Does Flipped Classroom Improve Student Cognitive and Behavioral Outcomes in STEM Subjects? Evidence from a Second-Order Meta-Analysis and Validation Study

Khe Foon Hew[(✉)], Shurui Bai, Weijiao Huang, Jiahui Du, Guoyuhui Huang, Chengyuan Jia, and Thankrit Khongjan

The University of Hong Kong, Pokfulam Road, Hong Kong SAR, China
`kfhew@hku.hk`

Abstract. The flipped classroom approach has attracted particular interest from educators in the Science-Technology-Engineering-Mathematics (STEM) subject disciplines. Despite its popularity, the effect of the flipped classroom approach on student outcomes compared with conventional learning has not yet been conconclusively determined. Although previous reviews have reported positive student *perceptions* of flipped courses, this does *not* necessarily imply improved student learning. This study used a *second-order* meta-analysis procedure to summarize more than 10 years of research examining the following question: "*Does the flipped classroom improve student cognitive and behavioral outcomes across STEM subjects as compared with non-flipped classroom?*" A total of 10 primary meta-analyses covering 217 unique STEM studies were analyzed. Results showed that flipped classroom significantly improves student cognitive learning ($g = 0.49$, $p < 0.001$), and student behavioral learning ($g = 1.70$, $p < 0.001$) as compared to conventional classroom. To validate the results of the second-order meta-analysis, we also conducted a study-level meta-analytic validation. Students' perceptions of using flipped classroom were also analyzed.

Keywords: Flipped classroom · Flipped learning · Meta-analysis · STEM

1 Introduction

1.1 Background

Many education classes now utilize the flipped classroom approach. In flipped classrooms, students are expected to learn the subject content before class usually through watching recorded video lectures [1] and completing some pre-set online exercises. Students then attend classes to complete individual and/or group activities under the instructor's supervision [2]. The whole idea of the flipped classroom approach is to foster more active learning opportunity for students.

The flipped classroom approach has made particular inroads across the science, technology, engineering, and mathematics (STEM) disciplines. Traditionally, the learning of

STEM disciplines has been characterized by the didactic teacher-lecture approach [3] as these disciplines generally involve the learning of complex concepts before students are able to move on to the practical applications [4]. Instructors have typically spent class time delivering content through lectures and then assigning supplementary homework exercises.

For this reason, increasing numbers of educators have begun experimenting with flipped classrooms. Previous reviews have mainly concentrated on specific subject areas such as nursing or medical education [5–7] and chemistry [8], or have focused on flipped classroom implementation in a particular context, such as Asian universities [9]. In these reviews the researchers generally summarize the characteristics of the primary studies, such as the types of research method used, the types of participants, and the advantages of using flipped classrooms. Some also describe the challenges or disadvantages of flipped classrooms.

To synthesize the findings about learning outcomes, some researchers would merely list the studies that showed an increase in academic performance, e.g. [2], describe the results of each comparison study e.g. [5], or generally report that "flipping the class is a way to improve learning performance" (p. 26) [2]. A few researchers simply used descriptive statistics to summarize the findings of their reviewed studies. For example, Seery [8] counted the flipped classrooms in higher education chemistry that "half of the studies showed no improvement in exam scores" (p. 763). In Bernard's [10] review, 15 out of 24 studies indicated that flipped classroom students outperformed traditional classroom students.

However, what is the overall effect of flipped classroom on STEM student learning achievement? To achieve this aim, several authors have begun to meta-analyze the published individual empirical studies. The main goal of primary meta-analysis, or first-order meta-analysis, is to estimate an overall mean effect from *individual empirical studies*, and to identify possible factors that moderate that effect [11].

Although primary meta-analyses can offer important advantages, it is important to note that they are limited in their scope and focus. Primary meta-analyses usually focus on specific aspects (e.g., subject matter, grade level). Greater generality could be achieved by examining multiple primary meta-analyses and synthesizing their results.

