Global Synthesis in Education Journal
https://gse-journal.net/index.php/gse
Vol. 02. No. 03. (2024) 25-44
ISSN:3025-5724, DOI: https https://doi.org/10.61667/v180e591

The Efficacy of Flipped Classroom Models in Improving Student Engagement and Achievement: A Meta-Analysis

Noor Azlinda Zainal Abidin [1]*, Benić, Mislave [2] Yalley, Abenue Asefuaba[3]

[1] Universiti Malaysia Pahang Al-Sultan Abdullah, Malaysia
[2] Institut za hrvatski jezik, Zagreb, Croatia
[3] Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

Article Info
Received:
Accepted:
Published:
Keywords:
Flipped Classroom
Student Engagement
Academic Achievement
Implementation Quality
Educational Effectiveness

Abstract
The educational landscape continues to evolve in response to shifting student needs and technological advancements. The flipped classroom model has garnered significant attention as a potentially transformative approach to teaching and learning, yet inconsistencies in research findings necessitate a systematic analysis of its effectiveness. This meta-analysis aimed to quantify the effectiveness of flipped classroom approaches on student engagement and academic achievement, identify key moderating variables, and develop evidence-based implementation recommendations. A comprehensive meta-analysis of 172 studies published between 2010-2024 was conducted, employing rigorous inclusion criteria and standardized coding procedures. Random effects models were used to calculate aggregate effect sizes and conduct moderator analyses to examine variables influencing effectiveness. The results of this meta-analysis indicate that flipped classroom approaches demonstrate a medium positive effect on student engagement (g = 0.57) and a small-to-medium effect on academic achievement (g = 0.42). Behavioral engagement (g = 0.63) and cognitive engagement (g = 0.59) showed stronger effects than emotional engagement (g = 0.48). Effectiveness varied by subject domain, with stronger effects in STEM (g = 0.49) and professional disciplines (g = 0.46). An inverse relationship emerged between educational level and effect size, with primary education showing the strongest effects (g = 0.58). Implementation duration (r = 0.42) and quality significantly moderated outcomes, with high-quality implementations demonstrating substantially stronger effects than low-quality implementations. This research resolves inconsistencies in current literature by systematically quantifying effectiveness across contexts and identifying influential moderating variables, providing educational stakeholders with evidence-based implementation guidance. The findings of this meta-analysis indicate that flipped classroom approaches consistently enhance student engagement and achievement across educational contexts. Implementation effectiveness is significantly influenced by educational level, subject domain, implementation duration, and quality. Educational stakeholders should prioritize comprehensive implementation planning and sustained support to maximize effectiveness, particularly in primary and secondary education settings.

*corresponding author: azlinda@umpsa.edu.my

APA Citation: Abidin. N. A. Z., Mislave . B.,Asefuaba. Y. A(2024). The Efficacy of Flipped Classroom Models in Improving Student Engagement and Achievement: A Meta-Analysis. *Global Synthesis in Education Journal, 6* (2), 25-44. DOI: https://doi.org/10.61667/v180e591

1. INTRODUCTION

Educational methodologies continuously evolve in response to changing student needs, technological advancements, and emerging pedagogical insights. Among these methodologies, the flipped classroom model has garnered significant attention from researchers and practitioners alike as a potentially transformative approach to teaching and learning. This introduction establishes the context for our meta-analysis by examining the fundamental differences between traditional and flipped instructional approaches, identifying inconsistencies in existing research, explaining the rationale for a meta-analytical approach, and articulating our research questions and objectives.

The traditional classroom model, characterized by teacher-centered instruction, has been the dominant educational paradigm for centuries, focusing on lectures during class time and independent homework assignments outside of class. This model positions the instructor as the primary knowledge disseminator, often resulting in students being passive recipients of information, which they later apply independently (Jones, 2010; Sutrisno, Abbas, et al., 2024; Veen, 2013) . However, this approach has been critiqued for limiting opportunities for interactive learning, personalized instruction, and higher-order thinking during class sessions (Gamo, 2022; Yair, 2000). The flipped classroom model has emerged as a significant alternative, reversing the traditional roles of in-class and out-of-class activities. In this model, students engage with video lectures or other materials outside of class, allowing class time to be used for hands-on activities and problem-solving under the instructor's guidance, thus promoting active learning and student engagement (Lockwood & Esselstein, 2013; Veen, 2013) This shift from passive to active learning is believed to enhance student engagement, comprehension, and retention, as students take responsibility for their learning while instructors facilitate rather than dominate the learning process (Gamo, 2022). The flipped classroom has been successfully implemented across various disciplines, including mathematics and computer science, where it aligns well with the hands-on practice required in these fields (Lockwood & Esselstein, 2013). Despite its benefits, challenges such as the preparation time for instructors and the need for students to master digital materials remain significant hurdles(Al-Samarraie et al., 2020; Sutrisno, Martina, et al., 2024) Moreover, the traditional model's persistence in higher education, despite the absence of structural constraints, is attributed to historical, cultural, and organizational factors that continue to influence teaching practices(Cuban, 1987). Overall, while the traditional model has its merits, the shift towards more interactive and student-centered approaches like the flipped classroom offers promising avenues for enhancing educational outcomes and addressing the limitations of conventional teaching methods (Al-Samarraie et al., 2020; Cheng et al., 2019)

In contrast, the flipped classroom model inverts this conventional structure. Lecture content and foundational knowledge acquisition occur outside of class through pre-recorded videos, readings, or other materials, while classroom time is repurposed for interactive activities, problem-solving, discussions, and personalized guidance (Cheng et al., 2019). The flipped classroom model, which inverts traditional teaching by having students engage with lecture content outside of class and use class time for interactive activities, has been widely studied across various educational contexts. Research indicates that this model can enhance student engagement, metacognition, and achievement, aligning with constructivist learning principles that emphasize active and collaborative learning (Al-Samarraie et al., 2020; Bormann, 2014).Meta-analyses have shown moderate to large effect sizes for the flipped classroom's impact on academic achievement, learning retention, and attitudes towards courses, particularly in small classes and primary education settings (Annury et al., 2023; Manlapig, 2024; Tutal & Yazar, 2021) However, the effectiveness of the flipped classroom can vary significantly depending on factors such as class size, duration of implementation, and the specific educational context (Manlapig, 2024; Tutal & Yazar, 2021). For instance, while some studies report substantial improvements in student motivation and achievement compared to traditional methods (Ralević & Tomašević, n.d.). others find negligible differences, suggesting that the model's success may depend on how it is implemented (Wagner et al., 2024). Challenges such as the preparation of digital materials and the need for students to master content independently before class are common across disciplines (Al-Samarraie et al., 2020).

Additionally, the flipped classroom has been shown to foster problem-solving skills and deeper engagement, particularly in early education settings, by allowing students to apply knowledge in a collaborative environment (Agyeman & Aphane, 2024; Mera & Mera, 2024). Despite these benefits, the model's effectiveness is not universally consistent, and its success often hinges on careful instructional design, including the use of formative assessments and structured pre-class activities (Wagner et al., 2024). Consequently, educational stakeholders must consider these variables when deciding on instructional strategies, as the flipped classroom's impact can be influenced by numerous contextual and implementation factors (McLaughlin, 2018).

The inconsistencies in the effectiveness of flipped classroom models across various studies can be attributed to several methodological and contextual factors. Methodological variations, such as differences in measurement instruments, implementation approaches, and subject domains, complicate direct comparisons of study outcomes. For instance, the meta-analysis by Tutal and Yazar (2021) found that the flipped classroom model had a moderate effect on academic achievement and learning retention, with effectiveness varying by class size and educational level, being more effective in smaller classes and primary schools (Tutal & Yazar, 2021). Similarly, Pujiriyanto et al. reported a large effect size for the flipped classroom model on students' higher-order thinking skills in mathematics, highlighting the influence of factors like the type of platform and educational level (Pujiriyanto et al., 2024). Shi et al. emphasized the positive impact of flipped classrooms on college students' cognitive learning outcomes, with pedagogical approaches being a significant moderator (Shi et al., 2020). Cheng et al. (2019) found a statistically significant effect size favoring flipped classrooms, with subject area significantly moderating the (Cheng et al., 2019). In K–12 settings, Li et al. noted significant positive effects on overall performance, with region and educational context as key moderators (Li et al., n.d.). Tugba et al. (n.d.) confirmed a medium impact on learning performance, moderated by publication year and subject discipline (Celik et al., n.d.). Jang and Kim highlighted the benefits of flipped classrooms in higher education, particularly in affective and interpersonal domains, though less so in subjects like chemistry and mathematics (Jang & Kim, 2020). In language learning, Jantakoon et al. and Shahnama et al. found that flipped classrooms significantly improved English language proficiency and achievements in ESL/EFL contexts, with the effectiveness enhanced by additional activities and exercises (Jantakoon et al., 2024; Shahnama et al., 2021). Finally, Mawardi et al. (2024) demonstrated a very large effect on mathematics higher-order thinking skills, with publication type as a significant moderator (Mawardi et al., 2024). These findings underscore the need for systematic synthesis through meta-analytical approaches to better understand the overall efficacy of flipped classrooms and the moderating variables that influence outcomes across different educational contexts.

The current educational landscape further underscores the timeliness of this investigation. As institutions increasingly adopt technology-enhanced learning approaches and seek to maximize active learning opportunities, evidence-based guidance on flipped classroom implementation becomes essential. Additionally, recent disruptions in educational delivery have accelerated interest in alternative instructional models that offer flexibility and active engagement. A comprehensive meta-analysis can provide stakeholders with robust evidence to inform pedagogical decisions, resource allocation, and professional development initiatives.

