Educational Research Review 22 (2017) 50–73

Contents lists available at ScienceDirect

Educational Research Review

journal homepage: www.elsevier.com/locate/edurev

Review

Toward a set of design principles for mathematics flipped classrooms: A synthesis of research in mathematics education

Chung Kwan Lo*, Khe Foon Hew, Gaowei Chen

Faculty of Education, The University of Hong Kong, Hong Kong

A R T I C L E I N F O

Article history:
Received 7 April 2017
Received in revised form 8 August 2017
Accepted 10 August 2017
Available online 24 August 2017

Keywords:
Flipped classroom
Inverted classroom
Mathematics education
Statistics education
Teaching approach

A B S T R A C T

This paper analyzed the journal publications of mathematics flipped classroom studies in K-12 and higher education contexts. We focused specifically on a set of flipped classroom studies in which pre-class instructional videos were provided prior to face-to-face class meetings. We examined the following four major issues: (a) the types of out-of-class and in-class instructional activities used, (b) the effect of flipped learning on student achievement, (c) the participant perceptions of flipped classroom benefits, and (d) the main challenges of flipped classroom implementations. A meta-analysis of 21 comparison studies showed an overall significant effect in favor of the flipped classroom over the traditional classroom for mathematics education (Hedges' $g = 0.298$, 95% CI [0.16, 0.44]), with no evidence of publication bias. A broader research synthesis of 61 studies revealed that the flipped classroom approach benefited student learning in three main aspects: increasing in-class time for task/practice, integrating new knowledge with existing beliefs, and real-time feedback. The two most frequently reported flipped classroom challenges were students' unfamiliarity with flipped learning and significant start-up effort on the part of instructors. We hence propose a set of design principles to help foster the transition to the flipped classroom and improve the out-of-class and in-class learning designs. This set of design principles can also provide a more focused agenda for future research to examine the effect of the flipped classroom approach on student learning and motivation.

Contents

* Corresponding author.
E-mail addresses: cklohku@hku.hk (C.K. Lo), kfhew@hku.hk (K.F. Hew), gwchen@hku.hks (G. Chen).

http://dx.doi.org/10.1016/j.edurev.2017.08.002

1. Introduction

Mathematics plays an important role in furthering the development of individuals and society (Hoyles, Morgan, & Woodhouse, 1999). Students' performance and learning experience in mathematics is critical in STEM (science, technology, engineering, and mathematics) education (Dove & Dove, 2015; McGivney-Burelle & Xue, 2013; Ogden, 2015; Petrillo, 2016; Van Sickle, 2016). However, Phillips and Phillips (2016) observed that students often experienced high levels of mathematics anxiety or "statisticophobia" (Dillon, 1982). As Petrillo (2016) and Ogden (2015) cautioned, students' negative experiences in learning mathematics could discourage them from choosing careers in STEM fields. Petrillo (2016) added that in some universities, more than 40% of engineering students eventually leave their disciplines largely because of their underperformance in mathematics courses. Weng (2015) highlighted that 70% of the U.S. college students taking remedial mathematics failed the course and thus could not continue their studies. Accordingly, there is a critical need to improve the teaching and learning of mathematics.

Some mathematics instructors (e.g., Ford, 2015; Ichinose & Clinkenbeard, 2016; McBride, 2015) have suggested that the flipped (or inverted) classroom approach has the potential to improve mathematics instruction. This technology-enhanced pedagogy is now frequently used in many undergraduate mathematics and statistics courses (Naccarato & Karakok, 2015). In a typical flipped classroom, instructors deliver lectures before class meetings using instructional videos or other multimedia learning materials (Bishop & Verleger, 2013; Lo & Hew, 2017). The in-class time is then freed up from instructor-centered teaching and can be spent on student-centered learning activities such as collaborative problem solving with the instructor's guidance (Bishop & Verleger, 2013; Giannakos, Krogstie, & Chrisochoides, 2014; O'Flaherty & Phillips, 2015).

With interest growing in the flipped classroom approach to mathematics education, practitioners want to know its impact in addition to the benefits and challenges of using this instructional approach (Naccarato & Karakok, 2015). Although several systematic review studies have been undertaken (e.g., Betihavas, Bridgman, Kornhaber, & Cross, 2016; Bishop & Verleger, 2013; Chua & Lateef, 2014; Giannakos et al., 2014; Lo & Hew, 2017; O'Flaherty & Phillips, 2015; Presti, 2016; Seery, 2015), none have specifically focused on the effects of flipped learning on student mathematics achievement using a meta-analytic approach.

Several articles (e.g., Ford, 2015; Gundlach, Richards, Nelson, & Levesque-Bristol, 2015; Overmyer, 2015; Yong, Levy, & Lape, 2015) have revealed that some mathematics educators lack the knowledge or experience to use the flipped classroom approach. Inexperienced instructors tend to design their flipped classrooms based on their intuitive beliefs, which can affect the efficacy of the approach (Overmyer, 2015; Yong et al., 2015). A recent study by Ronau et al. (2015) evaluated 1165 scholarly papers pertaining to the use of technology in mathematics education and found that only 28% of them (i.e., dissertations, journal articles, and other publications) were connected to empirically driven principles. This missing link could become a major obstacle to the success of flipped mathematics learning. What, then, are some empirically based design principles that mathematics instructors can use to design their flipped courses?

Other resources have been dedicated to developing design principles for the flipped classroom approach, including the Flipped Learning Network (http://flippedlearning.org), Flipped Learning Global Initiative (http://flglobal.org), and Bergmann and Sams's (2012) book on flipped learning. Although these resources have made significant contributions to flipped learning, not all of their recommendations have been grounded in the recent empirical studies on mathematics flipped classrooms. At the time of writing, we could only find Kim, Kim, Khera, and Getman's (2014) study that attempted to identify a set of design principles for the flipped classroom approach. However, the design principles were somewhat limited because they were based on the single context of one urban American university with only 41 students and three instructors. More importantly, their principles were developed outside the context of mathematics education. To advance the flipped classroom approach, there is a need to establish a set of design principles grounded in empirical research on mathematics education.

2. The purpose of this review

This review first examines the types of out-of-class and in-class instructional activities used in mathematics flipped classrooms. We then examine student achievement in mathematics flipped classrooms compared to traditional classrooms.

By the traditional classroom approach, we refer to an environment in which students attend class in which instructors use a range of strategies such as lectures, student group work and presentations; although a portion of in-class time may be spent doing homework, students complete most of their homework after school due to a packed lecture schedule (Dove & Dove, 2015; Jungić, Kaur, Mulholland, & Xin, 2015). Based on the comments of the instructors and students, we also investigate how the flipped classroom approach benefits student learning in mathematics and identify the challenge to flipping a mathematics course. The overarching goal of this review is to develop a set of empirically based design principles for mathematics flipped classrooms. The following specific questions guided our review:

1. What types of instructional activities are used outside and inside mathematics flipped classrooms?
2. What is the effect of mathematics flipped classrooms on student achievement compared to their traditional counterparts?
3. How does the flipped classroom approach benefit student learning in mathematics courses?
4. What are the challenges to implementing mathematics flipped classrooms?

3. Definition of the flipped classroom approach

EDUCAUSE, one of the leading associations focusing on instructional technology in higher education, defined the flipped classroom approach as "a pedagogical model in which the typical lecture and homework elements of a course are reversed … Short video lectures are viewed by students at home before the class session, while in-class time is devoted to exercises, projects, or discussions" (EDUCAUSE, 2012, p. 1). In a highly cited paper, Bishop and Verleger (2013) defined the flipped classroom approach as "an educational technique that consists of two parts: interactive group learning activities inside the classroom, and direct computer-based individual instruction outside the classroom" (p. 5). Essentially, the flipped classroom approach (Fig. 1) consists of some form of pre-class activity (e.g., viewing videos) before class meetings and complete individual or group activities during face-to-face lessons (Abeysekera & Dawson, 2015; Dove & Dove, 2015; He, Holton, Farkas, & Warschauer, 2016; Jungić et al., 2015). He et al. (2016) argued that a classroom "is not genuinely flipped" (p. 62) unless face-to-face class attendance is mandatory because in-class instruction is vitally important to promoting student learning.

There is currently no standard practice for the flipped classroom approach (Guerrero, Beal, Lamb, Sonderegger, & Baumgartel, 2015; Larsen, 2015; Love, Hodge, Grandgenett, & Swift, 2014; Ziegelmeier & Topaz, 2015). Naccarato and Karakok (2015) interviewed 19 mathematics instructors from 14 higher institutions and found that although all of the instructors provided pre-class video lectures, there were many different approaches to in-class activities, including a combination of short quizzes at the start of lessons, reviews of video lectures, small-group or large-group discussions, student presentations, and application projects. Some mathematics instructors still reserved in-class time for traditional lecture-

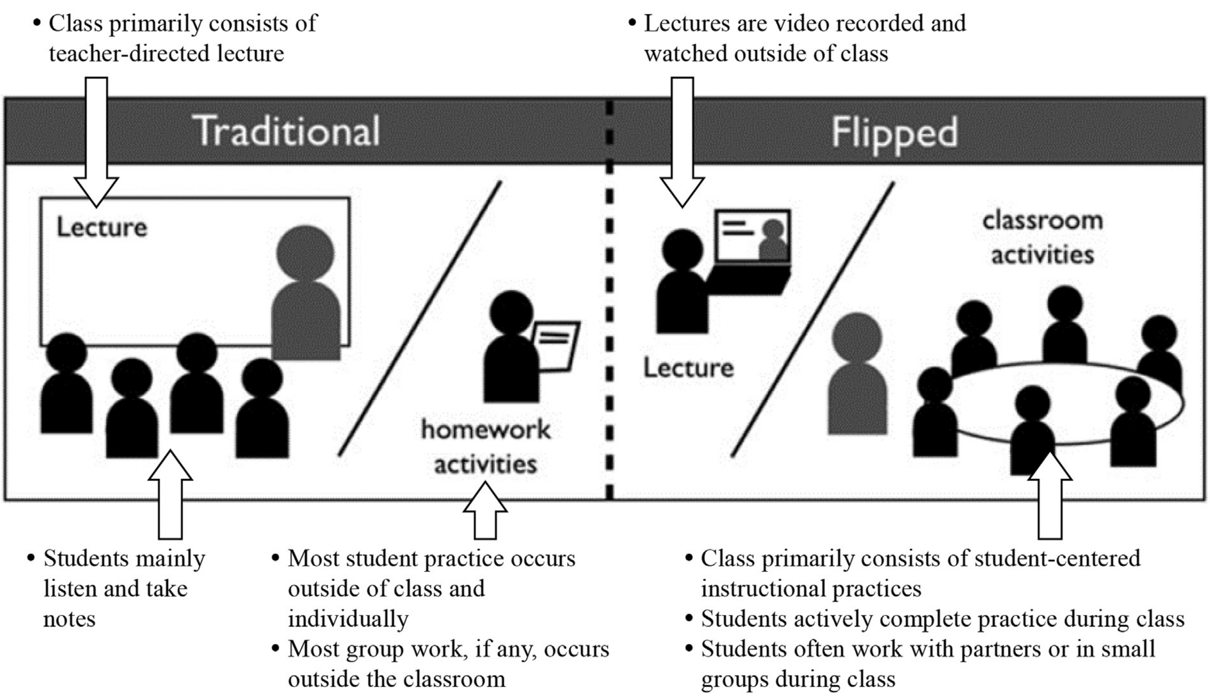

Fig. 1. A comparison of the traditional classroom and the flipped classroom approaches (Dove & Dove, 2015, p. 169; Jungić et al., 2015, p. 509).

based instruction in their flipped courses. Despite the variation of in-class learning activities, it appears that no mathematics instructors regarded the diversity of these activities as violations of the flipped classroom approach.

The major contention of both EDUCAUSE (2012) and Bishop and Verleger's (2013) definitions is the requirement that instructional videos be used in out-of-class learning sessions (Bernard, 2015). This definition excludes the *sole* use of pre-class reading materials as a form of flipped learning. Some educators, however, have objected to this constraint, arguing that "qualifying instructional medium is unnecessary and unjustified" (He et al., 2016, p. 61). The Flipped Learning Network (2014) offers the following definition:

> Flipped Learning is a pedagogical approach in which direct instruction moves from the group learning space to the individual learning space, and the resulting group space is transformed into a dynamic, interactive learning environment where the educator guides students as they apply concepts and engage creatively in the subject matter. (p. 1)

Without specifying the instructors' technological choices, this definition defines flipped learning in terms of student activity relative to whether they are acting individually or in groups. Although a wide-range of pre-class materials can be used under this definition, we believe that text-based materials cannot "closely mimic what students in a traditional setting would experience" (Love, Hodge, Corritore, & Ernst, 2015, p. 749) given that the use of these materials does not involve instructors' explanations or elaboration of the content (Lo & Hew, 2017). As a student of Muir and Geiger (2016) commented, "[a] book doesn't really walk through the steps on how to do something" (p. 164). In contrast, the use of instructional videos enables the instructors to elaborate on course content just as they would in a traditional lecture (Lo & Hew, 2017; Muir & Geiger, 2016).

Following the definitions of EDUCAUSE (2012) and Bishop and Verleger (2013), the scope of our discussion is confined to flipped classrooms that have used video-recorded lectures when shifting lecture-based instruction outside the classroom. We include various commonly used video styles (see Chen & Wu, 2015; Chorianopoulos & Giannakos, 2013; Guo, Kim, & Rubin, 2014), such as instructor-created classroom lectures, YouTube, Khan Academy, TED talks, screencast, PowerPoints with instructor talking head, and PowerPoints with instructor voiceover. For in-class learning, attendance is mandatory (He et al., 2016), and we include a variety of learning activities in the definition. We define the flipped classroom approach as a technology-enhanced pedagogy that delivers parts of the course materials through video resources before class, followed by the integrated use of assessments, mini-lectures, individual problem solving, and small-group learning activities inside the classroom.

4. Method

4.1. Search strategies

The process of selecting relevant literature followed the preferred-reporting of items for systematic reviews and meta-analyses (PRISMA) statement (Moher, Liberati, Tetzlaff, & Altman, 2009). To be as comprehensive as possible, the following seven electronic databases were searched: (1) Academic Search Complete, (2) British Education Index, (3) Education Research Complete, (4) ERIC, (5) Library, Information Science & Technology Abstracts, (6) Teacher Reference Center, and (7) TOC Premier. The search terms with Boolean operators used in this review were as follows: (flip* OR invert*) AND (class* OR learn*) AND (math* OR algebra OR trigonometry OR geometry OR calculus OR statistics). The asterisk was used as a wild card to include most of the common expressions of mathematics flipped classrooms (e.g., inverted classroom, flipped learning, flipping statistics class).

