Effects of Flipped Instruction on College Students' Learning in STEM Subject Domains: A Meta-Analysis

Chioma Ezeh Olusola Adesope Olasunkanmi Kehinde Emmanuel Jaiyeola
Washington State University

Abstract

The increasing need for innovative constructivist learning approaches in college STEM subject domains has necessitated the use of flipped instruction which infuses computer-assisted learning with face-to-face learning to achieve optimal learning. This meta-analysis investigates the effect of flipped instruction and the conditions under which it is beneficial or deleterious for college students learning in STEM subject domains. Overall, the meta-analysis results show that flipped instruction is an effective learning strategy for STEM subject domains. Moderator analyses reveal that pre-class instructional features, the structure of in-class activities, STEM subject domains, and test format moderated the overall effect size. Our findings indicate that flipped instruction yields robust learning benefits when the pre-class videos are accompanied by pre-class tasks for students to work on and collaborative in-class activities that allow them to interact with one another. These findings have implications for designing flipped instruction, particularly for STEM subject domains.

Keywords: flipped instruction, college students' learning, STEM subject domains, constructivist learning

Flipped instruction, also called inverted instruction, is a growing trend of instructional delivery that has drawn the attention of researchers in different fields and has been applied in classrooms to increase student engagement and learning outcomes ostensibly. Since the advancements in educational technology, teachers have continually sought ways to integrate technology into instruction to enhance learning. Flipped instruction (FI) was born out of technological advances in education. Although a minimal form of FI has existed for a while (Buch & Warren, 2017), its use in different fields and research on it have only peaked over the last few years (since 2012) as its potential benefits become more pronounced (Hao, 2016; Love et al., 2014; Ryan & Reid, 2015). In line with the growing research on FI, previous meta-analyses have been conducted to ascertain its learning benefits, among which there have been mostly positive findings on the learning effects of FI. However, there are mixed findings on whether subject domains moderate the effect of FI or not. Although FI has been used extensively in STEM domains, there is no comprehensive and systematic meta-analysis of the effects of FI on college students' learning in STEM subject domains. The lack of such evidence-based meta-analysis limits our understanding of the different conditions under which FI is beneficial or deleterious for learning in STEM subject domains. The present meta-analysis seeks to fill this gap.

Flipped Instruction

The meaning of the term "flipped instruction" has expanded beyond the intent of its pioneers to simply switch homework activities for the classroom and that of the classroom for homework to have more time for deep processing and learning of materials (Bergmann & Sams, 2009). FI now incorporates the use of digital videos and other technologies to carry out the intended classroom activities at home and integrates other activity-based/constructivist learning strategies such as problem-solving and collaborative learning during classes (Abeysekera & Dawson, 2015; Akçayır & Akçayır, 2018; Hao, 2016; Thai et al., 2017). Typically, in traditional instruction (TI), the teacher prepares and delivers a lecture while students are encouraged to tackle assignment problems independently. FI involves reversing what is done in a typical traditional classroom so that student-centered constructivist learning activities are employed in the classroom to ensure deeper learning of the material (Hao, 2016; He et al., 2016; Lage et al., 2000).

FI comprises two key sessions: pre-class and in-class sessions. In the pre-class session, students familiarize themselves with an upcoming lesson at home watching a video of the lesson (Cieliebak & Frie, 2016; He et al., 2016). In many cases, the videos are accompanied by other tasks that students are required to carry out, such as taking online quizzes or working on online modules (Jonsson, 2015), studying a part of an assigned textbook for the topic, and taking notes or preparing questions for class discussions (Aşıksoy & Özdamlı, 2016; Peterson, 2015). The pre-class and in-class activities implemented in the studies included in this meta-analysis are presented in Table 1.