1.2 Purpose of the Present Study

The current study used an approach known as *second-order* meta-analysis to synthesize a number of methodologically comparable primary meta-analyses [12]. A second-order meta-analysis combines the results of multiple primary meta-analyses. This method has two important advantages. First, by summarizing the findings of more than one primary meta-synthesis, a second-order meta-analysis can generate a more robust generalizable result [13]. Second, a second-order meta-analysis can provide a more accurate estimate of the true mean effect sizes in each primary meta-analysis [14].

In this study, we conducted the first-ever second-order meta-analysis to synthesize the results of prior primary meta-analyses on STEM subjects to gain a broader understanding of the effects of flipped classroom use on student learning performance. To validate the results, we also conducted a study-level synthesis of all available effect sizes reported

in the primary meta-analyses included in our second-order meta-analysis. Our study is guided by the following questions:

Research question 1: What is the overall effect of flipped classroom used on STEM students' cognitive and behavioral outcomes? Do subject discipline and student grade level moderate the effect of flipped classroom used on students' outcome?
Research question 2: How do participants perceive the use of flipped classroom?

Cognitive Outcomes refer to domain-specific knowledge of a subject. Cognitive outcomes are usually assessed using teacher-developed or standardized tests and exams. *Behavioral outcomes* refer to learners' motor skills or competences in performing a task. Behavioral outcomes are usually assessed through observations such as objective structured clinical examinations (OSCEs).

2 Method

To be as comprehensive as possible, we searched more than ten major academic databases were searched, including ACM Digital Library, all EBSCO host research databases (e.g., Academic Search Premier, British Education Index, ERIC, TOC Premier), Emerald Insight, IEEE Xplore, ProQuest Dissertations & Theses A&I, Science Direct, Scopus, Springer, Web of Science, JSTOR, PubMed, and Google Scholar. To ensure that our search was comprehensive, additional primary meta-analyses identified in Google Scholar were also included. The search string used in this review was: *"meta-analysis" AND ("flip*" OR "invert*") AND ("class*" OR "learn*").* No time restriction was applied to the search, which was completed in November 2019.

For this second-order meta-analysis, a primary meta-analysis was included if it met the following criteria. The primary meta-analysis must: (a) compare the effects of flipped classrooms with those of non-flipped classrooms in STEM specific subjects, including health sciences, general sciences (e.g., chemistry, biology, physics), engineering and/or technology, and mathematics, (b) focus on students' cognitive or behavioral outcome, (c) report an average effect size, (d) provide a list of the primary studies analyzed, and (e) be written in English and publicly available, or available through a library database subscription.

2.1 Data Extraction

For each primary meta-analysis, we extracted the following data: (a) study identification (e.g., author), (b) contextual features (e.g., subject matter), (c) methodological features (e.g., total number of primary studies reviewed, total number of participants), and (d) results (e.g., type of effect size, effect size data). The process of data extraction was conducted by the first author. To test the reliability of the coding, the second author coded five randomly selected meta-analyses independently. There was perfect agreement between the two coders.

2.2 Data Analysis

All of the primary meta-analyses reported the overall effect size using either Hedges' *g* or SMD (standard mean difference, Cohen's *d*). Following Young [15], we assumed that the differences between *g* and *d* would be negligible. Therefore, we used the effect sizes in the original metrics in which they were reported. We retrieved the standard errors (SEs) of effect sizes as directly reported by the authors when available, or computed (SEs) using confidence intervals:

$$SE = (95\% \, CI_{upper \, limt} - 95\% \, CI_{lower \, limit})/3.92 \qquad (1)$$

We used the Comprehensive Meta-Analysis software package [16] to conduct effect size synthesis, publication bias and moderator analyses. We conducted a classic fail-safe *N* test, plotted funnel plots [17], and applied the trim-and-fill method [18] to adjust for any possible publication bias.

2.3 Validation of the Second-Order Meta-Analysis

One important issue when conducting a second-order meta-analysis is the issue of study overlap across the various primary meta-analyses [19]. Study overlap occurs when the same empirical studies are included in more than one primary meta-analysis [12]. Although several approaches to addressing study overlap have been proposed, it is not clear which approach is the most appropriate [20].

