The Effectiveness of Flipped Classroom in English Language Learning: A Meta-Analysis

Thada Jantakoon[1], Kitsadaporn Jantakun[2], Thiti Jantakun[2], Atjana Noibuddee[2], Rungfa Pasmala[3], Panita Wannapiroon[3] & Prachyanun Nilsook[3]

[1] Faculty of Science and Technology, Rajabhat Maha Sarakham University, Maha Sarakham, Thailand

[2] Faculty of Education and Human Development, Roi Et Rajabhat University, Roi Et, Thailand

[3] Faculty of Technical Education, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand

Correspondence: Thiti Jantakun, Faculty of Education and Human Development, Roi Et Rajabhat University, Roi Et, Thailand. E-mail: thiti100@gmail.com

Received: June 16, 2024 Accepted: September 3, 2024 Online Published: December 20, 2024

doi:10.5430/wjel.v15n3p50 URL: https://doi.org/10.5430/wjel.v15n3p50

Abstract

The Flipped Classroom (FC) model, a teaching method used in various educational settings, including language learning, aims to improve student engagement and understanding. Its application in English language learning involves restructuring traditional teaching and learning methods. This study was meticulously designed to assess FC's effectiveness in improving English language proficiency. A comprehensive meta-analysis was conducted on research articles from January 2021 to November 2023, retrieved from ERIC and the Scopus Index. After a rigorous independent review and data extraction process by two reviewers, nine studies with a total of 705 participants were included. The methodological quality of the selected articles was evaluated using the Fail-Safe N for Publication Bias Assessment. The results, which showed that FC was more effective than conventional methods in enhancing overall English language proficiency (SMD=0.85, 95% CI -0.57 to 1.12, P<.001, I^2=65.45%), knowledge (SMD=0.84, 95% CI -0.55 to 1.12, P<.001, I^2=49.49%), and skills (SMD=0.70, 95% CI -0.30 to 1.11, P<.01, I^2=75.97%), instill confidence in the robustness of our findings. These results suggest that FC has the potential to significantly improve English language acquisition outcomes. However, further research with larger sample sizes is needed to confirm and strengthen these results.

Keywords: flipped learning, flipped classroom, teaching english, english educational, meta-analysis

1. Introduction

The flipped classroom model, using Internet technology and promoting active learning, has gained significant traction in education. This curriculum offers three distinct types of instruction: the training master-style flip class, the issue inquiry flip class, and the research construction flip class. It has been effectively executed at renowned institutions worldwide (Naing et al., 2023). Ouariach et al. (2023) found that this approach enhances student participation in interactive education and necessitates teachers consistently revising the instructional resources and the curriculum. Research has shown the advantages of using the flipped classroom approach in language acquisition, especially for non-native Spanish-speaking pupils. Garc á-All én and Taylor (2023) proposed instructing grammatical content before in-person lectures. Conversely, Jabeen et al. (2022) provides empirical data indicating that students who engage in flipped classrooms have superior academic achievement than their counterparts in traditional educational environments. Cao (2023) identifies two positive outcomes of high school English education: increased engagement and enhanced communication skills. Nevertheless, challenges such as insufficient data, student involvement, and technical issues persist as obstacles (Ramadhani et al., 2023; Jain et al., 2023; Ammade et al., 2023). While the flipped classroom approach enhances English language learning, there are challenges. Rullytama et al. (2023) and Han (2022) note issues with student adjustment and increased teacher workload. Despite this, studies by Wang et al. (2018) and Huang et al. (2023) demonstrate major improvements in writing, motivation, and speaking skills. Rahim and Wahi (2023) even found that pre-class videos plus in-class practice make students more comfortable with writing. It's clear that the flipped classroom is a powerful learning tool, especially when combined with activities like group work (Lin, 2022), which keeps students engaged (Awidi & Paynter, 2019).

The success of the flipped classroom model in English language learning (ELL) hinges on several factors. Firstly, it enhances learners' motivation, autonomy, and adaptability, improving pronunciation (Alzahrani & Alqurashi, 2023). Additionally, it facilitates self-directed learning, fosters active, communicative participation, and reduces anxiety associated with public speaking (Dariyemez, 2023). This approach addresses individual learner needs by integrating traditional methods with social networking platforms, enhancing knowledge acquisition, engagement, motivation, and overall student well-being (Han, 2022). Moreover, flexibility, efficiency, clarity, and resource wealth allow customized experiences and a conducive language-learning environment (Zhang, 2022). Research demonstrates the flipped classroom's effectiveness in increasing English proficiency and engagement (Qiang, 2022), supported by the "Zone of Proximal Development" theoretical framework. Furthermore, it strengthens grammatical competence, encourages instructional innovation, and

promotes learner autonomy, confidence, and active discussion participation (Puspita et al., 2023). It has proven instrumental in improving fundamental English skills across various contexts (Rahim & Wahi, 2023) and cultivates critical thinking, analytical, collaborative, and communicative competencies (Dana et al., 2023).

Notably, some studies indicate that the flipped classroom may not consistently outperform traditional methods in all areas, such as knowledge acquisition and performance scores (Jalili et al., 2020; Hosseini et al., 2023). The multifaceted nature of the flipped classroom's benefits warrants further research, leading to comprehensive evaluations specifically within English language education (Kernagaran & Abdullah, 2022). Therefore, a rigorous assessment of the flipped classroom model's impact on English language learning comprehension is necessary. This study aims to contribute through a meta-analysis assessing the effectiveness of the flipped classroom strategy in advancing knowledge and skills within the domain of ELL.

Meta-analysis study in English language acquisition has shown promising results about the efficacy of technology-enhanced language learning (Seyyedrezaei et al., 2022). Lebedeva et al. (2023) indicated that using modern technologies, such as mobile learning and information and communication technology (ICT) tools, may improve language acquisition and overall performance. This underscores the capacity of technology to provide customized learning experiences (Nurmala et al., 2023). Moreover, studies indicate that technology-enhanced language acquisition is as successful as traditional instruction. Research has shown that mobile learning significantly impacts student learning, particularly among students currently pursuing undergraduate programs (Han & Shin, 2016). Garzón et al. (2023) have shown that mobile learning achieves the highest effectiveness when combined with collaborative learning in semi-formal settings, such as field trips and activities conducted outside. Furthermore, research supports the notion that storytelling is advantageous for the acquisition and instruction of the English language, particularly in its cognitive and emotional dimensions.

Despite the growing body of evidence supporting the effectiveness of the Flipped Classroom (FC) model in enhancing English language learning, several gaps remain that warrant further investigation. First, there is variability in outcomes observed across studies, potentially due to factors such as student demographics, learning environments, and the integration of technology. As a result, additional research is needed to explore these influencing factors and better understand the specific conditions under which FC can most effectively enhance learning outcomes (Cao, 2023; Dincer & Polat, 2022). Furthermore, while the majority of studies indicate that FC improves both knowledge and skills acquisition, limitations related to sample size and study duration persist. Many studies employ small sample sizes and short-term interventions, raising concerns about the generalizability of the findings (Daulay et al., 2021; Han et al., 2022). Conducting research with larger samples and longer follow-up periods will help validate and solidify the current findings (Rahim & Wahi, 2023).

Another critical gap is the risk of bias present in some studies, particularly concerning participant randomization and selective reporting of results (Al-Jarrah et al., 2021; Purwanti & Suryawati, 2022). Addressing these methodological concerns, such as improving randomization processes and ensuring comprehensive reporting, is essential for producing more reliable results (Tang & Liu, 2000). Finally, the cultural applicability of the FC model remains underexplored, as the majority of research has been conducted in specific geographical regions with limited consideration of diverse cultural contexts (Han, 2022). Future studies should examine how the FC model can be adapted and applied in diverse cultural and educational settings to determine its broader applicability and effectiveness (Puspita et al., 2023). This study seeks to address these gaps through a comprehensive meta-analysis, contributing to a deeper understanding of the impact of FC on English language learning.