During the in-class session, students engage in constructivist learning activities that foster deep learning. This session is usually influenced by the outcomes of the pre-class activities, such as students' performance on the quizzes they took while watching the videos or the questions they bring to class (Jonsson, 2015). Some classes begin with the teacher answering and clarifying questions on challenging aspects of the topic and further incorporating activities like discussion, solving problems/exercises individually or collaboratively, and simulations (Cieliebak & Frie, 2016). In some classes, the teacher highlights some critical aspects of the topic for classroom discussion, alongside reviewing the more challenging homework problems (Buch & Warren, 2017). While the in-class activities go on under the teacher's supervision, students establish cognitive connections with the prior knowledge they have built during the pre-class session, unlike in traditional classrooms where students are left to achieve this task at home after the class (Cieliebak & Frie, 2016). Moreover, students can use their skills face-to-face with the teacher's support in a flipped classroom (Clark et al., 2016). Such teaching adaptation potentially meets the needs of all students and reflects just-in-time teaching (Jonsson, 2015).

An expanded view of FI involves a kind of blended learning where computer-assisted learning is infused with face-to-face learning. In most flipped courses, technology tools serve as online platforms for out-of-class learning (Thai et al., 2016; Zainuddin & Halili, 2016). In the majority of the studies on flipped instruction (e.g., Maneeratana et al., 2016; Mason et al., 2013; Timmerman et al., 2016), videos or audio presentations of an upcoming lesson are made available for students online through learning management systems or course websites, while they meet face-to-face with the teacher during in-class sessions. Not only do students positively perceive using videos to flip the class, but research also suggests that out-of-class video lectures positively affect learning as much as face-to-face lectures do in passing basic content information to learners (Fryling et al., 2016; Jonsson, 2015). Students have opportunities to rewatch videos or some parts of the videos to understand the topic at their own pace better.

Table 1. Description of Flipped Instruction Features in Individual Studies

Author	Pre-class instruction features	In-class strategy to assess pre-class learning	Structure of in-class activity	In-class teaching approach	Subject domain
Adams & Dove (2017)	Video	None	Collaboration	Activities	Math
Akkaraju (2016)	Video	Quizzes & review	Collaboration	Problem-solving, lecture, & notetaking	Science
Albalawi (2018)	Video	None	Noncollaborative	Lecture	Math
Aşıksoy & Özdamlı (2015)	Video + quiz, questions preparation	Question and answers	Collaboration	Discussions, problem-solving, & simulations	Science
Baepler et al. (2014)	Video	None	Collaboration	Problem-solving, simulations, & games	Science
Bakr et al. (2016)	Video + quiz	None	Noncollaborative	Lectures	Science
Barral et al. (2018)	Video	None	Noncollaborative	Activities	Science
Baytiyeh & Naja (2017)	Video + homework	Quiz	Collaboration	Activities	Engineering
Bradford et al. (2014)	Video	None	Noncollaborative	Problem-solving & topic exploration	Math
Braun et al. (2014)	Video + guiding questions	None	Noncollaborative	Practice	Math
Brooks (2014)	Video + summary	Summary writing	Collaboration	Activities	Technology
Buch & Warren (2017)	Video	None	Noncollaborative	Problem-solving	Math
Choi & Lee (2015)	Video + ebooks reviews	Reviews	Collaboration	Interactive activities	Technology
Cilli-Turner (2015)	Video	None	Collaboration	Lecture, quizzes, & problem-solving	Math
Clark et al. (2016)	Video	None	Noncollaborative	Problem-solving & student dialogue	Engineering
Crimmins & Midkiff (2017)	Video + reading + quiz	Quiz	Noncollaborative	Problem-solving & iclicker questions	Science
Davies et al. (2013)	Video	None	Noncollaborative	Lectures, homework	Technology
Eichler & Peeples (2016)	Video + reading + quiz	Quiz	Noncollaborative	Lectures	Science
El-Banna et al. (2017)	Video + text	None	Noncollaborative	Activities	Science
Elmaleh & Shankararaman (2017)	Video + quiz	None	Noncollaborative	Problem-solving, quiz & feedback	Technology
Entezari & Javdan (2016)	Video	None	Collaboration	Activities	Science

Table 1. (*Continued*) Description of Flipped Instruction Features in Individual Studies