4.2. Inclusion and exclusion criteria

Based on the existing flipped classroom reviews, the inclusion and exclusion criteria to be studied were developed (Table 1). To be included in this review, the flipped mathematics courses must have satisfied the aforementioned definition of the flipped classroom approach. In other words, we excluded studies that did not provide video instructional materials for students' class preparation or did not offer regular face-to-face lecture hours (Lo & Hew, 2017). Given that little flipped

Table 1
Inclusion and exclusion criteria for selection.

Criterion	Inclusion	Exclusion
Definition of the flipped classroom approach	Flipped classrooms should at least include (1) the use of video instructional materials for students' class preparation and (2) regular face-to-face class meetings.	Flipped classrooms that did not use multimedia learning materials in out-of-class learning activities or cancel regular lecture hours after flipping.
Subject area	All content areas in mathematics education such as algebra, trigonometry, geometry, calculus, and statistics.	All subject disciplines other than mathematics.
Time period	January 2012 to December 2016 (five years).	Studies outside the time period.
Type of article	Empirical research published in peer-reviewed journals.	Studies that were not peer-reviewed; non-empirical studies or articles that provided little empirical evidence.
Language	English.	Non-English reported studies.

classroom research was published before 2012 (Giannakos et al., 2014), the period of our search was from January 2012 to December 2016 (i.e., the most recent five years).

No constraints were imposed on the education contexts (e.g., K-12 or higher education); however, the studies must have been based on empirical research focused on the implementation of mathematics flipped classrooms. Non-empirical studies or articles that provided little empirical evidence were excluded (Betihavas et al., 2016; O'Flaherty & Phillips, 2015). No constraints were imposed on the language of instruction or the location of the studies; however, the manuscripts must have been written in English and published in peer-reviewed journals (Betihavas et al., 2016; Lo & Hew, 2017; O'Flaherty & Phillips, 2015; Seery, 2015). Peer review is a useful criterion for selecting studies of sufficient quality (Korpershoek, Harms, de Boer, van Kuijk, & Doolaard, 2016).

In the next stage, we selected eligible studies for further quantitative meta-analysis. To be included in the meta-analysis, selected studies must satisfy the following more restrictive criteria:

(a) Studies must have compared student achievement under a flipped classroom approach and a traditional classroom approach.
(b) Student achievement (i.e., the dependent measure) must have been based on some objective quantitative measures of mathematics performance such as post-tests, final exams, or other standardized tests. Such objective measures have commonly been used by other authors of meta-analytic studies (e.g. Li & Ma, 2010; Rakes, Valentine, McGatha, & Ronau, 2010; Zheng, Warschauer, Lin, & Chang, 2016). Self-reported measures such as survey data on self-perceived learning were excluded.
(b) Studies must have provided enough statistical data to compute the effect size.
(c) Studies that found a significant difference between students' initial knowledge must have used appropriate statistical tests to control for this difference and provided the adjusted means data.

4.3. Data extraction and analysis

The first two authors contributed to the extraction and categorization of the data. The data included author(s), year of publication, research context, instructional activities (for RQ1), results of student achievement (for RQ2), how the flipped classroom approach benefited student learning (for RQ3), and challenges of the flipped courses (for RQ4). For student achievement, we adopted a meta-analytic approach. Based on the comments of the instructors and students, we then synthesized the benefits and challenges using a thematic analysis.

To analyze participant perceptions of flipped classroom benefits, we drew upon the work of Kuiper, Carver, Posner, and Everson (2015). Relying on the cognitive learning theory (Lovett & Greenhouse, 2000), they stated that (1) sufficient time on task/practice, (2) integrating new knowledge with existing beliefs, and (3) real-time feedback are factors that contribute to learning. In this review, we adopted the framework of Kuiper et al. (2015) for the initial coding of flipped classroom benefits.

As for the flipped classroom challenges, the general framework for coding followed the three categories defined by Betihavas et al. (2016): (1) student-related challenges, (2) faculty challenges, and (3) operational challenges. This framework covered all major aspects of flipped learning and was used in other review studies (e.g., Lo & Hew, 2017). However, it is important to note that our analysis could only focus on what the authors reported in their articles. The absence of a theme did not necessarily imply the absence of a specific dimension; instead, it only indicated that the authors did not explicitly report that aspect in their articles. To establish coding reliability, 20% of the articles were randomly selected and coded by the first two authors. Inter-coder reliability was high (93%). In the event of disagreements, the two authors re-examined the studies in question together to come to a consensus.

5. Results

5.1. Study selection

Using the search terms, a total of 1469 peer-reviewed journal articles (published from January 2012 to December 2016) were found by early February 2017 (the time of writing). Some articles were removed due to replication across databases. Our search terms enabled us to capture various relevant studies. However, after reviewing their titles and abstracts, many articles in the search outcomes were found to be irrelevant, particularly those that did not report on empirical research or were not related to mathematics education. For example, many irrelevant search outcomes came from engineering (e.g., electric inverters).

Finally, 72 full-text articles were assessed for eligibility. Eleven were excluded: five studies did not provide pre-class videos, one did not mention the provision of face-to-face lessons, three reported minimal empirical data, and two (i.e., Eisenhut & Taylor, 2015; Naccarato & Karakok, 2015) merely interviewed mathematics instructors without reporting the details of their flipped courses. Nevertheless, some of the removed literature was used for background reference. The final selection yielded a total of 61 articles for qualitative synthesis (see Appendix A) and 21 articles for meta-analysis. Fig. 2 outlines the process of article selection.

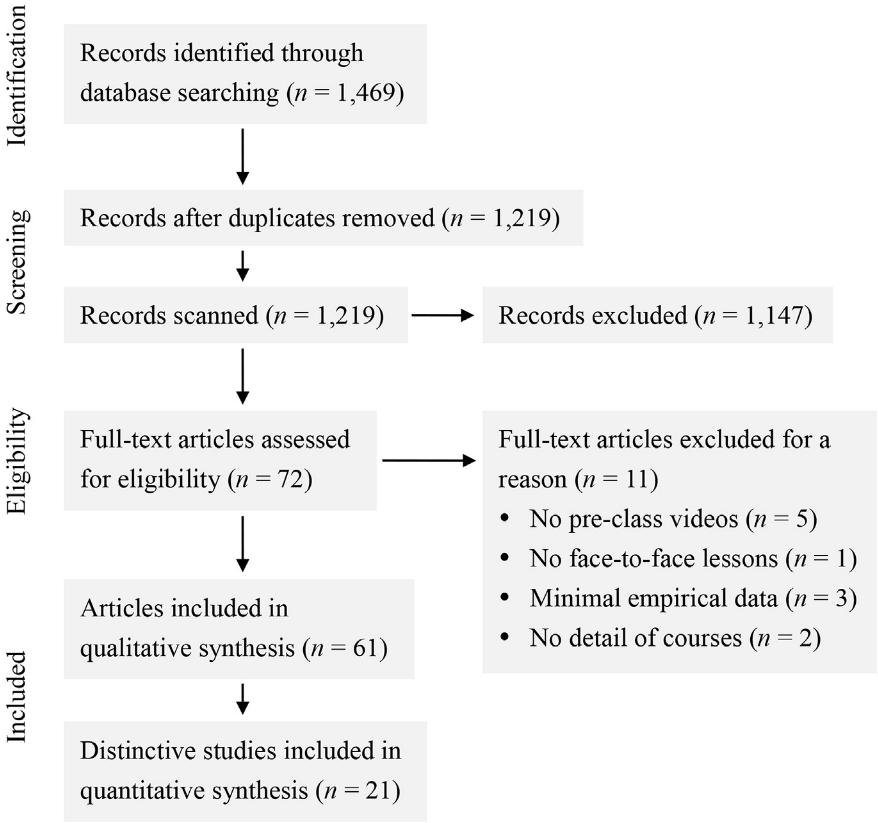

Fig. 2. PRISMA flow diagram of article selection.

5.2. Characteristics of the reviewed studies

The majority of the studies were conducted in the United States ($n = 51$), followed by Taiwan ($n = 4$) and Canada ($n = 3$). There was a scattering of studies from Australia ($n = 1$), Germany ($n = 1$), and New Zealand ($n = 1$). Most of the studies were carried out at the undergraduate level ($n = 50$), with a notable paucity of research in secondary ($n = 7$), elementary ($n = 2$), doctorial ($n = 1$), and adult ($n = 1$) education. Table 2 shows that among the 61 articles, 59 flipped classroom studies were reported. A total of 21 studies provided sufficient data for meta-analysis (RQ2). Grouping the findings on the same interventions in flipped classrooms (see Love et al. (2014, 2015) and Van Sickle (2015, 2016) for reviews) yielded a total of 59 distinct studies for the thematic analysis of flipped classroom benefits (RQ3) and challenges (RQ4). Eight articles (i.e., Braun, Ritter, & Vasko, 2014; Carney, Ormes, & Swanson, 2015; Jungić et al., 2015; Kuiper et al., 2015; McBride, 2015; McCallum, Schultz, Sellke, & Spartz, 2015; Schroeder, Xue, & McGivney, 2013; Zack, Fuselier, Graham-Squire, Lamb, & O'Hara, 2015) reported more than one flipped course. However, some of these courses were not in mathematics (e.g., McCallum et al., 2015) or did not satisfy the definition of the flipped classroom approach (e.g., Kuiper et al., 2015). A total of 72 flipped mathematics or statistics courses were involved in the analysis of instructional activities (RQ1). Most of the courses (84.7%) were undergraduate courses; secondary and elementary flipped classrooms only constituted 9.7% and 2.8% of the courses, respectively. Fig. 3 shows that calculus (43.1%), statistics (19.4%), and algebra (19.4%) were the three most common content areas of the flipped courses.

Table 2
Number of studies and flipped courses for data analysis and the related research questions.

	Count	Data analysis and research questions
Total number of publications reviewed	61	
◆ Number of distinct studies	59	Analysis of flipped classroom benefits (RQ3) and challenges (RQ4)
◆ Number of distinct studies providing sufficient data for a meta-analysis on student achievement	21	Meta-analysis of student achievement (RQ2)
Total number of flipped courses involved in the reviewed studies	72	Analysis of instructional activities (RQ1)

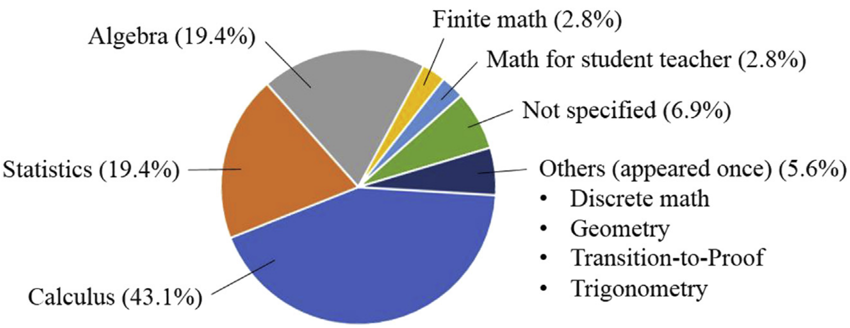

Fig. 3. Content areas of the flipped courses ($n = 72$).

5.3. RQ1: instructional activities of mathematics flipped classrooms

The instructional activities of the 72 flipped courses were analyzed (Fig. 4). All of the flipped courses provided instructional videos for students' out-of-class learning, and instructors from 52 flipped courses produced their own videos for students. Some instructors used existing videos such as those from the Khan Academy ($n = 4$) or textbook publishers (e.g., Cilli-Turner, 2015), or they made multiple types of videos available to their students (e.g., Scott, Green, & Etheridge, 2016; Van Sickle, 2015; Wright, 2015). The instructors in 13 flipped courses provided worksheets or content notes to guide students' video watching. In addition to the instructional videos, text-based materials (e.g., textbooks, readings, notes) were assigned as supplements in 22 of the flipped courses, and online quizzes were provided in 39 flipped courses. Only 10 flipped courses provided online discussion through online discussion boards or other technologies such as web-based media players (Schroeder & Dorn, 2016).

For in-class learning, the instructors in 23 of the flipped courses provided structured formative assessments such as quizzes at the start of face-to-face lessons. Small-group activities ($n = 69$) were generally used in the 72 flipped courses. Some specific approaches to small-group activities included Crouch and Mazur's (2001) peer instruction (e.g., Jungić et al., 2015; Phillips & Phillips, 2016; Saumier, 2016; Talbert, 2014; Touchton, 2015), cooperative learning (e.g., Chen, Chen, & Chen, 2015; Ogden, 2015; Overmyer, 2015), and a team-based approach (Carney et al., 2015). Some instructors also used individual practices ($n = 26$), student presentations ($n = 12$), or quizzes near the end of the lessons ($n = 3$) inside their flipped classrooms. Instructors in 14 flipped courses delivered in-class lectures on new materials. In other words, part of the course material was still presented inside the classroom.

5.4. RQ2: effect of mathematics flipped classrooms on student achievement

In this section, we report our findings on whether flipped learning improved student mathematics achievement compared to the traditional classroom. We previously defined the traditional classroom as one in which students come to class, where teachers use a range of strategies such as lectures, student group work and presentations, and then the students complete most of their homework after school (Dove & Dove, 2015; Jungić et al., 2015). A total of 21 articles yielding 22 independent effect sizes were meta-analyzed using student performance data (e.g., exam scores) as the outcome variable. The first and second authors independently checked the data, and any concerns or differences were resolved through discussion. These 22 effect sizes involved 3184 participants. Our coding protocol was as follows.

Coding: multiple assessments. In cases where articles reported multiple assessments, we selected the assessment that was most summative. By summative assessment, we mean *the most comprehensive assessment* such as course final exams. Course final exams were used because they encompassed all of the course content that students learned throughout the entire semester, unlike mid-terms or weekly tests. The most comprehensive assessment in our meta-analysis was consistent with other researchers' practice (e.g., Freeman et al., 2014).

Coding: student initial equivalence. To determine students' initial equivalence, we examined whether the comparison design was based on the following categories: (a) treatment-and-control studies in which the authors provided no data on student equivalence in terms of initial academic performance, or where the authors claimed that the students were equivalent but failed to provide a relevant statistical test; or (b) treatment-and-control studies in which the data indicated no statistical difference on a pre-test that was directly related to the topic, or on a metric of academic performance (e.g., college GPA), or where any significant difference on a pre-test had been handled by using a statistical test and providing adjusted data.