In this study, we followed the method employed by Young [15] and Tamim et al. [21] in validating the findings of the present second-order meta-analysis. To do this, *individual* effect sizes and sample sizes from available empirical studies reported in the primary meta-analyses were extracted. In this validation sample, all empirical study overlap was eliminated. We then compared the overall mean effect size from the present second-order meta-analysis with the overall mean effect size from the validation sample, where all study overlap had been removed, to determine whether the average effect sizes were similar.

3 Results

As shown in Fig. 1, the initial searches resulted in 42 article abstracts (after duplicates were removed). From these identified 42 documents, 15 were removed after a title and abstract review. Ultimately, 27 full-text primary meta-analyses were assessed for eligibility. Of these 27 full-text primary meta-analyses, 20 were excluded due to a number of reasons, including: no effect size was provided (n = 6); the outcome measure was not student learning performance (n = 6); and non-STEM subjects (n = 8).

The final dataset subsequently used in this study consisted of 10 primary meta-analyses comparing the effects of the flipped classroom and the non-flipped classroom on STEM students' academic performance. Seven primary meta-analyses examined student cognitive learning outcomes [22–28]. These meta-analyses covered a total of 176 unique primary studies with 37,775 participants (19,045 in a flipped class and 19,730 participants in a comparison class). Three primary meta-analyses covering 41
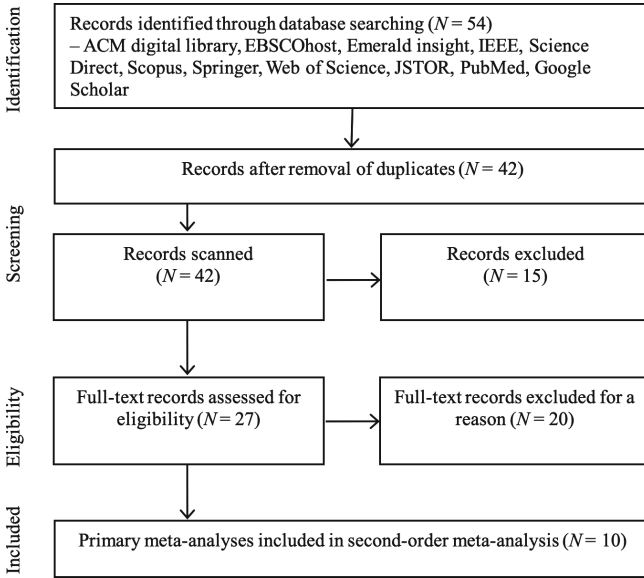
Records identified through database searching ($N = 54$)
– ACM digital library, EBSCOhost, Emerald insight, IEEE, Science
Direct, Scopus, Springer, Web of Science, JSTOR, PubMed, Google
Scholar

Records after removal of duplicates ($N = 42$)

Records scanned
($N = 42$)

Records excluded
($N = 15$)

Full-text records assessed for
eligibility ($N = 27$)

Full-text records excluded for a
reason ($N = 20$)

Primary meta-analyses included in second-order meta-analysis ($N = 10$)

Identification

Screening

Eligibility

Included

Fig. 1. PRISMA flowchart of article screening process.

unique primary studies with 9,473 examined student behavioral learning outcomes [24, 27, 29].

All 10 primary meta-analyses relied on the effect sizes extracted from primary studies published in peer-reviewed journals. However, four primary meta-analyses included the analysis of *additional* primary studies from conference proceedings and/or dissertations [25, 27–29]. Five meta-analyses focused solely on health sciences education [22–24, 27, 29]. Other meta-analyses focused on science education [28], mathematics education [26], and engineering education [25].

3.1 Effect Size Synthesis: Cognitive and Behavioral Outcomes

We chose the random-effects model to compute the overall effect size of this second-order meta-analysis. The random effects model was considered more appropriate for interpretation because of the wide diversity of settings and subject matter. Under the random-effects model, results revealed a small effect in favor of the flipped classroom approach on student assessed cognitive learning outcomes (Hedges's $g = 0.49$, 95% CI $= 0.326 - 0.656$, $p < 0.001$), and a large effect on student behavioral outcomes (Hedges's $g = 1.70$, 95% CI $= 1.395 - 1.998$, $p < 0.001$) (Fig. 2, Fig. 3 respectively). We conducted outlier analysis using the 'one-study-remove' method. Results revealed that all 7 effect sizes for cognitive outcomes fell within the 95% CI of the overall effect size ($0.326 - 0.656$). Thus, there was no need to remove any meta-analysis.