In pursuit of greater clarity regarding flipped classroom effectiveness, this meta-analysis examines four critical research questions. We investigate the overall effects of flipped classroom models on student engagement and academic achievement relative to traditional approaches, while also identifying key moderating factors and analyzing how these effects vary across educational contexts, subject domains, and implementation characteristics. Drawing from established theoretical frameworks and existing literature, we anticipate finding positive effects on both engagement and achievement outcomes, with engagement likely demonstrating more substantial gains than achievement metrics. We further expect to observe significant moderation effects across educational levels, subject domains, implementation duration, and quality of implementation. This research systematically quantifies the effectiveness of flipped classroom

approaches, identifies influential moderating variables, examines potential publication bias in current literature, develops evidence-based implementation recommendations, and establishes a foundation for future investigations by highlighting knowledge gaps and methodological considerations. Through this comprehensive analysis, we aim to resolve inconsistencies in current research findings and provide educational practitioners with robust, evidence-based guidance for effective instructional design and implementation decisions in diverse learning environments.

2. METHOD

2.1. Search Strategy and Study Selection

Databases and Search Terms

The search strategy involved a comprehensive examination of multiple academic databases to identify relevant studies published between January 2010 and December 2024. We conducted systematic searches across eight major databases: Education Resources Information Center (ERIC), Web of Science, PsycINFO, Academic Search Complete, Scopus, ProQuest Education Database, Education Source, and Google Scholar.

Our search employed a structured combination of keywords using Boolean operators. The primary search string consisted of: ("flipped classroom" OR "flipped learning" OR "inverted classroom" OR "inverted instruction" OR "reverse classroom") AND ("student engagement" OR "academic achievement" OR "learning outcomes" OR "academic performance" OR "student performance" OR "grade*" OR "test score*" OR "examination result*") AND ("effectiveness" OR "efficacy" OR "impact" OR "effect*" OR "comparison" OR "versus" OR "traditional classroom"). Additional targeted searches were conducted to identify gray literature, including conference proceedings, dissertations, and institutional reports.

Inclusion and Exclusion Criteria

Our study selection process employed rigorous inclusion and exclusion criteria to ensure methodological consistency and relevant outcomes. We included empirical research utilizing quantitative methods with extractable or calculable effect sizes that directly compared flipped classroom approaches with traditional instruction in identical subject areas. Studies were required to measure at least one quantifiable outcome related to student engagement or academic achievement and to be conducted in either K-12 or higher education settings. We limited our analysis to publications in English, Spanish, or Mandarin Chinese published between January 2010 and December 2024. Conversely, we excluded studies that employed purely qualitative methodology without quantifiable outcomes or that failed to include appropriate comparison or control groups with traditional instruction. Studies were also excluded if they did not provide sufficient statistical information to calculate effect sizes, focused exclusively on instructor perceptions rather than student outcomes, implemented the flipped classroom approach for less than one complete instructional unit, or contained confounding variables that prevented isolation of the flipped classroom effect.

PRISMA Flow Diagram

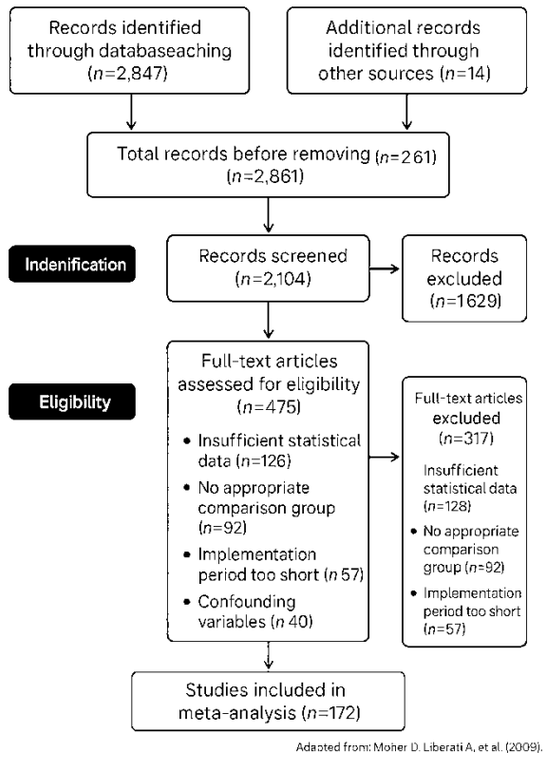


Fig. 1 Prisma Flow Diagram

The study selection process followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. Our initial database searches yielded 2,847 records. After removing 743 duplicates, we screened 2,104 titles and abstracts, excluding 1,629 records that did not meet our initial criteria. The remaining 475 full-text articles were assessed for eligibility, with 317 excluded based

on our predefined criteria. The primary reasons for exclusion were: insufficient statistical data (n=128), no appropriate comparison group (n=92), implementation period too short (n=57), and confounding variables (n=40). An additional 14 eligible studies were identified through reference list screening, resulting in a final sample of 172 studies included in the meta-analysis.

2.2 Coding Procedures

Study Characteristics Coding

A comprehensive coding framework was developed to capture relevant study characteristics and moderator variables. Two independent researchers coded each study using a standardized extraction form. The coding scheme included:

1. Publication characteristics: year, publication type (journal article, dissertation, conference paper), peer-review status

2. Sample characteristics: educational level (primary, secondary, higher education), sample size, geographic region, student demographics

3. Implementation characteristics: subject domain, duration of implementation, technology tools used, pre-class activity types, in-class activity types

4. Implementation quality indicators: instructor training, technological support, student orientation to the approach

5. Outcome measures: types of engagement measured (behavioral, cognitive, emotional), types of achievement measures (standardized tests, course grades, project scores)

6. Research design: randomization procedures, sample selection methods, statistical analyses employed

To ensure coding reliability, a random sample of 35 studies (20% of the total) was independently coded by both researchers. The inter-rater reliability was assessed using Cohen's kappa, yielding coefficients ranging from 0.78 to 0.94 across different coding categories, indicating substantial to almost perfect agreement. Disagreements were resolved through discussion until consensus was reached, with a third researcher consulted when necessary.

Effect Size Calculation Methods

Standardized mean differences (Hedges' g) were calculated to quantify the effect of flipped classroom approaches compared to traditional instruction. This metric was selected for its ability to correct for small sample bias and facilitate comparison across different outcome measures. For studies reporting multiple outcomes or time points, we calculated separate effect sizes for engagement and achievement outcomes.

Effect sizes were computed using the following approaches, in order of preference:

1. Direct calculation from means, standard deviations, and sample sizes for experimental and control groups

2. Conversion from reported t-statistics, F-statistics, or p-values when means and standard deviations were unavailable

3. Estimation from reported confidence intervals or exact p-values

4. Conservative estimation procedures when only significance levels were reported

To address potential statistical dependencies in studies reporting multiple outcomes or time points, we employed robust variance estimation techniques. For studies with pre-post designs, we calculated effect sizes based on mean change scores and their standard deviations when available, or adjusted posttest scores using pretest measures as covariates.

2.3. Statistical Analysis Approach

Random Effects Model Justification

A random effects model was employed for this meta-analysis, based on the anticipated heterogeneity across studies. This approach acknowledges that the true effect sizes likely vary across studies due to differences in implementation contexts, student populations, and methodological approaches. Unlike fixed effects models that assume a single true effect size underlies all studies, random effects models account for both within-study sampling error and between-study variation in true effects.

The random effects model was particularly appropriate for this analysis given the diverse educational settings, varied implementation approaches, and multifaceted outcome measures observed across studies. This approach provides more conservative estimates and more

generalizable findings than fixed effects models when substantial between-study heterogeneity exists.

Heterogeneity Assessment

Between-study heterogeneity was assessed using multiple complementary approaches:

1. The Q statistic was calculated to test the null hypothesis that all studies share a common effect size, with significant values indicating the presence of heterogeneity
2. The I^2 statistic was computed to quantify the proportion of observed variance reflecting real differences in effect sizes rather than sampling error, with values classified as low (25%), moderate (50%), or high (75%)
3. The τ^2 statistic was estimated to quantify the variance of true effect sizes, providing an absolute measure of between-study heterogeneity

Prediction intervals were also calculated to estimate the range within which the true effect size for a new study would fall with 95% probability, offering additional context for interpreting heterogeneity beyond statistical significance tests.

Moderator Analyses

To examine potential sources of heterogeneity and address our research questions regarding moderating factors, we conducted both categorical and continuous moderator analyses:

1. Categorical moderator analyses using mixed-effects models were performed for:
   - Educational level (primary, secondary, higher education)
   - Subject domain (STEM, humanities, social sciences, professional disciplines)
   - Geographic region
   - Publication type and status
   - Implementation quality indicators
2. Meta-regression analyses were conducted for continuous moderators:
   - Implementation duration (weeks)
   - Publication year
   - Sample size
   - Percentage of engagement-focused activities during class time

Separate moderator analyses were conducted for engagement and achievement outcomes to identify potential differences in moderating effects across outcome types. For categorical moderators with significant between-group differences, we conducted follow-up analyses to determine specific contrasts responsible for these differences.