Coding: instructor equivalence. Gundlach et al. (2015) cautioned that “the difference in student performance may depend on the particular instructor” (p. 24). In some studies (e.g., Maciejewski, 2016; Scott et al., 2016; Van Sickle, 2015), the flipped and the traditional classes were taught by different instructors. For example, the traditional classes in Maciejewski's (2016) study were taught by two graduate students and a postdoctoral research fellow, but the flipped classes were taught by two

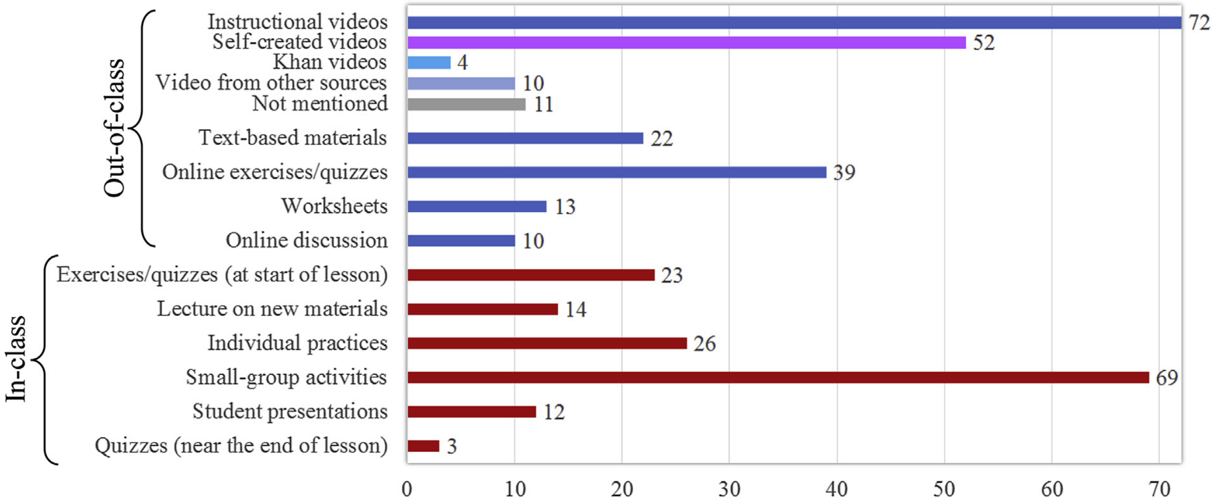

Fig. 4. Instructional activities of mathematics flipped classrooms ($n = 72$).

professors and a postdoctoral research fellow. Although Maciejewski (2016) argued that the graduate student instructors had completed their teacher training, the concern over instructors with different levels of experience highlighted the need to conduct a heterogeneity analysis on instructor equivalence. To evaluate instructor equivalence, we checked whether the study involved: (a) identical instructors for both the flipped and the traditional classes; or (b) different instructors for different classes, or if no data were provided.

We computed effect sizes using the comprehensive meta-analysis software package. All reported *p* values were two-tailed unless otherwise reported. To compare the effect sizes, we used a random effects model or random effects analysis (Gurevitch & Hedges, 1999) because conditions that could affect student achievement differed among studies in the analysis, including the frequency of the lessons flipped, student population, and course level. We computed effect sizes as Hedges' *g* from the means and standard deviations of the student achievement data (e.g., exam scores, post-test scores). If standard errors but not standard deviations were used in the empirical studies, we used the following formula (Altman & Bland, 2005) to calculate the standard deviations:

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

If the means and standard deviations were not reported in the previous empirical studies, the standardized mean difference was estimated using a variety of sources, including *t*-tests (or formulas; see Borenstein, Hedges, Higgins, & Rothstein, 2009; Lipsey & Wilson, 2001). Overall, as Fig. 5 indicates, the results show significant evidence in terms of student performance in favor of the flipped classroom (Hedges' $g = 0.298$, 95% CI [0.16, 0.44], $Z = 4.186$, $p < 0.001$).

In Table 3, the heterogeneity analyses show no statistically significant variation among studies based on specific content areas ($Q = 4.951$, $df = 6$, $p = 0.550$). Thus, the data suggest that the flipped classroom approach increased student performance across the various content areas of mathematics.

We also looked for heterogeneity in the effect sizes of the student performance data based on whether the studies controlled for student or instructor equivalence (Table 4). We found no evidence of heterogeneity between the studies that reported initial student equivalence and the studies that did not provide such data ($Q = 2.316$, $df = 1$, $p = 0.128$). Analyzing variation with respect to instructor equivalence also suggested no evidence of heterogeneity ($Q = 0.159$, $df = 1$, $p = 0.690$). Thus, the overall effect size for the student performance data appears to be robust in terms of the varying methodological rigor of the published studies (e.g., less well-controlled studies with different instructors or no data provided on student or instructor equivalence).

The heterogeneity analysis shown in Table 5 indicates that the effect size was significantly higher when instructors used a structured formative assessment such as a quiz at the *start* of face-to-face lessons to assess students' pre-class learning compared to instructors who did not ($Q = 6.199$, $df = 1$, $p = 0.013$). By the structured formative assessment, we mean specific questions that were developed by the instructor beforehand and used to assess student learning based on pre-class learning. These included questions specific to the learning items presented in the video lectures that had been answered by the students at the start of class (e.g., Adams & Dove, 2016); open-note quizzes given at the start of class asking students to write a definition, give a formula, explain a notation (e.g., Cilli-Turner, 2015), or solve a problem similar to the one given in the pre-class videos (e.g., Schroeder, McGivney-Burelle, & Xue, 2015); or multiple-choice questions using electronic devices (e.g., Clickers) at the start of class (e.g., Scott et al., 2016). The data suggested the importance of activating students' pre-class

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
	Statistics for each study						
Adams & Dove (2016)	0.051	0.335	0.112	-0.606	0.708	0.152	0.879
Anderson & Brennan (2015)	0.482	0.117	0.014	0.253	0.710	4.127	0.000
Bhagat et al. (2016)-high	0.054	0.352	0.124	-0.635	0.743	0.155	0.877
Bhagat et al. (2016)-low	0.591	0.484	0.234	-0.358	1.540	1.221	0.222
Braun et al. (2014)	0.018	0.124	0.015	-0.225	0.261	0.145	0.885
Cilli-Turner (2015)	1.139	0.188	0.035	0.772	1.507	6.075	0.000
Clark (2015)	0.033	0.219	0.048	-0.396	0.462	0.152	0.879
DeSantis et al. (2015)	-0.120	0.289	0.083	-0.686	0.446	-0.417	0.677
Dove & Dove (2015)	0.744	0.262	0.068	0.231	1.256	2.843	0.004
Guerrero et al. (2015)	0.294	0.242	0.059	-0.180	0.769	1.215	0.224
Gundlach et al. (2015)	-0.163	0.212	0.045	-0.579	0.252	-0.771	0.441
Kenneddy et al. (2015)	-0.114	0.161	0.026	-0.429	0.202	-0.707	0.480
Kirvan et al. (2015)	0.845	0.281	0.079	0.294	1.396	3.007	0.003
Maciejewski (2016)	0.389	0.083	0.007	0.225	0.552	4.663	0.000
Murphy et al. (2016)	0.629	0.249	0.062	0.140	1.118	2.523	0.012
Overmyer (2015)	0.220	0.116	0.013	-0.007	0.448	1.902	0.057
Peterson (2016)	0.731	0.312	0.097	0.120	1.342	2.347	0.019
Petrillo (2016)	0.406	0.163	0.026	0.087	0.724	2.496	0.013
Schroeder et al. (2015)	0.313	0.190	0.036	-0.060	0.686	1.645	0.100
Scott et al. (2016)	0.032	0.217	0.047	-0.393	0.457	0.146	0.884
Van Sickle (2015)	0.427	0.190	0.036	0.055	0.800	2.250	0.024
Yong et al. (2015)	-0.096	0.150	0.023	-0.391	0.198	-0.640	0.522
	0.298	0.071	0.005	0.159	0.438	4.186	0.000

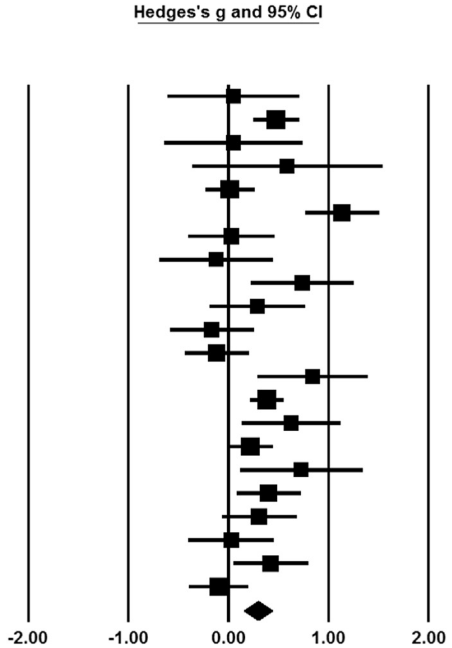

Fig. 5. Forest plot of effect sizes (Hedges' *g*) using the random effect model.

learning by implementing a structured formative assessment of the out-of-class learning content at the *beginning* of face-to-face lessons. This helps instructors quickly identify students' misconceptions of the pre-class materials and provide remedial action if necessary (Kirvan, Rakes, & Zamora, 2015; Talbert, 2015).

We also examined our data corpus of 21 studies to see whether the previous studies explicitly *described* and *assessed* the type of questions used (e.g., near transfer and far transfer questions). However, our efforts were hampered because most of the studies ($n = 17$) studies did not clearly describe their types of questions. In the following section, we briefly report on two studies (i.e., Kennedy, Beaudrie, Ernst, & St. Laurent, 2015; Kirvan et al., 2015) that provided clear descriptions of the types of questions used and another two studies (i.e., Gundlach et al., 2015; Maciejewski, 2016) that used some validated assessment tools (e.g., Calculus Concept Inventory) in their summative assessments.

In their undergraduate calculus course, Kennedy et al. (2015) reported data on their overall final exam together with the types of the questions used in the exam. The exam was comprised of computational and conceptual problems. For the *computational* problems, students were expected to use a known formula or algorithm to solve problems such as "Given the formula for arc length, find the length of the curve $y = 2x^{3/2}/3$ for $0 \leq x \leq 1$" (p. 897). The *Conceptual* problems required knowledge from a known formula or algorithm to be applied or extended, such as "If a car leaves at noon and travels along the curve $y = 2x^{3/2}/3$, at what time has the car gone a distance of $2(2^{3/2}-1)/3$ miles?" (p. 897). The conceptual problems required students to recognize that the formula for arc length was needed, whereas in the computational problem, students were directly told to use the formula for arc length (Kennedy et al., 2015). Kennedy et al. (2015) found that the levels of student achievement for both the computational questions (flipped class: $M = 79.9$, $SE = 2.0$; traditional class: $M = 82.1$, $SE = 2.0$, $p = 0.44$) and the conceptual questions (flipped class: $M = 77.0$, $SE = 2.1$; traditional class: $M = 78.7$, $SE = 2.1$, $p = 0.58$) were similar for both the flipped class and the traditional class.

In their high school algebra course, Kirvan et al. (2015) reported data on three types of questions: *analyze* (examine the meaning of constituent parts of a system and/or context), *model* (develop a system to represent a situation), and *solve* (compute solutions). Their results showed that students in the flipped classroom ($M = 0.402$, $SD = 0.136$) had significantly higher *solve* scores than those in the traditional classroom ($M = 0.284$, $SD = 0.139$), $t(52) = 3.14$, $p = 0.0028$. However, no significant difference was found with the *analyze* scores (flipped class: $M = 0.179$, $SD = 0.169$; traditional class: $M = 0.123$, $SD = 0.126$, $t(52) = 1.39$, $p = 0.169$) or *model* scores (flipped class: $M = 1.267$, $SD = 0.329$; traditional class: $M = 1.162$, $SD = 0.331$, $t(52) = 1.166$, $p = 0.249$).

Two other studies (i.e., Gundlach et al., 2015; Maciejewski, 2016) conducted assessments using validated concept inventories, namely the calculus concept inventory (Epstein, 2013) and the statistical reasoning assessment instrument (Garfield, 2003). Gundlach et al. (2015) found a non-significant difference between the flipped class ($M = 1.92$, $SD = 0.89$) and the traditional class ($M = 1.79$, $SD = 0.84$), $t(216) = 0.72$, $p = 0.47$ in terms of students' misconceptions of statistical concepts using the statistical reasoning assessment instrument. Administering the calculus concept inventory test, Maciejewski (2016) found that the flipped class ($M = 52.16$, $SD = 16.71$) significantly outperformed the traditional class ($M = 49.17$, 17.93), $t(275) = 1.76$, $p = 0.04$, $d = 0.18$.

Table 3
Comparing effect sizes among different content areas.

Content area	n	Hedges' g	SE	95% confidence interval	
				Lower limit	Upper limit
Algebra	6	0.316	0.136	0.049	0.584
Calculus	8	0.203	0.114	−0.021	0.427
Finite math	1	0.294	0.364	−0.420	1.008
Geometry	1	−0.120	0.397	−0.898	0.657
Trigonometry	2	0.264	0.347	−0.416	0.944
Math for preservice teachers	1	0.744	0.378	0.004	1.484
Statistics	3	0.570	0.205	0.167	0.972

Table 4
Comparing the effect sizes from well-controlled versus less well-controlled studies.

Type of control	n	Hedges' g	SE	95% confidence interval	
				Lower limit	Upper limit
For student equivalence					
No data provided	10	0.409	0.102	0.209	0.610
No statistical difference on pre-tests or other metric scores	12	0.191	0.101	−0.007	0.388
For instructor equivalence					
No data, or different instructors	8	0.332	0.111	0.115	0.550
Identical instructor	14	0.274	0.097	0.084	0.463

Table 5
Comparing the effect sizes of studies using structured formative assessments at the start of face-to-face lessons versus studies that did not.

Structured formative assessment	n	Hedges' g	SE	95% confidence interval	
				Lower limit	Upper limit
With a structured formative assessment (e.g., quiz) at the start of face-to-face lessons	6	0.572*	0.129	0.319	0.824
Without a structured formative assessment, or no data provided	16	0.198	0.077	0.048	0.349

*$p < 0.05$.

To sum up, a meta-analysis of 21 studies showed an overall significant effect in favor of the flipped classroom over the traditional classroom for mathematics education (Hedges' $g = 0.298$, 95% CI [0.16, 0.44]). However, our attempt to further investigate the possible effects of flipped learning on specific types of questions was hampered by the fact that most studies did not clearly describe the types of questions used in their method sections. We found only two studies that provided clear descriptions of their question types and two studies that used validated assessment tools in addition to their final exams.

So far, the effects of flipped learning on the types of question have been mixed. Some studies (Gundlach et al., 2015; Kennedy et al., 2015) have reported that students in both flipped and traditional classrooms performed equally well on conceptual questions. However, Maciejewski's (2016) flipped classroom students did better than his traditional classroom students on the calculus concept inventory test. For questions that required students to compute solutions, there was also no clear conclusion. Kennedy et al. (2015) found that their flipped and traditional classes did equally well on computational problems. Kirvan et al. (2015), however, reported that their flipped classroom students showed greater improvement in their ability to compute solutions to linear equations. Unfortunately, the small sample of studies limits the generalizability of the results on question types. None of the previous studies explicitly described and assessed the effects of flipped mathematics learning on near transfers (for similar problems that were sufficiently different from the practice problems) versus far transfers (for problems that are entirely different from the practice problems). We therefore urge future research to include the variable of question type in examining how flipped learning may benefit students.

