



# Exploring the Effectiveness of Flipped Classroom on STEM Student Achievement: A Meta-analysis

Jin Gong<sup>1</sup> · Sanfa Cai<sup>2</sup> · Ming Cheng<sup>3</sup>

Accepted: 22 October 2023 / Published online: 15 November 2023  
© The Author(s), under exclusive licence to Springer Nature B.V. 2023

## Abstract

A ‘flipped classroom’ creates a dynamic and effective learning environment where students learn basic knowledge at home and engage in active learning activities during the class. It has been widely used across different STEM (Science, Technology, Environment, Mathematics) disciplines. However, there is a lack of systematic and quantitative assessments of the effectiveness of the flipped classroom model on student performance. In this study, a meta-analysis was conducted with 53 research articles on the flipped classroom in STEM higher education. It showed an overall significant effect in favor of flipped learning over traditional learning models for STEM education (Hedges’  $g=0.263$ , 95% [0.190, 0.337],  $Z=7.03$ ,  $p<0.0001$ ), with no evidence of publication bias. Moderator analysis showed that student performance was associated with the implementation strategies of flipped learning and the research control methods. Testing of the moderating factors *link activity*, *pre-class activity*, *in-class activity*, and *post-class activity* suggests that link activities to gauge home learning at the beginning of class, a mixture of group activities and individual activities during class and setting a combination of quizzes and exercises after class, all serve to improve student achievement in STEM education.

**Keywords** Flipped classroom · Meta-analysis · STEM education · Student performance

---

✉ Jin Gong  
gong35597@gmail.com

Sanfa Cai  
csf@tongji.edu.cn

Ming Cheng  
chengming9934@hotmail.com

<sup>1</sup> School of Economics and Management, Tongji University, Tongji Building A, Siping Road 1500, Shanghai 200092, China

<sup>2</sup> Institute of Higher Education, Tongji University, 1239 Siping Road, Shanghai 200092, China

<sup>3</sup> Sheffield Institute of Education, College of Social Sciences and Arts, Sheffield Hallam University, City Campus, Howard Street, Sheffield S1 1WB, UK

## 1 Introduction

Global recognition of the increasing importance of STEM education has led to various interpretations of its meaning. STEM education can be understood as broadly encompassing all disciplines within science, mathematics, technology, and engineering. It can also refer to combining individual STEM subjects in an interdisciplinary manner (Li et al., 2020a, 2020b). For our study, we specifically refer to STEM education as science, mathematics, technology, and engineering.

Learning and mastering the STEM disciplines presents several challenges due to their complex, abstract, and multi-dimensional nature (Corredor et al., 2014; Sedig, 2008). Several methods of learning are believed to have significant potential for positively impacting students' learning outcomes and attitudes in STEM education, among which game-based learning, blended learning, and flipped classrooms are the most popular approaches. Game-based learning, which draws on students' sense of playfulness, can increase students' motivation to engage in learning activities (Alt, 2023) and to master knowledge (Chang et al., 2022). Blended learning recruits digital technology and combines the advantages of traditional digital educational tools (digital video support, projectors, lecture notes, and dictionaries, etc.) and more high-tech digital and online learning tools (interactive boards, scientific software, etc.) (Graham, 2013; Means et al., 2013; Moskal et al., 2013; Mu et al., 2023; Müller & Mildenerger, 2021). Five models in the blended classroom have been identified: the face-to-face instructor-led model, face-to-face collaboration model, online instructor-led model, online collaboration model, and online self-paced model (Alammary, 2019).

The flipped classroom approach, also known as an inverted classroom or flipped learning, occurs when students engage in content learning (e.g., watching online video lectures) before class and then use the in-class time to actively use and apply the learning and do instructor-guided exercises (e.g., engaging in group activities or presentations). This learning approach, first introduced by Lage et al. (2000), has been integrated into a wide variety of disciplines due to its advantages in developing student collaboration, critical thinking skills, and interpersonal skills (Hawks, 2014; National Academies of Sciences & Medicine, 2017).

Many courses in STEM disciplines require students to learn complicated principles and concepts before engaging in more practical and authentic applications (Huber, 2016). (Wong et al., 2014) introduced the flipped classroom into pharmaceutical education and found it significantly improved students' performance and perceptions. In courses that require practical applications (e.g., gross human anatomy), students can learn about the procedures and anatomical structures before class and spend class time using their knowledge in the clinical applications (Fleagle et al., 2018).

However, not all research implementing the flipped classroom has produced positive results. Some studies (Chiquito et al., 2020; McCabe et al., 2017; Muzyk et al., 2015; Senousy et al., 2017) found that the flipped classroom in STEM is less effective than the traditional classroom model in improving student grades.

There is increasing interest in using meta-analyses to explore the effects of the flipped classroom. Meta-analysis is an effective method for combining and summarizing conflicting mixed research results (Higgins et al., 2003). The advantages of meta-analysis include its potential to find relationships across studies and to protect against over-interpreting differences across studies (Lipsey & Wilson, 2001). Meta-analyses focusing on different disciplines have revealed that using a flipped classroom approach is associated

with better learning performance in the mathematics (Algarni, 2018; Lo et al., 2017), health professions education (Hew & Lo, 2018a), endodontic education (Nagendrababu et al., 2019), nursing (Li et al., 2020a, 2020b; Xu et al., 2019), language learning (Chen et al., 2020; Kubra & Lee, 2020; Vitta & Al-Hoorie, 2020) and chemistry (Rahman & Lewis, 2020).

(Freeman et al., 2014) conducted a meta-analysis of 225 published research articles in STEM education and found that use of the active learning can significantly increase examination performance and decrease the failure rate. While it represents the most comprehensive meta-analysis of STEM education to date, there remains uncertainty regarding whether the benefits of active learning on student performance extend to the flipped classroom approach. (Hew et al., 2020a, 2020b) investigated the effect of the flipped classroom on student cognitive and behavioral outcomes across STEM disciplines, but it used a second-order meta-analysis and did not cover original quantitative research studies. Despite the work of both these studies (Freeman et al., 2014; Hew et al., 2020a, 2020b), the influence of the flipped classroom on student outcomes and the key factors determining the success of the approach has yet to be well established.

Therefore, this paper reports the conduct of a meta-analysis to examine the effect of the flipped classroom on student learning outcomes compared with the traditional classroom in STEM disciplines. Traditional classroom here refers to the environment in which teachers use class time to teach the course materials, and students receive the instruction in the class and complete homework and practice exercises after class (Gonzalez-Gomez et al., 2016; Lin, 2019).

Specific conditions under which flipped classrooms are more or less effective than traditional classrooms will be identified. The moderator variables were chosen by examining the procedures of past studies on the flipped classroom and reviewing previous publications and meta-analyses conducted in the flipped learning context. The aim was to identify the variables commonly employed in flipped learning studies and to investigate how the difference in effect sizes between studies can be explained.

In short, this study will address two main research questions:

1. What is the effect of the flipped classroom on student achievement compared with the traditional classroom in STEM education?
2. What is the effect of each moderator variable on student learning outcomes in STEM education?

## 2 Moderator Variables

### 2.1 Education Context Characteristics

The following moderator variables were selected as the education context characteristics: *discipline*, *class level*, and *study duration*. Although many studies in different fields have investigated the possible effects of the flipped classroom, there is substantial variation across their results. Some researchers found that the applied STEM disciplines such as engineering and nursing were particularly suitable for the flipped classroom (Hu et al., 2018; Karabulut-Ilgu et al., 2018). However, (Cheng et al., 2019a) found that engineering courses saw less improvement in student achievement through the flipped classroom than in other disciplines, including the arts and humanities, natural and social sciences, and

mathematics. This study, therefore, examines the possible influence of the *discipline* in which the flipped classroom is applied, with different STEM disciplines as categories.

Moreover, we included *class level* in the moderator analysis to see whether flipped classrooms are more effective in upper-level or graduate courses than introductory courses, as proposed by (Bredow et al., 2021). In introductory lectures, the teaching goal is to transform students' knowledge, while in upper-level courses, the primary purpose of taught classes is for students to apply and transfer previously learned knowledge into practical use in different contexts. Students in upper-level courses may tend to be better at self-regulated learning, enabling them to benefit more from the flipped classroom (Wigfield et al., 2013).

Lastly, *study duration* was considered as a moderator variable. Only one article included intervention duration as a moderator variable and found no effect on (Cheng et al., 2019a). However, given that instructors and students both have to manage the difficulties in the transition to the flipped classroom due to unfamiliarity with the approach (Anderson & Brennan, 2015; DeSantis et al., 2015), it was considered that study duration might moderate the results if instructors and students need a longer time to adjust.