We conducted sub-group analyses to examine whether the subject discipline and student grade level might moderate the magnitude of the effect sizes (Table 1). The moderator analysis concerning subject discipline suggested no significant effect size
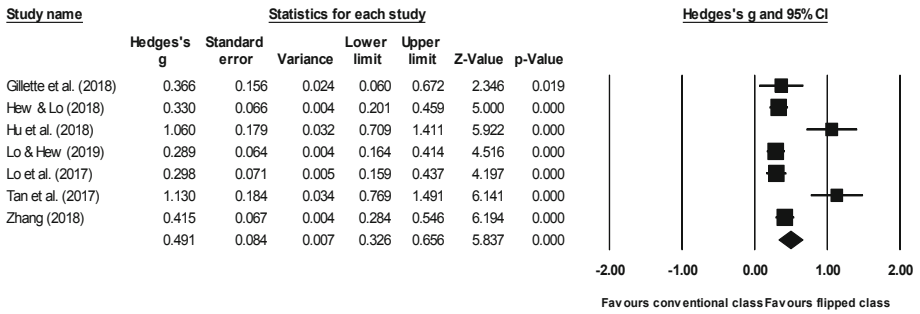
Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Gillette et al. (2018)	0.366	0.156	0.024	0.060	0.672	2.346	0.019
Hew & Lo (2018)	0.330	0.066	0.004	0.201	0.459	5.000	0.000
Hu et al. (2018)	1.060	0.179	0.032	0.709	1.411	5.922	0.000
Lo & Hew (2019)	0.289	0.064	0.004	0.164	0.414	4.516	0.000
Lo et al. (2017)	0.298	0.071	0.005	0.159	0.437	4.197	0.000
Tan et al. (2017)	1.130	0.184	0.034	0.769	1.491	6.141	0.000
Zhang (2018)	0.415	0.067	0.004	0.284	0.546	6.194	0.000
	0.491	0.084	0.007	0.326	0.656	5.837	0.000

Hedges's g and 95% CI

-2.00 -1.00 0.00 1.00 2.00

Favours conventional class   Favours flipped class

Fig. 2. Forest plot of effect size on student *cognitive* learning outcomes (random-effects)

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Hu et al. (2018)	1.400	0.480	0.230	0.459	2.341	2.917	0.004
Tan et al. (2017)	1.680	0.220	0.048	1.249	2.111	7.636	0.000
Xu et al. (2019)	1.790	0.240	0.058	1.320	2.260	7.458	0.000
	1.696	0.154	0.024	1.395	1.998	11.041	0.000

Hedges's g and 95% CI

-4.00 -2.00 0.00 2.00 4.00
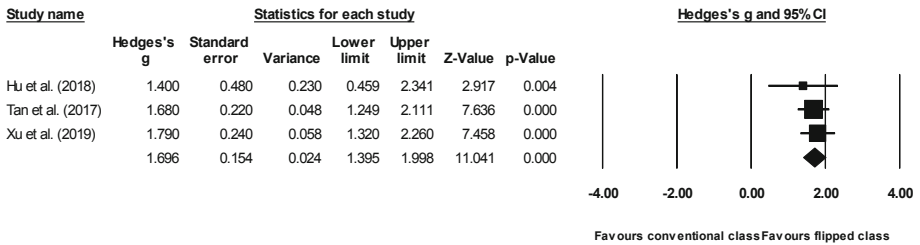
Favours conventional class   Favours flipped class

Fig. 3. Forest plot of effect size on student *behavioral* learning outcomes (random-effects)

difference among different disciplines ($Q_B = 1.411$, df $= 3$, $p = 0.703$). The moderator analysis concerning student grade level also suggested no significant effect size difference ($Q_B = 3.374$, df $= 1$, $p = 0.066$).