In the context of English language learning, knowledge and skills are fundamental components that contribute to language proficiency. Knowledge typically refers to learners' understanding of linguistic structures, including grammar, vocabulary, and language conventions. For instance, a strong grasp of grammar enables learners to construct correct sentences, while an extensive vocabulary equips them with the ability to comprehend and express complex ideas. Additionally, understanding cultural nuances, idioms, and language usage in different contexts further enhances learners' capacity to use English effectively (Alzahrani & Alqurashi, 2023). The Flipped Classroom (FC) model addresses these areas by allowing learners to first gain this knowledge through pre-class materials, such as videos or readings, where they can engage with content at their own pace before attending class. This pre-class preparation helps learners build a solid foundation in grammatical rules, vocabulary usage, and cultural understanding before they apply this knowledge in more interactive settings (Rahim & Wahi, 2023).

On the other hand, skills in English language learning refer to practical abilities, including speaking, listening, reading, and writing. These skills are essential for real-world communication and academic success. The FC model significantly contributes to the development of these skills through its focus on in-class active learning. For example, in-class activities such as group discussions and presentations offer students opportunities to practice speaking and listening in real time, building their confidence and fluency in English (Chen Hsieh, Wu, & Marek, 2017). Similarly, tasks such as writing assignments and peer reviews encourage learners to enhance their writing skills by applying the grammar and vocabulary, they acquired during pre-class sessions (Abedi, Keshmirshekan, & Namaziandost, 2019). The integration of these skills into the FC model ensures that learners not only understand the language theoretically but can also use it effectively in various communicative contexts (Dariyemez, 2023).

Therefore, the FC model provides an optimal environment for learners to acquire both knowledge and skills in English language learning. By focusing on pre-class knowledge acquisition and in-class skill application, the FC approach offers a more holistic and effective means of improving learners' overall English proficiency. This study, by analyzing the impact of the FC model on English language learners, aims to further explore how this approach enhances learners' knowledge of language structures and their practical skills in using the

language (Rahim & Wahi, 2023; Muntrikaeo & Poonpon, 2022).

Although various advantages are associated with using the flipped classroom (FC), certain studies have revealed that for specific outcomes, such as knowledge and performance scores, FC does not exhibit superior efficacy to conventional techniques. Notably, significant disparities exist in the findings of studies on FC's effectiveness. Kernagaran and Abdullah (2022) have highlighted that recent systematic reviews have comprehensively examined the efficacy of FC in the domains of education and English. Consequently, there exists a need to evaluate the effectiveness of FC in promoting the English language development of students. The primary objective of this study was to carry out a meta-analysis to determine the influence of FC usage on knowledge and skills acquisition in the English language.

2. Method

We conducted a meta-analysis to see how flipped classrooms impact students' learning of English. Our approach followed a well-established five-step process used in previous research. This process involved carefully searching for relevant studies, setting clear criteria for which studies to include, gathering the important data, checking for bias, and finally, analyzing everything we found.

2.1 Search Strategy

This meta-analysis followed the well-established PRISMA guidelines (Moher et al., 2010) to make sure our results were reliable and comprehensive. We focused on studies about flipped learning and English language education, searching through trusted databases like ERIC and Scopus for the most relevant research. We only looked at articles published between 2021 and 2023 to get the most up-to-date information.

2.2 Inclusion Criteria

This analysis incorporated investigations that utilized a quasi-experimental randomized design, such as randomized controlled trials (RCTs) or investigations that adhered to the PICO (Population-Intervention-Comparison-Outcome) framework. The PICO components of the study were as follows:

Population: students before and after registration.

Intervention: various forms of Flipped learning approaches.

Comparison: conventional educational approaches (including presentations, classes, written instructions, etc.) and non-Flipped learning approaches. Outcomes: knowledge and abilities.

Outcomes: Knowledge and skills

2.3 Data Extraction

Two reviewers (AN and KJ) independently extracted data on authors, publication date, country, sample size, participant type, intervention regimens, and outcomes using predetermined criteria.

2.4 Publication bias test

Numerous studies, such as those by Ferrer (1998), Tang and Liu (2000), Song et al. (2002), and Souza et al. (2007), have utilized the funnel plot as a means of bias assessment. Various forms of bias, including publication bias, where studies with positive findings are more likely to be published (Roehr, 2012) as well as English language bias, wherein negative studies might not be as frequently published in English language journals, and citation bias, where studies with positive conclusions receive more citations, potentially skewing their prominence, are critical factors influencing bias. Eyding et al. (2010) also highlight the potential for deliberate bias by pharmaceutical companies, suppressing research that contradicts their product's efficacy. A symmetrical plot should emerge in an unbiased scenario where all studies randomly assess the same mean value. Thus, asymmetry in the funnel plot, with fewer data points at the bottom (refer to Figure 1), may indicate bias. For instance, the funnel plot would exhibit such asymmetry if there is a scarcity of small studies reporting positive outcomes with large effect sizes. A symmetrical plot should emerge in an unbiased scenario where all studies randomly assess the same mean value. Thus, asymmetry in the funnel plot, with fewer data points at the bottom (refer to Figure 1), may indicate bias. For instance, the funnel plot would exhibit such asymmetry if there is a scarcity of small studies reporting positive outcomes with large effect sizes.

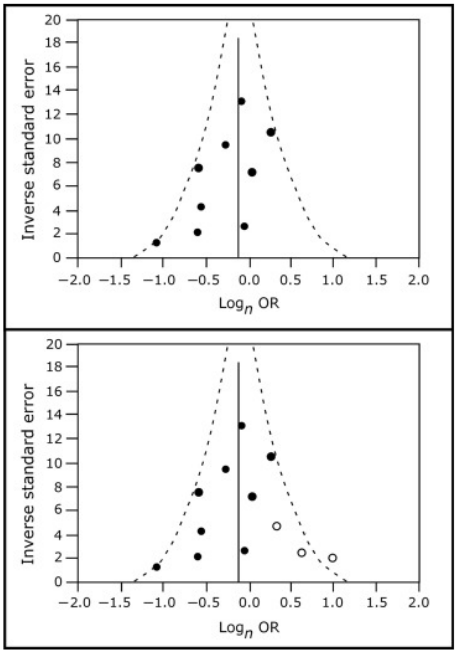

Figure 1. Top panel: deficit of points with large effect and small sample size. Bottom panel

2.5 Fail-Safe N

Fail-Safe N (Number) is a statistical concept employed in meta-analysis to evaluate the resilience of the results by determining the quantity of non-significant or neutral studies required to invalidate the observed effect. It offers insight into the potential influence of publication bias on the overall findings.

Calculation of Fail-Safe N:

Calculate the measured magnitude of the effect: Determine the magnitude of the effect obtained from the meta-analysis.

Establish a standard for determining statistical significance: Determine a cutoff point for statistical significance, often set at $p < 0.05$.

Compute the Z-score by utilizing the observed effect size and its standard error.

Calculate the Fail-Safe N: Apply Rosenthal's method (1979) to ascertain the minimum number of studies that must produce non-significant results in order to obliterate the overall effect.

Explanation:

A greater Fail-Safe N value signifies that the meta-analytic outcome is more robust, as a substantial amount of non-significant research would be required to invalidate the observed impact. Conversely, a smaller Fail-Safe N value instills concerns regarding the durability of the findings.

Example:

A meta-analysis yielding a p-value of 0.02 and a fail-safe N of 100 suggests that an additional 100 studies exhibiting a similar magnitude of effect would need to be conducted to render the overall outcome statistically nonsignificant. This finding engenders a sense of assurance and steadfastness in the observed effect.