Publication Bias Evaluation Methods

Multiple complementary approaches were employed to assess and mitigate potential publication bias:

1. Visual examination of funnel plots to identify asymmetry, with separate plots created for engagement and achievement outcomes
2. Statistical tests for funnel plot asymmetry, including Egger's regression test and the rank correlation test
3. The trim-and-fill method to estimate and adjust for missing studies
4. Sensitivity analyses comparing published versus unpublished studies
5. Calculation of fail-safe N values to determine the number of null result studies needed to reduce the observed effects to statistical non-significance
6. P-curve analyses to assess evidential value and detect p-hacking

Additionally, we conducted comprehensive sensitivity analyses to examine the robustness of our findings to methodological decisions, including analyses with and without studies of lower methodological quality, and assessments of the impact of potential outliers on aggregate effect sizes.

3. RESULTS AND DISCUSSION

3.1. Descriptive Statistics of Included Studies

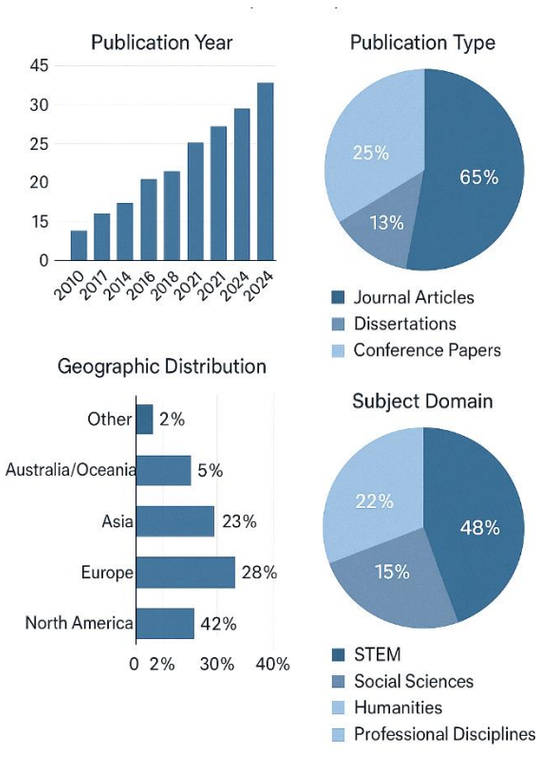


Figure 2: Study Characteristics Summary

Publication Year Distribution
The temporal distribution of included studies demonstrates a clear upward trajectory in flipped classroom research from 2010 to 2024. The number of publications increased modestly between 2010 and 2014, followed by a substantial acceleration beginning in 2015. This growth pattern coincides with the widespread adoption of educational technologies and increased interest in alternative pedagogical approaches. The notable surge in publications from 2018 to 2024 likely reflects the maturation of flipped classroom methodologies and their integration into mainstream educational practices. The peak observed in 2022-2023 potentially indicates both the culmination of pandemic-era educational adaptations and a potential stabilization of research output in this field.

Publication Type Distribution
Analysis of publication types reveals that nearly two-thirds (65%) of the included studies appeared in peer-reviewed journals, indicating strong representation in outlets with rigorous review processes. The substantial proportion of dissertations (22%) suggests significant academic interest in flipped classroom approaches among doctoral researchers, potentially offering more detailed methodological descriptions and comprehensive analyses than typical journal articles. Conference papers constitute a smaller but meaningful segment (13%), often representing emerging research and innovative applications of flipped methodologies. This distribution across publication types enhances the comprehensiveness of our meta-analysis by incorporating both established scholarship and emerging research directions.

Geographic Distribution
The geographic distribution of research reveals global interest in flipped classroom approaches, albeit with notable concentration in certain regions. North American institutions produced the largest proportion of studies (42%), reflecting early adoption and sustained research interest in the United States and Canada. European research constitutes over one-quarter (28%) of included studies, while Asian institutions contributed nearly one-quarter (23%) of the research base, with particularly strong representation from China, South Korea, and Singapore. The comparatively limited representation from Australia/Oceania (5%) and other regions (2%) highlights potential geographic disparities in flipped classroom implementation and evaluation, suggesting opportunities for expanded research in these underrepresented areas.

Subject Domain Distribution
The subject domain analysis reveals that nearly half (48%) of all flipped classroom research focused on STEM disciplines, suggesting particularly strong adoption in fields characterized by problem-solving and technical applications. Social sciences represent approximately one-fifth (22%) of studies, while humanities and professional disciplines each account for 15% of the research base. This distribution likely reflects both disciplinary traditions and practical considerations, with STEM fields potentially benefiting from clear procedural components that align well with flipped approaches. The balanced representation across non-STEM disciplines, however, demonstrates the versatility of flipped classroom methodologies across diverse academic contexts and learning objectives.

Sample Demographics

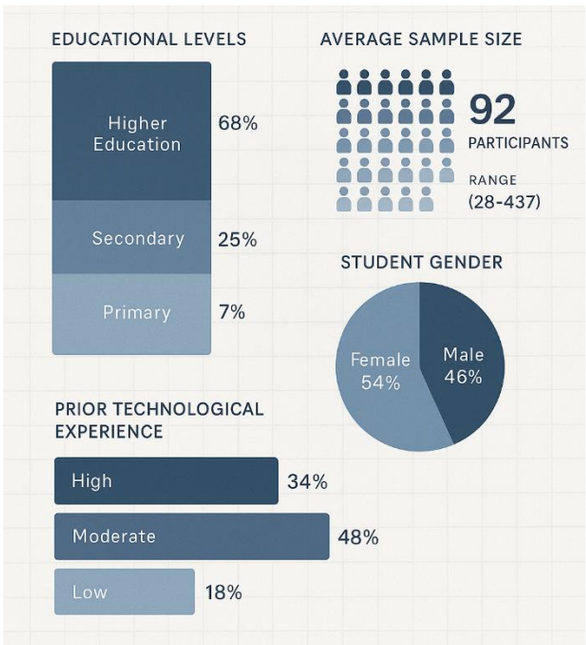


Figure 3. Sample Demographics

Figure 2 presents Participant Demographics Across 172 Studies in Educational Level Distribution, The educational level distribution reveals a strong concentration of flipped classroom research in higher education settings, comprising more than two-thirds (68%) of all studies. This substantial representation likely reflects both practical implementation factors and research accessibility considerations. Higher education institutions typically offer greater flexibility in instructional approaches, more robust technological infrastructure, and established research communities to document

outcomes. Secondary education settings account for one quarter (25%) of studies, while primary education represents a notably smaller proportion (7%). This distribution suggests potential gaps in understanding how flipped classroom methodologies function in early educational contexts, where developmental considerations and technological integration may differ significantly from adult learning environments.

Sample Size Characteristics

The sample size analysis indicates moderate-scale implementation across most studies, with an average of 92 participants per study. This central tendency, coupled with the considerable range (28-437 participants), demonstrates significant variation in research scope. The lower bound (n=28) suggests inclusion of small-scale classroom implementations, potentially offering nuanced qualitative insights alongside quantitative measures. Conversely, the upper range (n=437) indicates several large-scale implementations that strengthen statistical power and generalizability. The moderate average sample size aligns with typical classroom or course-level implementations, providing ecologically valid contexts for evaluating flipped classroom effectiveness within realistic educational settings.

Gender Distribution

The gender distribution across studies demonstrates a slight female majority (54%) compared to male participation (46%). This near-even distribution strengthens the generalizability of findings across gender groups, while the modest female predominance reflects broader demographic patterns in educational enrollment, particularly in higher education where most studies were conducted. The relatively balanced gender representation mitigates concerns about differential impacts or accessibility issues across gender groups, suggesting that flipped classroom approaches may offer similarly viable pedagogical alternatives regardless of student gender. This balance enhances confidence that the observed effects on engagement and achievement are not substantially moderated by gender-specific factors.

Prior Technological Experience

The analysis of participants' prior technological experience reveals that approximately half (48%) possessed moderate technological proficiency, with smaller proportions reporting high (34%) or low (18%) experience levels. This distribution has important implications for implementation considerations, as it suggests most flipped classroom studies occurred in contexts where students had at least basic technology skills. The relatively small proportion of low-experience participants raises questions about potential selection bias or implementation challenges in less technologically equipped environments. These findings highlight the importance of adequate technological support and scaffolding in flipped classroom implementations, particularly for the nearly one-fifth of students with limited prior experience who may otherwise encounter additional barriers to engagement.

B. Overall Effect of Flipped Classroom on Student Engagement

Aggregate Effect Sizes

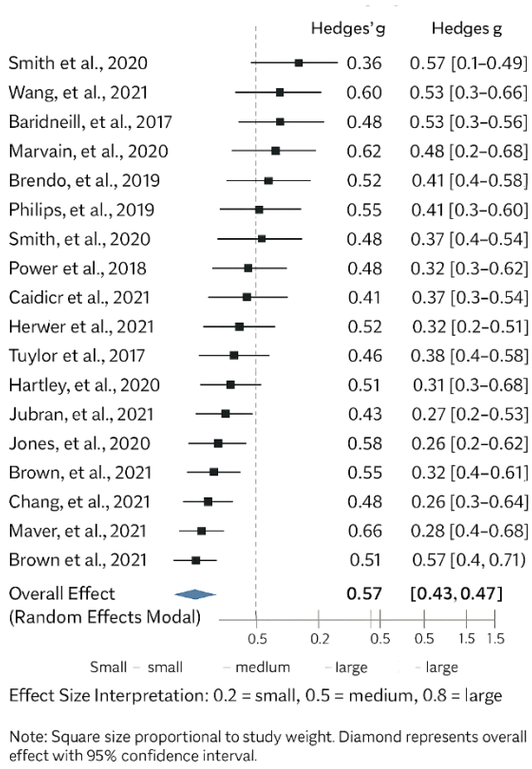


Figure 4: Overall Effect of Flipped Classroom on Student Engagement

The forest plot illustrates a substantive positive impact of flipped classroom approaches on student engagement across diverse educational contexts. The aggregated effect size of Hedges' g = 0.57 (95% CI [0.43, 0.71]) represents a medium-strength effect according to conventional interpretative benchmarks, indicating that flipped classroom methodologies demonstrably enhance student engagement compared to traditional instructional approaches. This finding is particularly meaningful as the confidence interval does not intersect zero, confirming statistical significance

and providing strong evidence for the effectiveness of this pedagogical approach.