Author	Pre-class instruction features	In-class strategy to assess pre-class learning	Structure of in-class activity	In-class teaching approach	Subject domain
Foldnes (2016)	Video + exercises	Question & answers	Collaboration	Discussions & problem-solving	Math
Fryling et al. (2016)	Video	Quiz	Collaboration	Discussions, activities	Technology
Harrigton et al. (2015)	Video	None	Noncollaborative	Lectures	Science
He et al. (2016)	Video	Quiz	Noncollaborative	Reviews	Science
He et al. (2019)	Video + homework	Quiz	Noncollaborative	Reviews	Science
Heuett (2017)	Video	None	Noncollaborative	Activities & quizzes	Math
Heyborne & Perrett (2016)	Video	None	Noncollaborative	Discussions & problem-solving	Science
Hotle & Garrow (2016)	Video	None	Noncollaborative	Lectures	Engineering
Hu et al. (2016)	Video	None	Noncollaborative	Discussions, quizzes	Technology
Ichinose & Clinkenbeard (2016)	Video + assessment	None	Noncollaborative	Problem-solving	Math
Jonsson (2015)	Video + quiz	None	Collaboration	Problem solving	Technology
Kennedy et al. (2015)	Video	None	Noncollaborative	Problem-solving	Math
Kim et al. (2014)	Video	None	Collaboration	Problem-solving	Engineering
	Video	None	Noncollaborative	Discussions & exams	Technology
Liebert et al. (2016)	Video + review	Review	Collaboration	Activities	Science
Lloyd & Ebener (2014)	Video + activities	Activities	Collaboration	Lecture & activities	Science
Love et al. (2014)	Video	None	Collaboration	Discussions, problem-solving, & quizzes	Math
Maciejewski (2016)	Video + quiz	Question & answer	Collaboration	Discussions	Math
Maneeratana et al. (2016)	Video + discussion	None	Noncollaborative	Lectures	Engineering
Mason et al. (2013)	Video + quiz	None	Collaboration	Problem-solving	Engineering
Mattis (2014)	Video	None	Noncollaborative	Lectures & problem-solving	Math
McLaughlin et al. (2013)	Video	None	Noncollaborative	Exercises & quizzes	Science
Missildine et al. (2013)	Video	None	Noncollaborative	Simulations, games & exercises	Science
Moffett & Mill (2014)	Presentation + reading + quizzes	Quiz	Noncollaborative	Activities & discussions	Science

Table 1. *(Continued)* Description of Flipped Instruction Features in Individual Studies

Author	Pre-class instruction features	In-class strategy to assess pre-class learning	Structure of in-class activity	In-class teaching approach	Subject domain
Morton et al. (2017)	Video + workbook + notes	Worksheet	Collaboration	Discussions	Science
Moudgalya et al. (2016)	Video	Quiz	Noncollaborative	Discussions, quizzes, & tutorials	Engineering
Munson & Pierce (2015)	Video	Test	Noncollaborative	Activities & discussions	Science
Murray et al. (2014)	Presentation+ notetaking + question preparation	Question & answer	Noncollaborative	Lectures, question & answers	Science
Ojennus (2016)	Presentation + exercises	None	Collaboration	Problem-solving	Science
Ozpinar et al. (2016)	Video	Question & answers	Collaboration	Discussions, problem-solving, & practices	Technology
Peterson (2016)	Video + questions preparation	Question & answers	Collaboration	Problem-solving & quizzes	Mathematics
Ranalli & Moore (2016)	Video	Quizzes	Noncollaborative	Problem-solving & homework	Engineering
Reid (2016)	Video + assignment	None	Noncollaborative	Discussions & problem-solving	Science
Rui et al. (2017)	Video + discussion	Question & answers	Noncollaborative	Discussions & problem solving	Science
Ryan & Reid (2016)	Video	None	Noncollaborative	Discussions & problem-solving	Science
Sahin et al. (2015)	Video	None	Noncollaborative	Problem-solving	Math
Salama et al. (2016)	Video		Noncollaborative	Lecture	Engineering
Saterbak et al. (2016)	Video + quiz	None	Collaboration	Discussion/ problem solving	Engineering
Sengel (2014)	Video + problem-solving	Question & answers	Collaboration	Discussion, problem-solving	Science
Şengel (2016)	Video	None	Noncollaborative	Discussion, problem-solving	Science
Street et al. (2014)	Video + quizzes	Quiz	Collaboration	Interactive activities	Science
Sun & Wu (2016)	Video	None	Collaboration	Activities	Science
Timmerman et al. (2016)	Video + quiz	None	Noncollaborative	Activities, quizzes	Engineering
Tsai et al. (2015)	Video + online discussions	None	Collaboration	Activities, discussions	Technology