2.6 Data Analysis

The meta-analysis was conducted utilizing the Jamovi desktop software (Caldwell, 2022). Standard mean difference (SMD) and 95% confidence intervals were presented for continuous data. Statistical heterogeneity among the studies was assessed using I^2 in each analysis. The fixed-effect model was selected when the P-value exceeded 0.1 and I^2 was below 50%. Conversely, the random-effects model was utilized when the P-value ranged from 0 to 0.1 and I^2 was 50% or higher (Higgins & Thompson, 2002).

3. Findings and Discussion

3.1 Results of the Literature Search

A total of 78 potential studies (n=78) were retrieved from two databases. After eliminating one duplicate and scrutinizing the remaining publications to ensure compliance with inclusion criteria, 77 articles underwent title and abstract evaluation. Subsequently, 66 articles were excluded based on this assessment. Two of the remaining 11 full-text articles were further excluded, leaving a final selection of nine studies comprising 705 participants for this investigation. The study selection process is illustrated in Figure 2 using a flowchart.

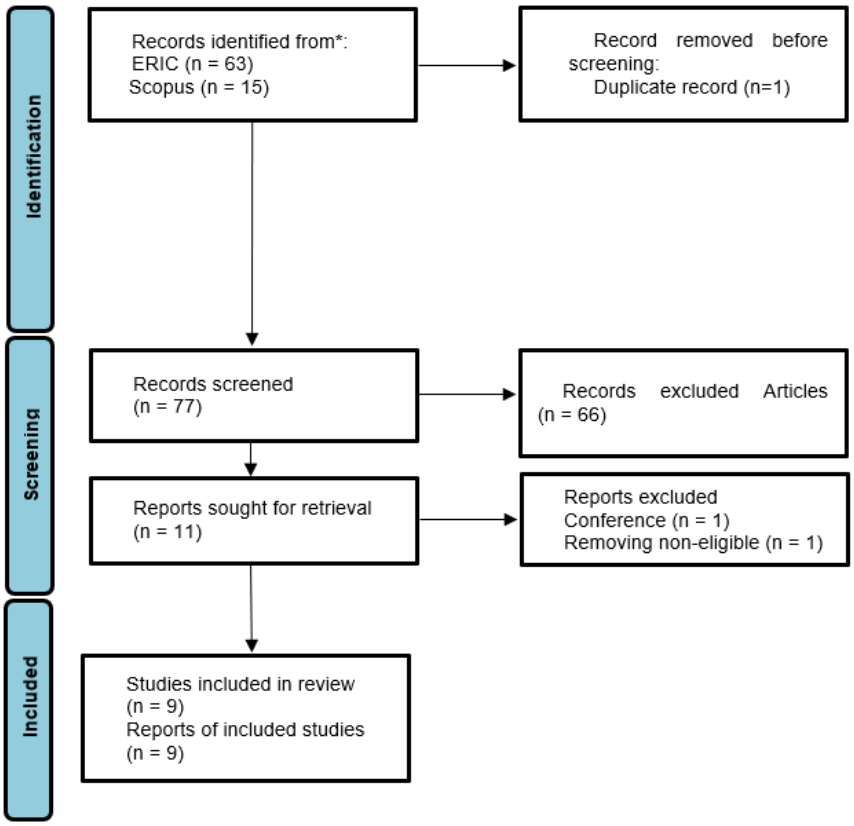

Figure 2. PRISMA article screening flowchart

The PRISMA diagram (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) you provided outlines the process of selecting studies for a systematic review. It tracks how studies were identified, screened, and ultimately included or excluded from the final review. Here's a detailed explanation of each phase in the diagram:

3.1.1 Identification

- A total of 78 records were identified from two databases: ERIC (63 records) and Scopus (15 records).
- After removing one duplicate record, 77 unique records remained for screening.

3.1.2 Screening

- During the screening phase, 77 records were initially reviewed.
- Based on the title and abstract, 66 articles were excluded, leaving 11 records for further examination.

3.1.3 Eligibility

- After closer review, two reports were excluded, one for being a conference paper and the other for being non-eligible according to predefined criteria.

3.1.4 Included

- Finally, nine studies were deemed eligible and included in the systematic review. These studies were also included in the final analysis, contributing to the meta-analysis findings.

This PRISMA flowchart visually represents the process used to ensure a rigorous and systematic selection of studies for inclusion in the review. The diagram highlights how irrelevant or non-eligible studies were systematically excluded at each step to maintain the quality and relevance of the research synthesis.

3.2 Study Characteristics

The studies encompassed trials conducted in eight countries: Jordan (Al-Jarrah et al., 2021), Indonesia (Purwanti & Suryawati, 2022; Daulay et al., 2021), Malaysia (Daulay et al., 2021; Han et al., 2022), Iraq (Dehham et al., 2022; Rad, 2023), Turkey (Dincer & Polat, 2022), Thailand (Muntrikaeo & Poonpon, 2022), and China (Zhiyong & Jiaying, 2022). One study implemented a 3-arm group design (Han et al., 2022), while 8 studies employed a 2-arm group design. The sample sizes varied between 37 and 200 participants. All participants in the trials were EFL students, undergraduate students, grade 8 students, and grade 9 students, with the exception of one study where the participants were Diploma in Maintenance Engineering (DME) students (Daulay et al., 2021). All 12 trials conducted a

comparison between flipped learning and traditional education. Table 1 displays the individuals' attributes, the intervention's specifics, and the metrics used to evaluate the results.

Table 1. Characteristics of the 9 included studies

Author (year)	Country	Type of participant	Number of participants			Outcomes
			Table (number of groups)	*Experimental (Flipped learning)*	*Control group (traditional education)*	
Al-Jarrah et al (2021)	Jordan	Students (grade 8)	50 (2)	25	25	Knowledge assessment
Daulay et al (2021)	Indonesia and Malaysia	Diploma in Maintenance Engineering (DME) students	80 (2)	40	40	Skills
Dehham et al (2022)	Iraq	EFL students	61 (2)	30	31	Knowledge assessment
Dincer & Polat (2022)	Turkey	EFL students	37 (2)	19	18	Knowledge assessment
Han et al (2022)	Malaysia	Undergraduate students	300 (3)	100	100	Knowledge assessment, Skills
Muntrikaeo & Poonpon (2022)	Thailand	Students (grade 9)	80 (2)	40	40	Skills
Purwanti & Suryawati (2022)	Indonesia	Undergraduate students	78 (2)	37	41	Knowledge assessment
Zhiyong & Jiaying (2022)	China	EFL students	69 (2)	34	35	Skills
Rad (2023)	Iran	EFL students	50 (2)	25	25	Knowledge assessment, Skills

S3.3 Risk of Bias

3.3.1 Risk of Bias Analysis

The study applied a comprehensive risk-of-bias assessment framework to categorize each included study as having low, some concerns, or high risk of bias across several domains. Here are the specific criteria and methods used to determine the level of bias in each domain:

1. Randomization Process (D1)

Low Risk: Studies were classified as having low risk of bias if participants were randomly assigned to intervention and control groups using an appropriate randomization technique (e.g., computer-generated random numbers) and if baseline characteristics between the groups were similar. Additionally, if the allocation was concealed (meaning that the person assigning participants to groups was unaware of the group to which the participant would be assigned), this further supported a low-risk rating.

Some Concerns: A study was rated with "some concerns" if randomization was performed, but there was inadequate information on whether allocation concealment was maintained, or if there were minor imbalances in baseline characteristics that could potentially influence the outcomes.

High Risk: Studies were rated as having a high risk of bias if randomization was clearly flawed (e.g., non-random assignment like alternating participants into groups), or if there was predictable randomization (e.g., participants were assigned based on order of enrollment), leading to significant imbalances between groups.