Table 1. Sub-group analyses.

Sub-group	N	g	SE	95% CI (LL)	95% CI (LL)	QB (p)
Subject discipline						1.411 (0.703)
Engineering	1	0.289	0.404	−0.502	1.080	
Health sciences	4	0.701	0.213	0.284	1.119	
Mathematics	1	0.298	0.405	−0.495	1.091	
Science	1	0.415	0.404	−0.377	1.207	
Student level						3.374 (0.066)
Post-secondary	4	0.662	0.127	0.414	0.910	
K-12 + post-secondary	3	0.334	0.126	0.088	0.580	

We also conducted outlier analysis through the 'one-study-remove' method. Results revealed that all three effect sizes for behavioral outcomes fell within the 95% CI of the overall effect size (1.395 – 1.998). Thus, there is no need to remove any meta-analysis.

3.2 Data Validation

We administered the data validation by extracting the available raw individual effect sizes used in the seven primary meta-analyses and running a regular meta-analysis. A total of 149 independent effect sizes and standard errors were extracted. We computed the overall mean effect size using the random-effects model. The results showed a significant medium effect size in favour of flipped classroom on students' cognitive learning outcome (Hedges's g $= 0.51$, CI $= 0.438$-0.599, $p < 0.001$). We also computed the overall mean effect size of student behavioral outcome using the random-effects model among 41 effect sizes. The results showed a significant medium effect size in favour of flipped classroom on students' behavioral learning outcome (Hedges's g $= 1.70$, CI $= 1.387 - 2.010$, $p < 0.001$).

In comparing the second-order meta-analysis with the validation study, it is evident that the mean effect sizes for the random-effects model for both student cognitive and behavioral outcomes are closely similar. For example, the mean effect size of the second-order meta-analyses for cognitive outcome was 0.49 (random-effects model), while the mean effect size of the validation study sample was 0.51 (random-effects model). There was only a difference of 0.02, a magnitude which can be deemed trivial [10]. The mean effect size of the second-order meta-analyses for behavioral outcome was 1.70 (random-effects model) which was similar to the mean effect size of the validation study sample (1.70, random-effects model) Thus, the results of the second-order meta-analysis were considered to be a valid representation of the cumulative effects of the primary meta-analyses.

3.3 Publication Bias

Figure 4 shows the funnel plot of the seven primary meta-analyses on student cognitive outcome. The classic fail-safe N test result showed that 300 additional studies would be required to invalidate the overall effect.
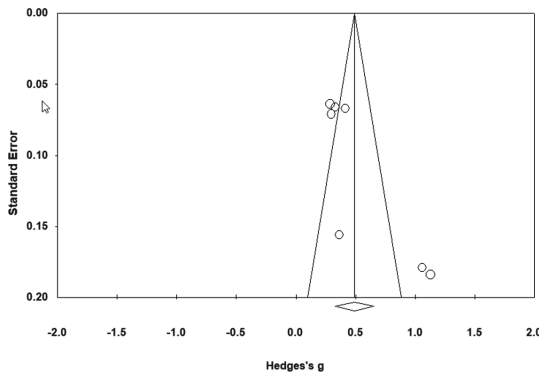

Fig. 4. Funnel plot of seven primary meta-analyses on cognitive learning outcome

We also conducted an examination of publication bias of the cognitive outcomes validation study which consists of 149 independent effect sizes used in the seven primary meta-analyses (see Fig. 5).
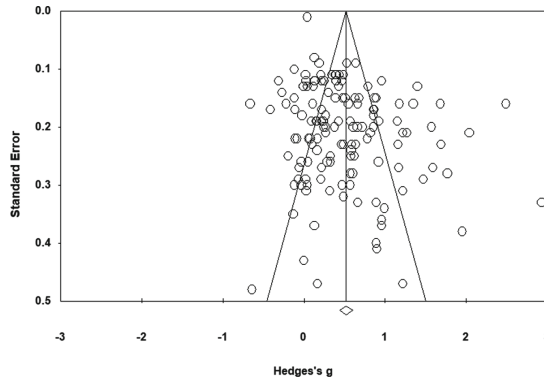

Fig. 5. Funnel plot of the validation study (n = 149 effect sizes) on cognitive learning outcomes

The classic fail-safe N was 1,524, indicating that 1,524 null effect studies are needed to invalidate the overall effect for cognitive outcomes. Kendall's Tau was 0.006 (one-tailed $p = 0.456$) which suggested no presence of publication bias. The trim and fill-method, using random-effects model, suggests that 0 studies are missing from the left side of the mean effect.