Table 1. *(Continued)* Description of Flipped Instruction Features in Individual Studies

Author	Pre-class instruction features	In-class strategy to assess pre-class learning	Structure of in-class activity	In-class teaching approach	Subject domain
Tutrang & Schenke (2017)	Video	Quiz	Noncollaborative	Problem-solving	Science
Van Sickle (2015)	Video + notetaking	Quiz	Collaboration	Problem-solving	Math
Velegol et al. (2015)	Video + assessment	Question & answers	Noncollaborative	Problem-solving, field trips, & quizzes	Engineering
Wasserman et al. (2017)	Video + homework	None	Noncollaborative	Lectures & problem-solving	Math
Webster et al. (2016)	Video	None	Collaboration	Problem solving & quiz	Engineering
Whillier et al. (2015)	Video + readings + quizzes	None	Collaboration	Tutorials & problem-solving	Science
Yelamarthi & Drake (2015)	Video + quiz	Reflection	Collaboration	Problem-solving, quizzes	Engineering

Flipped Instruction as a Constructivist Learning Approach

Core principles of constructivist learning support FI methods. A fundamental underpinning of constructivist learning is that learners create meaning by interacting with meaningful ideas from a sensory input presented to them (Dewey et al., 1997; Hein, 1991; Richardson, 2003). That is, learning emanates from actions/situations that spur students to act on and construct knowledge from experiences that are meaningful to them (Bhattacharjee, 2015). As FI lies on the premise that students' active engagement with materials and responsibility for constructing knowledge in all stages of FI increases learning outcomes (Ranalli & Moore, 2016), at the center of constructivist learning is that deep learning is attained with the use of active learning techniques. (Bhattacharjee, 2015). One fundamental principle of constructivism is that the teacher's role changes from an all-knowing teacher to a facilitator who supervises students' work and activities and guides them to learn better. Unlike the traditional teaching approaches, FI advances this learner-centered pedagogical approach.

Constructivists view learning as a gradual process that benefits from prior knowledge and interaction. Students need opportunities to revisit, think about, play with, reflect on, and apply ideas; as repeated exposure to content, a learning material deepens learning (Hein, 1991). FI presents the pre-class opportunity to build prior knowledge that will accelerate in-class learning and the opportunity to relearn, practice, and demonstrate those skills in class (Cieliebak & Frie, 2016; Clark et al., 2016). These opportunities to make cognitive connections with prior knowledge facilitate deeper learning. Moreover, the FI method of engaging students in class/group discussions, problem-solving, and projects where they interact and share ideas creates the social setting that the constructivist learning approach encourages (Richardson, 2003).

Learning Benefits of Flipped Instruction

Repeated research shows that FI enhances student learning (e.g., Feng et al., 2016; Foldnes, 2016; Jonsson, 2015; Kostaris et al., 2017). The critical factor that has contributed to the success of FI is inherently its delivery style (Elmaleh & Shankararaman, 2017). It advances independent learning as the learning responsibility is shared between students and teachers (Abdelaziz, 2014). For FI to be effective, students must watch videos and engage in related work from home. The teacher primarily supports students during class as they develop higher order and meta-cognitive skills such as reflection, self-assessment, think-aloud, and summarizing and synthesizing information (Clark et al., 2016). In class, students have ample opportunities to critically analyze relevant aspects of course topics, taking ownership of their learning (Bradford et al., 2014).