2. Deviations from Intended Intervention (D2)

Low Risk: This domain was rated low risk if participants received the interventions exactly as planned, with no deviations that could affect outcomes. This includes adherence to intervention protocols, such as ensuring that the flipped classroom (FC) model was implemented consistently and that participants followed the designated learning activities.

Some Concerns: A study received a rating of "some concerns" if there were minor deviations from the planned interventions, but these were unlikely to have a significant impact on the results (e.g., slight changes in lesson timing or content delivery).

High Risk: A study was classified as high risk if there were significant deviations from the intended intervention, such as substantial differences in how the FC model was implemented across participants, or if there was contamination between the intervention and control groups (e.g., control group participants inadvertently accessing flipped classroom resources).

3. Missing Outcome Data (D3)

Low Risk: A study was considered low risk if there was minimal missing data (e.g., less than 5% of participants) or if the missing data

was balanced across groups and unlikely to affect the overall findings. This would apply if participants who dropped out of the study were distributed evenly between the intervention and control groups, and reasons for missing data were unrelated to the intervention.

Some Concerns: A study was rated with "some concerns" if there was a moderate amount of missing data, but it was unlikely to have a substantial effect on the overall results, or if missing data was imputed using accepted statistical methods.

High Risk: Studies were rated high risk if there was a significant amount of missing outcome data (e.g., more than 20%), especially if the missing data was not accounted for or if the dropout rates were significantly higher in one group than the other, which could skew the results

4. Measurement of the Outcome (D4)

Low Risk: A study was classified as low risk if the methods used to measure outcomes were objective (e.g., standardized tests or validated scales), and the same methods were used across all groups. Additionally, the outcome assessors were blind to the group assignments, reducing the risk of performance or detection bias.

Some Concerns: A study received a "some concerns" rating if the measurement of outcomes had minor potential for bias, such as when outcome assessors were not blinded to group assignments, but the outcomes were still objectively measured.

High Risk: A study was rated high risk if the outcome measurement was subjective and assessors were not blinded, or if different methods were used across groups to measure outcomes, potentially leading to biased results.

5. Selection of the Reported Result (D5)

Low Risk: A study was considered low risk if all outcomes specified in the study's protocol or methodology were reported, and there was no evidence of selective reporting. This would mean that both positive and negative outcomes were disclosed.

Some Concerns: A study was rated with "some concerns" if there was uncertainty about whether all outcomes were reported or if some results appeared to be selectively reported, but the overall impact on the findings was considered minimal.

High Risk: Studies were rated high risk if there was clear evidence of selective reporting, where only favorable results were reported, while other prespecified outcomes were omitted, potentially skewing the study's conclusions.

Overall Bias Assessment:

After assessing each domain, the studies were classified into an overall risk-of-bias category:

Low Risk: If all domains were rated as low risk.

Some Concerns: If one or more domains had "some concerns" but none were classified as high risk.

High Risk: If any of the domains were classified as high risk, indicating a significant threat to the reliability of the study's results.

The risk-of-bias assessment provided a clear understanding of the quality of the included studies, allowing the researchers to interpret the findings with caution and adjust for potential biases where necessary.

The diagram illustrates that four out of the nine studies exhibit a significant risk of bias. The bias primarily originates from the randomization process (D1), with four studies (Al-Jarrah et al., 2021; Daulay et al., 2021; Dehham et al., 2022; Dincer & Polat, 2022) identified as having a substantial bias in this domain. This suggests potential issues in how participants were assigned to intervention or control groups, such as predictable randomization sequences or baseline characteristic imbalances.

Bias related to deviations in the planned intervention (D2) was noted in two studies (Purwanti & Suryawati, 2022; Zhiyong & Jiaying, 2022), indicating potential problems in delivering the intervention as intended, like variations in intervention dosage or contamination between groups.

One study (Dehham et al., 2022) showed a high risk of bias due to missing outcome data (D3), indicating significant missing data that could impact results, potentially skewing the perceived effectiveness of the intervention.

No studies were found to have a high risk of bias in outcome measurement (D4). However, two studies (Purwanti & Suryawati, 2022; Zhiyong & Jiaying, 2022) were identified as having a high risk of bias in selecting reported results (D5), suggesting selective reporting that may favor the researchers' hypothesis.

Overall, the diagram highlights that several studies in the review carry a high risk of bias, raising questions about the reliability of their findings and emphasizing the need for caution in their interpretation.

3.3.2 How Bias Was Managed

To enhance the robustness of the explanation of how bias was managed in the study, here are some specific examples of the types of bias that were considered, as well as the methods used to minimize them:

1. Publication Bias

Issue: Publication bias occurs when studies with positive or statistically significant results are more likely to be published than studies with null or negative results. This can lead to an overestimation of the intervention's effectiveness.

Solution: The study addressed publication bias by calculating the Fail-Safe N value, which helps determine how many additional studies with null results would be needed to negate the observed effect. The higher the Fail-Safe N, the more robust the results. In this study, the Fail-Safe N was calculated to be 333, meaning that 333 null studies would be needed to overturn the conclusion that the flipped classroom (FC) model is effective. Additionally, funnel plots were used to visually assess publication bias by looking for asymmetry. An asymmetrical funnel plot could indicate that smaller studies with negative results were not published, but no significant asymmetry was found, adding to the robustness of the results.

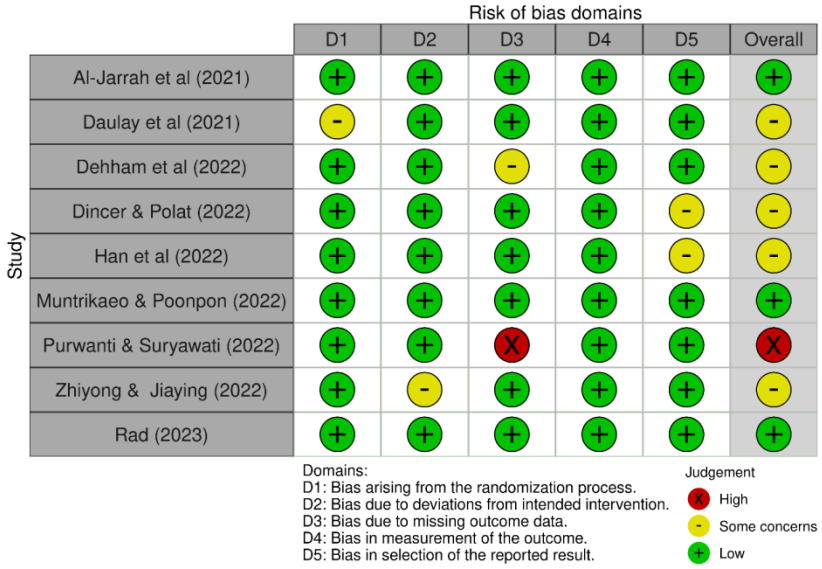

Figure 3. Risk of bias analysis of each included study

2. Selection Bias

Issue: Selection bias can occur if the study participants were not randomly selected or if the intervention and control groups were not comparable at baseline, potentially skewing the results in favor of the intervention.

Solution: Four of the included studies (Al-Jarrah et al., 2021; Daulay et al., 2021; Dehham et al., 2022; Dincer & Polat, 2022) were noted to have a significant risk of bias due to issues with the randomization process. These issues were carefully considered in the meta-analysis, and adjustments were made where necessary to account for baseline imbalances in the intervention and control groups

3. Reporting Bias

Issue: Reporting bias occurs when studies selectively report favorable results while omitting unfavorable outcomes. This can distort the overall findings of a meta-analysis.