With regard to student behavioural learning outcome, the classic fail-safe N analysis revealed that 82 null effect studies are needed to invalidate the overall effect. Kendall's Tau was 0.000 (one-tailed $p = 0.500$) which suggested no presence of publication bias. The trim and fill-method, using random-effects model, suggests that 0 studies are missing from the left side of the mean effect. We also conducted an examination of publication bias of the behavioral learning outcomes validation study which consists of 41 independent effect sizes used in the three primary meta-analyses (see Fig. 6).
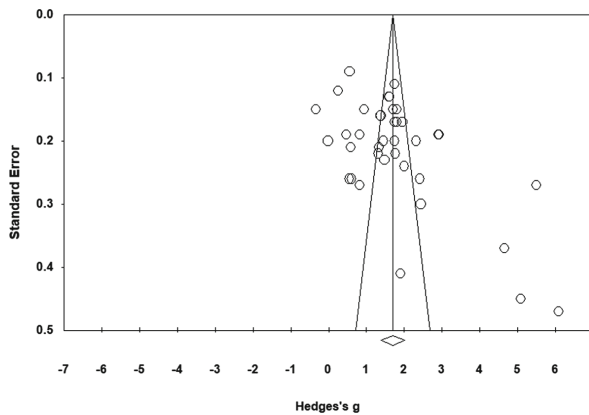

Fig. 6. Funnel plot of the validation study on behavioral outcomes

The classic fail-safe N was 7,940, indicating that 7,940 null effect studies are needed to invalidate the overall effect for behavioral outcomes. Kendall's Tau was 0.215 (one-tailed $p = 0.024$) which suggested a slight presence of publication bias. The trim and fill-method, using random-effects model, however, suggests that 0 studies are missing from the left side of the mean effect. In summary, the results of the classic fail-safe Ns and funnel plots suggest no obvious publication bias for both student cognitive and behavioral outcomes.

3.4 Student Perception of Flipped Classroom

Three primary meta-analyses [25, 26, 28] reported student perception of using flipped classroom. In this section, we briefly synthesize the main findings according to two main categories: benefits of flipped classroom, and challenges of flipped classroom.

Six main benefits of flipped classroom were identified: (a) promote self-paced learning, (b) allow multiple exposures to course materials, (c) help prepare students for class, (d) more opportunity for knowledge application (e.g., problem solving) activities, (e) provide more opportunity for peer-assisted learning, and (f) allow more timely instructor intervention and support. Compared to non-flipped classroom, flipped classroom students have a greater opportunity to do self-paced learning due to the availability of pre-class activity. Students can choose to watch the video or read the course materials at any time and in whatever pace they desire. Flipped classroom also provided students with more than one exposure to the course materials. Students are first exposed to the course materials before class during the pre-class activity. Students engaged with the course materials again later during the in-class session. Multiple exposure to course materials can help improve student understanding of the lesson.

Five main challenges of flipped classroom were also identified. These include: (a) student unfamiliarity with the new approach, (b) student not willing to complete the pre-class activity due to perceived additional workload, (c) student inability to ask questions during the pre-class activity, (d) significant start-up effort on the part of the instructor, and (e) instructor unfamiliarity with the flipped classroom approach.