Moreover, FI provides a form of individualized education every learner's needs can be addressed (Abdelaziz, 2014; Bergmann & Sams, 2012; Thai et al., 2017). Students and teachers interact in a "self-paced "pause and rewind" capability" (Clark et al., 2016, p. 5). In order words, ample time is given in pre-class for convenient self-paced learning and during in-class relearning of the materials where students come prepared to interact with the teacher and other students. This kind of individualized learning could enhance a deep understanding of content for all categories of learners (Abdelaziz, 2014).

STEM Subject Domains and Flipped Instruction

STEM subject domains demand constructivist learning methods such as FI. STEM subjects are often identified with complex topics that are challenging for students to understand (Clark et al., 2016). Achieving optimal learning of such subjects requires learning methods that help learners construct their understanding of the topics (Bradford et al., 2014; Smith et al., 2014). However, STEM subject domains are known for using lecture-based instruction during class and leaving application activities for students to complete at home (Love et al., 2014; Margulieux et al., 2015). Recently, research has shown that such instructional structure does not provide the support that students need to deeply understand and acquire application such knowledge in STEM subjects (Margulieux et al., 2015). Constructivist learning methods like FI, which allow students to construct knowledge and receive feedback during application activities, are called for in STEM subject domains (Clark et al., 2016; Margulieux et al., 2015). For instance, a meta-analysis that investigated the impact of active learning on students' performance in STEM subject domains by Freeman et al. (2014) showed that students' learning outcomes in active learning environments increased significantly compared to traditional learning environments.

In many STEM subject domains where FI has been employed, student learning outcomes increased compared to when they were taught with TI (e.g., Akkaraju, 2016; Ichinose & Clinkenbeard, 2016; Maciejewski, 2016; Missildine et al., 2013; Olakanmi, 2017; Peterson, 2015). With such positive findings, an important question to answer is about the role of FI in STEM subject domains. Pre-class sessions of FI give students opportunities to familiarize themselves with complex topics and abstract concepts, establishing some prior knowledge that will promote deeper understanding in class (Margulieux et al., 2015). The structure of FI allows students to identify challenging concepts and incomprehensible problems during pre-class sessions. At the same time, the teacher facilitates extensive discussions and hands-on activities that focus on such issues. By engaging students in activities stimulating them to process and construct knowledge deeply, FI contributes to teaching innovations in college STEM subject domains as lecture-based TI continues to phase out. Building prior knowledge before class and focusing on more challenging topics is a virtuous approach that could make up for differences in students' backgrounds (Clark et al., 2016).

Previous Reviews and Meta-Analyses

The effect of FI on college student learning has been repeatedly studied over the last five years. Generally, previous reviews have provided some insights on its use in different/specific subject domains, its research trends, benefits, and challenges. For example, Akçayır and Akçayır (2018) reviewed its advantages and challenges. Rahman et al. (2014) conducted a descriptive review of studies on FI comprising the instruments used to measure learning achievement, the disciplines where it has been applied, and brief explanations of individual studies, including the contexts, focus, and results. Zainuddin and Halili (2016) reviewed 20 studies on the trends in FI research in different subject domains within the period 2013-2015. O'Flaherty and Phillips (2015) reviewed 28 studies on the use of FI in higher education with a focus on the types of resources and development of resources used in flipped classrooms, the activities in pre-class and in-class sessions, perception of staff and students, and the design and learning outcomes of FI.

Regarding the effect of FI on college students' learning, some reviewers have reported positive effects of FI in studies (e.g., Rahman et al., 2014; Zainuddin & Halili, 2016), while some have reported mixed findings (e.g., Hughes & Lyons, 2017; Kozikoğlu, 2019; Margulieux et al., 2015; O'Flaherty & Phillips, 2015). Two other reviews on FI related to STEM-subject domains reported differing findings. In the more subject-specific one, Karabulut-Ilgu et al. (2018) qualitatively analyzed 62 studies on the use of FI in engineering education and reported mixed results on the effect of FI. On the contrary, Huber and Werner (2016) authored a review of the use of FI in STEM subject domains, examining 58 studies in higher education, and found no statistical difference between FI and TI methods. This heterogeneous trend of reports on the effect of FI on college students' learning is not limited to previous literature reviews.