To evaluate the possibility of publication bias in this review, we performed four analyses: (a) assessing the funnel plot, (b) computing Begg and Mazumdar rank correlation, (c) calculating Egger's regression, and (d) calculating the classic fail-safe N test. Visual inspection of Fig. 6 suggests there is no presence of publication bias. This is supported by two statistical analyses: Kendall's Tau with continuity correction 0.10, one-tailed $p = 0.249$; and Egger's regression intercept 0.208, one-tailed $p = 0.415$.

We also conducted a classic fail-safe N test to determine the number of null effect studies needed to raise the p value associated with the mean effect above an arbitrary alpha level ($\alpha = 0.05$). The results showed that 289 additional missing studies with zero mean effect size would be required to make the overall effect statistically insignificant. There would therefore have to be an unreasonably large number of undetected studies with zero effect to bring the effect sizes reported in this review to values that might be statistically insignificant. Based on the visual inspection of funnel plot, statistical analyses, and fail-safe Ns, we believe that our overall mean effect size is not inflated by publication bias.

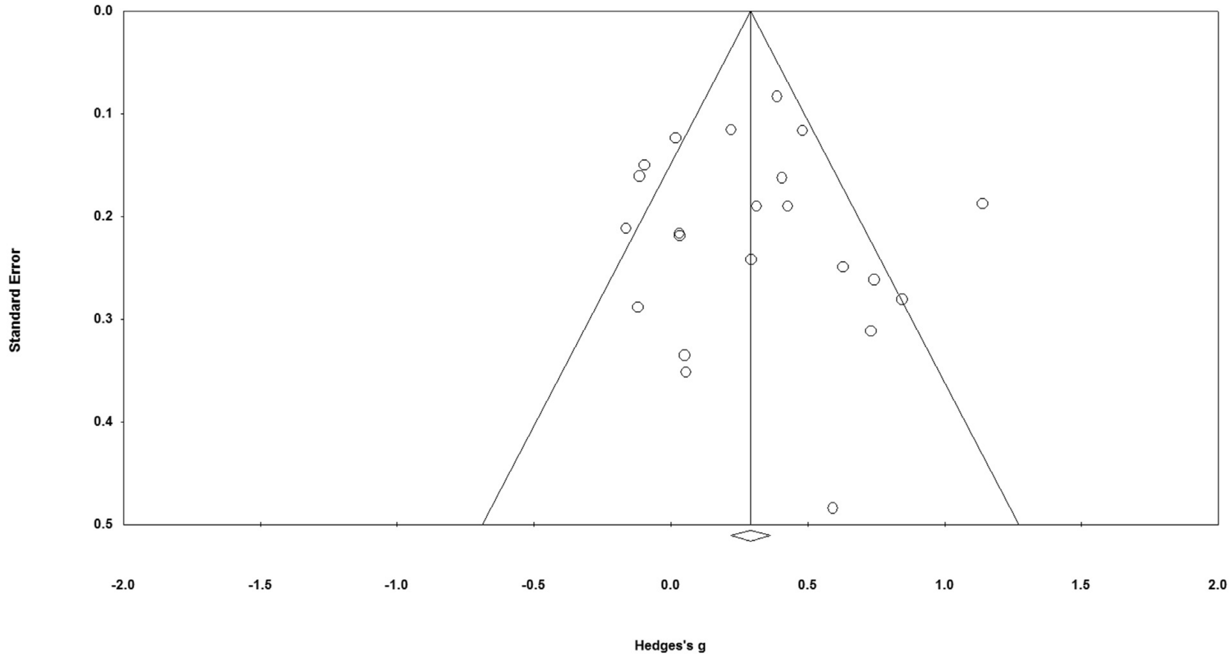

Fig. 6. Funnel plot of standard error by Hedges' g.

5.5. RQ3: how mathematics flipped classrooms benefit student learning

The top three most frequently reported benefits of flipped learning were instructor feedback ($n = 40$), peer-assisted learning ($n = 33$), and more in-class time to apply concepts during activities ($n = 30$). These three benefits pertained to the in-class learning experiences of flipped classrooms. Overall, the flipped classroom approach helped to increase the students' interactions with their instructors and classmates during in-class sessions. Many of the in-class activities, such as group discussions, promoted the students' interactions with their peers (e.g., Jungić et al., 2015; Kennedy et al., 2015; McCallum et al., 2015). Instructors also indicated that they had greater opportunities to provide students with more feedback during in-class sessions (e.g., Bhagat, Chang, & Chang, 2016; Clark, 2015; Guerrero et al., 2015). Further, there were more opportunities for students to apply their knowledge to solve problems in flipped classes (e.g., Kirvan et al., 2015; Overmyer, 2015; Ziegelmeier & Topaz, 2015).

Another two commonly reported benefits were the on-demand accessibility of video lectures ($n = 27$) and preparing students for class ($n = 17$). The students reported that watching video lectures before class helped to prepare them for class activities (e.g., Kraut, 2015; McGivney-Burelle & Xue, 2013; Schroeder et al., 2013). They had time to review the videos as often as they wished (e.g., Grypp & Luebeck, 2015; Love et al., 2014; Tawfik & Lilly, 2015). Some students also revisited the videos when they encountered problems during the problem-solving process (Tawfik & Lilly, 2015) or when they studied for exams (Zack et al., 2015). Two less frequently stated benefits were that the flipped classroom approach allowed some instructors to use differentiated instructional activities in the classroom ($n = 10$) (e.g., Kirvan et al., 2015; McBride, 2015; Muir & Geiger, 2016) and that this instructional approach enabled them to make adjustments to in-class teaching by analyzing the students' completed pre-class activities ($n = 9$) (e.g., Love et al., 2014; Strayer, Hart, & Bleiler, 2015; Talbert, 2015).

Together with the other benefits identified, the findings can be organized into three themes from cognitive learning theory (Kuiper et al., 2015; Lovett & Greenhouse, 2000): (1) sufficient time on task/practice, (2) integrating new knowledge with existing beliefs, and (3) real-time feedback. Several sub-themes emerged during data analysis. Representative citations from the studies are provided in Table 6 to exemplify each sub-theme.

5.6. RQ4: challenges of mathematics flipped classrooms

Notwithstanding the positive participant satisfaction, there are several concerns regarding the use of the flipped classroom approach in practice. In this review, the top two major challenges to implementing flipped mathematics courses were the students' unfamiliarity with flipped learning ($n = 26$) and the instructors' significant start-up effort ($n = 21$). These two challenges largely occurred because the flipped classroom approach was new to both the students and the instructors. In a traditional classroom, students typically learn about the subject matter through a teacher-led lecture format during class time; however, in a flipped classroom, students are required to complete some pre-class learning tasks (e.g., watching video lectures, doing online exercises) before class. Not every student responded favorably to this new learning model (e.g., Braun

Table 6
Identified benefits of mathematics flipped classrooms.

Themes and sub-themes (Count)	Representative citation
Sufficient time on task/practice	
◆ More in-class time for important concepts/activities ($n = 30$)	"In the flipped model, by having the mathematical content explained through online video instruction, the instructors were able to have sufficient face-to-face time to utilize collaborative and inquiry-based instruction" (Overmyer, 2015, p. 799).
◆ On-demand accessibility of video lectures ($n = 27$)	"The learners would often revisit the videos as needed. This on-demand aspect allowed students to seek out answers to questions that arose at various stages of the problem-solving process" (Tawfik & Lilly, 2015, p. 310).
◆ The use of differentiated instructional activities ($n = 10$)	"Mr. Hill rarely engaged in whole class teaching, adopting instead a pedagogical approach that allowed him to differentiate individual instruction" (Muir & Geiger, 2016, p. 168).
Integrating new knowledge with existing beliefs	
◆ Preparing students for class ($n = 17$)	"The videos prompted students to see the big ideas of the section before class" (McGivney-Burelle & Xue, 2013, p. 482).
◆ Adjustment to teaching based on pre-class analysis ($n = 9$)	"Students were required to submit their daily [pre-class] readiness test responses by 2 pm on those days, so the instructor had two hours to read through the answers to question 3 and determine what to focus upon in the upcoming class" (Love et al., 2014, p. 321).
Real-time feedback	
◆ Instructor feedback ($n = 40$)	"In the flipped classroom, the teacher was able to speak with every student in every class and address unique concerns or questions about the current topic being studied" (Clark, 2015, p. 106).
◆ Peer-assisted learning ($n = 33$)	"Sometimes, the teacher or professor isn't always able to explain it the way you are thinking about it, and your partner may be able to explain it a different way" (McCallum et al., 2015, p. 49).

et al., 2014; DeSantis, Van Curen, Putsch, & Metzger, 2015; Van Sickle, 2016). As for the instructors, common complaints included the large amount of time needed to create and edit the video lectures in addition to preparing the in-class activities (e.g., Anderson & Brennan, 201; McGivney-Burelle & Xue, 2013; Talbert, 2015).

Another three frequently reported challenges included the students' unpreparedness for pre-class learning tasks ($n = 14$) (e.g., Kraut, 2015; Sahin, Cavlazoglu, & Zeytuncu, 2015; Scott et al., 2016), being unable to ask questions during out-of-class learning ($n = 13$) (e.g., Anderson & Brennan, 2015; Guerrero et al., 2015; Zack et al., 2015), and being unable to understand the pre-class video content ($n = 11$) (e.g., Palmer, 2015; Scott et al., 2016; Tawfik & Lilly, 2015). These three challenges can be categorized as student-related challenges specifically related to their out-of-class learning. Together with the other challenges identified (Table 7), the findings can be organized into three categories defined by Betihavas et al. (2016): (1) student-

Table 7
Identified challenges to mathematics flipped classrooms.

Themes and sub-themes	Representative citations
Student-related challenges	
◆ Unfamiliarity with flipped learning ($n = 26$)	"This study took place in a setting where traditional lesson planning was routine. Students might have reacted negatively to the change in their learning routines" (DeSantis et al., 2015, p. 50).
◆ Unpreparedness for pre-class learning tasks ($n = 14$)	"39% of participants indicated that they did not do any preparation for the class" (Sahin et al., 2015, p. 146).
◆ Unable to ask questions during out-of-class learning ($n = 13$)	"The most frequent response from students in the flipped sections was that they could not ask questions directly to the instructor while taking notes from the videos" (Zack et al., 2015, p. 807).
◆ Unable to understand video content ($n = 11$)	"The last chapter was difficult to learn via video. I had to go over these concepts again when they [the students] arrived in class." (Scott et al., 2016, p. 261).
◆ Increased workload ($n = 9$)	"Half of the eight students who would be unwilling to take another flipped class cited the additional time required to succeed as a reason" (Murphy, Chang, & Suaray, 2016, p. 668).
◆ Disengaged from watching videos ($n = 3$)	Student: "When I watch the videos, if I'm not writing down what you're lecturing about I did find myself spacing out a lot more" (Ogden, 2015, p. 786).
Faculty challenges	
◆ Significant start-up effort ($n = 21$)	"Creating, editing, and posting videos are time-consuming endeavors as is the development of the in-class quizzes and problem sets. On average, for every class meeting, it took us about 1.5 h to make one short video and an additional 45 min to prepare the quiz and in-class problem set" (McGivney-Burelle & Xue, 2013, p. 484).
◆ Not accustomed to flipping ($n = 10$)	"One flipped section instructor, who had previously been successful in traditional lecture courses, found himself to be very uncomfortable with the format, and despite his best efforts his students did worse than other sections" (Anderson & Brennan, 2015, p. 867).
◆ Ineffectiveness of using others' videos ($n = 4$)	"In statistics courses, notation varies dramatically between textbooks and instructors, making it difficult to use videos prepared by other instructors" (Kuiper et al., 2015, p. 603).
Operational challenges	
◆ Instructors' lacking IT skills ($n = 3$)	"we hit our first major obstacle – one of us (Anderson) found the technology maddeningly unwieldy" (Anderson & Brennan, 2015, p. 866).
◆ Students' lacking IT resources ($n = 3$)	"They did not have access to hardware/software at home and thus had to practice the skills at school" (Chen et al., 2015, p. 628).

related challenges, (2) faculty challenges, and (3) operational challenges. Several sub-themes emerged during data analysis. Representative citations from the studies are provided in Table 7 to exemplify each sub-theme.

6. Discussion

Based on the above findings, we propose a framework (Fig. 7) with a set of 10 design principles for mathematics flipped classrooms. As shown in Table 8, these principles emerged and were established based on the benefits and challenges identified in addition to the findings of our meta-analysis. This section discusses these resulting principles, focusing on three aspects, namely the transition to the flipped classroom (Principles 1 and 2), out-of-class learning design (Principles 3 to 5), and in-class learning design (Principles 6 to 10).

First, there is a need to manage the transition to the flipped classroom because unfamiliarity with this instructional approach was the top challenge for both the students (Principle 1) and the instructors (Principle 2).

Second, instructors can consider presenting introductory materials and providing online support to manage students' problems during class preparation (Principle 3). Then, they can produce short instructional videos (Principle 4) and motivate student learning by using online exercises with grades (Principle 5).

Third, in-class learning experiences are critical to the success of flipped classrooms. Once students have finished the online pre-class learning tasks, instructors can design face-to-face lessons based on the students' performance of the tasks (Principle 6). At the beginning of the lessons, instructors should activate students' pre-class learning by using a structured formative assessment such as a quiz (Principle 7). After that, students should have a chance to apply their knowledge to solve varied tasks and real-world problems (Principle 8). Throughout the learning process, instructor feedback (Principle 9) and peer support (Principle 10) are essential in promoting student learning.

6.1. Transition to the flipped classroom (Principles 1 and 2)

Principle 1. Manage the transition to the flipped classroom for students.

Twenty-six studies reported that not all students were familiar with flipped learning. A few students even wrongly perceived their flipped course to be an online independent study course and “felt like they did not have a teacher for the class” (Zack et al., 2015, p. 807). Instructors should therefore promote students' understanding of this new instructional approach at the beginning of flipping, and the following issues should be articulated:

1. The rationale for using the flipped classroom approach (Carney et al., 2015; Eager, Peirce, & Barlow, 2014; Talbert, 2015);
2. The potential benefits and challenges of flipped learning (Eager et al., 2014; Van Sickle, 2015);
3. The syllabus and how the flipped course is implemented (Carney et al., 2015; Cilli-Turner, 2015; Lai & Hwang, 2016);
4. The tasks that students are required to do, especially in out-of-class learning (Cilli-Turner, 2015; Palmer, 2015; Talbert, 2014).