4   Conclusion

Many individual empirical studies that examined the effects of flipped classroom on student outcomes have been conducted and published in journals, conference proceedings, and dissertations. Along with this growing number of individual empirical studies, there has also been a corresponding increase of meta-analytic studies of the flipped classroom approach in a variety of contexts. The present study used a second-order meta-analysis method to synthesize the findings of ten primary meta-analyses on STEM student cognitive and behavioral outcomes. Overall, we found a significant mean effect size of 0.49, and effect size of 1.70 under the random effects model in favor of the flipped classroom approach in enhancing student cognitive and behavioral outcomes respectively. Findings of our validation study suggest that the results of the second-order meta-analysis can be considered valid representations of the cumulative effects of the primary meta-analyses.

The overall effect size of 0.49 for cognitive outcome may be considered a medium effect [30] and is closely similar to the typical value of 0.40 for student achievement that is reported elsewhere [31]. The overall effect size of 1.70 for behavioral outcome is considered a large effect [30]. An interesting finding is that the effect size of behavioral outcome is much higher than that of the effect size of cognitive outcome. One possible reason why flipped classroom lends itself particularly well to improving behavioral outcome is that it promotes more time for application of knowledge than traditional classroom.

We conclude by highlighting two limitations of previous primary meta-analyses. A majority of the primary meta-analyses did not report whether or how they control for potential student initial equivalence. Student initial differences can result in substantial bias in the outcome measures. If students have different initial knowledge or skill about the subject matter, it becomes unclear whether it is the actual flipped learning pedagogy that caused the effect, or the student's initial knowledge or skill that influenced the outcome. A majority of previous meta-analyses also did not report whether or how they control for instructor equivalence. Since different instructors have different teaching styles (although the content materials remain similar), it becomes unclear whether the actual flipped learning pedagogy caused the effect, or the presence of the instructor confounding variable influenced the outcome. We therefore urge future primary meta-analyses to carefully consider these two issues.

References

1. Abeysekera, L., Dawson, P.: Motivation and cognitive load in the flipped classroom: definition, rationale and a call for research. High Educ. Res. Dev. **34**, 1–14 (2014)
2. O'Flaherty, J., Phillips, C.: The use of flipped classrooms in higher education: a scoping review. Internet High. Educ. **25**, 85–95 (2015). https://doi.org/10.1016/j.iheduc.2015.02.002
3. Bates, S., Galloway, R.: The inverted classroom in a large enrolment introductory physics course: a case study (2012)
4. Huber, E., Werner, A.: A review of the literature on flipping the STEM classroom: preliminary findings. In: Barker, S., Dawson, S., Pardo, A., Colvin, C. (eds.). Show me the learning. Proceedings ASCILITE 2016, Adelaide, pp. 267–274 (2016)
5. Betihavas, V., Bridgman, H., Kornhaber, R., Cross, M.: The evidence for "flipping out": a systematic review of the flipped classroom in nursing education. Nurse Educ. Today **38**, 15–21 (2016)
6. Chen, K.S., et al.: Academic outcomes of flipped classroom learning: a meta analysis. Med. Educ. **52**, 910–924 (2018)
7. Presti, C.R.: The flipped learning approach in nursing education: a literature review. J. Nurs. Educ. **55**(5), 252–257 (2016)
8. Seery, M.K.: Flipped learning in higher education chemistry: Emerging trends and potential directions. Chem. Educ. Res. Pract. **16**(4), 758–768 (2015)
9. Chua, J., Lateef, F.: The flipped classroom: viewpoints in Asian universities. Educ. Med. J. **6** (2014). https://doi.org/10.5959/eimj.v6i4.316