## 2.2 Levels of Control

Studies with a rigorous experimental study design are likely to yield significantly more reliable results than studies that do not employ good design (Cheung & Slavin, 2016). The control condition for study groups is one of the most critical design factors. Therefore, *group equivalence* and *instructor equivalence* were included in this study as moderator variables to investigate if levels of control influence the effects of the flipped classroom. Concerning group equivalence, the meta-analysis by (Lag & Saele, 2019) showed positive effect sizes for flipped classroom studies with randomly allocated groups, while other researchers (Bredow et al., 2021; Hew & Lo, 2018a) found no significant differences. In terms of instructor equivalence, the meta-analysis by (Strelan et al., 2020) found that the flipped classroom had a more substantial effect on student performance when the flipped classroom and traditional classroom had different instructors, compared with studies where the instructor was the same for both groups.

## 2.3 Course Design Characteristics

The following moderator variables concerning course design characteristics were selected: *pre-class activity*, *link activity*, *in-class activity*, and *post-class activity*. Video lectures and reading materials are frequently implemented in pre-class activities to prepare students for active learning during class. There is a belief that video lectures are distinct from other forms of pre-class content materials because they tap into both visual and auditory modes of information processing (Bredow et al., 2021; Jensen et al., 2018). However, most student learning involves the use of text-based materials such as textbooks (Besser et al., 1998), which offer quick and straightforward content delivery that is easy to skim and readily searchable, unlike a video lecture. With different types of materials having their unique advantages, it is possible that the delivery of different combinations of pre-class materials can have varying effects on student learning and retention.

*Link activity* allows students to make connections between the lesson at home and the in-class activity. It has three purposes: (1) to check the degree of assimilation of the concepts in activities before class; (2) to get students actively engaged in the pre-class learning process; (3) to detect deficits in learning materials. Previous meta-analyses

have found that the use of quizzes before in-class activity positively affected learning outcomes as it can help students recall knowledge learned at home (Dirkx et al., 2014; Hew & Lo, 2018a; Roediger & Karpicke, 2006). Some articles also used a questionnaire or treated the instructor's interpretation of students' misunderstandings about pre-class content materials as a link activity (Castedo et al., 2019; Chiquito et al., 2020).

*In-class activity* often includes individual tasks and group activities that involve active, constructive, and interactive engagement. (Lo et al., 2017) provided a general summary of characteristics of in-class activities used in their reviewed studies but did not conduct moderator analysis on these characteristics. Only one article (Lo & Hew, 2019) included in-class activity as a moderator variable and found that small-group activities helped to enhance student performance, but this study was limited to the engineering field.

Finally, *post-class activity* is considered as a moderator variable. There is reason to believe that different post-class activities applied in the flipped classroom might explain some of the variance in effects between studies. One of the ultimate goals of learning is storing knowledge in one's memory: post-class assignments including quizzes, surveys, and reflection exercises can help students reflect on and remember previously learned content. To date, only one meta-analysis (Lo & Hew, 2019), confined to engineering, used the post-class activity as a moderator variable and found no differences in the effect sizes.

## 3 Methodology

### 3.1 Search Strategies and Data Sources

A keyword search was conducted in the following five electronic databases: PubMed, Scopus, ERIC, Web of Science, and Cochrane, covering education research from STEM disciplines. The literature selection process followed the Preferred Reporting of Items for Systematic Reviews and Meta-analyses (PRISMA) Statement (Moher et al., 2009). The key search terms used in the electronic databases were: "flipped classroom" or "flipped class" or "flipped teaching" or "flipped learning" or "flipped instruction" or "flipped lecture" or "flipped course".

### 3.2 Inclusion and Exclusion Criteria

The inclusion and exclusion criteria used to select studies were based on recent flipped classroom reviews (Table 1). We selected empirical studies from January 2010 to July 2023, because very few research studies on this subject were published before 2010. Only peer-reviewed journal studies were included, as their results are believed to be of sufficient quality (Cheng et al., 2019a; Hew & Lo, 2018a; Korpershoek et al., 2016; Lo et al., 2017). The education context in our review was limited to higher education (undergraduate and postgraduate courses) (Shi et al., 2020a), and the subject areas were confined to disciplines in STEM. Moreover, eligible research must have used designs that enabled the comparison of learning outcomes of students in the flipped classroom with those in the traditional classroom, such as quasi-experimental design, historical cohort-controlled research design, and randomized controlled trials (Hew & Lo, 2018a; Shi et al., 2020b).

**Table 1** Inclusion and exclusion criteria for selection

Criteria	Inclusion	Exclusion
Time period	January 2010 to July 2023	Studies published outside this period
Language	Studies reported in English	Studies reported in languages other than English
Article type	Peer-reviewed studies	Studies that are not peer-reviewed
Subject area	All disciplines in STEM	All other disciplines not in STEM
Education Level	Higher education	K-12 education
Statistical Information	Adequate data for calculating the effect size, which includes N, M and SD, or N and t, for both control group and experiment group	No sufficient data
Educational Outcomes	Objective quantitative measures of performance such as post-tests, final tests, or other standardized tests	Student survey of self-perceived learning (e.g. interview or student satisfaction questionnaire)
Research design	Studies measuring student performance in the flipped classroom (experiment group) and traditional classroom (control group)	Studies without a traditional class for comparison

### 3.3 Data Extraction

The following data were extracted from each selected article: title of the study, author's name, publication year, location in which the study was conducted, discipline, class level, study duration, group equivalence, instructor equivalence, delivery type, and details of the experimental implementation including the types of pre-class, in-class, and post-class activities. To calculate effect sizes, we recorded the *M* (mean), *SD* (standard deviation), and *N* (sample size) for the control group and experimental group. In the case of articles that did not provide enough data for calculating effect sizes or gave vague statements about the pre-class, in-class, and post-class interventions, we contacted the original study researchers to get the required details. Two authors independently extracted the data, and any discrepancies between their extracted data were reviewed, discussed, and resolved.

### 3.4 Meta-analysis

We used the random-effects model (Gurevitch & Hedges, 1999) to compute effect sizes using the metafor package for the software program R 4.1.2. All reported *p* values are two-tailed unless otherwise reported. Hedges' *g* is useful in adjusting small sample size bias, so we computed effect sizes as Hedges' *g* from the means and standard deviations of the student achievement data, such as exam scores (Hedges, 1981). If the empirical studies used the standard errors, not the standard deviation, we used the following formula to compute the standard deviation (Hedges, 1982):

$$SE = \frac{SD}{\sqrt{\text{sample size}}}$$

If there were more than one test or exam to estimate student performance, the most appropriate results were chosen, such as final exams instead of mid-term exams. If the empirical studies divided students into groups based on irrelevant categories such as gender, the results were combined into one group (DerSimonian & Laird, 1986).