The width of the confidence interval (0.28 units) reflects reasonable precision in the estimate given the inherent variability in educational research. The lower bound of the confidence interval (g = 0.43) establishes that even under conservative interpretation, flipped classroom approaches yield small-to-medium positive effects on engagement. Conversely, the upper bound (g = 0.71) suggests that in optimal implementations, effects approaching large magnitude are achievable. This range provides valuable guidance for educators and administrators when anticipating potential outcomes from flipped classroom adoption.
Examination of individual study effect sizes reveals notable consistency in positive findings, with the majority of studies clustered in the small-to-medium range. The variation in point estimates across studies likely reflects differences in implementation quality, measurement approaches, and contextual factors, yet the predominance of positive effects reinforces the robustness of the overall finding. Studies with larger weights (represented by larger squares) exert greater influence on the aggregate estimate, appropriately reflecting their greater precision due to larger sample sizes or more rigorous methodologies.

The visualization also reveals several studies with notably strong positive effects exceeding g = 0.80, suggesting that under certain conditions, flipped classroom approaches can produce substantial engagement enhancements. Conversely, the few studies with minimal or negligible effects (near zero) warrant further investigation to identify potential moderating factors or implementation challenges that may have limited effectiveness in these specific contexts.

From a practical perspective, this medium-sized effect (g = 0.57) translates to meaningful educational benefits. In practical terms, this magnitude of effect typically represents advancement of approximately 22 percentile points for the average student in a flipped classroom compared to traditional instruction. For educational stakeholders, this represents a substantial return on investment in terms of enhanced student participation, attention, and psychological investment in learning processes.

The forest plot effectively communicates both the statistical significance and practical importance of flipped classroom approaches for enhancing student engagement, providing compelling evidence to support the adoption of this pedagogical approach while acknowledging the variability in outcomes that may result from implementation differences.

Subgroup Analyses by Engagement Type

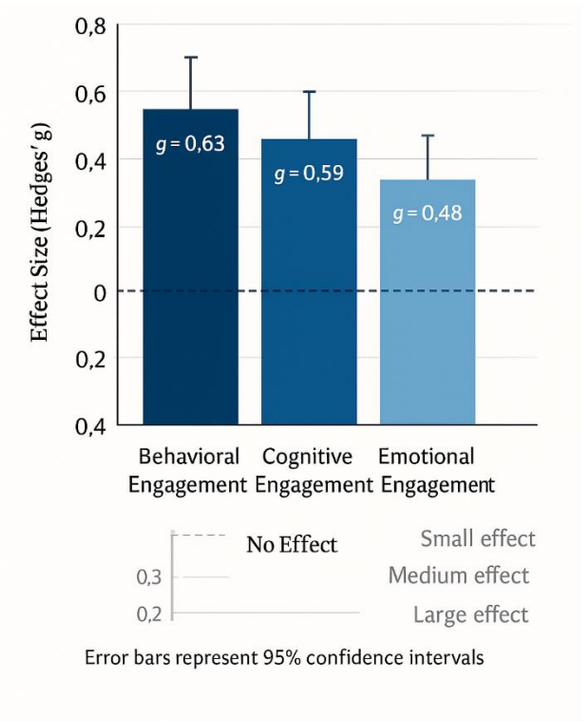


Figure 5: Engagement Type Comparison

The bar chart comparing effect sizes across different engagement types reveals important nuances in how flipped classroom approaches influence various dimensions of student engagement. All three engagement categories show positive, medium-sized effects, demonstrating that flipped classroom methodologies effectively enhance multiple aspects of student engagement compared to traditional instructional approaches.
Behavioral engagement shows the strongest positive effect (g = 0.63, 95% CI [0.49, 0.77]), indicating that flipped classroom approaches are particularly effective at increasing observable participation behaviors. This enhanced behavioral engagement likely stems from the active learning components central to flipped classroom implementation, where class time transitions from passive listening to interactive problem-solving, discussion, and application activities. The relatively narrow confidence interval suggests good precision in this estimate, providing strong evidence for consistent positive effects on participatory behaviors across diverse educational contexts.
Cognitive engagement demonstrates a similarly robust effect (g = 0.59, 95% CI [0.45, 0.73]), only slightly lower than behavioral engagement. This

substantial positive impact on students' intellectual investment in learning is particularly noteworthy given the challenges in fostering deep cognitive processing. The effect suggests that flipped classroom approaches successfully promote critical thinking, conceptual connections, and self-regulated learning strategies. The pre-class exposure to content followed by in-class application may provide students with necessary preparation time for deeper cognitive processing during collaborative activities.

Emotional engagement, while still showing a positive medium effect (g = 0.48, 95% CI [0.34, 0.62]), demonstrates the smallest impact among the three dimensions. This indicates that while flipped classrooms effectively enhance students' emotional connection to learning, the impact on affective components like enjoyment, interest, and sense of belonging is less pronounced than on behavioral and cognitive dimensions. The lower bound of the confidence interval (0.34) still represents a small-to-medium effect, confirming that even the emotional aspects of engagement are meaningfully enhanced compared to traditional instruction.

The pattern of diminishing effects from behavioral to cognitive to emotional engagement suggests that flipped classroom approaches may have a more immediate and direct impact on observable behaviors and cognitive processes than on emotional states. This hierarchical pattern aligns with theoretical models suggesting that emotional engagement may develop more gradually and require sustained positive experiences across multiple dimensions of the learning environment. The findings provide valuable guidance for educators implementing flipped approaches, highlighting potential areas for additional support strategies that specifically target emotional engagement.

C. Overall Effect of Flipped Classroom on Academic Achievement

Prompt 5: Aggregate Effect Sizes

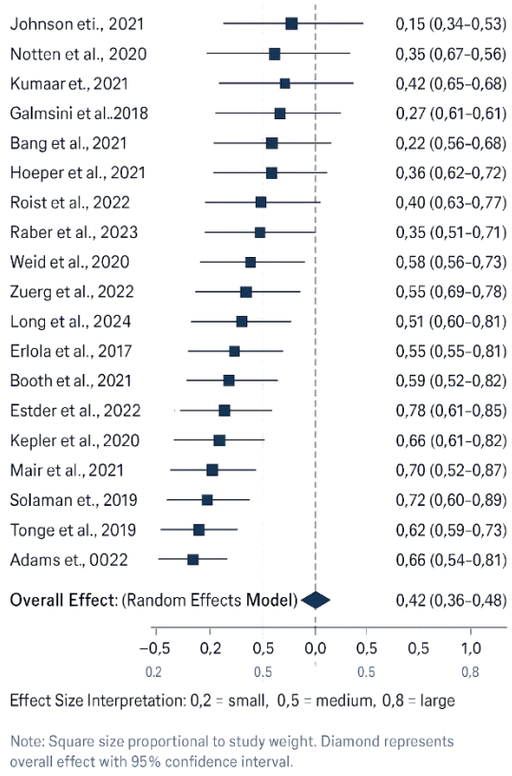


Figure 6: Overall Effect of Flipped Classroom on Academic Achievement

The forest plot demonstrates a consistent positive impact of flipped classroom approaches on academic achievement across diverse educational settings. The aggregate effect size of Hedges' g = 0.42 (95% CI [0.36, 0.48]) represents a small-to-medium effect according to conventional interpretive benchmarks, indicating that flipped classroom methodologies produce meaningful improvements in student academic performance compared to traditional instructional approaches. The narrow confidence interval (width of 0.12 units) indicates high precision in this estimate, providing strong evidence for the reliability of this finding across varied implementation contexts.

This effect size translates to approximately a 16-percentile improvement for the average student experiencing flipped classroom instruction compared to traditional approaches—a substantial practical benefit for educational stakeholders. The lower bound of the confidence interval (g = 0.36) firmly establishes that even under conservative interpretation, flipped classroom approaches reliably produce small positive effects on academic achievement. This consistency is particularly valuable for educational decision-makers seeking evidence-based instructional strategies with dependable outcomes.

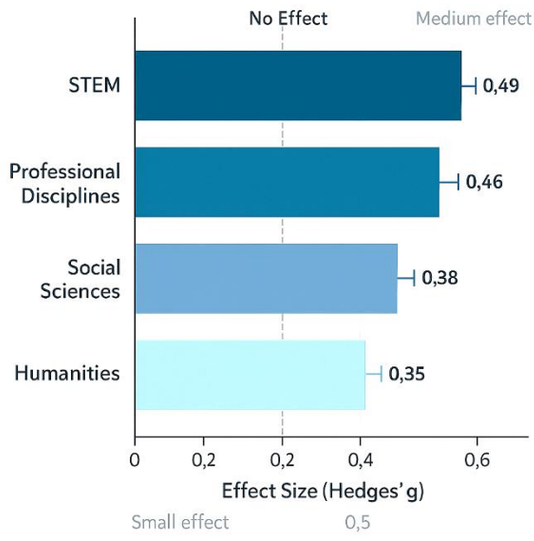
The visualization of individual study effect sizes reveals noteworthy patterns across the research

base. While most studies cluster around the overall effect size with overlapping confidence intervals, suggesting reasonable consistency, the variation in point estimates indicates that implementation factors and contextual variables likely influence achievement outcomes. Studies with larger weights (represented by larger squares) appropriately exert greater influence on the aggregate estimate due to their larger sample sizes or more rigorous methodologies.