10. Bernard, J.S.: The flipped classroom: fertile ground for nursing education research. Int. J. Nurs. Educ. Scholarsh. **12**(1), 99–109 (2015)
11. Gurevitch, J., Koricheva, J., Nakagawa, S., Stewart, G.: Meta-analysis and the science of research synthesis. Nature **555**(7695), 175–182 (2018). https://doi.org/10.1038/nature25753
12. Cooper, H., Koenka, A.C.: The overview of reviews: unique challenges and opportunities when research syntheses are the principal elements of new integrative scholarship. Am. Psychol. **67**(6), 446–462 (2012). https://doi.org/10.1037/a0027119
13. Busch, T., Friede, G.: The robustness of the corporate social and financial performance relation: a second-order meta-analysis. Corp. Soc. Responsib. Environ. Manag. **25**(4), 583–608 (2018). https://doi.org/10.1002/csr.1480
14. Schmidt, F., Oh, I.-S.: The crisis of confidence in research findings in psychology: is lack of replication the real problem? or is it something else? Arch. Sci. Psychol. **4**, 32–37 (2016). https://doi.org/10.1037/arc0000029
15. Young, J.: Technology-enhanced mathematics instruction: a second-order meta-analysis of 30 years of research. Educ. Res. Rev. **22**, 19–33 (2017). https://doi.org/10.1016/j.edurev.2017.07.001
16. Borenstein, M., Hedges, L.V., Higgings, J.P.T., Rothstein, H.R.: Introduction to Meta-Analysis. Wiley, Chichester (2009)
17. Egger, M., Smith, G.D., Schneider, M., Minder, C.: Bias in meta-analysis detected by a simple, graphical test. Br. Med. J. **315**(7109), 629–634 (1997)
18. Duval, S., Tweedie, R.: Trim and fill: a simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. Biometrics **56**(2), 455–463 (2000). https://doi.org/10.1111/j.0006-341x.2000.00455
19. Polanin, J.R., Tanner-Smith, E.E., Hennessy, E.A.: Estimating the difference between published and unpublished effect sizes: a meta-review. Rev. Educ. Res. **86**(1), 207–236 (2016). https://doi.org/10.3102/0034654315582067
20. Steenbergen-Hu, S., Makel, M.C., Olszewski-Kubilius, P.: What one hundred years of research says about the effects of ability grouping and acceleration on K–12 students' academic achievement: findings of two second-order meta-analyses. Rev. Educ. Res. **86**(4), 849–899 (2016). https://doi.org/10.3102/0034654316675417
21. Tamim, R.M., Bernard, R.M., Borokhovski, E., Abrami, P.C., Schmid, R.F.: What forty years of research says about the impact of technology on learning: a second-order meta-analysis and validation study. Rev. Educ. Res. **81**(1), 4–28 (2011). https://doi.org/10.3102/0034654310393361
22. Gillette, C., Rudolph, M., Kimble, C., Rockich-Winston, N., Smith, L., Broedel-Zaugg, K.: A meta-analysis of outcomes comparing flipped classroom and lecture. Am. J. Pharm. Educ. **82**(5), 433–440 (2018)
23. Hew, K.F., Lo, C.K.: Flipped classroom improves student learning in health professions education: a meta-analysis. BMC Med. Educ. **18**, 38 (2018)
24. Hu, R., Gao, H., Ye, Y., Ni, Z., Jiang, N., Jiang, X.: Effectiveness of flipped classrooms in Chinese baccalaureate nursing education: a meta-analysis of randomized controlled trials. Int. J. Nurs. Stud. **79**, 94–103 (2018)
25. Lo, C.K., Hew, K.F.: The impact of flipped classrooms on student achievement in engineering education: a meta-analysis of 10 years of research. J. Eng. Edu **108**(4), 523–546 (2019)
26. Lo, C.K., Hew, K.F., Chen, G.: Toward a set of design principles for mathematics flipped classrooms: a synthesis of research in mathematics education. Educ. Res. Rev. **22**, 50–73 (2017)
27. Tan, C., Yue, W.G., Fu, Y.: Effectiveness of flipped classrooms in nursing education: systematic review and meta-analysis. Chin. Nurs. Res. **4**(4), 192–200 (2017)
28. Zhang, S.: A systematic review and meta-analysis on flipped learning in science education. Unpublished Master thesis. The University of Hong Kong (2018)

29. Xu, P., et al.: The effectiveness of a flipped classroom on the development of Chinese nursing students' skill competence: a systematic review and meta-analysis. Nurse Educ. Today **80**, 67–77 (2019)
30. Cohen, J.: Statistical Power Analysis for the Behavioral Sciences, 2nd edn. Lawrence Erlbaum Associates, New York (1988)
31. Hattie, J.: Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement. Routledge, London (2009)