Solution: Two studies (Purwanti & Suryawati, 2022; Zhiyong & Jiaying, 2022) had a high risk of reporting bias, meaning that the results reported might have been selectively chosen to favor the hypothesis. The study incorporated this into the risk-of-bias analysis and took steps to interpret the results cautiously, reducing the potential for skewed conclusions

4. Attrition Bias

Issue: Attrition bias occurs when a significant amount of data is missing, such as when participants drop out of a study, potentially affecting the comparability of the intervention and control groups.

Solution: One study (Dehham et al., 2022) was found to have a high risk of attrition bias due to a large proportion of missing outcome data. The researchers noted this bias and incorporated it into the analysis to ensure that the overall results were not disproportionately affected by studies with incomplete data.

5. Heterogeneity in Results

Issue: Heterogeneity refers to the variability or differences in study outcomes, which can arise due to different methodologies, participant characteristics, or intervention implementations across studies.

Solution: The study assessed heterogeneity using the I^2 statistic and tau ^{2}values. An I^2 value of 65.45% was observed, indicating moderate heterogeneity among the studies. By using a random-effects model, the researchers accounted for this heterogeneity, ensuring that the variability between studies was considered in the final results.

These detailed strategies to manage and mitigate various types of bias significantly contributed to the study's reliability and validity, making the findings more robust and generalizable.

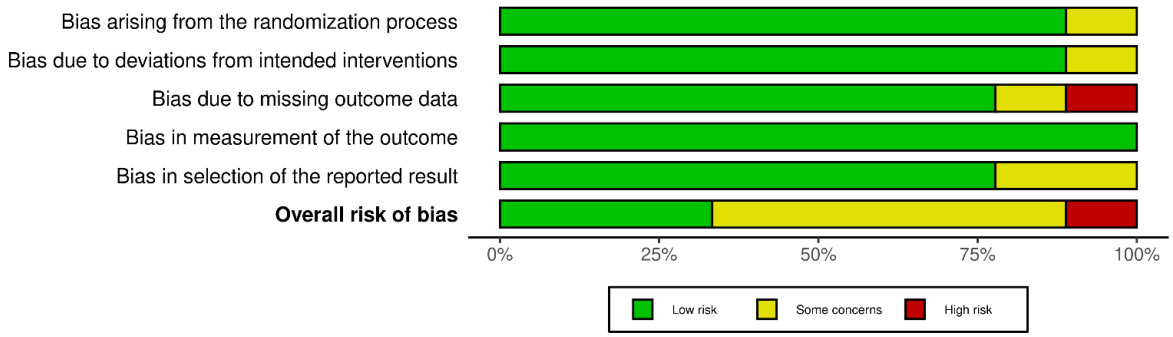

Figure 4. The overall risk of bias analysis of included studies

The green lines in the graph represent the proportion of bases absent from the foundation process, while the yellow lines depict both the proportion of bases missing from the foundation process and the overall risk of bias. On the other hand, the red lines show the percentage of bases missing from the foundation process, with the green lines representing both the percentage of bases missing from the foundation process and the overall risk of bias.

3.4 Results of the Meta-analysis

The study employed the standardized mean difference as the final measurement and analyzed the findings using a random-effects model. The level of heterogeneity (tau ²) was determined using the restricted maximum-likelihood estimate (Viechtbauer, 2005), alongside the I^2 measure and the Q-test for heterogeneity (Cochran, 1954), with the value of tau ²provided. In cases where heterogeneity was observed (tau ²> 0), a range of predictions for real outcomes was generated. Cooked distances and studentized residuals were utilized to identify potential, influential studies and outliers. Studies with studentized residuals exceeding the 100 x (1 - 0.05/(2 X k))th percentile, calculated using a Bonferroni correction and an alpha value of 0.05 for the k studies involved in the meta-analysis, were considered likely outliers. Furthermore, Cook's distance values of two research papers surpassed six times the interquartile range of those distances, indicating their significance. The asymmetry of the funnel plot was examined using the rank correlation test and the regression test, with the standard error of the measured findings serving as a predictor.

Table 2. Random-Effect Model (k=9)

	Estimate	*se*	*Z*	*p*	*CI Lower Bound*	*CI Upper Bound*
Intercept	0.845	0.140	6.02	<.001	0.570	1.120
.	.	.	.	.	.	

Table 3. Heterogeneity Statistics

Tau	*Tau²*	*I²*	*H²*	*R²*	*df*	*Q*	*p*
0.334	0.1112(SE=0.0879)	65.45%	2.895	.	8.000	22.667	0.004

In the analysis, a total of nine studies were included. The calculated standardized mean differences ranged from 0.2478 to 1.4909, with predominantly positive estimates (100 percent). Utilizing the random effects model, the estimated average standardized mean difference was determined to be 0.8454 (95% confidence interval: 0.5704 to 1.1204). Consequently, the mean result indicated a significant deviation from zero (z= 6.0249, p < 0.0001). Based on the Q-test, the observed results exhibited some variability (Q(8) = 22.6668, p = 0.0038, *tau* ²= 0.1112, I^2= 65.4536%). The 95% prediction interval for the actual outcomes ranged from 0.1362 to 1.5545. Therefore, despite the potential for some heterogeneity, the overall outcomes of the studies generally support the estimated average outcome.

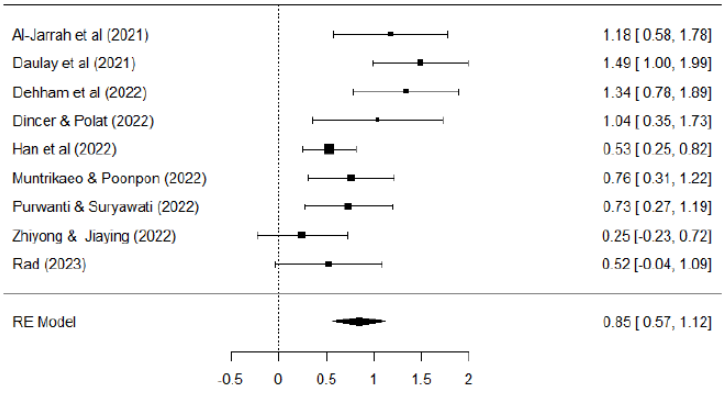

Figure 5. Overall effect forest map

Table 4. Publication Bias Assessment

Test Name	*value*	*p*
Fail-Safe N	333.000	<.001
Begg and Mazumdar Rank Correlation	0.222	0.477
Egger's Regression	1.347	0.178

Note. Fail-safe N Calculation Using the Rosenthal Approach

After examining the studentized residuals, it has been determined that none of the studies had a value exceeding ± 2.7729. Consequently, there is no evidence suggesting the presence of outliers within this model. Cook's distance analysis indicates that none of the studies exert excessive influence in the field. Additionally, both the rank correlation and the regression test failed to identify any asymmetry in the funnel plot (p = 0.4767 and p = 0.1780, respectively).

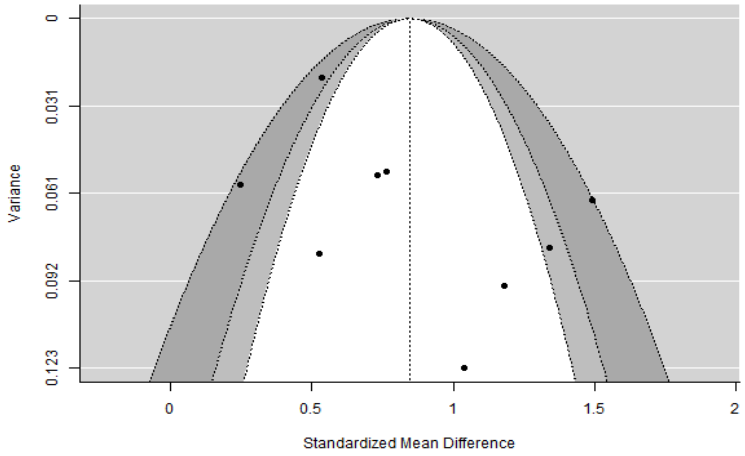

Figure 6. Funnel Plot

Knowledge: In this study, knowledge is defined as the acquisition of information or understanding through learning, study, or experience. Six individual investigations (Al-Jarrah et al., 2021; Dehham et al., 2022; Dincer & Polat, 2022; Han et al., 2022; Purwanti & Suryawati, 2022; Rad, 2023) utilized knowledge scores as their primary measure. The findings from these studies suggested that flipped learning resulted in significantly greater improvements in participants' knowledge compared to control conditions (standardized mean difference [SMD] = 0.84, 95% confidence interval [CI] 0.55–1.12, P<.001, I^2 = 49.49%, Figure 7).