Instructors can also demonstrate the learning activities for their flipped courses during the first few lessons. For example, Kirvan et al. (2015) and Van Sickle (2015) played instructional videos during class time. In the lessons, they demonstrated

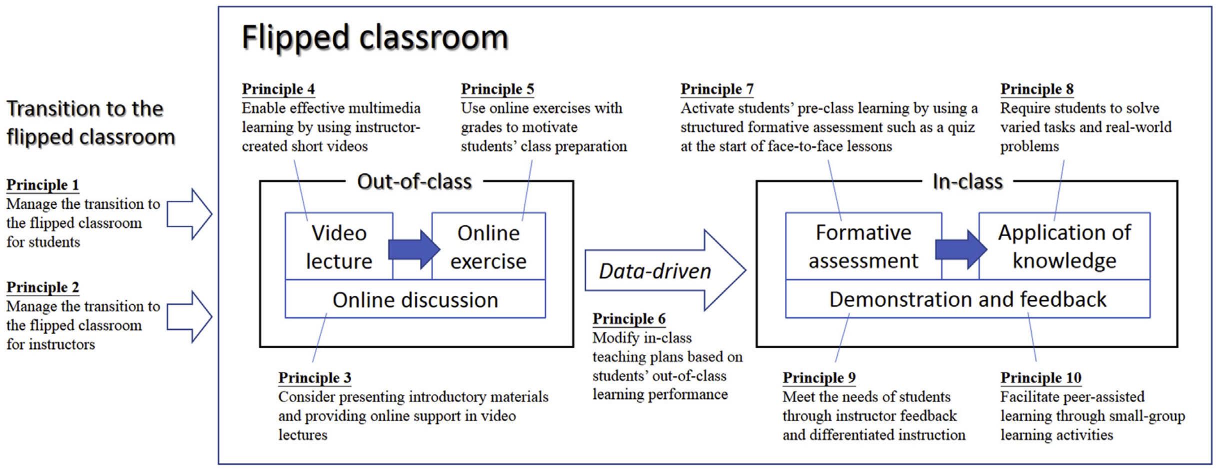

Fig. 7. Framework for the flipped classroom design.

Table 8
Overview of the design principles for the flipped classroom approach.

Theme	Design principles	Empirical rationale
Transition to the flipped classroom	Principle 1: Manage the transition to the flipped classroom for students	◆ Student-related challenge: Unfamiliarity with flipped learning ◆ Operational challenge: Students' lacking IT resources
	Principle 2: Manage the transition to the flipped classroom for instructors	◆ Faculty challenge: Significant start-up effort ◆ Faculty challenge: Not accustomed to flipping ◆ Operational challenge: Instructors' lacking IT skills
Out-of-class learning design	Principle 3: Consider presenting introductory materials and providing online support in video lectures	◆ Integrating new knowledge with existing beliefs: Preparing students for class ◆ Student-related challenge: Unable to understand video content ◆ Student-related challenge: Unable to ask questions during out-of-class learning
	Principle 4: Enable effective multimedia learning by using instructor-created short videos	◆ Sufficient time for task/practice: On-demand accessibility of video lectures ◆ Faculty challenge: Ineffectiveness of using others' videos ◆ Student-related challenge: Disengaged from watching videos ◆ Student challenge: Increased workload
	Principle 5: Use online exercises with grades to motivate students' class preparation	◆ Student-related challenge: Unpreparedness for pre-class learning tasks
In-class learning design	Principle 6: Modify in-class teaching plans based on students' out-of-class learning performance	◆ Integrating new knowledge with existing beliefs: Adjustment to teaching based on pre-class analysis
	Principle 7: Activate students' pre-class learning by using a structured formative assessment such as a quiz at the start of face-to-face lessons	◆ Findings of meta-analysis: A structured formative assessment of pre-class learning at the start of face-to-face lessons promoted student achievement
	Principle 8: Require students to solve varied tasks and real-world problems	◆ Sufficient time for task/practice: More in-class time for important concepts/activities
	Principle 9: Meet the needs of students through instructor feedback and differentiated instruction	◆ Real-time feedback: Instructor feedback ◆ Sufficient time for task/practice: The use of differentiated instructional activities
	Principle 10: Facilitate peer-assisted learning through small-group learning activities	◆ Real-time feedback: Peer-assisted learning

some note-taking strategies and reminded their students to revisit the videos whenever necessary. If students forget to visit the video lectures, instructors can send a reminder to them a day before the pre-class learning tasks are due (Schroeder et al., 2013).

Three studies (i.e., Clark, 2015; D'addato & Miller, 2016; Chen et al., 2015) reported that a few students encountered technical problems when accessing flipped learning materials. In this regard, instructors can extend their office hours in the first few weeks to offer needed assistance for such things as setting up accounts, downloading materials, or submitting assignments (Kraut, 2015). Instructors can prepare extra IT resources (e.g., instructional DVDs or flash drives) for students who do not have Internet access at home (Clark, 2015). D'addato and Miller (2016) reserved computer facilities on campus to support the implementation of their flipped classrooms. In this way, even socioeconomically disadvantaged students can benefit from flipped learning.

Principle 2. Manage the transition to the flipped classroom for instructors.

For mathematics instructors, the significant start-up effort was a major challenge when they changed their course format to a flipped classroom. For example, Adams and Dove (2016) estimated that a total of 70 working hours was invested to flip their undergraduate calculus course. In Talbert's (2015) experience, the time spent on producing an instructional video roughly followed a 6:1 ratio (i.e., a 5-min video required 30 min of production time). In addition to the videos, various flipped learning materials had to be prepared, such as online follow-up quizzes, in-class problems, and assessment tasks (Adams & Dove, 2016; Grypp & Luebeck, 2015; McGivney-Burelle & Xue, 2013). Apart from their start-up effort, the instructors' unfamiliarity with this new instructional approach and their dearth of IT skills were barriers reported in the studies reviewed.

These challenges can be addressed by taking small steps in flipping (Grypp & Luebeck, 2015; Naccarato & Karakok, 2015; Palmer, 2015) and working with other instructors (Carney et al., 2015; Naccarato & Karakok, 2015). First, instructors can convert their courses progressively and produce a bearable amount of flipped learning materials in each semester. In this way, they can gradually accumulate instructional resources and gain experience using the flipped classroom approach. Second, as Carney et al. (2015) highlighted, there is a benefit to working as a team when preparing flipped courses. Instructors can, for example, discuss their video production tips, activity ideas, and even teaching philosophies. This kind of knowledge sharing can enhance their technique in implementing the flipped classroom approach.

6.2. Out-of-class learning design (Principles 3 to 5)

Principle 3. Consider presenting introductory materials and providing online support in video lectures.

Eleven studies reported that some students did not understand the materials presented in the video lectures. As a result, several mathematics instructors (e.g., Anderson & Brennan, 2015; Kirvan et al., 2015; Scott et al., 2016) had to re-teach the concepts during class meetings. In addition, the inability to ask questions during out-of-class learning was a flipped classroom challenge reported in 13 studies. Therefore, a deliberate selection of flipped content (Kuiper et al., 2015) and instructors' online support (Bhagat et al., 2016) are essential to managing student problems during class preparation. As Scott et al. (2016) stated, “some topics were easy to learn through video while others were too complex for students to understand” (p. 262). Thus, instructors can consider covering the introductory materials for each topic through instructional videos, with the details and more advanced content being handled inside the classroom (Anderson & Brennan, 2015; Carney et al., 2015; Cilli-Turner, 2015; Talbert, 2014; Wright, 2015). For example, Young (2015) introduced the definition of inverse functions and its associated properties through instructional videos, but the more complicated materials (e.g., the proofs of the product and quotient rules for differentiation) were presented during in-class mini-lectures. In this way, instructors can explain and elaborate these materials in a more interactive environment and prevent students from having too many questions during their independent study (Anderson & Brennan, 2015). In addition, instructors can provide support through online discussion boards or other social networking sites (e.g., Twitter) in their flipped courses. These kinds of technologies also make it possible for instructors and peers to provide timely feedback and assistance outside the classroom.

Principle 4. Enable effective multimedia learning by using instructor-created short videos.

Our findings suggest that it is more desirable to use instructor-created instructional videos as the primary instructional materials in mathematics flipped classrooms, and to use existing resources such as the Khan Academy videos as supplementary resources (Schwartz, 2014; Zack et al., 2015). From the instructor perspective, mathematics symbols and notations inevitably vary among videos (Kuiper et al., 2015; Schwartz, 2014), and videos produced by others may not feature the host instructor (DeSantis et al., 2015). Several studies have reported that students preferred their own instructor to a stranger in the video lectures (Jungić et al., 2015; Palmer, 2015; Van Sickle, 2015; Zack et al., 2015). By using self-created videos, instructors can establish a cohesive relationship between their out-of-class and in-class learning activities (Carney et al., 2015; Peterson, 2016; Van Sickle, 2015).

To produce instructional videos in mathematics, we suggest that instructors consider two main issues: the video's duration and style. First, Mayer's (2014) cognitive theory of multimedia learning recommends dividing a long video into a series of short video segments. According to Mayer (2014), breaking video lectures into short segments can reduce students' cognitive load and thus facilitate student learning. In the reviewed studies, we found several instructors (e.g., Maciejewski, 2016; Saumier, 2016) who used the six-minute guideline of Guo et al. (2014) in their video productions. Guo et al. (2014) analyzed 6.9 million video-watching sessions across four edX massive open online courses. They found that the duration of the videos was the most significant predictor of engagement, and the median engagement time was six minutes. Their findings suggest that instructors should keep each video within six minutes. In addition to the duration of individual videos, the preliminary findings suggest that a total of 20–25 min for all combined video segments would be a bearable workload for most students in flipped classrooms (Kennedy et al., 2015; McGivney-Burelle & Xue, 2013). Otherwise, they may feel overwhelmed by the pre-class work requirements.

Second, we recommend using a write-while-speaking video style in mathematics flipped classrooms because the natural motion of human handwriting is more engaging than static computer-generated fonts (Guo et al., 2014). To show the materiality of mathematics, Greiffenhagen (2014) argued that even a sequence of well-prepared slides might not be as effective as a blackboard drawing that “makes visible the process of mathematical reasoning” (p. 521). Echoed in the research of flipped learning, Ford (2015) recommended showing the mathematics process by “animating” (p. 374) the writing in videos. This kind of instructional video could give students a step-by-step approach to a problem and thus enhance their learning (Grypp & Luebeck, 2015; McGivney-Burelle & Xue, 2013; Muir & Geiger, 2016).

Principle 5. Use online exercises with grades to motivate students' class preparation.

To promote student learning, students should have a chance to apply what they have learned and receive feedback (Gagné, 1985; Merrill, 2002). The use of online exercises allows students to apply the knowledge they have learned in video lectures. Computerized feedback from these online exercises enables students to monitor their own learning progress so that they can purposely review the video lectures to gain a deeper understanding (Petrillo, 2016; Wright, 2015). However, we found that only about half (54.2%) of the flipped courses used online pre-class quizzes/exercises in students' out-of-class learning. In future flipped classroom practice, we suggest providing online exercises with computerized feedback for student practice and self-checking.

The use of pre-class online exercises that become a small portion of the students' course grade can also motivate them to watch the video lectures (Eager et al., 2014; Kuiper et al., 2015; Scott et al., 2016; Van Sickle, 2015). Students' unpreparedness for out-of-class learning tasks was a flipped classroom challenge reported in 14 studies. In some cases, more than 70% of the students skipped the pre-class video lectures (Palmer, 2015). By using online exercises with grades, instructors can motivate students to complete the assigned learning tasks before class and make their efforts accountable (Kennedy et al., 2015; Talbert, 2014; Van Sickle, 2015).

6.3. In-class learning design (Principles 6 to 10)

Principle 6. Modify in-class teaching plans based on students' out-of-class learning performance.

The flipped classroom approach enables the use of a data-driven decision-making process to design in-class learning activities (Schroeder & Dorn, 2016). The out-of-class online exercises can serve as formative assessments providing information for instructors to modify their teaching plans to meet the needs of their students (Black, Harrison, Lee, Marshall, & Wiliam, 2003; Schroeder & Dorn, 2016). In response to students' out-of-class learning performance, Jungić et al. (2015) proposed four kinds of possible in-class follow-up actions: reviewing prior knowledge, reinforcing new concepts, modifying individual activities, and revising the entire lesson. For example, some mathematics instructors (Jungić et al., 2015; Love et al., 2014; Schroeder & Dorn, 2016) reviewed students' answers to the online exercises to determine the focus of upcoming lessons. In the practice of Jungić et al. (2015), the instructor prepared a list of in-class discussion questions centered around students' misconceptions of their out-of-class learning. This data-driven teaching plan allowed the instructors to better address the students' difficulties and misunderstandings (D'addato & Miller, 2016; Schroeder & Dorn, 2016).

Principle 7. Activate students' pre-class learning by using a structured formative assessment such as a quiz at the start of face-to-face lessons.

At the beginning of face-to-face lessons, we suggest that instructors use a structured formative assessment such as a quiz to assess students' mastery of the out-of-class learning materials. McBride (2015) reported that having a quiz at the start of face-to-face lessons improved attendance at class meetings in his flipped course. Most importantly, we found in our meta-analysis that the use of a structured formative assessment at the start of face-to-face lessons significantly promoted student achievement in mathematics flipped classrooms. By performing this assessment, students were prompted to recall and apply the knowledge they had learned outside the classroom. According to Merrill (2002), the activation of their previous learning experience facilitates students' learning because it lays a solid foundation for them to learn the new material presented inside the classroom. Further, the quiz results provide information on the students' readiness to assume the in-class learning tasks (Kirvan et al., 2015). Based on student performance, instructors can decide whether to review the pre-class video lectures or make adjustments to their in-class teaching plans (Talbert, 2015).

Principle 8. Require students to solve varied tasks and real-world problems.

Thirty studies reported that the flipped classroom approach benefited student learning because the instructors could spend more in-class time handling important concepts or problem-solving activities. From a cognitive psychology perspective, students learn better when they are engaged in solving problems (Mayer, 1992). Merrill (2002) recommended that instructors require their students to apply their new knowledge and skills to solve a sequence of varied problems. The in-class learning tasks of flipped classrooms should thus begin with a few basic exercises and progress to some challenging problems (Eisenhut & Taylor, 2015; Petrillo, 2016; Van Sickle, 2015). In this way, the problem-solving activities can reinforce and extend the content presented in the video lectures (Kennedy et al., 2015; Kirvan et al., 2015; McCallum et al., 2015; Strayer et al., 2015; Yong et al., 2015).

