning. *Innovations in Education and Teaching International*, 56(4), 493–504.
- Müller, C., & Mildenerger, T. (2021). Facilitating flexible learning by replacing classroom time with an online learning environment: A systematic review of blended learning in higher education. *Educational Research Review*, 34, 1–16.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *British Medical Journal*, 372. <https://doi.org/10.1136/bmj.n71>. Article n71.


- Peng, R., & Fu, R. (2021). The effect of Chinese EFL students' learning motivation on learning outcomes within a blended learning environment. *Australasian Journal of Educational Technology*, 37(6), 61–74.
- Poláková, P. (2022). Use of a mobile learning application in the process of foreign vocabulary learning. *Procedia Computer Science*, 207, 64–70.
- Proitz, S. (2010). Learning outcomes: What are they? Who defines them? When and where are they defined? *Educ Asse Eval Acc*, 22, 119–137.
- Quarmley, M., Feldman, J., Grossman, H., Clarkson, T., Moyer, A., & Jarcho, M. (2022). Testing effects of social rejection on aggressive and prosocial behavior: A meta-analysis. *Aggressive Behavior*, 48(6), 529–545.
- Ramalingam, S., Yunus, M., & Hashim, H. (2022). Blended learning strategies for sustainable English as a second Language Education: A systematic review. *Sustainability*, 14(13), 8051.
- Ran, G., Li, J., Zhang, Q., & Niu, X. (2022). The association between social anxiety and mobile phone addiction: A three-level meta-analysis. *Computers in Human Behavior*, 130, 107198. <https://doi.org/10.1016/j.chb.2022.107198>
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86, 638–641. <https://doi.org/10.1037/0033-2909.86.3.638>
- Schramm, W. (1954). *The process and effects of mass communication*. University of Illinois Press.
- Starkey, L. (2019). Three dimensions of student-centred education: A framework for policy and practice. *Critical Studies in Education*, 60(3), 375–390.
- Suranakkharin, T. (2017). Using the flipped model to foster Thai learners' second language collocation knowledge. *3L the Southeast Asian Journal of English Language Studies*, 23(3), 1–20.
- Veeraiyan, N., Varghese, S., Rajasekar, A., Karobari, I., Thangavelu, L., Marya, A., Messina, P., & Scardina, A. (2022). Comparison of interactive teaching in Online and Offline platforms among Dental undergraduates. *International Journal of Environmental Research and Public Health*. <https://doi.org/10.3390/ijerph19063170>
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1–48.
- Vo, M., Chang, Z., & Diep, N. (2017). The effect of blended learning on student performance at course-level in higher education: A meta-analysis. *Studies in Educational Evaluation*, 53, 17–28.
- Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., & van Doorn, J. (2017). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, 1–19.
- Wang, C., & Hu, Y. (2018). Can flipped classroom really improve learning performance? A Meta-analysis of 38 experimental and quasi-experimental studies. *Open Education Research*, 4, 72–80.
- Wang, Y. H., & Liao, H. C. (2017). Learning performance enhancement using computer-assisted Language Learning by Collaborative Learning groups. *Symmetry*. <https://doi.org/10.3390/sym9080141>
- Wu, A., Xiong, C., & Huang, B. (2011). Research on the change of teacher-student interaction under the environment of network communication. *Journal of Distance Education*, (3): 60–65.
- Xu, M., Wang, C., Chen, X., Sun, T., & Ma, X. (2022). Improving self-efficacy beliefs and English language proficiency through a summer intensive program. *System*. DOI10.1016/j.system.2022.102797
- Yu, Z., Xu, W., & Sukjairungwattana, P. (2022). Meta-analyses of differences in blended and traditional learning outcomes and students' attitudes. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2022.926947>
- Zainuddin, Z., & Perera, J. (2017). Exploring students' competence, autonomy and relatedness in the flipped classroom pedagogical model. *Journal of Further and Higher Education*, 43(1), 115–126.
- Zhang, J. (2019). The current situation and effect of teacher-student interaction in college teaching. *Research in Educational Development*, 23, 10–17.
- Zhang, W., & Zhu, C. (2018). Comparing learning outcomes of blended learning and traditional face-to-face learning of university students in ESL courses. *International Journal on E-Learning*, 17(2), 251–273.
- Zhou, C. (2018). Empirical study on the effectiveness of teaching model of college English writing within blended learning mode. *Educational Sciences: Theory and Practice*, 18(5), 1060–1076.
- Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, 25, 3–17.
- Zimmerman, B. J. (2015). Self-regulated learning: Theories, measures, and outcomes. In Smelser, N. J., & Baltes, P. B. (Eds.), *International encyclopedia of the social & behavioral sciences* (Second edition) (pp. 541–546). Amsterdam: Elsevier.

Zuzovsky, R. (2013). What works where? The relationship between Instructional Variables and schools' Mean scores in Mathematics and Science in Low-, Medium-, and high- achieving countries. *Large-scale Assessments in Education*, 1(01), 2.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Gang Yang¹ · Qiongqiong Kuang² · Renfeng Jiang¹ · Renfeng Jiang¹ 

✉ Renfeng Jiang
jiangr@sdu.edu.cn

Gang Yang
yanggang@sdu.edu.cn

Qiongqiong Kuang
veronica_kuang@outlook.com

¹ School of Foreign Languages and Literature, Shandong University, Jinan, Shandong 250100, China

² Faculty of Foreign Languages, Guangdong Baiyun University, Guangzhou, Guangdong 510080, China