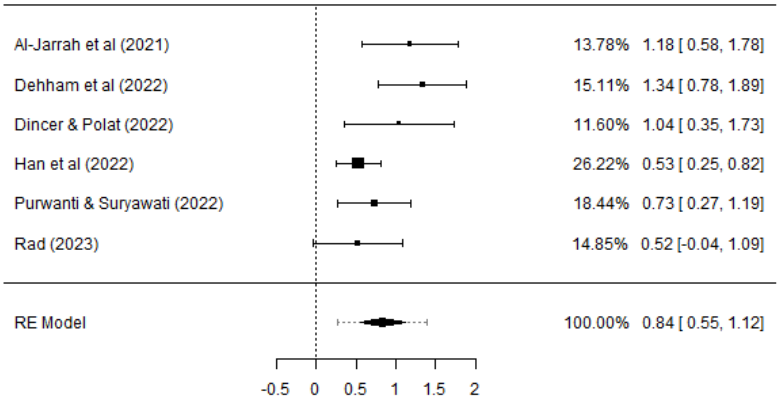

Figure 7. Forest plot of Flipped classroom on knowledge

Skills: Competence was assessed as the outcome measure in five separate experiments (Daulay et al., 2021; Han et al., 2022; Muntrikaeo & Poonpon, 2022; Zhiyong & Jiaying, 2022; Rad, 2023). The standardized mean difference (SMD) was 0.70, with a 95% confidence interval (CI) ranging from -0.30 to 1.11, and a p-value of 0.01. Additionally, the I ²value was calculated to be 75.97% (Figure 8). Examination of the results revealed significant variation in skills enhancement when comparing the flipped classroom approach to other instructional methodologies.

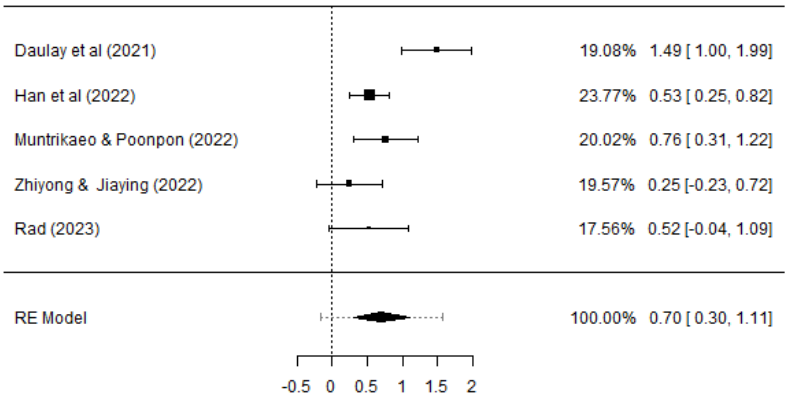

Figure 8. Forest plot of the Flipped classroom on skills

4. Discussion

The findings from this meta-analysis demonstrate the significant impact of the FC approach on knowledge acquisition in English language learning, with the results indicating a clear advantage over traditional teaching methods (SMD = 0.84, 95% CI 0.55–1.12, P < .001, I^2 = 49.49%). The FC model allows for active engagement through pre-class learning and in-class interactive activities, which seem to foster deeper understanding of material. This is consistent with previous studies, which suggest that pre-class preparation, coupled with active learning during class, enhances comprehension and retention of knowledge (Nguyen, 2018). The meta-analysis revealed that learners in FC environments demonstrated a higher ability to grasp complex concepts compared to their peers in conventional classrooms. This could be attributed to the increased opportunity for reflection and application of knowledge in class, as well as the flexibility to review materials at their own pace before classroom discussions. As such, the FC model effectively leverages students' prior exposure to content, allowing classroom time to be dedicated to higher-order thinking and problem-solving tasks (Chen Hsieh et al., 2017).

In terms of skills development, the FC model was found to have a similarly positive effect, although with greater variability across studies (SMD = 0.70, 95% CI -0.30 to 1.11, P = .01, I^2 = 75.97%) . The FC's focus on active and collaborative learning may explain this improvement in skill acquisition, particularly in areas such as speaking and writing (Han et al., 2022). In these cases, the interactive, student-centered nature of the FC encourages learners to engage in communicative activities, peer discussions, and problem-solving exercises, all of which are crucial for developing practical language skills. This finding aligns with the Zone of Proximal Development (ZPD) framework, where students' skills are enhanced through guided interaction and collaborative learning (Puspita et al., 2023). However, the higher level of heterogeneity in skills-related outcomes suggests that not all contexts and implementations of FC yield the same benefits. Factors such as the design of in-class activities, teacher facilitation, and students' engagement levels could influence the degree of skill improvement observed. Studies by Zhiyong and Jiaying (2022) and Rad (2023) highlighted that variations in instructional design and execution play a critical role in determining the extent to which students' skills improve in FC settings.

Moreover, while the FC model seems to offer substantial benefits for language skills, there is a need for further research to examine long-term retention of these skills, as most studies in this meta-analysis were short-term interventions. Research focusing on longitudinal outcomes would provide valuable insights into whether the observed improvements in skills are sustainable over time. Additionally, studies should explore how FC can be tailored to meet the specific skill development needs of diverse learner populations, including those from different cultural and educational backgrounds (Han, 2022; Rahim & Wahi, 2023). The high variability observed in this meta-analysis suggests that contextual factors, such as class size, student motivation, and access to technology, may moderate the effectiveness of FC in developing language skills.

5. Conclusion

This meta-analysis provides robust evidence supporting the effectiveness of the Flipped Classroom (FC) approach in English language learning, with significant positive effects observed in both knowledge acquisition (SMD=0.84, 95% CI 0.55-1.12, *P*<.001) and skills development (SMD=0.70, 95% CI -0.30-1.11, *P*<.01). The findings, derived from nine studies encompassing 705 participants, demonstrate that FC implementation consistently outperforms traditional teaching methods in fostering English language proficiency.

The study's reliability is substantiated by rigorous methodological approaches, including comprehensive bias assessment (Fail-Safe N=333), independent data extraction, and thorough heterogeneity testing (I^2=65.45%). While moderate heterogeneity was observed across studies, particularly in skills development outcomes (I^2=75.97%), this variability suggests that contextual factors and implementation strategies play crucial roles in FC effectiveness.

Several important implications emerge from these findings. First, the strong positive effect on knowledge acquisition indicates that FC's structure of pre-class preparation and in-class active learning effectively enhances understanding of English language concepts. Second,

the positive impact on skills development, albeit with greater variability, suggests that FC can successfully facilitate practical language application when properly implemented. These results align with contemporary theories of language acquisition that emphasize the importance of active engagement and learner autonomy.

However, certain limitations warrant consideration. The included studies' relatively short duration and varying implementation approaches suggest a need for longitudinal research examining long-term retention and skill sustainability. Additionally, the observed heterogeneity in skills development outcomes indicates that success may be contingent upon factors such as technological infrastructure, teacher preparation, and student engagement levels.