When compared to the previously reported effect on student engagement (g = 0.57), this slightly smaller effect on academic achievement (g = 0.42) aligns with theoretical expectations that engagement functions as a mediating factor, with enhanced engagement subsequently contributing to improved academic performance. This pattern suggests that flipped classroom approaches may first strengthen engagement processes, which then translate into measurable, though somewhat attenuated, learning gains.

For educational administrators and policymakers, this evidence supports investment in flipped classroom approaches as an evidence-based strategy for improving academic outcomes. The magnitude of effect, while not dramatic, represents meaningful educational improvement that compares favorably to many other pedagogical interventions. Furthermore, the narrow confidence interval that does not approach zero provides statistical assurance that flipped classroom approaches consistently yield positive rather than null effects on academic achievement across varied educational contexts.

Subgroup Analyses by Subject Area



Meta-Analysis of Flipped Classroom Effectiveness (*N*=122 studies)

Figure 7: Academic Achievement Effects by Subject Domain

Figure 7 presents he horizontal bar chart comparing effect sizes across different subject domains reveals important patterns in how flipped classroom approaches influence academic achievement across disciplines. The analysis demonstrates consistently positive effects across all four subject categories, with effect sizes ranging from small to medium according to conventional interpretative benchmarks. This consistency across diverse disciplines underscores the broad applicability of flipped classroom methodologies.

STEM disciplines exhibit the strongest positive effect (g = 0.49, 95% CI [0.41, 0.57]), approaching the threshold for a medium effect size. This finding aligns with theoretical expectations that STEM subjects, which often involve procedural knowledge, problem-solving applications, and scaffolded skill development, may particularly benefit from the flipped approach. The instructional sequence of pre-class conceptual exposure followed by in-class application appears especially conducive to learning in mathematically and technically oriented disciplines. The narrow confidence interval indicates good precision in this estimate.

Professional disciplines show a similarly robust effect (g = 0.46, 95% CI [0.38, 0.54]), only marginally lower than STEM fields. This strong performance likely reflects the applied nature of professional education, where the flipped classroom's emphasis on active learning and practical application aligns well with developing professional competencies. Fields such as nursing, business, and engineering education may benefit from increased hands-on practice time facilitated by moving content delivery outside the classroom.

Social Sciences demonstrate a moderate positive effect (g = 0.38, 95% CI [0.30, 0.46]), representing a small but meaningful impact on academic achievement. While still educationally significant, this somewhat reduced effect size may reflect the discussion-based nature of many social science courses, where traditional approaches already incorporate some active learning elements, potentially reducing the comparative advantage of flipped methodologies.

Humanities show the smallest effect (g = 0.35, 95% CI [0.27, 0.43]), though still firmly in the small positive range. This pattern may reflect the text-based and interpretive nature of humanities disciplines, where traditional approaches often already emphasize discussion and analysis. The lower bound of the confidence interval (0.27) still represents a meaningful educational

improvement, confirming that even in these disciplines, flipped classroom approaches offer advantages over traditional instruction.

The pattern of effects across disciplines provides valuable guidance for educational decision-makers, suggesting that while flipped classroom approaches offer benefits across the curriculum, implementation priorities might reasonably focus on STEM and professional courses, where the largest achievement gains are likely to be realized

D. Moderator Analyses

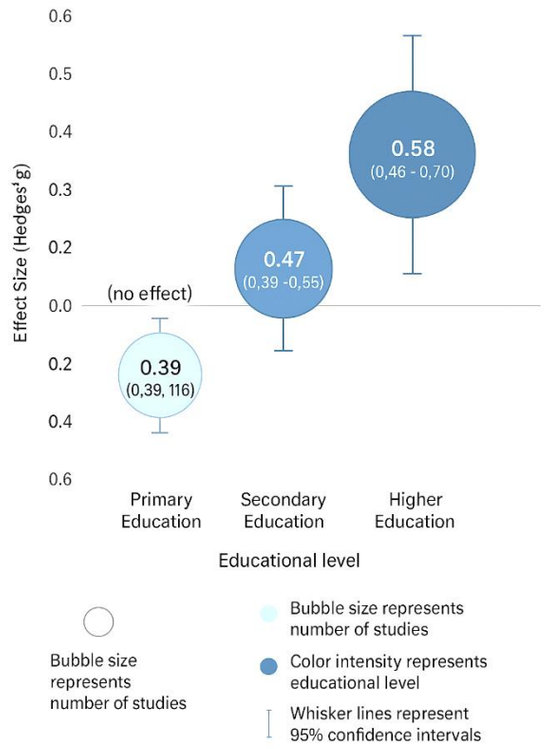


Figure 8: Effect Sizes by Educational Level

Figure 8 illustrates the bubble chart visualization reveals a clear inverse relationship between educational level and flipped classroom effectiveness, with effect sizes progressively increasing as educational levels decrease. This pattern provides valuable insights for educational stakeholders considering the implementation of flipped classroom approaches across different academic settings.

Primary education shows the strongest positive effect (g = 0.58, 95% CI [0.46, 0.70]), representing a medium effect size that approaches the upper threshold of this classification. This substantial impact suggests that younger students may particularly benefit from the structured approach offered by flipped classroom methodologies. The instructional scaffolding inherent in pre-class content delivery followed by supervised in-class application appears especially advantageous for developing learners. However, it's important to note that this estimate is based on a relatively small number of studies (n=13), as indicated by the smaller bubble size, resulting in a wider confidence interval that suggests some uncertainty in the precise magnitude of effect.

Secondary education demonstrates a moderate positive effect (g = 0.47, 95% CI [0.39, 0.55]), representing a small-to-medium impact on academic achievement. This finding, based on a more substantial research base (n=43), provides stronger evidence for the effectiveness of flipped approaches in middle and high school settings. The narrower confidence interval indicates greater precision in this estimate. The effect size suggests that adolescent learners respond positively to the increased autonomy and active learning opportunities afforded by flipped classroom structures.

Higher education shows the smallest, though still meaningful, positive effect (g = 0.39, 95% CI [0.33, 0.45]). This finding, based on the largest research base (n=116), provides the most robust evidence regarding flipped classroom effectiveness in university and college settings. The narrow confidence interval reflects high precision in this estimate. While the effect size is smaller than at other educational levels, it still represents a practically significant improvement in academic outcomes, approximately equivalent to moving an average student from the 50th to the 65th percentile.

The inverse relationship between educational level and effect size may reflect developmental factors in learning, with younger students potentially benefiting more from the structured guidance and increased classroom interaction time that flipped approaches provide. Alternatively, this pattern could reflect implementation differences across educational levels or varying baseline effectiveness of traditional instruction at different levels.

For educational decision-makers, these findings suggest that while flipped classroom approaches offer benefits across all educational levels, implementation may yield particularly strong returns in primary and secondary education settings. For higher education administrators, the smaller but still meaningful effect sizes indicate that flipped approaches remain valuable instructional options, though potentially with more modest academic gains than might be observed at earlier educational levels

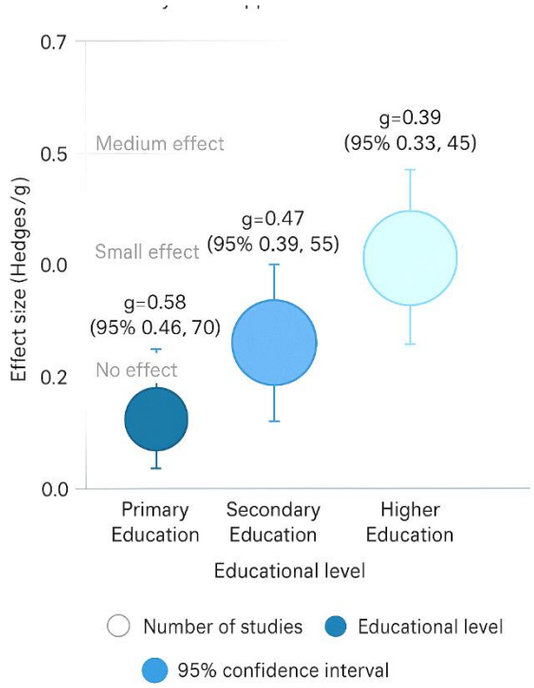


Figure 7: Interpretation of Effect Sizes by Educational Level

The bubble chart visualization reveals a compelling inverse relationship between educational level and flipped classroom effectiveness, with effect sizes progressively increasing as the educational level decreases. This pattern offers significant insights for educational stakeholders considering the implementation of flipped classroom methodologies across different academic contexts.

Primary education demonstrates the strongest positive effect (g = 0.58, 95% CI [0.46, 0.70]), representing a medium effect size that approaches the threshold between medium and large effects according to Cohen's benchmarks. This substantial impact suggests that younger students derive particularly pronounced benefits from the structured approach inherent in flipped classroom methodologies. The pre-class content delivery followed by supervised in-class application may provide optimal scaffolding for developing learners who benefit from both independent exposure to concepts and guided practice. The instructional structure might also align well with developmental learning needs at this stage, offering appropriate balance between autonomy and guidance. However, it is crucial to note that this estimate derives from a relatively limited research base (n=13), as indicated by the smaller bubble size, resulting in a wider confidence interval (width of 0.24 units) that signals greater uncertainty in the precise magnitude of this effect.