Merrill (2002) pointed out that engaging students in solving real-world problems helps to promote student learning. Consistent with Merrill (2002), Tawfik and Lilly (2015) found that students were interested and motivated to learn when the course materials were similar to real-world problems. However, we found that the use of real-world problems has not been emphasized in all mathematics flipped classrooms. Based on the studies reviewed, we provide an example of using real-world problems in calculus, algebra, and statistics (i.e., the three most common content areas) flipped classrooms:

- Calculus (introductory differential equations course): Using differential equations to model real-world situations (Yong et al., 2015).
- Algebra (linear algebra course): Using least-squares methods to handle a system arising from real-world situations (Talbert, 2014).
- Statistics (statistical literacy course): Using real data and stories, and discussing materials from everyday life such as advertisements, medical advice, and legislation (Gundlach et al., 2015).

Principle 9. Meet the needs of students through instructor feedback and differentiated instruction.

One widely reported benefit of flipped learning is its ability to free up more in-class time for individualized feedback from the instructor (e.g., Bhagat et al., 2016; Larsen, 2015; Muir & Geiger, 2016; Van Sickle, 2015; Weng, 2015). Because certain parts of the course materials are delivered outside the classroom, instructors can spend more time answering students' questions when demonstrating difficult concepts or examples. Instructors can also circulate around the classroom to provide assistance. For example, Clark (2015) offered individualized instruction to confirm students' understanding, clarify concepts, and direct their further study; Strayer et al. (2015) scaffolded students' mathematical thinking and supported their sense-making of course materials.

In addition to individualized feedback, differentiated instruction can be used to meet the other needs of students in flipped classrooms (Chen, Yang, & Hsiao, 2016; Kuiper et al., 2015; Tawfik & Lilly, 2015). As Talbert (2015) explained, although some students finish assigned learning tasks within minutes, underperforming students may even struggle with basic terminology. Instructors can therefore prepare different levels of in-class learning tasks. For the high-achieving students, more challenging questions can be provided to strengthen their learning (Talbert, 2015; Young, 2015). For the underperforming students, instructors can provide more simple exercises at the beginning of lessons to help them acquire the basic concepts (Anderson & Brennan, 2015; Talbert, 2015).

Principle 10. Facilitate peer-assisted learning through small-group learning activities.

In this review, 33 studies reported on the benefits of peer-assisted learning. Topping and Ehly (1998) defined peer-assisted learning as "the acquisition of knowledge and skill through active helping and supporting among status equals or matched companions" (p. 1). They argued that although students could benefit from peer assistance, those who offered help could also benefit because helping others stimulates cognition. In the context of flipped learning, some mathematics instructors (e.g., Love et al., 2014; Van Sickle, 2015) observed that students could gain a deeper understanding by explaining a problem or an idea to their peers. Sometimes the students were even better than the instructors at explaining concepts in a way their peers could understand (McCallum et al., 2015; Wright, 2015). Therefore, instructors should encourage their students to ask questions and discuss them with their classmates (McCallum et al., 2015; Touchton, 2015).

In a peer-assisted learning environment, the students of Jungić et al. (2015) generally agreed that "[h]earing other students explain their understanding of a problem helps [them] learn better" (p. 518) and that "[h]aving to explain [their] own understanding of a problem to other students helps [them] to learn better" (p. 518). The specific model that Jungić et al. (2015) used in their flipped courses was peer instruction (Crouch & Mazur, 2001). The peer instruction approach has also been adopted in other flipped mathematics courses (e.g., Phillips & Phillips, 2016; Saumier, 2016; Talbert, 2014; Touchton, 2015). Jungić et al. (2015) and Saumier (2016) provided detailed procedures showing how they implemented it. First, the instructor posed a question for the students to work on independently. After a few minutes of individual thinking, the students submitted their answers through an electronic device (i.e., Clickers). If the rate of correct answers was unsatisfactory, the students would be asked to convince, discuss, or seek help from their neighbors and then submit the answer again. If the rate of correct answers upon re-submission was still unsatisfactory (e.g., lower than 80%), the instructor would give a hint or demonstrate how to solve the problem. After that, the instructor would either move on to the next question or provide students with a similar question so they could make another attempt. As Jungić et al. (2015) and Talbert (2014) noted, these kinds of activities facilitated peer interactions and furthered the students' understandings. In their experience, the students' responses to the questions usually converged to the correct answer following group discussion with minimal instructor input. Mathematics instructors can thus consider using the peer instruction approach in flipped learning.

7. Conclusion

This review highlighted the need to improve mathematics teaching and learning. Whereas the flipped classroom approach has become increasingly popular in mathematics education, we need empirically based principles rather than mere instructor intuition to guide the design and implementation of flipped courses. A systematic literature review was conducted, aimed at examining the effects of flipped learning, synthesizing the findings of how this instructional approach benefits student learning, and identifying the challenges encountered in flipped mathematics courses. The results showed an overall significant effect in favor of the flipped classroom over the traditional classroom for mathematics education (Hedges' $g = 0.298$, 95% CI [0.16, 0.44]), with no evidence of publication bias. In addition, we found that the top three benefits reported were all related to the students' in-class learning experiences, namely the instructor's individualized feedback, peer-assisted learning, and more in-class time for important concepts/activities. Finally, we found that the students' unfamiliarity with flipped learning and the instructors' significant start-up effort were the two major challenges to implementing flipped classrooms. Based on the findings, we proposed a set of 10 design principles focusing on three aspects of flipped classroom practices, including the transition to the flipped classroom, out-of-class learning design, and in-class learning design.

This review contributes to the literature by offering a rudimentary set of design principles for mathematics flipped classrooms. All of the principles were established on the basis of relevant empirical evidence and aimed at addressing the major challenges and highlighting the practices that could benefit student learning in mathematics flipped classrooms. This set of principles provides insight for mathematics instructors wanting to offer a rigorously designed flipped classroom. For example, Principle 6 suggests modifying in-class teaching plans based on students' out-of-class learning performance, which had only been discussed in nine of the reviewed studies. In another example, Principle 7 recommends using a structured formative assessment such as a quiz at the start of face-to-face lessons to assess and recall students' out-of-class learning. Using these principles may enhance the effectiveness and efficiency of the flipped classroom approach. This set of design principles thus provide a potential agenda for future research to examine the effect of the flipped classroom approach on student learning and motivation. We hope mathematics instructors and instructors from other disciplines (e.g., science, technology, engineering) adopt and test this set of principles in their educational contexts so that the generalizability of the principles can be enhanced.

8. Limitations and recommendations for future research

Before the principles proposed in this review are applied, several of its limitations must be acknowledged. First, although we searched for publications across various databases, the articles included in this review largely focused on higher education in the United States. Most of the research participants were undergraduate students from Western cultures. Therefore, one should consider whether the principles are context-specific. Modifying or extending them may be necessary before applying them to other educational contexts (e.g., secondary education in Asian countries).

Second, this review focused on a specific set of flipped classroom studies in which pre-class videos were provided (Bishop & Verleger, 2013; EDUCAUSE, 2012) and class attendance was mandatory (He et al., 2016). To broaden the scope of review, future studies can adopt other definitions of flipped learning (e.g., Flipped Learning Network, 2014) that do not specify instructors' technological choices or physical locations.

Third, although all of the proposed principles emerged from the studies reviewed and were supported by empirical evidence, the quality and quantity of empirical support largely depended on the researchers' methodology and the focus of their reports. For example, some studies (e.g., Jungić et al., 2015; Muir & Geiger, 2016; Sahin et al., 2015) relied on students' self-reported data of out-of-class learning efforts instead of their online learning data. It is somewhat difficult to objectively investigate student performance outside the classroom. In future research, we recommend that instructors use a learning management system (e.g., Moodle, Blackboard) that can trace students' online behavior so that instructors can have a more comprehensive understanding of how students learn in flipped courses.

Fourth, the dependent measures used in our meta-analysis included objective quantitative measures of mathematics performance such as post-tests, final exams, or other standardized tests. Unfortunately, we cannot determine from our data how many previous flipped classroom studies specifically “taught to the test” because no study explicitly mentioned it.

Apart from the limitations to this review, there are four limitations associated with the reviewed articles. First, the duration of the existing studies was usually short and not more than one semester. Several researchers (e.g., Clark, 2015; Guerrero et al., 2015; Haughton & Kelly, 2015; Phillips & Phillips, 2016) acknowledged that a possible novelty effect could result in a short-term boost to student achievement and perceptions of flipped learning. Initially, students might be excited by this new instructional approach (Guerrero et al., 2015; Haughton & Kelly, 2015). Clark (1983) argued that students' tendency to pay increased attention to newer media could be a confounding variable of learning gains. As Guerrero et al. (2015) observed, “[o]nce the novelty of the videos wore off, fatigue and boredom with the same instructional approach day after day became a factor” (p. 827). Therefore, one should exercise caution when viewing our results because some of the effects found in this review could be the result of novelty (Clark, 1983). Further research with a longer duration is needed to examine the effects of flipped learning when students have more experience with it (Clark, 2015; Phillips & Phillips, 2016).

Second, although various studies have attempted to compare student achievement in flipped classrooms with traditional lecture-based classrooms, about half have provided insufficient data for a meta-analysis. Specifically, several studies provided only the letter grades of their research groups. Without detailed descriptive statistics, further quantitative analysis (e.g., effect size) of these studies is unfeasible.

Third, students' knowledge and skills acquired in flipped learning have not been adequately evaluated. Only two reviewed studies (i.e., Kennedy et al., 2015; Kirvan et al., 2015) described and assessed some specific types of questions used (e.g., conceptual problems and computational problems). Future research can, for example, examine how flipped mathematics learning may affect student performance in both near transfers and far transfers.

Fourth, the flipped classroom designs in the reviewed studies were not always clearly reported. We suggest that future research provide more detail on flipped courses, including the duration (e.g., average, range) of the videos used, the time allocation for different instructional activities (e.g., individual practices, small-group activities), the intensity of flipping (e.g., the percentage of course materials flipped, the exact duration of flipped courses), and the actual description of small-group learning activities. Specifically, with regard to small-group learning activities, there has been no consensus about what various activities actually meant in practice (Freeman et al., 2014). As a result, we could not really differentiate between activities when the authors merely indicated the use of group discussion, in-class collaboration, or group problem solving, without providing details on the actual tasks involved in the group activities (because group discussions, collaboration, and problem-solving all involved discussion; and problem-solving can also be a form of collaboration). Therefore, the meta-analysis in this review is limited, based on the information available in the reviewed studies.

Appendix A

Summary of the reviewed studies

Author(s) and year	Location	Subject area (Grade level)	Course duration	Number of students	Major findings
Adams and Dove (2016)	USA	Calculus I (UG)	1 semester (10 weeks)	TC = 41# FC = 20	FC was not found to have any significant impact on students different from TC; students showed appreciation for FC and wished to take more math courses that used it.
	USA	Calculus 1 (UG)	1 semester	TC = 186# FC = 126#	

(continued on next page)

(*continued*)

Author(s) and year	Location	Subject area (Grade level)	Course duration	Number of students	Major findings
Anderson and Brennan (2015)					The quantitative analysis found moderate benefits to FC over TC; FC students were more engaged and the instructors were able to give students more individualized attention.
Bhagat et al. (2016)	Taiwan	Trigonometry (HS)	6 weeks	TC = 41 FC = 41	FC students had significantly higher learning achievement and motivation than TC students; the performance of low achievers in FC was better than those in TC.
Braun et al. (2014)	Germany	Course I (Algebra) (UG) Course III (Calculus) (UG)	6 lessons 3 lessons	FC = 190# TC = 116# FC = 19	FC students appreciated the increased amount of practice in the classroom and the possibility of learning at their own pace; exam performance remained constant in the topics taught using FC compared to TC.
Carney et al. (2015)	USA	Introductory linear algebra (UG) Introductory linear algebra (UG)	1 semester (15 weeks) 1 semester (10 weeks)	FC = 159#	FC students had generally positive responses; most students believed that the videos helped them understand new and confusing concepts and helped them prepare for class, and they found the format enjoyable.
Chen et al. (2015)	Taiwan	Introductory statistics (UG)	1 semester (16 weeks)	FC = 45	Cooperative learning was an effective instructional strategy in FC; individual levels of support should be provided in FC.
Chen et al. (2016)	Taiwan	Precalculus (HS)	5 weeks	FC = 632	Feelings predicted the final grades for males, whereas the course design predicted the final grades for females; even when females and males showed interest in different topics, they performed equally well.
Cilli-Turner (2015)	USA	Introductory statistics (UG)	1 semester	TC = 56# FC = 78#	Course grades improved significantly when the course was taught using FC; however, the majority of students did not enjoy FC or they preferred TC.
Clark (2015)	USA	Algebra I (HS)	7 weeks	TC = 40# FC = 42#	FC students responded favorably to FC, experienced an increase in their engagement and communication, recognized improvements in the quality of instruction and the use of class time; no significant changes in terms of academic performance were demonstrated between FC and TC.
D'addato and Miller (2016)	USA	Mathematics (ES)	Appeared to be half year	FC = 27	FC shifted the role of the teacher to become a classroom facilitator; FC created an engaging instructional environment, which afforded students the opportunity to experience a greater sense a responsibility over their learning process; parents responded to FC in mixed ways.
DeSantis et al. (2015)	USA	Geometry (HS)	1 lesson	TC = 21 FC = 26	No significant differences in the learning outcomes between TC and FC students; TC students reported significantly higher satisfaction with their own learning than those in FC.
Dove and Dove (2015)	USA	Mathematics Content for Teachers I (UG)	1 semester	TC = 27 FC = 35	Both TC and FC had significantly decreased mathematics anxiety scores; FC showed significantly higher overall achievement in overall course grades; the post-course math anxiety was significantly negatively correlated with the overall course grades in TC, whereas no correlational relationship was found for FC.
Eager et al. (2014)	USA	Mathematical biology (UG)	1 semester (14 weeks)	FC = 48#	FC students performed well in the course and their appreciation for FC increased as the semester progressed; the mean and median final grades for the class in spring 2014 cohort were 83% and 89% respectively.
Ford (2015)	USA	Math content courses for pre-service elementary school teachers (UG)	1 semester (16 weeks)	FC = 32	Compared to TC, more senseful explanations carried over to FC students' performance on their final exams; the instructor received high ratings on teaching evaluations in FC.
Grypp and Luebeck (2015)	USA	Calculus (HS)	3 weeks	FC = 21	FC could create meaningful calculus lessons and at-home assignments while meeting curriculum goals; students could successfully learn challenging material through FC; FC students felt supported in their mathematics learning.
Guerrero et al. (2015)	USA	Finite Math (UG)	1 semester	TC = 37 FC = 39	FC allowed instructors to repurpose class time for more student-centered interaction and problem solving; FC had positive effects on student attitudes toward mathematics, but had no significant impact on student learning over TC.
Gundlach et al. (2015)	USA	Statistics literacy (UG)	1 semester (16 weeks)	TC = 330 FC = 56 OS = 74	TC students scored higher on average on all three exams, but there were no significant differences among groups on homework, projects, or university evaluations of the course or instructor.
Haughton and Kelly (2015)	USA	Introductory business statistics (UG)	1 semester	TC seems to be ~380 FC seems to be ~200	FC students performed better on the common final exam; however, there were no significant differences in the final grades or student satisfaction between TC and FC.
Ichinose and		Clinkenbeard (2016)	USA	Algebra (UG)	1 semester (16 weeks)
TC = 536# FC = 133#	FC students had	consistently higher levels of achievement throughout the course than TC students; FC students reported greater gains			