### 3.5 Heterogeneity Analysis and Publication Bias

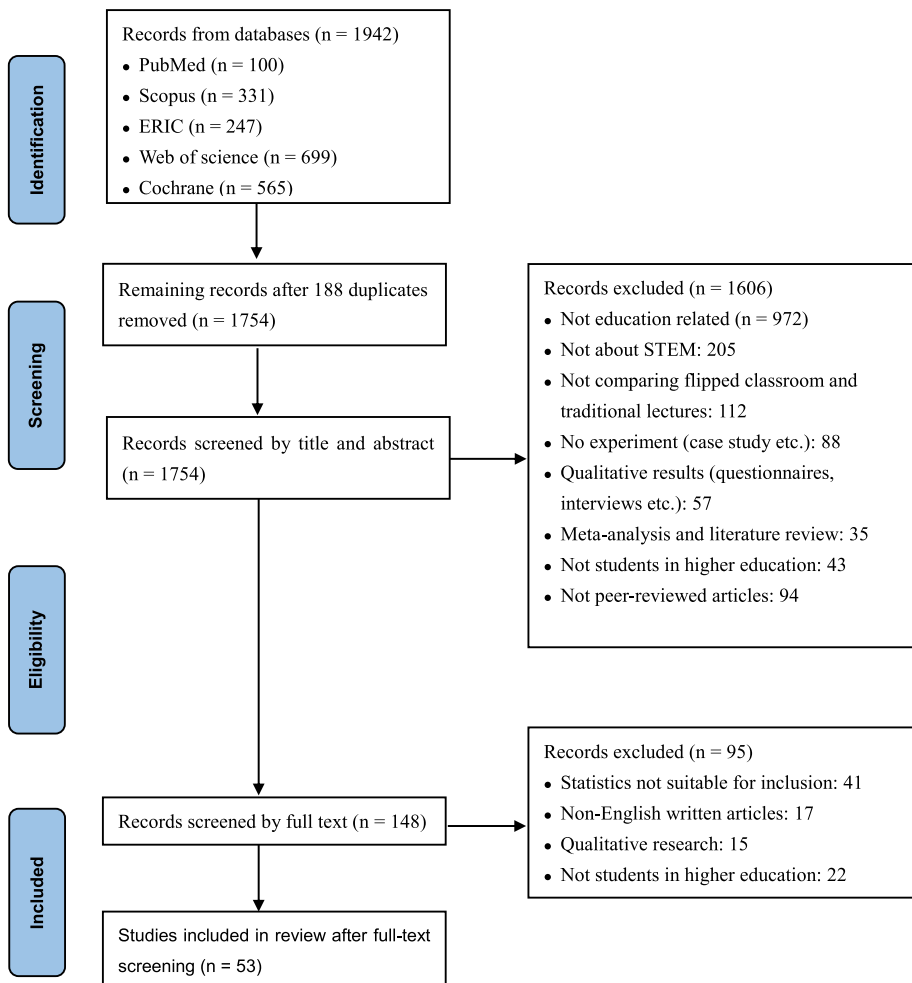
To determine the level of heterogeneity in the 53 articles included in this review, we used the *Q* statistics (Borenstein et al., 2010) and the *I*<sup>2</sup> test (Borenstein et al., 2009). The *I*<sup>2</sup> is the percentage of variation across studies attributable to heterogeneity. Values of *I*<sup>2</sup> lower than 25% indicate low heterogeneity, values between 25 and 50% indicate moderate heterogeneity and values higher than 50% indicate high heterogeneity (Higgins et al., 2003).

Publication bias occurs when the nature of the results makes it more likely that the study will be published (Peplow, 2014). Therefore, the following standard tests were used to identify whether the present review suffered from publication bias, including: (1) assessing the funnel plot (Light & Pillemer, 1986); (2) computing the Begg & Mazumdar rank correlation (Begg & Mazumdar, 1994); (3) calculating Egger's linear regression (Egger et al., 1997); (4) conducting the classic fail-safe *N* test (Rosenthal, 1979).

## 4 Results

### 4.1 Study Selection

Figure 1 shows the PRISMA flow diagram of the detailed literature search process, setting out the number of articles screened, excluded, and finally included in the review. In the first stage of our search process, a total of 1942 articles were identified from five databases, and after removing 188 duplicates, 1754 articles remained. After reviewing their titles and abstracts during the second phase, many articles were excluded for not meeting our inclusion criteria. We carefully examined whether research articles compared the flipped classroom and the traditional classroom, paying particular attention to terms such as comparison, traditional class, control group and experiment, and historical performance. At the end



**Fig. 1** PRISMA flowchart showing how studies of flipped learning were identified and screened for the meta-analysis



of this stage, 1606 articles were excluded, leaving 148 articles for further consideration. During the third phase, full-text articles were carefully read and assessed for eligibility, especially focusing on the methodological design and data provided by the articles. We removed articles that did not provide adequate data for calculating effect sizes. Articles that did not give a detailed description of the research process were also excluded. In the end, the literature selection process yielded a total of 53 articles for computing effect sizes and further analysis.

### 4.2 The Overall Effect of the Flipped Classroom

As Fig. 2 indicates, a random-effects meta-analysis of the 53 included articles involving 3740 students exposed to flipped classrooms and 3793 students exposed to traditional

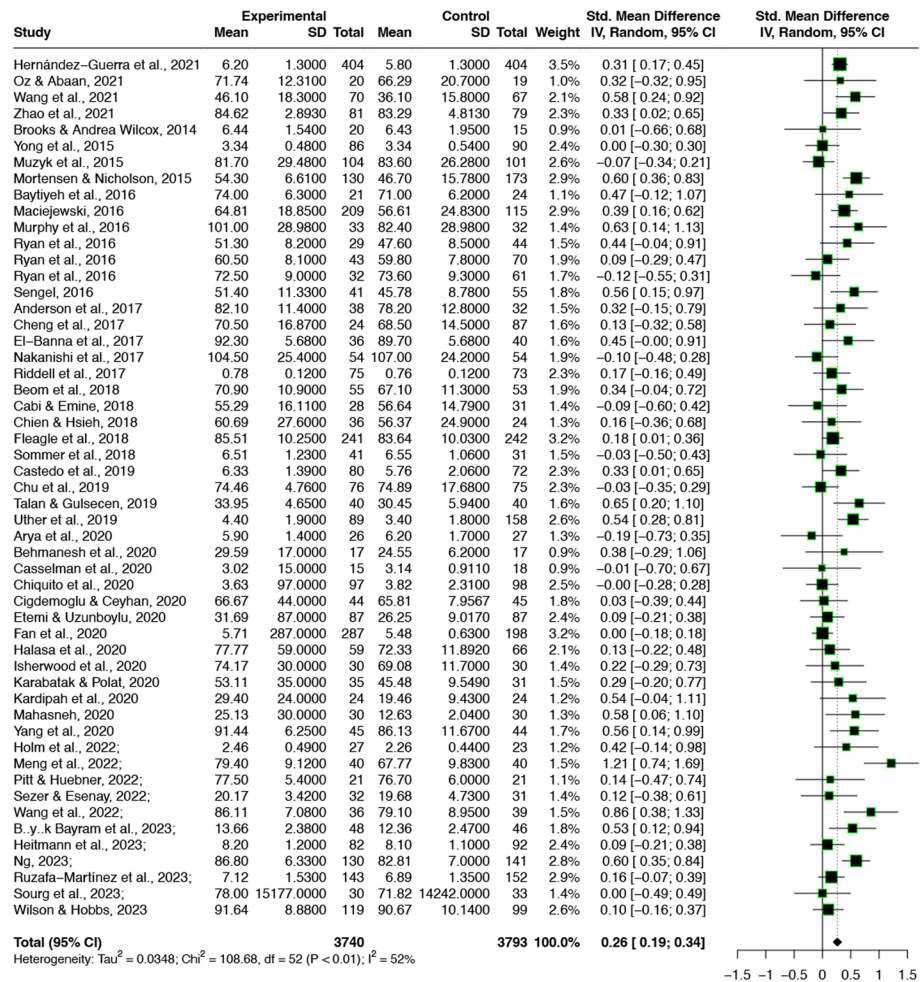


Fig. 2 Forest plot of effect sizes (Hedges' g) using the random effect size model showing the distribution of 53 studies' effect sizes

classrooms showed an overall significant effect in favor of the flipped classroom in terms of student performance (Hedges'  $g=0.263$ , 95% [0.190, 0.337],  $Z=7.03$ ,  $p<0.0001$ ). Nine of the studies had negative effect sizes for the flipped classroom. Forty-four studies were in favor of the flipped classroom condition but only 17 of these had statistically significant effect sizes. The heterogeneity analyses showed statistically significant variation among the 53 studies ( $Q=108.68$ ,  $df=52$ ,  $p=0.0001$ ), and the level of observed heterogeneity was moderate ( $I^2=52.2\%$ ). Subgroup analyses were conducted using the random-effects model on nine categorical moderator variables to identify possible sources of variation in the effect sizes.

### 4.3 Moderator Analysis

#### 4.3.1 Education Context Characteristics

The first set of moderator analyses focused on differences related to education context characteristics. The results indicated that the between-level difference was not statistically significant for *discipline* ( $Q=3.01$ ,  $df=9$ ,  $p=0.964$ ), *class level* ( $Q=0.87$ ,  $df=1$ ,  $p=0.351$ ), and *study duration* ( $Q=0.71$ ,  $df=1$ ,  $p=0.398$ ), which suggested no evidence of heterogeneity across studies for these three moderators.

Table 2 summarizes effect size according to *discipline*. The flipped classroom tended to be beneficial to students regardless of discipline. There were only three significant effects: mathematics ( $g=0.316$ ), medical science ( $g=0.273$ ), and technology ( $g=0.210$ ). The results for medical science were based on 26 studies, while the results for mathematics and technology were based on only five and three studies respectively. The results for the other seven disciplines were based on no more than six studies, which indicated that more research was needed to increase confidence in these findings.

The distribution of effect size by *class level* is illustrated in Table 3. The flipped classroom was associated with better student performance across all class levels. The flipped classroom was likely to have a stronger effect on student performance if the class was taught at the introductory level ( $g=0.281$ ), compared with the upper level ( $g=0.216$ ).