Previous meta-analyses have shown mixed findings on the moderator effect of subject domains on the effect of FI. In a prior meta-analysis on the effectiveness of FI on college students' learning in diverse subject domains, Shi et al. (2020) analyzed the effect sizes of 33 studies published from 2013 to 2017. They showed an overall positive effect size ($Z = 6.22$, $P < 0.05$) for FI. As the subject domain was one of the moderator variables analyzed, it was found to be a non-significant moderator. Cheng et al. (2019) examined FI's effects on students' learning in diverse subject areas in both K-12 and college education, analyzing 115 effect sizes. The results showed an overall statistically significant effect size ($g = 0.193$, $p < .001$, 95% CI [0.113–0.274]). Further analysis of moderator variables, including subject areas (mathematics, science, social sciences, engineering, arts and humanities, health, and business), showed that subject domains significantly moderated the effect of FI.

Likewise, Låg and Sæle (2019) conducted a meta-analysis on FI for both K-12 and college levels and diverse subject domains. They found overall positive learning effect of FI ($g = 0.35$, $p < .001$, 95% CI [0.31, 0.40]) while subject domains significantly moderated the effect. In a similar meta-analysis of 174 studies, Strelan et al. (2020) found an overall moderate positive effect of FI ($g = 0.50$) on student performance across education and subject areas. However, levels of education and subject areas significantly moderated the effect size. In addition, findings from Jang and Kim's (2020) meta-analysis of 43 FI studies on students' cognitive, affective, and interpersonal outcomes in different college subject domains showed that FI had a lower effect on students' cognitive outcomes (ES = 0.24) than affective (ES = 0.59) and interpersonal (ES = 0.53) outcomes. Moreover, higher positive effects were found for some subject domains than others (Computer Science, ES = 0.96, Education, ES = 0.80, Psychology, ES = 0.45, Nursing, ES = 0.44, Mixed subject, ES = 0.44, Science, ES = 0.43, Medicine, ES = 0.37, Pharmaceutical Medicine, ES = 0.29, Physiology, ES = 0.22, Management, ES = 0.19, Engineering, ES = 0.17, Physics, ES = 0.15, Math, ES = 0.13, English, ES = 0.12, Chemistry, ES = 0.10, and Business, ES = −0.93). Chen, Monrouxe, et al. (2018) also analyzed 46 health and non-health professions education studies and found a statistically significant positive effect size in favor of FI on students' examination scores ($g = 0.47$), while subject domains were a significant moderator. While Chen, Monrouxe, et al. (2018), Cheng et al. (2019), Jang and Kim (2020); Låg and Sæle (2019), and Strelan et al. (2020) found subject domains to be a significant moderator of the effect of FI, Shi et al. (2020) reported it as a non-significant moderator. Such findings neither provide us with an adequate understanding of the effect of FI nor the circumstances under which FI is advantageous or not to college students learning in STEM subject domains.

In addition, mixed results were found among some other meta-analyses that specifically focused on the use of FI in STEM-related subject domains, including pharmaceutical education (Gillette et al., 2018), health professions education (Hew & lo, 2018), and nursing education (Hu, Gao, et al., 2018; Tan et al., 2017). Hew and Lo (2018) analyzed 28 studies that compared FI with TI in health professions (e.g., medicine, pharmacy, and nursing). The results showed a statistically significant effect size for FI ($SMD = 0.33$, 95% CI 0.21–0.46, $p < 0.001$) while starting in-class sessions with quizzes that assessed students' pre-class learning was the only significant moderator variable ($Q = 5.34$, $df = 1$, $p = 0.02$). Gillette et al. (2018) conducted a meta-analysis synthesizing five studies in pharmaceutical education and found no significant effect for FI. Tan et al. (2017) conducted a meta-analysis on the effect of FI with 29 studies in nursing education, while Hu, Gao, et al. (2018) conducted a like study with 11 studies. The two studies found positive effects, with Tan et al. (2017) reporting increased students' academic performance ($SMD = 1.13$), skills ($SMD = 1.68$), and self-learning abilities ($SMD = 1.51$). Hu et al. (2018) reported higher theoretical knowledge and skills for the FI group ($SMD = 1.06$, 95% CI: 0.70–1.41, $p < 0.001$). However, these subject-specific meta-analyses provide no full grasp of the effect of FI on college students' learning in STEM subject domains in general. A lack of comprehensive, evidence-based meta-analysis that shows the conditions under which FI enhances or inhibits college students' learning in STEM subject domains calls for further meta-analysis in this regard.