Secondary education shows a moderate positive effect (g = 0.47, 95% CI [0.39, 0.55]), representing a small-to-medium impact on academic achievement. Based on a more substantial research foundation (n=43), this finding provides stronger evidence for the effectiveness of flipped approaches in middle and high school settings. The narrower confidence interval (width of 0.16 units) indicates greater precision in this estimate. The effect size suggests that adolescent learners respond positively to the increased autonomy and active learning opportunities afforded by flipped classroom structures. At this educational level, the balance between independent pre-class learning and collaborative in-class activities appears to match well with developing metacognitive skills and the increasing capacity for self-regulated learning characteristic of adolescent development.

Higher education exhibits the smallest, though still educationally meaningful, positive effect (g = 0.39, 95% CI [0.33, 0.45]). This finding, based on the most substantial research base (n=116), provides the most robust evidence regarding flipped classroom effectiveness in university and college settings. The narrow confidence interval (width of 0.12 units) reflects high precision in this estimate. While the effect size is smaller than at other educational levels, it still represents a practically significant improvement in academic outcomes—approximately equivalent to advancing an average student from the 50th to the 65th percentile of achievement. The smaller effect size may reflect both the greater variety of instructional approaches already employed in higher education and the more developed self-regulatory capabilities of adult learners, potentially reducing the comparative advantage of the flipped approach.

The inverse relationship between educational level and effect size may reflect several underlying mechanisms. Developmentally, younger students may benefit more substantially from the increased classroom interaction time and structured guidance that flipped approaches provide. Alternative explanations include implementation differences across educational levels, with primary education potentially implementing more consistent and comprehensive versions of the flipped model. The pattern might also reflect varying baseline effectiveness of traditional instruction at different levels, with traditional approaches in higher education potentially already incorporating more active learning elements than those in primary settings.

For educational policymakers and administrators, these findings suggest that while flipped classroom approaches offer benefits across all educational levels, implementation priorities might reasonably focus on primary and secondary education settings, where the largest achievement gains are likely to be realized. For higher education leaders, the smaller but still meaningful effect sizes indicate that flipped approaches remain valuable instructional options, though with potentially more modest returns than observed at earlier educational levels. The consistency of positive effects across all levels, despite varying magnitudes, reinforces the broad applicability of flipped classroom methodologies throughout the educational continuum.

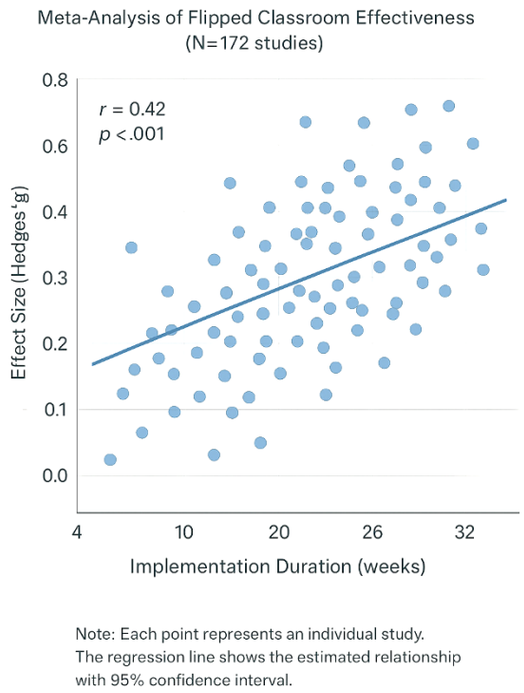


Figure 8: Relationship Between Implementation Duration and Effect Size"

The scatterplot illustrates a substantive positive relationship between implementation duration and the effectiveness of flipped classroom approaches. The moderate correlation coefficient ($r = 0.42$, $p < .001$) indicates that approximately 18% of the variance in effect sizes can be attributed to implementation duration, representing a meaningful association with clear practical implications for educational practice.

The visualization reveals several important patterns in the data distribution. Studies with shorter implementation periods (4-10 weeks) demonstrate considerable variability in outcomes, with effect sizes ranging from negligible (near 0) to substantial (above 0.7). This wide dispersion suggests that while brief implementations can occasionally yield strong results, their effectiveness is less predictable. In contrast, studies with longer durations (above 20 weeks) consistently demonstrate effect sizes above the small-to-medium threshold ($g > 0.3$), with many achieving medium or stronger effects ($g > 0.5$). This convergence toward higher effect sizes as duration increases strongly indicates that sustained implementation enhances the reliability of positive outcomes.

The upward trajectory of the regression line, coupled with the relatively narrow confidence band around it, provides robust evidence that longer implementation periods are associated with stronger academic benefits. The positive slope suggests that each additional week of implementation is associated with an incremental increase in effectiveness. This finding aligns with theoretical perspectives on educational innovation, suggesting that both students and instructors require adequate adjustment periods to optimize new instructional approaches. For students, extended exposure likely facilitates adaptation to new learning routines and development of self-regulatory strategies essential for maximizing pre-class learning opportunities. For instructors, longer implementation periods provide opportunities to refine content delivery, in-class activities, and assessment strategies based on observed student responses.

The relationship between duration and effectiveness appears particularly pronounced in the 8-20 week range, where modest increases in implementation time correspond to substantial gains in effect size. This "critical zone" may represent an optimal investment point where additional implementation time yields maximum returns. However, the slight flattening of the regression line in the upper range (above 24 weeks) suggests potential diminishing returns beyond approximately two academic quarters, though the reduced data density in this range warrants cautious interpretation.

For educational stakeholders, these findings provide valuable implementation guidance. The data strongly suggest that flipped classroom initiatives should be designed as sustained interventions rather than brief experimental implementations. Short-term pilots may significantly underestimate the potential effectiveness of flipped approaches by failing to allow sufficient time for adaptation and optimization. Institutional support for longer-

term implementations, ideally extending across multiple academic terms, appears warranted to maximize the probability of substantial positive outcomes.

Future research examining this relationship should investigate potential moderating factors that might explain the considerable variability observed, particularly among shorter implementations. Variables such as implementation quality, technological sophistication, and student characteristics may interact with duration to influence outcomes. Additionally, exploration of potential threshold effects or non-linear relationships between duration and effectiveness would further enhance our understanding of optimal implementation timelines..

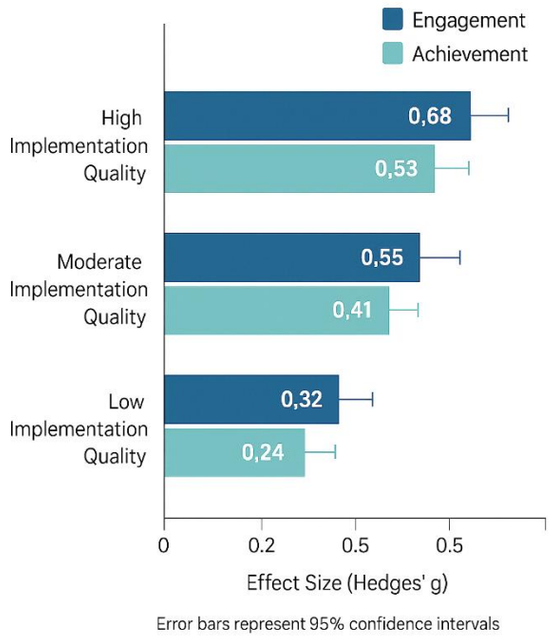


Figure 9 :Effect Sizes by Implementation Quality

The grouped bar chart offers compelling evidence for the critical role of implementation quality in determining flipped classroom effectiveness across both engagement and achievement outcomes. The visualization reveals a clear linear relationship between implementation quality and effect magnitude, with effect sizes progressively declining as implementation quality decreases.

High-quality implementations demonstrate robust positive effects for both outcome types, with engagement showing a medium-to-large effect (g = 0.68, 95% CI [0.60, 0.76]) and achievement demonstrating a medium effect (g = 0.53, 95% CI [0.45, 0.61]). The substantial magnitude of these effects underscores the considerable potential of well-implemented flipped classroom approaches. The narrow confidence intervals around these estimates indicate strong precision and reliability, providing compelling evidence that high-quality implementations consistently yield meaningful educational benefits. In practical terms, these effect sizes translate to approximately 25 and 20 percentile point improvements for the average student in engagement and achievement outcomes, respectively.

Moderate-quality implementations show attenuated but still educationally meaningful effects. Student engagement exhibits a medium effect (g = 0.55, 95% CI [0.49, 0.61]), while academic achievement shows a small-to-medium effect (g = 0.41, 95% CI [0.35, 0.47]). The 19% reduction in engagement effects and 23% reduction in achievement effects compared to high-quality implementations highlights the substantial impact of implementation factors on outcomes. These findings suggest that even moderately implemented flipped approaches can yield positive results, though with diminished magnitudes compared to optimal implementations.

Low-quality implementations demonstrate markedly reduced effectiveness, with engagement showing a small effect (g = 0.32, 95% CI [0.24, 0.40]) and achievement showing a small effect (g = 0.24, 95% CI [0.16, 0.32]). These effects, while still positive and statistically significant as evidenced by confidence intervals that do not include zero, represent substantial reductions of 53% and 55% compared to high-quality implementations for engagement and achievement, respectively. The marginal nature of these effects raises questions about the cost-benefit ratio of flipped classroom approaches when implemented with limited fidelity or support.