**Table 2** Effect sizes by subgroup related to discipline on flipped classroom student performance

Discipline	Effect size and 95% confidence interval					Heterogeneity	
	N	Hedges' g	Lower limit	Upper limit	Tau	Q	I <sup>2</sup> (%)
Medical Science	26	0.273	0.158	0.388	0.226	67.370	62.90
Technology	3	0.210	0.059	0.361	0.000	1.050	0.00
Mathematics	5	0.316	0.107	0.525	0.149	6.440	37.90
General Science	2	0.295	-0.371	0.961	0.439	5.690	82.40
Engineering	3	0.301	-0.090	0.692	0.271	5.730	65.10
Chemistry	3	0.168	-0.131	0.467	0.077	1.940	0.00
Physics	1	0.088	-0.292	0.468	-	0.000	-
Computer Science	6	0.279	-0.011	0.569	0.271	11.710	57.30
Health Science	3	0.242	0.000	0.484	0.000	1.150	0.00
Programming	1	0.088	-0.210	0.385	-	0.000	-

**Table 3** Effect sizes by subgroup related to class level on flipped classroom student performance

Class level	Effect size and 95% confidence interval					Heterogeneity	
	N	Hedges' g	Lower limit	Upper limit	Tau	Q	I <sup>2</sup> (%)
introductory	37	0.281	0.186	0.376	0.209	83.46	56.90
upper-level	16	0.216	0.116	0.315	0.103	23.14	35.20

As shown in Table 4, the study duration of a semester or more represented the majority of studies ( $k=40$ ), with an overall trivial to weak effect size ( $g=0.242$ ). The study duration of less than a semester ( $k=13$ ) had a slightly stronger effect size ( $g=0.343$ ). When the study duration for flipped classrooms was less than a semester, students in the flipped classroom condition significantly outperformed students in the traditional classroom.

### 4.3.2 Levels of Control

The second set of moderator analyses focused on the control between the traditional and flipped classroom groups. Some reviewed articles were well-controlled (e.g., using the same instructor for both groups and conducting a pre-test to ensure both groups' students were at the same level), while other articles were less well-controlled or gave little information about control methods. Our review examined two control levels: group equivalence and instructor equivalence. Both group equivalence ( $Q=4.36$ ,  $df=2$ ,  $p=0.1130$ ) and instructor equivalence ( $Q=5.30$ ,  $df=2$ ,  $p=0.0705$ ) did not result in any difference in the effect sizes.

Tables 5 and 6 summarize effect sizes according to group equivalence and instructor equivalence. The flipped classroom was likely to have a weaker effect on student performance if the students were at equivalent levels across the traditional and flipped modes of teaching ( $g=0.237$ ), compared with when the students were not equal at the beginning ( $g=0.326$ ) or studies where this information was not given ( $g=0.375$ ). However, the group equivalence of 'not equal' and 'not known' represented only a small number of studies ( $k=1$  and  $k=6$ , respectively), so these results should be treated with caution. Studies using different instructors across the traditional and flipped classroom conditions yielded

**Table 4** Effect sizes by subgroup related to study duration on flipped classroom student performance

Study duration	Effect size and 95% confidence interval					Heterogeneity	
	N	Hedges' g	Lower limit	Upper limit	Tau	Q	I <sup>2</sup> (%)
A semester or more	40	0.242	0.169	0.314	0.142	64.43	39.50
Less than a semester	13	0.343	0.120	0.566	0.336	42.44	71.70

**Table 5** Effect sizes by subgroup related to group equivalence on flipped classroom student performance

Group equivalence	Effect size and 95% confidence interval					Heterogeneity	
	N	Hedges' g	Lower limit	Upper limit	Tau	Q	I <sup>2</sup> (%)
Not known	6	0.375	0.276	0.474	0.012	3.2	0.00
Equal	46	0.237	0.153	0.321	0.203	97.61	53.90
Not equal	1	0.326	0.006	0.647	–	0	–

**Table 6** Effect sizes by subgroup related to instructor equivalence on flipped classroom student performance

Instructor equivalence	Effect size and 95% confidence interval					Heterogeneity	
	N	Hedges' <i>g</i>	Lower limit	Upper limit	<i>Tau</i>	Q	<i>I</i> <sup>2</sup> (%)
Equal	28	0.227	0.116	0.338	0.222	65.53	58.80
Not known	22	0.270	0.164	0.375	0.155	35.67	41.10
Not equal	3	0.471	0.293	0.649	0.000	1.33	0.00

a moderate effect size ( $g=0.471$ ), although this result was based on only three studies. The flipped classroom had a weak effect on student performance where the instructor was the same between the two models ( $g=0.227$ ) or where this information was not reported ( $g=0.270$ ).

### 4.3.3 Course Design Characteristics

As shown in Table 7, the flipped classroom effect was slightly stronger when the study used a link activity (pre-class quiz or questionnaire) ( $g=0.233$ ), compared to when the study used no link activity ( $g=0.108$ ). The distribution of effect sizes by pre-class activity is illustrated in Table 8. The flipped classroom had a moderate effect on student performance when students used video lectures only ( $g=0.252$ ) or both video lectures and text-based materials ( $g=0.286$ ). The effect of text-based materials alone ( $k=2$ ) was non-significant ( $g=0.213$ ).

Table 9 provides the effect sizes by in-class activity for the flipped classroom versus the traditional classroom. The flipped classroom effect was stronger when students were engaged in both group activities and individual tasks ( $g=0.380$ ), compared to when they participated only in group activities ( $g=0.115$ ) or only in individual tasks ( $g=0.278$ ).

Table 10 provides the effect sizes broken down by post-class activity. We classified post-class engagement as exercises ( $k=15$ ), quizzes ( $k=6$ ), both ( $k=5$ ) and not known ( $k=15$ ). The flipped classroom effect was highest in studies using both quizzes and exercises

**Table 7** Effect sizes by subgroup related to link activity on flipped classroom student performance

Link activity	Effect size and 95% confidence interval					Heterogeneity	
	N	Hedges' <i>g</i>	Lower limit	Upper limit	<i>Tau</i>	Q	<i>I</i> <sup>2</sup> (%)
Yes	20	0.338	0.233	0.443	0.157	36.26	47.60
No	30	0.208	0.108	0.307	0.202	68.76	53.50

**Table 8** Effect sizes by subgroup related to pre-class activity on flipped classroom student performance

Pre-class activity	Effect size and 95% confidence interval					Heterogeneity	
	N	Hedges' <i>g</i>	Lower limit	Upper limit	<i>Tau</i>	Q	<i>I</i> <sup>2</sup> (%)
Both	23	0.286	0.158	0.413	0.239	61.27	64.10
Text-based materials	2	0.213	-0.082	0.507	0.110	1.34	25.50
Video lectures	28	0.252	0.156	0.348	0.163	46.01	41.30

**Table 9** Effect sizes by subgroup related to in-class activity on flipped classroom student performance

In-class activity	Effect size and 95% confidence interval					Heterogeneity	
	N	Hedges' g	Lower limit	Upper limit	Tau	Q	I <sup>2</sup> (%)
Both	24	0.380	0.273	0.487	0.179	45.99	50.00
Group activities	21	0.115	0.032	0.199	0.078	25.6	21.90
Individual tasks	8	0.278	0.096	0.460	0.161	10.95	36.10

**Table 10** Effect sizes by subgroup related to post-class activity on flipped classroom student performance

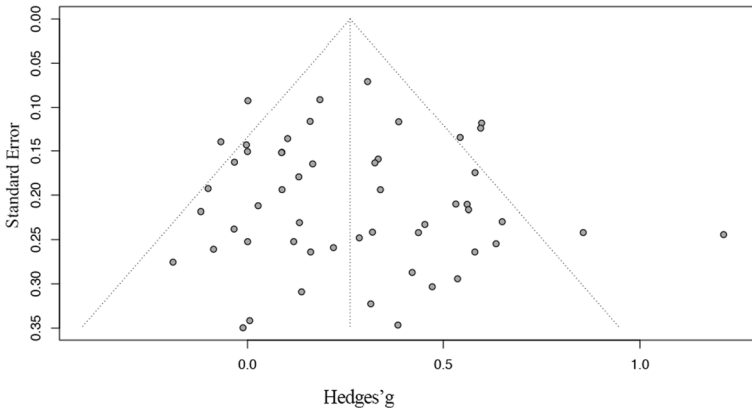
Post-class activity	Effect size and 95% confidence interval					Heterogeneity	
	N	Hedges' g	Lower limit	Upper limit	Tau	Q	I <sup>2</sup> (%)
Quizzes	6	0.275	0.088	0.462	0.141	7.8	35.90
Nothing	12	0.279	0.168	0.391	0.071	12.38	11.10
Both	5	0.578	0.226	0.929	0.333	12.78	68.70
Exercises	15	0.220	0.065	0.375	0.241	46.54	69.90

( $g=0.578$ ) compared with when only one of these activities was used after class. The flipped classroom had a moderate effect on student performance when students engaged in either exercises or quizzes after class ( $g=0.220$ ,  $g=0.275$ ).