Challenges of Flipped Instruction

While FI yields numerous learning benefits, it presents some challenges as well. Some of the practical difficulties of designing FI include students' failure to watch pre-class videos, poor quality of videos, and time constraints for instructors to produce quality videos (Akçayır & Akçayır, 2018; Peterson, 2015; Zainuddin & Halili, 2016). To ensure that students watch videos before class, some researchers suggest that videos be accompanied by tasks such as quizzes, discussion, note-taking, or question preparation (Aşıksoy & Özdamlı, 2016; Jonsson, 2015; Peterson, 2015). Moreover, a suspected challenge to implementing FI could be the inaccessibility of technology to all students, although this has not been reported in the literature.

Studies have indicated that students find it challenging to adapt to FI and have negative responses at its introductory stage (Missildine et al., 2013; Zainuddin & Halili, 2016). However, they develop good perceptions of it after a while of getting familiar with its structure, which explains the high positive perceptions among students reported in studies (e.g., Chen, Yang, et al., 2016; Gullayanon, 2014; Khan & Ibrahim, 2017; Prashar, 2015). O'Flaherty and Phillips (2015) suggested the introduction of FI to students earlier in their programs to get them accustomed to it. As research on FI has produced mixed findings (Bradford et al., 2014; Fryling et al., 2016; Maneeratana et al., 2016; Ryan & Reid, 2015; Saterbak et al., 2016; Strayer, 2012), some scholars have attributed this inconsistency of findings to its design (e.g., de Araujo et al., 2017; Peterson, 2015; Strayer, 2012). To address this challenge, Peterson (2015) emphasized the need for the same teachers to be responsible for the content creation of both the pre-class and in-class sections to ensure that pre-class content meshes well with the in-class activities.

Method

We strictly followed published guidelines for conducting meta-analyses outlined by Adesope et al. (2010), Adesope et al., 2017, Nesbit and Adesope (2006), and Lipsey and Wilson (2001), as well as the PRISMA 2009 checklist (Moher et al., 2009).

Purpose of the Present Meta-Analysis

The present meta-analysis focuses on studies that assessed college student learning outcomes with FI in STEM subject domains. In this study, learning outcomes are the results associated with instructional experiences, as measured and reported in the FI studies (Anderson et al., 2005). Studies in this meta-analysis reported measuring students' learning outcomes as their performance on tests of knowledge retention, problem-solving, and conceptual and procedural knowledge of learned materials. Tests used in the studies varied in formats, including open-ended, short-answer, and multiple-choice items. While FI is shown to yield increased learning outcomes, less is known about its effects on different populations and specific subject domains due to the mixed findings in previous research. Its impact on college students' learning has been investigated in earlier meta-analyses in various subject domains (Chen et al., 2018; Cheng et al., 2019; Gillette et al., 2018; Hew & Lo, 2018; Hu, Gao, et al., 2018; Låg & Sæle, 2019; Shi et al., 2020; Tan et al., 2017). Interestingly, analyses of moderator variables in the diverse-subject-area meta-analyses showed mixed findings on whether the subject domain was a significant moderator.

Moreover, none of the previous meta-analyses illuminates the learning effect of FI in STEM subject domains in general and the conditions under which FI is beneficial in STEM subject domains, such as the instructional features of FI in the studies. The present meta-analysis investigates the learning effects of FI in STEM subject domains and seeks to uncover the different conditions under which it is effective. Hence, this meta-analysis aims to answer the following research questions:

1. What are the effects of flipped instruction compared to traditional education in STEM subject domains?
2. How do the effects of flipped instruction vary across different STEM subject domains?
3. How are the effects of flipped instruction moderated by instructional features, design features, and context?