Future research should focus on identifying optimal implementation strategies across different educational contexts, investigating long-term learning outcomes, and examining how FC can be effectively adapted for diverse learner populations. Additionally, studies exploring the integration of emerging technologies and pedagogical innovations within the FC framework would be valuable for advancing understanding of this teaching approach.

In conclusion, while acknowledging the need for continued research, this meta-analysis provides strong evidence supporting FC as an effective pedagogical approach for English language instruction. The findings suggest that educational institutions should consider implementing FC strategies, while carefully attending to implementation factors that may influence success.

Acknowledgments

We greatly appreciate the valuable contributions of all members who participated in this study.

Authors' contributions

Asst. Prof. Dr. Thada Jantakoon contributed to the conceptualization, methodology, writing, original draft preparation, investigation, edition, and paper revision. Asst. Prof. Dr. Kitsadaporn Jantakun conceptualized and developed objectives, collected data, analyzed data, interpreted data, presented the arguments in the article, and wrote the manuscript. Asst. Prof. Dr. Thiti Jantakun contributed to the paper's edition and revision. Atjana Noibuddee contributed to the paper's edition and revision. Rungfa Pasmala contributed to the paper's edition and revision. Prof. Dr. Panita Wannapiroon and Prof. Dr. Prachyanun Nilsook provided critical insights and ensured the overall clarity of the paper. All authors read and approved the final manuscript.

Funding

Not applicable

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Sciedu Press.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

References

Abedi, P., Keshmirshekan, M. H., & Namaziandost, E. (2019). The comparative effect of flipped classroom instruction versus traditional instruction on Iranian intermediate EFL learners' English composition writing. *Journal of Applied Linguistics and Language Research*, *6*(4), 43-56.

Al-Jarrah, F. I. M., Ayasreh, M., Ahmad, F. B., & Mansour, O. (2021). The Effect of using Flipped Learning Strategy on the Academic Achievement of Eighth Grade Students in Jordan. *International Journal of Advanced Computer Science and Applications*, *12*(8). https://doi.org/10.14569/IJACSA.2021.0120862

Alzahrani, S. A., & Alqurashi, H. S. (2023). Using the flipped classroom model to improve Saudi EFL learners' English pronunciation. *Linguistics and Culture Review*, *7*(S1), 51-71. https://doi.org/10.21744/lingcure.v7nS1.2260

Ammade, S., Rahman, A. W., & Nurhasanah, N. (2023). Students' cognitive engagement and academic performance in a flipped classroom model in outstanding class program. *JINoP (Jurnal Inovasi Pembelajaran)*, *9*(1), 98-111. https://doi.org/10.22219/jinop.v9i1.23884

Awidi, I. T., & Paynter, M. (2019). The impact of a flipped classroom approach on student learning experience. *Computers & education*, *128*, 269-283. https://doi.org/10.1016/j.compedu.2018.09.013

Caldwell, A. R. (2022). SimplyAgree: an R package and jamovi module for simplifying agreement and reliability analyses. *Journal of Open Source Software*, *7*(71), 4148. https://doi.org/10.21105/joss.04148

Cao, Z. (2023). Exploration on the Application Status of Flipped Classroom in High School English Teaching in the Era of" Internet+". *International Journal of Education and Humanities*, *6*(3), 25-27. https://doi.org/10.54097/ijeh.v6i3.4128

Chen Hsieh, J. S., Wu, W. C. V., & Marek, M. W. (2017). Using the flipped classroom to enhance EFL learning. *Computer Assisted Language Learning*, *30*(1-2), 1-21. https://doi.org/10.1080/09588221.2015.1111910

Cochran, W. G. (1954). The combination of estimates from different experiments. *Biometrics*, *10*(1), 101-129. https://doi.org/10.2307/3001666

Dana, I. K. W., Nitiasih, K., & Santosa, H. (2022). A Review of Flipped Classroom Model for Promoting Students' Basic English Skills. IDEAS: *Journal on English Language Teaching and Learning, Linguistics and Literature*, *10*(2), 1505-1515. https://doi.org/10.24256/ideas.v10i2.3117

Dariyemez, T. (2023). Teaching Speaking Skills through Flipped Classroom Model: EFL Students' Autonomy, Willingness to Communicate, and Anxiety. *i-Manager's Journal on English Language Teaching*, *13*(2), 35. https://doi.org/10.26634/jelt.13.2.19444

Daulay, S., Sanjaya, D., Pin, T. L., Khazin, K. M., & Babar, M. Y. (2021). The Effect of Flipped Learning Instruction on Tertiary English Learners' Writing Achievement. *TESOL International Journal*, *16*(1), 232-252.

Dehham, S. H., Bairmani, H. K., & Shreeb, M. A. (2022). Developing Iraqi EFL preparatory students' performance in reading comprehension by flipped learning strategy. *Journal of Language and Linguistic Studies*, *18*(S1), 640-651.

Dincer, N., & Polat, M. (2022). The use of flipped learning in EFL grammar instruction. *Asian Journal of Distance Education*, *17*(1).

Eyding, D., Lelgemann, M., Grouven, U., Härter, M., Kromp, M., Kaiser, T., ... Wieseler, B. (2010). Reboxetine for acute treatment of major depression: systematic review and meta-analysis of published and unpublished placebo and selective serotonin reuptake inhibitor-controlled trials. *bmj*, *341*. https://doi.org/10.1136/bmj.c4737

Ferrer, R. L. (1998). Graphical methods for detecting bias in meta-analysis. *FAMILY MEDICINE-KANSAS CITY-*, *30*, 579-583.

García-Allén, A., & Taylor, S. K. (2023). Seeing innovation from different prisms: university students' and instructors' perspectives on flipping the Spanish language classroom. *Language Learning in Higher Education*, *13*(1), 105-125. https://doi.org/10.1515/cercles-2023-2004

Garzón, J., Lampropoulos, G., & Burgos, D. (2023). Effects of Mobile Learning in English Language Learning: A Meta-Analysis and Research Synthesis. *Electronics*, *12*(7), 1595. https://doi.org/10.3390/electronics12071595

Han, I., & Shin, W. S. (2016). The use of a mobile learning management system and academic achievement of online students. *Computers & Education*, *102*, 79-89. https://doi.org/10.1016/j.compedu.2016.07.003

Han, S. (2022). Flipped classroom: Challenges and benefits of using social media in English language teaching and learning. *Frontiers in Psychology*, *13*, 996294. https://doi.org/10.3389/fpsyg.2022.996294

Han, Y., Shuo, Z., & Nair, A. B. (2022). Effects of Content and Language Integrated Learning Instruction Model on Learning Outcomes in Chinese STEM Program: A Quasi-Experimental Comparative Study. *International TESOL & Technology Journal*, *17*(2), 6-30.

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in medicine*, *21*(11), 1539-1558. https://doi.org/10.1002/sim.1186

Hosseini, H. M., Ejtehadi, A., & Hosseini, M. M. (2020). Flipping Microlearning-Based EFL Classroom to Enhance Learners' Self-Regulation. *Language Teaching Research Quarterly*, *20*, 43-59. https://doi.org/10.32038/ltrq.2020.20.03

Huang, P. W., Hwang, Y. L., & Huang, K. S. (2023). The Investigation on College Students' English Writing Achievements and Perceptions towards a Flipped Instructional Model. *English Language Teaching*, *16*(5), 1-65. https://doi.org/10.5539/elt.v16n5p65

Jabeen, N., Afaque, H., & Sakha, T. (2022). Flipped Classroom Learning A Modern Strategy: An Insight. *Pakistan Journal of Social Research*, *4*(4), 286-291. https://doi.org/10.52567/pjsr.v4i04.811

Jain, J., Luaran, J., & Panayachi, S. (2023). *Challenges and Issues of Flipping Humanities Lessons: Investigating Teachers' Experiences in a Malaysian International School.* https://doi.org/10.58946/apjfes-2.1.P2

Jalili, S., Khalaji, H., & Ahmadi, H. (2020). Vocabulary learning in the mobile-assisted flipped classroom in an Iranian EFL context. *Teaching English with Technology, 20*(4), 82-95.