The effects of between-level difference of the link activity ( $Q=3.13$ ,  $df=1$ ,  $p=0.077$ ), pre-class activity ( $Q=0.28$ ,  $df=2$ ,  $p=0.869$ ), and post-class activity ( $Q=4.39$ ,  $df=4$ ,  $p=0.3562$ ) were not statistically significant while the between-level difference was statistically significant for the in-class activity ( $Q=15.08$ ,  $df=2$ ,  $p=0.0005$ ). This heterogeneity analysis indicated that the effect size is significantly higher when the flipped classroom instructor employed both group activities and individual activities as in-class activities to help students master knowledge.

#### 4.4 Analysis of Publication Bias

The visual inspection of the funnel plot (Fig. 3) generated from the meta-analysis generally shows symmetrical distributions around the weighted mean effect sizes. We conducted two statistical analyses which supported the presence of symmetry: the Begg and Mazumdar rank correlation test ( $z=0.74$ , one-tailed  $p=0.4615$ ); and Egger's Linear regression test ( $t=0.62$ , one-tailed  $p=0.5350$ ). Both tests indicated that publication bias was not significant. We also conducted a classic fail-safe N test which showed that 1034 additional missing studies with zero mean effect size would be required to make the overall effect statistically insignificant. Based on the visual inspection of the funnel plot, the statistical analyses, and the fail-safe N test, it can be concluded that the overall mean effect size was not inflated by publication bias.



**Fig. 3** Funnel plot of effect sizes assessing publication bias

## 5 Discussion

### 5.1 Flipped Classroom Promotes Student Achievement

Overall, we found a small positive effect ( $g=0.26$ ) of the flipped classroom on assessed student learning outcomes. In other words, students in the flipped classroom outperformed their counterparts in the traditional classroom by 0.26 standard deviations. The significant small effect size is close to those reported in previous meta-analyses in mathematics ( $es=0.298$ ,  $k=21$ ) (Lo et al., 2017) and engineering ( $es=0.289$ ,  $k=29$ ) (Lo & Hew, 2019), but substantially smaller than those observed for nursing education ( $es=1.06$ ,  $k=11$ ) (Hu et al., 2018) and radiology education ( $es=1.12$ ,  $k=19$ ) (Ge et al., 2020). According to analysis standards for influences on student outcomes (Hattie, 2017), an effect size of 0.26 is comparable to those associated with student-focused interventions such as individualized instruction, matching style of learning, and student-centered teaching, as well as those associated with instructional strategies such as collaborative learning, discovery-based learning, and adjunct aids. The effect size we found is significantly larger than the effect sizes due to strategies emphasizing feedback, such as different types of testing, the learning hierarchies-based approach, as well as those for approaches using technologies such as one-on-one laptops and web-based learning. As illustrated in (van Alten et al., 2019), an effect size of 0.26 on learning performance may be regarded as small on the face of it, but in the context of education, it is more meaningful.

In our meta-analysis, there was an advantage of the flipped classroom in medical science, mathematics, and technology, but no significant effect was found in the other seven STEM disciplines analyzed, although one should be cautious about interpreting effect sizes based on small numbers of studies. According to some previous reviews (Chen et al., 2018; Hew & Lo, 2018a; Lo et al., 2017), the effect sizes were small under a random effect sizes model across three disciplines (engineering, mathematics, and health profession education). Our results were inconsistent with these findings, with the effect size of 0.316 for mathematics being the largest of the disciplines analyzed. It is possible that in STEM courses the advantages of the flipped classroom approach translate to only modest impacts on student performance partly because the teaching and learning is already well structured

and incorporates scaffolding for students. If this were not the case, the added value of the flipped classroom might be more significant.

After performing a subgroup comparison based on class level, the analysis revealed that there was no statistically significant difference between the summary effect sizes of the various groups. We found that where flipped classroom courses lasted less than a semester, the students performed better ( $g=0.343$ ) than those in flipped classroom courses extending for a semester or more ( $g=0.242$ ). Most of the studies in our meta-analysis were conducted within 8 weeks (Cabi & Emine, 2018; Karabatak & Polat, 2020; Meng et al., 2022; Muzyk et al., 2015; Senousy et al., 2017; Wang et al., 2022; Wilson & Hobbs, 2023), and one studies lasted for half a semester (Chien & Hsieh, 2018) and two studies only lasted for a few days (Arya et al., 2020; Casselman et al., 2020). For both shorter and longer durations of study intervention, students in the flipped classroom demonstrated a significant performance advantage over those in the traditional classroom. However, it was observed that the effect sizes of the flipped classroom intervention were smaller in studies with longer durations than those with shorter durations. This result was consistent with the findings of previous meta-analyses. In meta-analyses that compared studies that were shorter and longer than a semester (Cheng et al., 2019b; Shi et al., 2020a, 2020b) and compared studies that lasted 1–4, 5–8, 9–12, 13–16 weeks, and more than 16 weeks (Karagöl & Esen, 2019; Tatal & Yazar, 2021), the effects gradually decreased as the durations increased.

The majority of studies in our meta-analysis ( $k=46$ ) reported using a pre-test to ensure initial equivalence between the control and experimental groups. Consequently, the overall effect size remained relatively consistent ( $g=0.237$ ). However, we observed a slightly smaller flipped classroom effect ( $g=0.271$ ,  $k=28$ ) when the instructors were the same across both groups. This difference might be attributed to the instructors applying the skills utilized in the flipped classroom approach, such as motivating students towards more active learning, even within the traditional classroom setting, which could moderate the effects.

One beneficial aspect of the flipped classroom is that it allows students to study the content at their own pace. With pre-recorded videos, students can pause, review, and reflect as many times as they want before class. Our review found that the use of video lectures as pre-class learning materials yielded a stronger effect on student performance than the use of reading materials only. It can be concluded that video lectures are a better form of pre-class learning materials in STEM education. For example, in the software engineering education (Etemi & Uzunboylu, 2020), the MATLAB toolkit and design exercises can be better demonstrated by teachers directly in videos than through text-based materials.

Link activity refers to a variety of tasks assigned to students between the pre-class home-learning phases and the in-class face-to-face activities. These link tasks aim to verify that students have used the pre-recorded videos and/or reading materials, as well as to inform teachers about student misunderstandings. (Hew & Lo, 2018a) found in their meta-analysis that using quizzes at the start of face-to-face courses can help students recall knowledge and identify possible misconceptions and so improve learning outcomes. In our review, the flipped classroom studies that used link activities ( $g=0.338$ ) had a bigger influence on student performance than those that did not ( $g=0.208$ ). This finding is in line with previous meta-analyses (Hew & Lo, 2018b; Låg & Sæle, 2019; Lo & Hew, 2019; van Alten et al., 2019) showing that formative assessments (e.g., quizzes or reviews) used between pre-class activities and in-class activities can greatly enhance students' learning. These assessments are often followed by feedback from instructors during class (Ng, 2023; Sezer & Esenay, 2022). Given that feedback is one of the most potent factors affecting learning and achievement (Hattie & Timperley, 2007; Shute, 2008), we suggest that it



may play a crucial role in driving the effectiveness of the flipped classroom approach. It is plausible that feedback that students receive during the face-to-face session on their home activities serves as a fundamental mechanism of the flipped classroom's efficacy (Hew et al., 2021).

Concerning in-class learning activities, the results of our meta-analysis indicate that studies that combined individual activities and group activities have a higher average effect ( $g=0.380$ ) than studies that reported using only one of these activity types. This result is consistent with the study by (Lo & Hew, 2019) which also found that the combination of individual tasks and small-group activities produced a larger effect size. The effect may be due to interactions among students during small-group activities which can have additional cognitive benefits for their individual learning performance. It may be that students can acquire a deeper understanding of the subject through arguing, collaborating, and knowledge sharing during group activities, which can also enhance their learning when they study individually.

Similarly, concerning post-class activities, it was the combination of different tasks—quizzes, and exercises—which produced a larger effect size in our review. This finding aligns with a study by (Murphy et al., 2016) who required students to do online quizzes and hand in assigned homework problems from the textbook. The researchers found that this combined approach to post-class activity was an effective way for students to reflect on knowledge learned in class and to help instructors evaluate student learning performance and reflect on the flipped classroom model as a whole.