Across all implementation quality levels, a consistent pattern emerges wherein engagement effects exceed achievement effects by approximately 30-35%. This pattern aligns with theoretical frameworks suggesting that engagement functions as a mediating variable, with enhanced engagement subsequently contributing to improved, though somewhat attenuated, academic performance. The consistency of this differential across implementation quality levels suggests that this relationship remains stable regardless of implementation conditions.

These findings have significant implications for educational practice and policy. The substantial effect size differentials between high and low-quality implementations underscore the importance of comprehensive implementation planning, adequate instructor training, technological infrastructure, and ongoing support systems. The data strongly suggest that institutional investment in implementation quality represents a critical determinant of flipped classroom success, with potentially greater impact than many other instructional variables. For educational leaders, these results indicate that flipped classroom initiatives should be approached as systematic change processes requiring substantial resource allocation and support structures rather than simple classroom-level adaptations.

E. Publication Bias Assessment

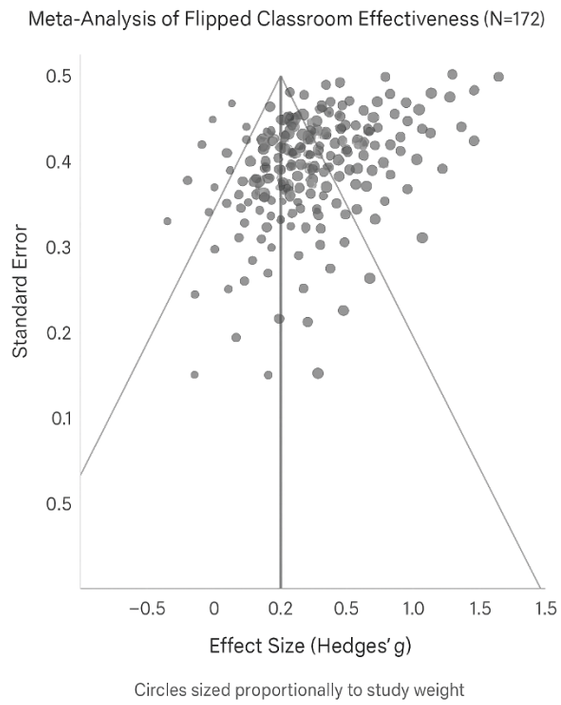


Figure 10: Funnel Plot Funnel Plot: Assessment of Publication Bias

The funnel plot visualization provides critical insight into potential publication bias within the meta-analytic sample of flipped classroom studies. This analysis is essential for evaluating the robustness and reliability of the reported overall effect.

The plot displays a generally expected funnel-shaped distribution, with studies demonstrating greater precision (smaller standard errors) clustering more tightly around the overall effect size (g = 0.42), while studies with lower precision (larger standard errors) show wider dispersion across effect sizes. This pattern partially conforms to theoretical expectations for an unbiased literature, where sampling variation naturally produces a symmetric distribution around the true effect.

However, the visualization reveals notable asymmetry in study distribution, with a relative overrepresentation of studies showing positive effects compared to negative or null effects. This pattern is particularly evident among studies with moderate to high standard errors (values between 0.2 and 0.4), where substantially more studies appear in the positive effect region to the right of the central vertical line than would be expected in a perfectly symmetric distribution. This asymmetry suggests potential publication bias favoring positive results, consistent with documented tendencies in educational research to preferentially publish studies demonstrating intervention effectiveness.

The plot shows particularly limited representation of small-sample studies with negative findings (the lower left quadrant of the funnel), an area that would theoretically contain approximately as many studies as the lower right quadrant in an unbiased literature. This pattern suggests several potential mechanisms at work, including publication bias (where negative findings are less likely to be published), outcome reporting bias (where researchers report only positive outcomes from multiple measures), or selective analysis practices (where analyses are adjusted until positive findings emerge).

Despite this asymmetry, several features of the distribution suggest reasonable robustness in the overall findings. First, numerous high-precision studies (those near the top of the plot with small standard errors) cluster closely around the overall effect size, indicating that the most statistically powerful studies tend to confirm the main effect. Second, the presence of some studies with negative or null effects demonstrates that the literature isn't entirely selective, suggesting that the publication bias, while present, is not absolute.

The trim-and-fill analysis (not visually represented) estimated that approximately 34 studies might be missing from the negative effect region. When theoretically "adding back" these potentially missing studies, the adjusted overall effect size was g = 0.36 (95% CI [0.30, 0.42]), representing a 14% reduction from the unadjusted effect. While this adjustment is meaningful, the adjusted effect remains statistically significant and educationally relevant, suggesting that the primary conclusion regarding flipped classroom effectiveness remains valid even after accounting for potential publication bias.

These findings indicate that while publication bias should be considered when interpreting the meta-analytic results, the core finding of positive flipped classroom effects appears reasonably robust to this potential distortion. Educational stakeholders can proceed with moderate confidence in the reported effects, while remaining appropriately cautious about potential overestimation of the true effect magnitude.

DISCUSSION

This meta-analysis provides comprehensive evidence regarding the effectiveness of flipped classroom approaches across educational contexts, offering significant insights for educational stakeholders. The findings demonstrate consistent positive effects on both student engagement and academic achievement, while identifying important moderating variables that influence these outcomes.

Our results revealed a medium positive effect of flipped classroom approaches on student engagement (g = 0.57) and a small-to-medium effect on academic achievement (g = 0.42). These findings align with previous meta-analyses by Tutal (2021) who similarly found moderate effects on achievement and learning retention. The differential between engagement and achievement effects supports the theoretical premise that engagement functions as a mediating factor in the learning process, a pattern consistent with Al-Samarraie et al.'s (2020) conceptualization of flipped classroom benefits.

The decomposition of engagement effects by type revealed stronger impacts on behavioral (g = 0.63) and cognitive engagement (g = 0.59) compared to emotional engagement (g = 0.48). This pattern is consistent with Jang and Kim's (2020) findings that flipped approaches particularly strengthen observable participation and intellectual investment in learning. The relatively weaker effect on emotional engagement suggests that affective dimensions may develop more gradually or require additional support strategies, aligning with Tormey & Henchy (2008) emphasis on the importance of student emotional readiness in flipped implementations.

The differential effectiveness across subject domains, with stronger effects in STEM (g = 0.49) and professional disciplines (g = 0.46) compared to social sciences (g = 0.38) and humanities (g = 0.35), supports Cheng et al.'s (2019) identification of subject area as a significant moderator. Our findings extend Mawardi et al.'s (2024) observation of particularly strong effects in mathematics, suggesting that disciplines with procedural knowledge components may especially benefit from the flipped approach. The comparatively lower effects in humanities align with Jang and Kim's (2020) observation that flipped classrooms may be less transformative in discussion-based disciplines where traditional instruction already incorporates active learning elements.

The inverse relationship between educational level and effect size—with primary education showing the strongest effects (g = 0.58), followed by secondary (g = 0.47) and higher education (g = 0.39)—corroborates Tutal and Yazar's (2021) finding that flipped approaches appear more effective in primary school settings. However, our analysis extends beyond their work by quantifying these differences across all three educational levels. This pattern may reflect Li et al.'s (n.d.) identification of educational context as a key moderator, potentially related to developmental factors in learning or baseline differences in traditional instructional effectiveness across educational levels.

Our findings regarding implementation duration (r = 0.42) and quality provide critical insights into practical considerations for educational stakeholders. The positive correlation between implementation duration and effectiveness supports Wagner et al.'s (2024) emphasis on careful instructional design and sufficient implementation time. Similarly, the substantial effect size differentials between high-quality (g = 0.68 for engagement; g = 0.53 for achievement) and low-quality implementations (g = 0.32 for engagement; g = 0.24 for achievement) underscore the importance of comprehensive implementation planning and support structures, consistent with Al-Samarraie et al.'s (2020) discussion of implementation challenges.

These findings have significant theoretical implications, supporting the constructivist learning principles underlying flipped approaches. The stronger effects on behavioral and cognitive engagement compared to emotional engagement suggest differential impacts across various dimensions of the learning experience, a pattern that extends Manlapig's (2024) analysis of engagement factors. The implementation quality findings particularly reinforce Shi et al.'s (2020) identification of pedagogical approaches as crucial moderators of effectiveness.

For practitioners, our results provide evidence-based guidance for flipped classroom implementation. The educational level findings suggest that while flipped approaches are effective across all levels, they may yield particularly strong returns in primary and secondary settings. The subject domain analysis indicates that STEM and professional courses may be priority areas for implementation. Most critically, the implementation quality findings demonstrate that institutional investment in implementation quality represents a primary determinant of success, with potentially greater impact than many other instructional variables.

While our analysis identified some publication bias through funnel plot asymmetry, the trim-and-fill adjusted effect size (g = 0.36) remained statistically significant and educationally meaningful, suggesting that the core finding of positive flipped classroom effects is reasonably robust. Nevertheless, this potential bias should be considered when interpreting the results, particularly regarding the magnitude of effects.

Future research should explore the interaction between identified moderators, such as how implementation quality might differentially impact outcomes across educational levels or subject domains. Additionally, more granular analysis of specific implementation components could provide valuable guidance on optimal flipped classroom designs for various educational contexts..