Kernagaran, V., & Abdullah, A. (2022). A Systematic Literature Review of Flipped Learning in English as Second Language (ESL) Context. *International Journal of Instruction, 15*(2), 793-814. https://doi.org/10.29333/iji.2022.15243a

Kim, C., & Storer, B. E. (1996). Reference values for Cook's distance. *Communications in Statistics-Simulation and Computation, 25*(3), 691-708. https://doi.org/10.1080/03610919608813337

Lebedeva, M., Taranova, M., & Beketov, V. (2023). Assessment of academic achievements in m-learning. *Education and Information Technologies, 28*(5), 5945-5965. https://doi.org/10.1007/s10639-022-11423-8

Li, Z., & Li, J. (2022). Using the flipped classroom to promote learner engagement for the sustainable development of language skills: A mixed-methods study. *Sustainability, 14*(10), 5983. https://doi.org/10.3390/su14105983

Lichvar, A. B., Hedges, A., Benedict, N. J., & Donihi, A. C. (2016). Combination of a flipped classroom format and a virtual patient case to enhance active learning in a required therapeutics course. *American journal of pharmaceutical education, 80*(10). https://doi.org/10.5688/ajpe8010175

Lin, M. (2022). Effects of the Cooperative Learning-based Flipped Classroom Model on Chinese Students' English Proficiency in the EFL Context. *JCTE, 6*(1). https://doi.org/10.58444/jcte.v6i1.727

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2010). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *International journal of surgery, 8*(5), 336-341. https://doi.org/10.1016/j.ijsu.2010.02.007

Mori, T. (2018). The flipped classroom: An instructional framework for promotion of active learning. *Deep Active Learning: Toward Greater Depth in University Education*, 95-109. https://doi.org/10.1007/978-981-10-5660-4_6

Muntrikaeo, K., & Poonpon, K. (2022). The Effects of Task-Based Instruction Using Online Language Games in a Flipped Learning Environment (TGF) on English Oral Communication Ability of Thai Secondary Students. *English Language Teaching, 15*(3), 9-21. https://doi.org/10.5539/elt.v15n3p9

Naing, C., Whittaker, M. A., Aung, H. H., Chellappan, D. K., & Riegelman, A. (2023). The effects of flipped classrooms to improve learning outcomes in undergraduate health professional education: A systematic review. *Campbell systematic reviews, 19*(3), e1339. https://doi.org/10.1002/cl2.1339

Nguyen, T. (2018). Implementation of English flipped classrooms: Students' perceptions and teacher's reflection. *International Journal of Research Studies in Language Learning, 7*(3), 87-108. https://doi.org/10.5861/ijrsll.2017.1876

Nurmala, I., Irianto, S., Franchisca, S., Amsa, H., & Susanti, R. (2023). Technology-Enhanced Language Learning: A Meta-Analysis Study On English Language Teaching Tools. *Journal on Education, 6*(1), 2188-2195. https://doi.org/10.31004/joe.v6i1.3221

Ouariach, S., Khaldi, M., Mohamed, E., & Khaldi, M. (2023). *The Flipped Classroom: From Passive Information Absorption to Active Learning. In Handbook of Research on Language Teacher Identity* (pp. 269-293). IGI Global. https://doi.org/10.4018/978-1-6684-7275-0.ch015

Purwanti, I. T., & Suryawati, E. (2022). Video Lectures in Online EFL Flipped-Classroom: Effectiveness, Students' Evaluation and Experiences. *European Journal of Educational Research, 11*(2), 885-898. https://doi.org/10.12973/eu-jer.11.2.885

Puspita, L., Wuwur, E. S. P. O., Setyosari, P., & Alfan, M. (2023). Analysis of the Flipped Classroom Model on Students' Critical Thinking Skills. *KnE Social Sciences*, 158-162. https://doi.org/10.18502/kss.v8i10.13441

Qiang, Y. (2022). *An Empirical Study on College English Flipped Classroom Teaching Model*. In Proceedings of the 6th International Conference on Digital Technology in Education (pp. 31-35). https://doi.org/10.1145/3568739.3568745

Rad, H. S. (2023). FLIPPING IN A MICROLEARNING WAY: EFFECTS ON EFL LEARNERS'ACHIEVEMENT AND MOTIVATION IN A GRAMMAR COURSE. *Teaching English with Technology, 23*(1). https://doi.org/10.56297/BKAM1691/DFGF8748

Rahim, S., & Wahi, W. (2023). Improving Low Proficiency ESL Primary School Students' Writing Skills using Flipped Classroom. *International Journal of Social Science And Human Research, 6*(04). https://doi.org/10.47191/ijsshr/v6-i4-56

Ramadhani, R., Syahputra, E., & Simamora, E. (2023). Merging Flipped Classroom Model with the Ethnomathematics Approach: A New Flexibility Learning Model. *FWU Journal of Social Sciences, 17*(2).

Rosenthal, R. (1979). The 'file drawer problem' and tolerance for null results. *Psychological Bulletin, 86*(3), 638-641. https://doi.org/10.1037/0033-2909.86.3.638

Rullytama, M. F., Hafirdin, E. K., & Darmawangsa, D. (2023). Penerapan flipped classroom dalam pembelajaran tata bahasa: Sebuah tinjauan pustaka tematik. Aksara: *Jurnal Bahasa dan Sastra, 24*(1), 118-136. https://doi.org/10.23960/aksara/v24i1.pp118-136

Seyyedrezaei, M. S., Amiryousefi, M., Gimeno-Sanz, A., & Tavakoli, M. (2022). A meta-analysis of the relative effectiveness of technology-enhanced language learning on ESL/EFL writing performance: Retrospect and prospect. *Computer Assisted Language Learning*, 1-34. https://doi.org/10.1080/09588221.2022.2118782

Song, F., Khan, K. S., Dinnes, J., & Sutton, A. J. (2002). Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. *International journal of epidemiology*, *31*(1), 88-95. https://doi.org/10.1093/ije/31.1.88

Tang, J. L., & Liu, J. L. (2000). Misleading funnel plot for detection of bias in meta-analysis. *Journal of clinical epidemiology*, *53*(5), 477-484. https://doi.org/10.1016/S0895-4356(99)00204-8

Viechtbauer, W. (2005). Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics*, *30*(3), 261-293. https://doi.org/10.3102/10769986030003261

Wang, J., An, N., & Wright, C. (2018). Enhancing beginner learners' oral proficiency in a flipped Chinese foreign language classroom. *Computer Assisted Language Learning*, *31*(5-6), 490-521. https://doi.org/10.1080/09588221.2017.1417872

Zarinfard, S., Rahimi, M., & Mohseni, A. (2021). The impact of flipped classroom on learning outcome in a general English course: grammar and vocabulary gains in focus. *International Journal of Foreign Language Teaching and Research*, *9*(38), 61-75. https://doi.org/10.52547/JFL.9.38.65

Zhang, J. (2022). Research on College English Flipped Classroom Teaching Model Based on Micro-class. Advances in Education, *Humanities and Social Science Research*, *1*(2), 421-421. https://doi.org/10.56028/aehssr.1.2.421

Zhiyong, Li., & Jiaying, Li. (2022). Using the flipped classroom to promote learner engagement for the sustainable development of language skills: A mixed-methods study. *Sustainability*, *14*(10), 5983. https://doi.org/10.3390/su14105983