## 6 Limitations and Recommendations for Future Research

There are several limitations to this review that should be acknowledged. First, the descriptions of the interventions and course designs in the reviewed articles varied in quality and completeness, making them difficult to code accurately. Many articles only mentioned the use of small group activities or collaborative team-based learning but did not specify how the activities were organized and how long they lasted. In future empirical studies, it is necessary that authors ensure that they provide clear and complete accounts of the interventions and the study process, including the duration of the videos used, the time allocation for different instructional activities, and detailed descriptions of the small-group activities. Comprehensive reporting will help future reviewers to identify the mechanisms that make the flipped classroom approach effective.

Second, although we searched for publications across different databases, the articles in this review only covered a few subsets of STEM disciplines, and it is not clear whether the conclusions can be applied to all STEM education contexts. Although the reviewed articles are from across different continents, the selection was limited to those published in English. The impact of potential language bias was not taken into consideration. To broaden the scope of this review, future reviews can incorporate conference studies, theses, and cover articles written in languages other than English. It is also noticeable that even though 28 of the studies included in our meta-analysis were published after 2020, only two of them considered the influence of Covid-19 on the delivery mode of in-class activities (Ng, 2023; Sezer & Esenay, 2022). It would be worth considering whether the flipped classroom approach could still improve students' performance when in-class face-to-face interaction is decreased or eliminated in favor of an online environment (Hew et al., 2020a, 2020b). Such an inquiry could consider what strategies in the remote flipped classroom design



would make a significant difference in students' learning experiences and achievements, in comparison to the regular flipped class approach where only the first preparation phase was online (Paul et al., 2023; Widodo, 2022; Zhong et al., 2022).

Finally, educational goals encompass more than academic attainment. They extend to areas including ethics, lifelong career orientation, civic engagement, and communication (National Academies of Sciences & Medicine, 2017). Therefore, academic performance is not the only dimension that should be taken into consideration when exploring the merits of the flipped classroom approach. Future reviews can evaluate the impacts of flipped classrooms in STEM education on other domains including interpersonal skills, perceived learning outcomes, and student satisfaction.

## 7 Conclusions

The meta-analysis of 53 studies covering various STEM disciplines revealed that the flipped classroom in higher education had a moderate positive effect on student performance. The results of the moderator analysis imply that the flipped classroom can significantly improve students' academic performance in the fields of engineering, mathematics, and health profession education. Our study offers valuable insights into the utility of practical design features of a flipped classroom. We recommend that the duration of a flipped classroom course should be capped within a semester. Educators should consider involving more video materials in the pre-class activities and incorporating link activities at the beginning of the face-to-face class to help students reinforce the knowledge acquired during the pre-class home-learning session. Combining group activities and individual tasks during class and providing quizzes and exercises after class are beneficial components that can improve student achievement in STEM education, offering promising directions for the future implementation of flipped classroom courses.

**Author contributions** JG performed the statistical analysis and wrote the first draft. JG and SC carried out data collection and data coding. MC helped to revise the manuscript. All authors read and approved the final manuscript.

**Funding** Supported by National Natural Science Foundation of China (62177036).

**Data availability and materials** All data, materials and code are available and can be provided. This manuscript has associated data in a data repository.

## Declarations

**Conflict of interest** The authors have no relevant financial or non-financial interests to disclose.

**Ethical approval** Not applicable.

**Consent for publication** Not applicable.

## References

Alammary, A. (2019). Blended learning models for introductory programming courses: A systematic review. *PLoS ONE*, 14(9), e0221765. <https://doi.org/10.1371/journal.pone.0221765>

- Algarni, B. (2018). A meta-analysis on the effectiveness of flipped classroom in mathematics education. *Edulearn Proceedings* [Edulearn18: 10th international conference on education and new learning technologies]. 10th International Conference on Education and New Learning Technologies (EDULEARN), Palma, Spain.
- Alt, D. (2023). Assessing the benefits of gamification in mathematics for student gameful experience and gaming motivation. *Computers and Education*, 200, 104806. <https://doi.org/10.1016/j.compedu.2023.104806>
- Anderson, L., & Brennan, J. P. (2015). An experiment in “Flipped” teaching in freshman calculus. *Primus*, 25(9), 861–875. <https://doi.org/10.1080/10511970.2015.1059916>
- Arya, V., Gehlawat, V. K., Rana, R., & Kaushik, J. (2020). Flipped classroom versus traditional lecture in training undergraduates in pediatric epilepsy. *Journal of Family Medicine and Primary Care*, 9(9), 4805–4808. [https://doi.org/10.4103/jfmpe.jfmpe\\_816\\_20](https://doi.org/10.4103/jfmpe.jfmpe_816_20)
- Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, 50(4), 1088–1101. <https://doi.org/10.2307/2533446>
- Besser, D., Stone, G., & Nan, L. (1998). Textbooks and teaching: A lesson from students. *Journalism and Mass Communication Educator*, 53(4), 4–17. <https://doi.org/10.1177/107769589805300401>
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. <https://doi.org/10.1002/9780470743386>
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, 1(2), 97–111. <https://doi.org/10.1002/jrsm.12>
- Bredow, C. A., Roehling, P. V., Knorp, A. J., & Sweet, A. M. (2021). To flip or not to flip? A meta-analysis of the efficacy of flipped learning in higher education. *Review of Educational Research*, 91(6), 878–918. <https://doi.org/10.3102/00346543211019122>
- Cabi, E. (2018). The impact of the flipped classroom model on students’ academic achievement. *International Review of Research in Open and Distributed Learning*, 19(3), 202–221.
- Casselman, M. D., Atit, K., Henbest, G., Guregyan, C., Mortezaei, K., & Eichler, J. F. (2020). Dissecting the flipped classroom: Using a randomized controlled trial experiment to determine when student learning occurs. *Journal of Chemical Education*, 97(1), 27–35. <https://doi.org/10.1021/acs.jchemed.9b00767>
- Castedo, R., Lopez, L. M., Chiquito, M., Navarro, J., Cabrera, J. D., & Ortega, M. F. (2019). Flipped classroom comparative case study in engineering higher education. *Computer Applications in Engineering Education*, 27(1), 206–216. <https://doi.org/10.1002/cae.22069>
- Chang, Y.-S., Hu, S. H., Kuo, S.-W., Chang, K.-M., Kuo, C.-L., Nguyen, T. V., & Chuang, Y.-H. (2022). Effects of board game play on nursing students’ medication knowledge: A randomized controlled trial. *Nurse Education in Practice*, 63, 103412. <https://doi.org/10.1016/j.nepr.2022.103412>
- Chen, K.-S., Monrouxe, L., Lu, Y.-H., Jenq, C.-C., Chang, Y.-J., Chang, Y.-C., & Chai, P.Y.-C. (2018). Academic outcomes of flipped classroom learning: A meta-analysis. *Medical Education*, 52(9), 910–924. <https://doi.org/10.1111/medu.13616>
- Chen, Z., Chen, W., Jia, J., & An, H. (2020). The effects of using mobile devices on language learning: A meta-analysis. *ETR&D-Educational Technology Research and Development*, 68(4), 1769–1789. <https://doi.org/10.1007/s11423-020-09801-5>
- Cheng, L., Ritzhaupt, A. D., & Antonenko, P. (2019a). Effects of the flipped classroom instructional strategy on students’ learning outcomes: A meta-analysis. *ETR&D-Educational Technology Research and Development*, 67(4), 793–824. <https://doi.org/10.1007/s11423-018-9633-7>
- Cheng, L., Ritzhaupt, A. D., & Antonenko, P. (2019b). 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>
- Cheung, A. C. K., & Slavin, R. E. (2016). How methodological features affect effect sizes in education. *Educational Researcher*, 45(5), 283–292. <https://doi.org/10.3102/0013189x16656615>
- Chien, C.-F., & Hsieh, L.-H.C. (2018). Exploring university students’ achievement, motivation, and receptivity of flipped learning in an engineering mathematics course. *International Journal of Online Pedagogy and Course Design*, 8(4), 22–37. <https://doi.org/10.4018/ijopcd.2018100102>
- Chiquito, M., Castedo, R., Santos, A. P., Lopez, L. M., & Alarcon, C. (2020). Flipped classroom in engineering: The influence of gender. *Computer Applications in Engineering Education*, 28(1), 80–89. <https://doi.org/10.1002/cae.22176>
- Corrador, J., Gaydos, M., & Squire, K. (2014). Seeing change in time: Video games to teach about temporal change in scientific phenomena. *Journal of Science Education and Technology*, 23(3), 324–343. <https://doi.org/10.1007/s10956-013-9466-4>

- DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7(3), 177–188. [https://doi.org/10.1016/0197-2456\(86\)90046-2](https://doi.org/10.1016/0197-2456(86)90046-2)
- DeSantis, J., Van Curen, R., Putsch, J., & Metzger, J. (2015). Do students learn more from a flip? An exploration of the efficacy of flipped and traditional lessons. *Journal of Interactive Learning Research*, 26(1), 39–63.
- Dirkx, K. J. H., Kester, L., & Kirschner, P. A. (2014). The testing effect for learning principles and procedures from texts. *The Journal of Educational Research*, 107(5), 357–364. <https://doi.org/10.1080/00220671.2013.823370>
- Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315(7109), 629–634. <https://doi.org/10.1136/bmj.315.7109.629>
- Etemi, B. P., & Uzunboyulu, H. (2020). The effects of flipped learning method on students' perception and learning of java programming. *International Journal of Engineering Education*, 36(4), 1372–1382.
- Fleagle, T. R., Borchering, N. C., Harris, J., & Hoffmann, D. S. (2018). Application of flipped classroom pedagogy to the human gross anatomy laboratory: Student preferences and learning outcomes. *Anatomical Sciences Education*, 11(4), 385–396. <https://doi.org/10.1002/ase.1755>
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23), 8410–8415. <https://doi.org/10.1073/pnas.1319030111>
- Ge, L., Chen, Y., Yan, C., Chen, Z., & Liu, J. (2020). Effectiveness of flipped classroom vs traditional lectures in radiology education A meta-analysis. *Medicine*, 99(40), e22430. <https://doi.org/10.1097/md.00000000000022430>
- Gonzalez-Gomez, D., Su Jeong, J., Airado Rodriguez, D., & Canada-Canada, F. (2016). Performance and perception in the flipped learning model: An initial approach to evaluate the effectiveness of a new teaching methodology in a general science classroom. *Journal of Science Education and Technology*, 25(3), 450–459. <https://doi.org/10.1007/s10956-016-9605-9>
- Graham, C. R. (2013). Emerging practice and research in blended learning. *Handbook of Distance Education*, 3, 333–350.
- Gurevitch, J., & Hedges, L. V. (1999). Statistical issues in ecological meta-analyses. *Ecology*, 80(4), 1142–1149. [https://doi.org/10.1890/0012-9658\(1999\)080\[1142:SIHEMA\]2.0.CO;2](https://doi.org/10.1890/0012-9658(1999)080[1142:SIHEMA]2.0.CO;2)
- Hattie, J. (2017). *Visible learning plus: 250+ Influences on student achievement*. Visible learning plus.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Hawks, S. J. (2014). The flipped classroom: Now or never? *AANA Journal*, 82(4), 264–269.
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, 6(2), 107–128.
- Hedges, L. V. (1982). Estimation of effect size from a series of independent experiments. *Psychological Bulletin*, 92(2), 490–499. <https://doi.org/10.1037/0033-2909.92.2.490>
- Hew, K. F., Bai, S., Dawson, P., & Lo, C. K. (2021). Meta-analyses of flipped classroom studies: A review of methodology. *Educational Research Review*, 33, 100393. <https://doi.org/10.1016/j.edurev.2021.100393>
- Hew, K. F., Bai, S., Huang, W., Du, J., Huang, G., Jia, C., & Khongjan, T. (2020). Does flipped classroom improve student cognitive and behavioral outcomes in stem subjects? Evidence from a second-order meta-analysis and validation study. *Blended Learning. Education in a Smart Learning Environment*.
- Hew, K. F., Jia, C. Y., Gonda, D. E., & Bai, S. R. (2020b). Transitioning to the “new normal” of learning in unpredictable times: Pedagogical practices and learning performance in fully online flipped classrooms. *International Journal of Educational Technology in Higher Education*, 17(1), 57. <https://doi.org/10.1186/s41239-020-00234-x>
- Hew, K. F., & Lo, C. K. (2018a). Flipped classroom improves student learning in health professions education: A meta-analysis. *BMC Medical Education*, 18, 38. <https://doi.org/10.1186/s12909-018-1144-z>
- Hew, K. F., & Lo, C. K. (2018b). Flipped classroom improves student learning in health professions education: A meta-analysis. *BMC Medical Education*, 18(1), 38. <https://doi.org/10.1186/s12909-018-1144-z>
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, 327(7414), 557. <https://doi.org/10.1136/bmj.327.7414.557>
- Hu, R., Gao, H., Ye, Y., Ni, Z., Jiang, N., & Jiang, X. (2018). Effectiveness of flipped classrooms in Chinese baccalaureate nursing education: A meta-analysis of randomized controlled trials. *International Journal of Nursing Studies*, 79, 94–103. <https://doi.org/10.1016/j.ijnurstu.2017.11.012>

- Huber, E., & Werner, A. (2016). A review of the literature on flipping the STEM classroom: Preliminary findings.
- J. Light, R., & B. Pillemer, D. (1986). Summing up: The science of reviewing research harvard university press: Cambridge, MA, 1984, xiii+191 pp. *Educational Researcher*, 15(8), 16–17. <https://doi.org/10.3102/0013189X015008016>
- Jensen, J. L., Holt, E. A., Sowards, J. B., Heath Ogden, T., & West, R. E. (2018). Investigating strategies for pre-class content learning in a flipped classroom. *Journal of Science Education and Technology*, 27(6), 523–535. <https://doi.org/10.1007/s10956-018-9740-6>
- Karabatak, S., & Polat, H. (2020). The effects of the flipped classroom model designed according to the ARCS motivation strategies on the students' motivation and academic achievement levels. *Education and Information Technologies*, 25(3), 1475–1495. <https://doi.org/10.1007/s10639-019-09985-1>
- Karabulut-Ilgu, A., Jaramillo Cherez, N., & Jahren, C. T. (2018). A systematic review of research on the flipped learning method in engineering education. *British Journal of Educational Technology*, 49(3), 398–411. <https://doi.org/10.1111/bjet.12548>
- Karagöl, İ., & Emrullah, E. S. E. N. (2019). The effect of flipped learning approach on academic achievement: A meta-analysis study. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 34(3), 708–727.
- Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A Meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Review of Educational Research*, 86(3), 643–680. <https://doi.org/10.3102/0034654315626799>
- Kubra, A. H., & Lee, J. (2020). Flipped learning research trends in english as a foreign language classrooms: Systematic reviews. *Global Creative Leader*, 10(3), 109–133.
- Lag, T., & Saele, R. G. (2019). Does the flipped classroom improve student learning and satisfaction? A systematic review and meta-analysis. *Aera Open*, 5(3), 2332858419870489. <https://doi.org/10.1177/2332858419870489>
- Låg, T., & Sæle, R. G. (2019). Does the flipped classroom improve student learning and satisfaction? A systematic review and meta-analysis. *Aera Open*, 5(3), 1–17.
- Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *The Journal of Economic Education*, 31(1), 30–43. <https://doi.org/10.2307/1183338>
- Li, B.-Z., Cao, N.-W., Ren, C.-X., Chu, X.-J., Zhou, H.-Y., & Guo, B. (2020a). Flipped classroom improves nursing students' theoretical learning in China: A meta-analysis. *PLoS ONE*, 15(8), e0237926. <https://doi.org/10.1371/journal.pone.0237926>
- Li, Y., Wang, K., Xiao, Y., & Froyd, J. E. (2020b). Research and trends in STEM education: A systematic review of journal publications. *International Journal of STEM Education*, 7(1), 11. <https://doi.org/10.1186/s40594-020-00207-6>
- Lin, Y.-T. (2019). Impacts of a flipped classroom with a smart learning diagnosis system on students' learning performance, perception, and problem solving ability in a software engineering course. *Computers in Human Behavior*, 95, 187–196. <https://doi.org/10.1016/j.chb.2018.11.036>
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE publications.
- Lo, C. K., & Hew, K. F. (2019). The impact of flipped classrooms on student achievement in engineering education: A meta-analysis of 10 years of research. *Journal of Engineering Education*, 108(4), 523–546. <https://doi.org/10.1002/jee.20293>
- Lo, C. K., Hew, K. F., & Chen, G. (2017). Toward a set of design principles for mathematics flipped classrooms: A synthesis of research in mathematics education. *Educational Research Review*, 22, 50–73. <https://doi.org/10.1016/j.edurev.2017.08.002>
- McCabe, C., Smith, M. G., & Ferreri, S. P. (2017). Comparison of flipped model to traditional classroom learning in a professional pharmacy course. *Education Sciences*, 7(3), 73. <https://doi.org/10.3390/educsci7030073>
- Means, B., Toyama, Y., Murphy, R., & Baki, M. (2013). The effectiveness of online and blended learning: A meta-analysis of the empirical literature. *Teachers College Record*, 115(3), 1–47. <https://doi.org/10.1177/016146811311500307>
- Meng, X.-H., Xu, X.-Y., Chen, H.-L., & Zhang, L. (2022). The effectiveness of combining e-learning, peer teaching, and flipped classroom for delivering a physiology laboratory course to nursing students. *Advances in Physiology Education*, 46(1), 21–26. <https://doi.org/10.1152/advan.00062.2020>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Altman, D., Antes, G., Atkins, D., Barbour, V., Barrowman, N., Berlin, J. A., Clark, J., Clarke, M., Cook, D., D'Amico, R., Deeks, J. J., Devereaux, P. J., Dickersin, K., Egger, M., Ernst, E., et al. (2014). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Revista Espanola De Nutricion Humana y Dietetica*, 18(3), 172–181. <https://doi.org/10.1371/journal.pmed.1000097>