(*continued*)

Author(s) and year	Location	Subject area (Grade level)	Course duration	Number of students	Major findings
		in affective variables related to mathematics than TC students.			
Jungić et al. (2015)	Canada	Calculus 1 (UG)	1 semester (12 weeks)	FC = 682#	FC encouraged students to take a more active role in the learning process before and during class time, including interactions among students and between students and the instructor; with a pre-class quiz and in-class clicker questions, instructors were able to address misconceptions immediately as they arose.
		Calculus 2 (UG)	1 semester (12 weeks)	FC = 256#	
Kennedy et al. (2015)	USA	Calculus II (UG)	1 semester	TC = 86# FC = 87#	TC students significantly outperformed FC students on conceptual portions of some exams; the overall Motivation score for FC students significantly dropped from the pre-test to the post-test; there was an increase in both the rehearsal score and peer learning score for FC students.
Kirvan et al. (2015)	USA	Algebra (HS)	1 topic	TC = 25 FC = 29	There were comparable statistically significant learning gains in both TC and FC; explicit attention to the substance of FC videos and in-class activities was needed to shift the instructional focus from procedural to conceptual understanding.
Kraut (2015)	USA	Introductory statistics (UG)	1 semester	FC = ~25	FC promoted confidence-building, a feeling of success and personal growth; FC delivered considerably more than foundational knowledge of statistics, helping students find a new approach to learning how to learn.
Kuiper et al. (2015)	USA	Interdisciplinary data analysis (UG)	1 semester (14 weeks)	FC = 20	Successful FC incorporated collaborative work that created space to experiment and struggle with deep concepts; FC students had an increased interest in statistics, thought more deeply about the concepts and were better at transferring knowledge to real-world situations.
		Introductory statistics (UG)	1 semester (15 weeks)	FC = 30	
		Introduction to business statistics (UG)	1 semester (14 weeks)	FC = 25	
Lai and Hwang (2016)	Taiwan	Mathematics (ES)	4 weeks	FC = 24 SRFC = 20	The post-test score of SRFC was significantly higher than FC; integrating the self-regulated strategy into FC could improve students' self-efficacy in addition to their planning strategies and use of study time, thereby promoting their learning achievements.
Larsen (2015)	Canada	Mathematics upgrading (Adult Education)	1 semester (14 weeks)	FC = 25	FC could bifurcate student interaction into two types: completely engaged and self-paced; key interrelated factors in this bifurcation included adoption of cognitive autonomy support, goal orientation, and attendance.
Love et al. (2014, 2015)	USA	Applied linear algebra (UG)	1 semester	TC = 28 FC = 27	FC students experienced a more significant increase between the sequential exams compared to TC students but performed similarly in the final exam; FC students were very positive about their experience in the course and appreciated the student collaboration and instructional video components.
Maciejewski (2016)	New Zealand	Calculus (UG)	1 semester	TC = 222# FC = 650#	FC students on average outperformed TC students on the final exam by approximately 8%; those with high basic mathematical ability and low initial calculus knowledge were the true beneficiaries of FC.
McBride (2015)	USA	Calculus 1 (UG) Precalculus (UG) Introduction to statistics (UG)	1 semester 1 semester 1 semester	FC = 65 FC = 20 FC = 62	PowerPoint slides, videos, applets, and other online tools were things that FC students were more willing to spend time with; there was an amazing transformation in FC students over a short period of time; they did better and were more engaged with FC.
McCallum et al. (2015)	USA	Calculus with Precalculus I (UG)	1 semester (16 weeks)	FC = 25	FC students' academic involvement was seen through note taking, viewing video lectures, active in-class learning and collaboration; peer-to-peer and student-faculty engagement were essential to relationship building, peer learning, and meaningful involvement with faculty.
		Calculus with Precalculus II (UG)	1 semester (16 weeks)	FC = 26	
McGivney-Burelle and Xue (2013)	USA	Calculus II (UG)	1 unit	TC = 29 FC = 31	The average score for FC students in the flipped unit (application of integration) was four points higher than the average score for TC students; FC students appreciated the way class time was used.
Muir and Geiger (2016)	Australia	Mathematics (HS)	Appeared to be 1 year	FC = 27	The teacher and students were positive about their experiences with FC; students were motivated to engage with the teacher-created online mathematics resources.
Murphy et al. (2016)	USA	Linear algebra (UG)	1 semester	TC = 40 FC = 37	FC students performed better in the overall comprehension of content with a 21% increase in the median final exam score; FC students felt more confident in their ability to learn mathematics independently, showed better retention of materials over time, and enjoyed the FC experience.
Ogden (2015)	USA	Algebra (UG)	1 semester	FC = 117#	FC students felt that the teaching approach enabled them to ask more questions in class, the course components worked together to foster increased student learning, and the course design facilitated self-paced instruction.
Overmyer (2015)	USA	Algebra (UG)	1 semester	TC = 165# FC = 136#	For the FC instructors who had experience with inquiry-based and cooperative learning methods, their sections had significantly higher

(*continued on next page*)

(*continued*)

Author(s) and year	Location	Subject area (Grade level)	Course duration	Number of students	Major findings
					common final exam scores than TC; otherwise, no significant difference was found.
Palmer (2015)	USA	Calculus I (UG)	1 semester	FC = 17	60% of the students preferred FC; there was no significant difference between FC and TC students in subsequent course (Calculus II) grades; there was no association between their grades and learning styles or between their preferences and personalities.
Peterson (2016)	USA	Introductory statistics (UG)	1 semester (10 weeks)	TC = 19 FC = 24	FC students outperformed TC students by more than a letter grade on the final exam; FC students were more satisfied with the course overall; the result was probably due to the strong cohesion between the in-class and the out-of-class content.
Petrillo (2016)	USA	Calculus I (UG)	1 semester	TC = 70# FC = 105#	There were positive and substantial effects of the flipped course on failure rates, scores on the common final exam, student opinion of calculus, teacher impact on measurable outcomes, and success in second-semester calculus.
Phillips and Phillips (2016)	USA	Business statistics (UG)	1 semester	TC = 35 FC = 33	Student response to FC was somewhat bimodal. Comprehensive final exam grades in FC were significantly higher on average by almost one full letter grade and course pass rates also rose.
Rufatto et al. (2016)	USA	Mathematics and its applications (UG)	1 semester	Not reported	An analysis of course grades indicated an increased performance for FC students; there were mixed feelings about FC; FC required students to complete more preparation work before coming to class, which was unnatural to them.
Sahin et al. (2015)	USA	Calculus (UG)	3 lessons	FC = 96	Students preferred watching videos (44%) to reading from textbooks (17%); students achieved significantly higher quiz scores in the flipped lessons than the non-flipped; 83% of the students stated that the flipped lessons better prepared them.
Saumier (2016)	Canada	Multivariable calculus (UG)	1 semester	TC = 102 FC = 52	The appreciation and effectiveness of choice in videos (a traditional lecture speed and a shorter version where the instructor commented on prewritten text and figures) were positive; FC students felt more engaged and did not perceive an increase in their workload.
Schroeder and Dorn (2016)	USA	Calculus I (UG)	1 semester	FC = 23	Modifications to FC lessons based on formative assessments fell into four different categories: review of prior knowledge, reinforce new concepts, modify individual activities, revise entire lesson; student survey data largely echoed the positive impact of such changes.
Schroeder et al. (2013)	USA	Precalculus with trigonometry (UG)	6 lessons	FC = 28	The positive effects of FC were generally independent of the students' abilities and the topics being studied; FC students worked on mathematics more attentively for longer periods of time than before; FC students valued the close attention given to them by instructors.
		Calculus II (UG)	1 chapter	Not reported	
		Discrete Math II (UG)	1 semester (14 weeks)	FC = 16	
Schroeder et al. (2015)	USA	Calculus I (UG)	1 semester	TC = 49# FC = 63#	FC students performed better than TC students on the common final exam; a follow-up examination of the subsequent course (Calculus II) grades suggested the benefits from enrolling in FC extended into performance in the course.
Schwartz (2014)	USA	Statistics (Doctorial)	1 semester	FC = ~12	FC was highly effective at supporting Ph.D. nursing students' learning of statistical materials; the improved efficiency to the use of classroom instructional time was of great importance.
Scott et al. (2016)	USA	Calculus 2 (UG)	1 semester (14 weeks)	TC = 45# FC = 51#	FC students had similar content knowledge gains to TC students; the use of online homework and in-class quizzes were critical motivating factors that likely contributed to increased student performance.
Strayer et al. (2015)	USA	Precalculus (UG)	1 semester	FC seems to be ~28	One of the greatest pedagogical affordances of FC appeared when instructors garnered information from students; instructors could identify gaps in students' understandings or misconceptions, when they could probe during the next class.
Talbert (2014)	USA	Linear algebra (UG)	1 semester	FC seems to be > 100	69.8% of the students indicated they preferred FC because the daily homework, which had previously been given as an out-of-class assignment, was converted to an in-class assignment.
Talbert (2015)	USA	Transition-to-Proof (UG)	1 semester	TC = 939# FC = 39#	The motivated strategies for the learning questionnaire (MSLQ) data between TC and FC were inconclusive; a marginal decrease in non-passing grades and a marginal increase in the top grades occurred with FC.
Tawfik and Lilly (2015)	USA	Psychological statistics (UG)	1 semester (16 weeks)	FC = 24	FC students had positive perceptions regarding the problem-based learning experience and described themselves as motivated to solve the ill-structured problems; FC supported the learners in different ways compared to other forms of scaffolds; on-demand access to the videos was important to answer questions when the instructor was not present.
Touchton (2015)	USA	Advanced statistics (UG)	1 semester	TC = 40 FC = 43	FC gave students a statistically significant advantage in difficult, applied areas emphasized in class; FC students said they learned more and enjoyed the course more than TC students.
	USA	Algebra (UG)	1 semester		

(continued)

Author(s) and year	Location	Subject area (Grade level)	Course duration	Number of students	Major findings
Van Sickle (2015, 2016)				TC = 54# FC = 58#	FC students achieved higher scores than TC students in the final exam; however, their perception of a number of measures decreased significantly, including how interested they were in the course and whether the instructor effectively facilitated learning.
Weng (2015)	USA	Introduction to mathematical thinking (UG)	1 semester (15 weeks)	TC = 63# FC = 62#	The learning outcome in FC was better than TC; student satisfaction with FC was high.
Wright (2015)	USA	Linear algebra (UG)	1 semester	Not reported	FC enabled the instructor to work with individuals and small groups of students, customizing the questions and examples for students, while keeping the large class on the same track.
Yong et al. (2015)	USA	Introductory differential equations (UG)	Half semester (7 weeks)	TC = 90# FC = 86#	Pre-test and post-test assessments showed no differences between FC and TC; there were no statistically significant differences between TC and FC in terms of attitudes toward STEM and motivation strategies for learning the questionnaire (MSLQ) data.
Young (2015)	USA	Calculus I (UG)	1 semester	FC seems to be 33#	The instructor was able to better serve the students without strong math backgrounds, while still challenging those who had stronger backgrounds; there were low-cost choices, in terms of money and faculty time, that made flipping a course manageable for a single instructor.
Zack et al. (2015)	USA	Finite Math (UG)	1 semester	Not reported	No statistical difference was found in the test scores of FC and TC students; qualitative data indicated potential problems with implementing FC; many FC students had negative opinions of FC and their attitudes toward math tended to decline in general.
		Precalculus (UG)	1 semester	Not reported	
		Business calculus (UG)	1 semester	Not reported	
		Calculus 1 (UG)	1 semester	Not reported	
Ziegelmeier and Topaz (2015)	USA	Calculus I (UG)	1 semester	TC = 23 FC = 22	Students in both TC and FC scored similarly on the graded components of the course; the majority of students were comfortable with the format of each section.

total number of students from multiple sections or cohorts.
ES: elementary school; FC: flipped classroom; HS: High school; OS: Online study; SRFC: self-regulated flipped classroom; TC: traditional classroom; UG: undergraduate.

References

*References marked with an asterisk indicate studies included in the review.
References marked with a plus were further included in the meta-analysis.

Abeysekera, L., & Dawson, P. (2015). Motivation and cognitive load in the flipped classroom: Definition, rationale and a call for research. *Higher Education Research and Development, 34*(1), 1–14.

*+ Adams, C., & Dove, A. (2016). Flipping calculus: The potential influence, and the lessons learned. *Electronic Journal of Mathematics & Technology, 10*(3), 154–164.

Altman, D. G., & Bland, J. M. (2005). Standard deviations and standard errors. *BMJ: British Medical Journal, 331*(7521), 903.

*+ Anderson, L., & Brennan, J. P. (2015). An experiment in "flipped" teaching in freshman calculus. *PRIMUS, 25*(9–10), 861–875.

Bergmann, J., & Sams, A. (2012). *Flip your classroom: Reach every student in every class every day*. Washington, DC: International Society for Technology in Education.

Bernard, J. S. (2015). The flipped classroom: Fertile ground for nursing education research. *International Journal of Nursing Education Scholarship, 12*(1), 99–109.

Betihavas, V., Bridgman, H., Kornhaber, R., & Cross, M. (2016). The evidence for 'flipping out': A systematic review of the flipped classroom in nursing education. *Nurse Education Today, 38*, 15–21.

*+ Bhagat, K. K., Chang, C. N., & Chang, C. Y. (2016). The impact of the flipped classroom on mathematics concept learning in high school. *Educational Technology & Society, 19*(3), 134–142.

Bishop, J. L., & Verleger, M. A. (2013). The flipped classroom: A survey of the research. In *120th ASEE national conference and exposition, Atlanta, GA (paper ID 6219)*. Washington, DC: American Society for Engineering Education.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Maidenhead: Open University Press.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, England: Wiley.

*+ Braun, I., Ritter, S., & Vasko, M. (2014). Inverted classroom by topic: A study in mathematics for electrical engineering students. *International Journal of Engineering Pedagogy, 4*(3), 11–17.

* Carney, D., Ormes, N., & Swanson, R. (2015). Partially flipped linear algebra: A team-based approach. *PRIMUS, 25*(8), 641–654.