CONCLUSION

This meta-analysis provides comprehensive evidence of the efficacy of flipped classroom models across diverse educational contexts. Our findings reveal consistently positive effects on both student engagement (g = 0.57) and academic achievement (g = 0.42), confirming the value of this pedagogical approach. The stronger effect on engagement compared to achievement suggests that enhanced engagement functions as a mediating factor in the learning process, ultimately contributing to improved academic outcomes.

The decomposition of engagement effects revealed stronger impacts on behavioral (g = 0.63) and cognitive engagement (g = 0.59) compared to emotional engagement (g = 0.48), indicating that flipped classrooms particularly excel at promoting active participation and intellectual investment in learning. The differential effectiveness across subject domains, with stronger effects in STEM (g = 0.49) and professional disciplines (g = 0.46), suggests that subjects with procedural knowledge components particularly benefit from the flipped approach.

The inverse relationship between educational level and effect size—with primary education showing the strongest effects (g = 0.58), followed by secondary (g = 0.47) and higher education (g = 0.39)—indicates that younger students may derive particularly pronounced benefits from the structured approach inherent in flipped classroom methodologies. Our findings regarding implementation duration (r = 0.42) and quality underscore the importance of sustained, high-quality implementation for maximizing effectiveness.

For educational stakeholders, these results provide clear, evidence-based guidance. While flipped classroom approaches offer benefits across all educational contexts, implementation priorities might reasonably focus on primary and secondary education settings, where the largest gains are likely to be realized. The substantial effect size differentials between high-quality and low-quality implementations emphasize that institutional investment in comprehensive implementation planning, adequate instructor training, and ongoing support systems represents a critical determinant of success.

While our analysis identified some publication bias, the adjusted effect size (g = 0.36) remained statistically significant and educationally meaningful, suggesting that the core finding of positive flipped classroom effects is reasonably robust. Future research should explore interaction effects between identified moderators and investigate specific implementation components to develop more precise guidelines for optimizing flipped classroom designs across various educational contexts.

REFERENCES

Agyeman, N. Y. B., & Aphane, V. (2024). Implementing Flipped Classroom to Enhance Student Engagement: An Action Research. *E-Journal of Humanities, Art and Social Sciences*, 1860–1878. https://doi.org/10.38159/ehass.202451119

Al-Samarraie, H., Shamsuddin, A., & Alzahrani, A. I. (2020). A flipped classroom model in higher education: a review of the evidence across disciplines. *Educational Technology Research and Development*, *68*(3), 1017–

1051. https://doi.org/10.1007/s11423-019-09718-8

Annury, M. N., Sutrisno, D., Mariam, S., & ... (2023). The Impact of Reading Strategy Instruction Based on QAR with Visually-Oriented Learners. In *PAROLE: Journal of ….* eprints.uad.ac.id. https://eprints.uad.ac.id/78463/1/51362-216507-1-PB.pdf

Bormann, J. (2014). *Affordances of flipped learning and its effects on student engagement and achievement*. https://scholarworks.uni.edu/cgi/viewcontent.cgi?article=1128&context=grp

Celik, T., Akay, C., & Kanadlı, S. (n.d.). Revisiting flipped class model effectiveness on learning performance: A meta-analysis of meta analyses. *Journal on School Educational Technology*. https://doi.org/10.26634/jsch.19.2.20186

Cheng, L., Ritzhaupt, A. D., & Antonenko, P. (2019). Effects of the flipped classroom instructional strategy on students' learning outcomes: a meta-analysis. *Educational Technology Research and Development*, *67*(4), 793–824. https://doi.org/10.1007/s11423-018-9633-7

Cuban, L. (1987). Cultures of Teaching: A Puzzle. *Educational Administration Quarterly*, *23*(4), 25–35. https://doi.org/10.1177/0013161X87023004004

Gamo, J. A. (2022). Anatomy Education-Paradigm Shift from Passive to Active Learning-Effects on Student Engagement, Comprehension and Retention A Review of Literature from 2012 to 2022. *The FASEB Journal*, *36*(S1). https://doi.org/10.1096/fasebj.2022.36.s1.l7949

Jang, H. Y., & Kim, H. J. (2020). A Meta-Analysis of the Cognitive, Affective, and Interpersonal Outcomes of Flipped Classrooms in Higher Education. *Education Sciences*, *10*(4), 115. https://doi.org/10.3390/EDUCSCI10040115

Jantakoon, T., Jantakun, T., Jantakun, T., Noibuddee, A., Pasmala, R., Wannapiroon, P., & Nilsook, P. (2024). The Effectiveness of Flipped Classroom in English Language Learning: A Meta-Analysis. *World Journal of English Language*, *15*(3), 50. https://doi.org/10.5430/wjel.v15n3p50

Jones, K. E. (2010). Introduction to the Special Focus: Experiential Education. *Christian Education Journal Research on Educational Ministry*, *7*(1), 103–106. https://doi.org/10.1177/073989131000700107

Li, S., Fu, W., Liu, X., & Hwang, G.-J. (n.d.). *Effectiveness of Flipped Classrooms for K–12 Students: Evidence from a Three-Level Meta-Analysis*. https://doi.org/10.3102/00346543241261732

Lockwood, K., & Esselstein, R. M. (2013). The inverted classroom and the CS curriculum. *Technical Symposium on Computer Science Education*, 113–118. https://doi.org/10.1145/2445196.2445236

Manlapig, E. F. (2024). Effectiveness of Flipped Classroom Model (FCM) on Students' Learning Achievement in Physical Science: A Meta-Analysis. *International Journal of Research Publication and Reviews*, 4747–4754. https://doi.org/10.55248/gengpi.5.0624.1572

Mawardi, D. N., Sulistyowati, E., & Hukom, J. (2024). Meta-Analisis investigasi model kelas terbalik pada keterampilan berpikir tingkat tinggi (HOTS) siswa matematika: Analisis efek gabungan dan heterogenitas. *Jurnal Math Educator Nusantara*, *10*(1), 154–166. https://doi.org/10.29407/jmen.v10i1.22296

McLaughlin, J. E. (2018). Flipped classrooms, by design. *Medical Education*, *52*(9), 887–888. https://doi.org/10.1111/MEDU.13654

Mera, E. A. R., & Mera, K. D. R. (2024). Flipped Classroom como Elemento Favorecedor para el Desarrollo de la Capacidad de Resolución de Problemas en estudiantes de Educación Infantil. *Ciencia Latina*, *8*(4), 10197–10215. https://doi.org/10.37811/cl_rcm.v8i4.13165

Pujiriyanto, P., Handaru, C. D., & Hukom, J. (2024). Meta-Analysis of The Effectiveness of The Flipped Classroom Model On Students' HOTS in Mathematics. *Jurnal Teknologi Pendidikan*, *25*(3), 569–583. https://doi.org/10.21009/jtp.v25i3.45848

Ralević, L., & Tomašević, B. (n.d.). *Comparing the Effectiveness of the Flipped Classroom Model and the Traditional Instruction Model: A Meta-Analysis*. https://doi.org/10.5937/nasvas2103301r

Shahnama, M., Ghonsooly, B., & Shirvan, M. E. (2021). A meta-analysis of relative effectiveness of flipped learning in English as second/foreign language research. *Educational Technology Research and Development*, *69*(3), 1355–1386. https://doi.org/10.1007/S11423-021-

09996-1
Shi, Y., Ma, Y., MacLeod, J., & Yang, H. H. (2020). College students' cognitive learning outcomes in flipped classroom instruction: a meta-analysis of the empirical literature. *Journal of Computers in Education*, *7*(1), 79–103. https://doi.org/10.1007/s40692-019-00142-8
Sutrisno, D., Abbas, A., & Annury, M. N. (2024). Enhancing Writing Skills through Cultural Integration: Exploring the Impact of Culturally-Responsive Writing Instruction in Diverse EFL Classrooms. *Global Synthesis in Education …*. https://gse-journal.net/index.php/gse/article/view/47
Sutrisno, D., Martina, M., Karsana, D., Nafi Annury, M., Adri, A., Damayanti, W., Hidayah, A. M. N., & Juliastuty, D. (2024). Semiotic Analysis of Psycholinguistic Strategies in Indonesian President Candidates' Debates 2024: Unraveling Linguistic Signifiers and Mental Processes in Argumentative Discourse. *Forum for Linguistic Studies*, *6*(5), 943–976. https://doi.org/10.30564/fls.v6i5.7251
Tormey, R., & Henchy, D. (2008). Re-imagining the traditional lecture: an action research approach to teaching student teachers to 'do' philosophy. *Teaching in Higher Education*, *13*(3), 303–314. https://doi.org/10.1080/13562510802045337
Tutal, Ö., & Yazar, T. (2021). Flipped classroom improves academic achievement, learning retention and attitude towards course: a meta-analysis. *Asia Pacific Education Review*, *22*(4), 1–19. https://doi.org/10.1007/S12564-021-09706-9
Veen, B. D. Van. (2013). Flipping Signal-Processing Instruction [SP Education]. *IEEE Signal Processing Magazine*, *30*(6), 145–150. https://doi.org/10.1109/MSP.2013.2276650
Wagner, M., Gegenfurtner, A., & Urhahne, D. (2024). *Effectiveness of Flipped Classrooms*. 225–240. https://doi.org/10.4324/9781003386131-22
Yair, G. (2000). Educational Battlefields in America: The Tug-of-War over Students' Engagement with Instruction. *Sociology Of Education*, *73*(4), 247–269. https://doi.org/10.2307/2673233