- Moskal, P., Dziuban, C., & Hartman, J. (2013). Blended learning: A dangerous idea? *The Internet and Higher Education*, 18, 15–23. <https://doi.org/10.1016/j.iheduc.2012.12.001>
- Mu, S., Hu, J., Liu, F., Zhao, C., & Sun, Y. (2023). New innovations and educational process in undergraduate neurology education in blended learning. *Postgraduate Medicine*. <https://doi.org/10.1080/00325481.2023.2234273>
- Müller, C., & Mildnerberger, T. (2021). Facilitating flexible learning by replacing classroom time with an online learning environment: A systematic review of blended learning in higher education. *Educational Research Review*, 34, 100394. <https://doi.org/10.1016/j.edurev.2021.100394>
- Murphy, J., Chang, J.-M., & Suaray, K. (2016). Student performance and attitudes in a collaborative and flipped linear algebra course. *International Journal of Mathematical Education in Science and Technology*, 47(5), 653–673. <https://doi.org/10.1080/0020739x.2015.1102979>
- Muzyk, A. J., Fuller, S., Jiroutek, M. R., Grochowski, C. O. C., Butler, A. C., & May, D. B. (2015). Implementation of a flipped classroom model to teach psychopharmacotherapy to third-year Doctor of Pharmacy (PharmD) students. *Pharmacy Education*, 15(1), 44–53.
- Nagendrababu, V., Pulikkotil, S. J., Sultan, O. S., Jayaraman, J., Soh, J. A., & Dummer, P. M. H. (2019). Effectiveness of technology-enhanced learning in Endodontic education: A systematic review and meta-analysis. *International Endodontic Journal*, 52(2), 181–192. <https://doi.org/10.1111/iej.12995>
- National Academies of Sciences, E., & Medicine. (2017). *Supporting students; college success: The role of assessment of intrapersonal and interpersonal competencies*. The National Academies Press. <https://doi.org/10.17226/24697>
- Ng, E. I. L. (2023). Flipped versus traditional classroom and student achievement and cognitive engagement in an associate degree nursing fundamental course [Article]. *Nurse Education in Practice*, 68, 103567. <https://doi.org/10.1016/j.nepr.2023.103567>
- Paul, A., Leung, D., Salas, R. M. E., Cruz, T. E., Abras, C., Saylor, D., Gugliucciello, V., Nunn, J., Gamaldo, C. E., & Strowd, R. E. (2023). Comparative effectiveness study of flipped classroom versus online-only instruction of clinical reasoning for medical students. *Medical Education Online*, 28(1), 2142358. <https://doi.org/10.1080/10872981.2022.2142358>
- Peplow, M. (2014). Social sciences suffer from severe publication bias. *Nature*. <https://doi.org/10.1038/nature.2014.15787>
- Rahman, T., & Lewis, S. E. (2020). Evaluating the evidence base for evidence-based instructional practices in chemistry through meta-analysis. *Journal of Research in Science Teaching*, 57(5), 765–793. <https://doi.org/10.1002/tea.21610>
- Roediger, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, 1(3), 181–210. <https://doi.org/10.1111/j.1745-6916.2006.00012.x>
- 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>
- Sedig, K. (2008). From play to thoughtful learning: A design strategy to engage children with mathematical representations. *Journal of Computers in Mathematics and Science Teaching*, 27(1), 65–101.
- El-Senousy, H., & Alquda, J. (2017). The effect of flipped classroom strategy using blackboard mash-up tools in enhancing achievement and self-regulated learning skills of university students. *World Journal on Educational Technology: Current Issues*, 9(3), 144–157.
- Sezer, T. A., & Esenay, F. I. (2022). Impact of flipped classroom approach on undergraduate nursing student's critical thinking skills. *Journal of Professional Nursing*, 42, 201–208. <https://doi.org/10.1016/j.profnurs.2022.07.002>
- Shi, Y., Ma, Y., MacLeod, J., & Yang, H. H. (2020a). 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>
- Shi, Y., Yang, H., MacLeod, J., Zhang, J., & Yang, H. H. (2020b). College students' cognitive learning outcomes in technology-enabled active learning environments: A meta-analysis of the empirical literature. *Journal of Educational Computing Research*, 58(4), 791–817. <https://doi.org/10.1177/0735633119881477>
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189. <https://doi.org/10.3102/0034654307313795>
- Strelan, P., Osborn, A., & Palmer, E. (2020). The flipped classroom: A meta-analysis of effects on student performance across disciplines and education levels. *Educational Research Review*, 30, 100314. <https://doi.org/10.1016/j.edurev.2020.100314>
- Total, Ö., & Yazar, T. (2021). Flipped classroom improves academic achievement, learning retention and attitude towards course: A meta-analysis. *Asia Pacific Education Review*, 22(4), 655–673. <https://doi.org/10.1007/s12564-021-09706-9>

- van Alten, D. C. D., Phielix, C., Janssen, J., & Kester, L. (2019). Effects of flipping the classroom on learning outcomes and satisfaction: A meta-analysis. *Educational Research Review*, 28, 100281. <https://doi.org/10.1016/j.edurev.2019.05.003>
- Vitta, J. P., & Al-Hoorie, A. H. (2020). The flipped classroom in second language learning: A meta-analysis. *Language Teaching Research*. <https://doi.org/10.1177/1362168820981403>
- Wang, A., Xiao, R., Zhang, C., Yuan, L., Lin, N., Yan, L., Wang, Y., Yu, J., Huang, Q., Gan, P., Xiong, C., Xu, Q., & Liao, H. (2022). Effectiveness of a combined problem-based learning and flipped classroom teaching method in ophthalmic clinical skill training. *BMC Medical Education*, 22(1), 487. <https://doi.org/10.1186/s12909-022-03538-w>
- Widodo, W. (2022). Online flipped classroom: Developing postgraduate science education students' critical thinking skills. *Journal of Science Learning*, 5(2), 469–477.
- Wigfield, A., Klauda, S. L., & Cambria, J. (2013). Influences on the development of academic self-regulatory processes. In *Handbook of self-regulation of learning and performance*. <https://doi.org/10.4324/9780203839010.ch3>
- Wilson, K. E., & Hobbs, J. R. (2023). Innovative use of a flipped-classroom approach to teach fundamental nursing skills. *Teaching and Learning in Nursing*, 18(1), 144–147. <https://doi.org/10.1016/j.teln.2022.08.002>
- Wong, T. H., Ip, E. J., Lopes, I., & Rajagopalan, V. (2014). Pharmacy students' performance and perceptions in a flipped teaching pilot on cardiac arrhythmias. *American Journal of Pharmaceutical Education*, 78(10), 185–185. <https://doi.org/10.5688/ajpe7810185>
- Xu, P., Chen, Y., Nie, W., Wang, Y., Song, T., Li, H., Li, J., Yi, J., & Zhao, L. (2019). The effectiveness of a flipped classroom on the development of Chinese nursing students' skill competence: A systematic review and meta-analysis. *Nurse Education Today*, 80, 67–77. <https://doi.org/10.1016/j.nedt.2019.06.005>
- Zhong, J., Li, Z., Hu, X., Wang, L., & Chen, Y. (2022). Effectiveness comparison between blended learning of histology practical in flipped physical classrooms and flipped virtual classrooms for MBBS students. *BMC Medical Education*, 22(1), 795. <https://doi.org/10.1186/s12909-022-03740-w>

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

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