* Chen, L., Chen, T. L., & Chen, N. S. (2015). Students' perspectives of using cooperative learning in a flipped statistics classroom. *Australasian Journal of Educational Technology, 31*(6), 621–640.

Chen, C. M., & Wu, C. H. (2015). Effects of different video lecture types on sustained attention, emotion, cognitive load, and learning performance. *Computers & Education, 80*, 108–121.

* Chen, S. C., Yang, S. J., & Hsiao, C. C. (2016). Exploring student perceptions, learning outcome and gender differences in a flipped mathematics course. *British Journal of Educational Technology, 47*(6), 1096–1112.

Chorianopoulos, K., & Giannakos, M. N. (2013). Usability design for video lectures. In *Proceedings of the 11th european conference on interactive TV and video* (pp. 163–164). ACM.

Chua, J. S. M., & Lateef, F. (2014). The flipped classroom: Viewpoints in Asian universities. *Education in Medicine Journal, 6*(4), 20–26.

*+ Cilli-Turner, E. (2015). Measuring learning outcomes and attitudes in a flipped introductory statistics course. *PRIMUS, 25*(9–10), 833–846.

Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research, 53*(4), 445–459.

*+ Clark, K. R. (2015). The effects of the flipped model of instruction on student engagement and performance in the secondary mathematics classroom. *Journal of Educators Online, 12*(1), 91–115.

Crouch, C. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics, 69*(9), 970–977.

*+ 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.
Dillon, K. M. (1982). Statisticophobia. *Teaching of Psychology, 9*(2), 117.
*+ Dove, A., & Dove, E. (2015). Examining the influence of a flipped mathematics course on preservice elementary teachers' mathematics anxiety and achievement. *Electronic Journal of Mathematics & Technology, 9*(2), 166–179.
*+ D'addato, T., & Miller, L. R. (2016). An inquiry into flipped learning in fourth grade math instruction. *The Canadian Journal of Action Research, 17*(2), 33–55.
* Eager, E. A., Peirce, J., & Barlow, P. (2014). Math bio or biomath? Flipping the mathematical biology classroom. *Letters in Biomathematics, 1*(2), 139–155.
EDUCAUSE. (2012). 7 things you should know about flipped classrooms. Retrieved from https://net.educause.edu/ir/library/pdf/eli7081.pdf.
Eisenhut, L. A., & Taylor, C. E. (2015). In-class purposes of flipped mathematics educators. *Journal of Mathematics Education at Teachers College, 6*(2), 17–25.
Epstein, J. (2013). The calculus concept inventory: Measurement of the effect of teaching methodology in mathematics. *Notices of the American Mathematical Society, 60*(8), 1018–1026.
Flipped Learning Network. (2014). What is flipped learning?. Retrieved from http://flippedlearning.org/wp-content/uploads/2016/07/FLIP_handout_FNL_Web.pdf.
* Ford, P. (2015). Flipping a math content course for pre-service elementary school teachers. *PRIMUS, 25*(4), 369–380.
Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., et al. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences, 111*(23), 8410–8415.
Gagné, R. M. (1985). *The conditions of learning and theory of instruction* (4th ed.). New York: Holt, Rinehart and Winston.
Garfield, J. B. (2003). Assessing statistical reasoning. *Statistics Education Research Journal, 2*(1), 22–38.
Giannakos, M. N., Krogstie, J., & Chrisochoides, N. (2014). Reviewing the flipped classroom research: Reflections for computer science education. In *Proceedings of the computer science education research conference* (pp. 23–29). ACM.
Greiffenhagen, C. (2014). The materiality of mathematics: Presenting mathematics at the blackboard. *The British Journal of Sociology, 65*(3), 502–528.
* Grypp, L., & Luebeck, J. (2015). Rotating solids and flipping instruction. *Mathematics Teacher, 109*(3), 186–193.
*+ Guerrero, S., Beal, M., Lamb, C., Sonderegger, D., & Baumgartel, D. (2015). Flipping undergraduate finite mathematics: Findings and implications. *PRIMUS, 25*(9–10), 814–832.
*+ Gundlach, E., Richards, K. A. R., Nelson, D., & Levesque-Bristol, C. (2015). A comparison of student attitudes, statistical reasoning, performance, and perceptions for web-augmented traditional, fully online, and flipped sections of a statistical literacy class. *Journal of Statistics Education, 23*(1), 1–33.
Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement: An empirical study of MOOC videos. In *Proceedings of the first ACM conference on Learning@ scale conference* (pp. 41–50). ACM.
Gurevitch, J., & Hedges, L. V. (1999). Statistical issues in ecological meta-analysis. *Ecology, 80*(4), 1142–1149.
* Haughton, J., & Kelly, A. (2015). Student performance in an introductory business statistics course: Does delivery mode matter?. *Journal of Education for Business, 90*(1), 31–43.
He, W., Holton, A., Farkas, G., & Warschauer, M. (2016). The effects of flipped instruction on out-of-class study time, exam performance, and student perceptions. *Learning and Instruction, 45*, 61–71.
Hoyles, C., Morgan, C., & Woodhouse, G. (Eds.). (1999). *Rethinking the mathematics curriculum*. London: Falmer Press.
* Ichinose, C., & Clinkenbeard, J. (2016). Flipping college algebra: Effects on student engagement and achievement. *Learning Assistance Review (TLAR), 21*(1), 115–129.
* Jungić, V., Kaur, H., Mulholland, J., & Xin, C. (2015). On flipping the classroom in large first year calculus courses. *International Journal of Mathematical Education in Science and Technology, 46*(4), 508–520.
*+ Kennedy, E., Beaudrie, B., Ernst, D. C., & St. Laurent, R. (2015). Inverted pedagogy in second semester calculus. *PRIMUS, 25*(9–10), 892–906.
Kim, M. K., Kim, S. M., Khera, O., & Getman, J. (2014). The experience of three flipped classrooms in an urban university: An exploration of design principles. *Internet and Higher Education, 22*, 37–50.
* Kirvan, R., Rakes, C. R., & Zamora, R. (2015). Flipping an algebra classroom: Analyzing, modeling, and solving systems of linear equations. *Computers in the Schools, 32*(3–4), 201–223.
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.
* Kraut, G. L. (2015). Inverting an introductory statistics classroom. *PRIMUS, 25*(8), 683–693.
* Kuiper, S. R., Carver, R. H., Posner, M. A., & Everson, M. G. (2015). Four perspectives on flipping the statistics classroom: Changing pedagogy to enhance student-centered learning. *PRIMUS, 25*(8), 655–682.
* Lai, C. L., & Hwang, G. J. (2016). A self-regulated flipped classroom approach to improving students' learning performance in a mathematics course. *Computers & Education, 100*, 126–140.
* Larsen, J. (2015). Adult students' experiences of a flipped mathematics classroom. *Adults Learning Mathematics, 10*(1), 50–67.
Li, Q., & Ma, X. (2010). A meta-analysis of the effects of computer technology on school students' mathematics learning. *Educational Psychology Review, 22*(3), 215–243.
Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
Lo, C. K., & Hew, K. F. (2017). A critical review of flipped classroom challenges in K-12 education: Possible solutions and recommendations for future research. *Research and Practice in Technology Enhanced Learning, 12*(1). article 4.
* Love, B., Hodge, A., Corritore, C., & Ernst, D. C. (2015). Inquiry-based learning and the flipped classroom model. *PRIMUS, 25*(8), 745–762.
* Love, B., Hodge, A., Grandgenett, N., & Swift, A. W. (2014). Student learning and perceptions in a flipped linear algebra course. *International Journal of Mathematical Education in Science and Technology, 45*(3), 317–324.
Lovett, M. C., & Greenhouse, J. B. (2000). Applying cognitive theory to statistics instruction. *The American Statistician, 54*(3), 196–206.
*+ Maciejewski, W. (2016). Flipping the calculus classroom: An evaluative study. *Teaching Mathematics and Its Applications, 35*, 187–201.
Mayer, R. E. (1992). *Thinking, problem solving, cognition* (2nd ed.). New York: W. H. Freeman.
Mayer, R. E. (2014). *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press.
* McBride, C. (2015). Flipping advice for beginners: What I learned flipping undergraduate mathematics and statistics classes. *PRIMUS, 25*(8), 694–712.
* McCallum, S., Schultz, J., Sellke, K., & Spartz, J. (2015). An examination of the flipped classroom approach on college student academic involvement. *International Journal of Teaching and Learning in Higher Education, 27*(1), 42–55.
* McGivney-Burelle, J., & Xue, F. (2013). Flipping calculus. *PRIMUS, 23*(5), 477–486.
Merrill, M. D. (2002). First principles of instruction. *Educational Technology Research and Development, 50*(3), 43–59.
Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Reprint-preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Physical Therapy, 89*(9), 873–880.
* Muir, T., & Geiger, V. (2016). The affordances of using a flipped classroom approach in the teaching of mathematics: A case study of a grade 10 mathematics class. *Mathematics Education Research Journal, 28*(1), 149–171.
*+ 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.
Naccarato, E., & Karakok, G. (2015). Expectations and implementations of the flipped classroom model in undergraduate mathematics courses. *International Journal of Mathematical Education in Science and Technology, 46*(7), 968–978.
* Ogden, L. (2015). Student perceptions of the flipped classroom in college algebra. *PRIMUS, 25*(9–10), 782–791.

*+ Overmyer, J. (2015). Research on flipping college algebra: Lessons learned and practical advice for flipping multiple sections. *PRIMUS, 25*(9–10), 792–802.
O'Flaherty, J., & Phillips, C. (2015). The use of flipped classrooms in higher education: A scoping review. *Internet and Higher Education, 25*, 85–95.
* Palmer, K. (2015). Flipping a calculus class: One instructor's experience. *PRIMUS, 25*(9–10), 886–891.
*+ Peterson, D. J. (2016). The flipped classroom improves student achievement and course satisfaction in a statistics course: A quasi-experimental study. *Teaching of Psychology, 43*(1), 10–15.
*+ Petrillo, J. (2016). On flipping first-semester calculus: A case study. *International Journal of Mathematical Education in Science and Technology, 47*(4), 573–582.
* Phillips, L., & Phillips, M. (2016). Improved student outcomes in a flipped statistics course. *Administrative Issues Journal, 6*(1), 88–98.
Presti, C. R. (2016). The flipped learning approach in nursing education: A literature review. *Journal of Nursing Education, 55*(5), 252–257.
Rakes, C. R., Valentine, J. C., McGatha, M. B., & Ronau, R. N. (2010). Methods of instructional improvement in algebra: A systematic review and meta-analysis. *Review of Educational Research, 80*(3), 372–400.
Ronau, R. N., Rakes, C. R., Bush, S. B., Driskell, S. O., Niess, M. L., & Pugalee, D. (2015). The quality of mathematics education technology literature. *Journal of MultiDisciplinary Evaluation, 11*(24), 12–36.
* Rufatto, R., Dickin, H., Florescu, A., Lorch, C., Bremigan, R., & Lorch, J. (2016). Learner-centered pedagogy in a liberal arts mathematics course. *PRIMUS, 26*(9), 863–877.
* Sahin, A., Cavlazoglu, B., & Zeytuncu, Y. E. (2015). Flipping a college calculus course: A case study. *Educational Technology & Society, 18*(3), 142–152.
* Saumier, L. P. (2016). Improvements on the peer-instruction method: A case study in multivariable calculus. *Electronic Journal of Mathematics & Technology, 10*(3), 137–153.
* Schroeder, L. B., & Dorn, B. (2016). Enabling and integrating online formative assessment in a flipped calculus course. *PRIMUS, 26*(6), 585–602.
*+ Schroeder, L. B., McGivney-Burelle, J., & Xue, F. (2015). To flip or not to flip? An exploratory study comparing student performance in calculus I. *PRIMUS, 25*(9–10), 876–885.
* Schroeder, L., Xue, F., & McGivney, R. (2013). Flipping three different mathematics courses: Common conclusions and plans for the future. *MathAMATYC Educator, 4*(2), 63–67.
* Schwartz, T. A. (2014). Flipping the statistics classroom in nursing education. *Journal of Nursing Education, 53*(4), 199–206.
*+ Scott, C. E., Green, L. E., & Etheridge, D. L. (2016). A comparison between flipped and lecture-based instruction in the calculus classroom. *Journal of Applied Research in Higher Education, 8*(2), 252–264.
Seery, M. K. (2015). Flipped learning in higher education chemistry: Emerging trends and potential directions. *Chemistry Education Research and Practice, 16*(4), 758–768.
* Strayer, J. F., Hart, J. B., & Bleiler, S. K. (2015). Fostering instructor knowledge of student thinking using the flipped classroom. *PRIMUS, 25*(8), 724–735.
* Talbert, R. (2014). Inverting the linear algebra classroom. *PRIMUS, 24*(5), 361–374.
* Talbert, R. (2015). Inverting the transition-to-proof classroom. *PRIMUS, 25*(8), 614–626.
* Tawfik, A. A., & Lilly, C. (2015). Using a flipped classroom approach to support problem-based learning. *Technology, Knowledge and Learning, 20*(3), 299–315.
Topping, K. J., & Ehly, S. W. (Eds.). (1998). *Peer-assisted learning*. Mahwah, NJ: Lawrence Erlbaum Associates.
* Touchton, M. (2015). Flipping the classroom and student performance in advanced statistics: Evidence from a quasi-experiment. *Journal of Political Science Education, 11*(1), 28–44.
*+ Van Sickle, J. (2015). Adventures in flipping college algebra. *PRIMUS, 25*(8), 600–613.
* Van Sickle, J. (2016). Discrepancies between student perception and achievement of learning outcomes in a flipped classroom. *Journal of the Scholarship of Teaching and Learning, 16*(2), 29–38.
* Weng, P. (2015). Developmental math, flipped and self-paced. *PRIMUS, 25*(9–10), 768–781.
* Wright, S. E. (2015). Linear algebra and the experiences of a “flipper”. *PRIMUS, 25*(8), 627–640.
*+ Yong, D., Levy, R., & Lape, N. (2015). Why no difference? A controlled flipped classroom study for an introductory differential equations course. *PRIMUS, 25*(9–10), 907–921.
* Young, A. (2015). Flipping the calculus classroom: A cost-effective approach. *PRIMUS, 25*(8), 713–723.
* Zack, L., Fuselier, J., Graham-Squire, A., Lamb, R., & O'Hara, K. (2015). Flipping freshman mathematics. *PRIMUS, 25*(9–10), 803–813.
Zheng, B., Warschauer, M., Lin, C. H., & Chang, C. (2016). Learning in one-to-one laptop environments: A meta-analysis and research synthesis. *Review of Educational Research, 86*(4), 1052–1084.
* Ziegelmeier, L. B., & Topaz, C. M. (2015). Flipped calculus: A study of student performance and perceptions. *PRIMUS, 25*(9–10), 847–860.