Inclusion Criteria

To carry out this meta-analysis, some inclusion criteria were established. For a study to be included in this meta-analysis, it must:

- compare the effect of FI with TI in STEM subject domains
- include at least a visual-audio presentation of a les son for the FI group
- the report measured cognitive learning outcomes
- report sufficient data (basic statistics such as means, standard error, and standard deviation) for effect size extraction or calculation

- be publicly available either online or in library archives
- be peer-reviewed

Location and Selection of Studies

We used the search terms "flip OR flipped OR inverted classroom OR flipped instruction/classroom" to retrieve relevant studies. In IEEE, ERIC, PsycINFO, EBSCOhost, and Google Scholar, studies were searched. In addition, we explored the reference lists of selected studies referred to as other sources in Figure 1. Studies were searched from fall 2018 to fall 2020. A total of 73 papers that met the inclusion criteria were included in this meta-analysis. Figure 1 shows the location, screening, and selection of studies.

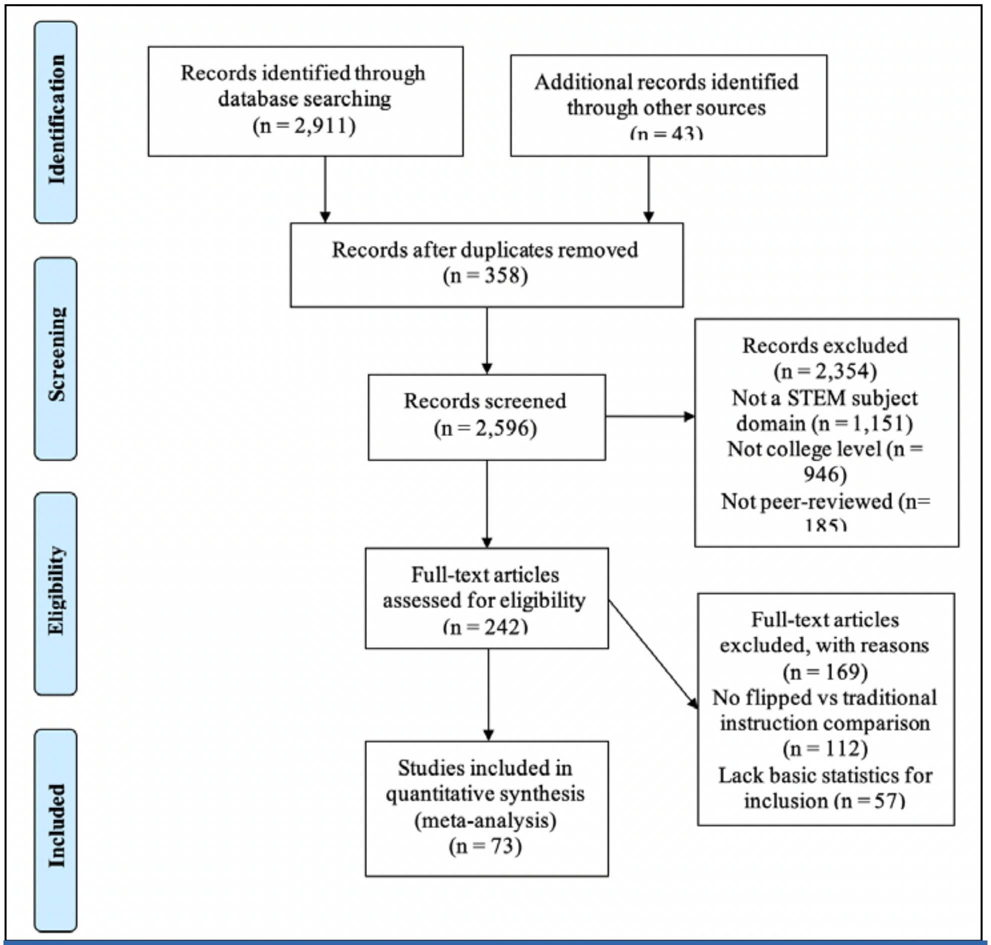


Figure 1. PRISMA Flow Chart of Location and Selection of Studies

Extraction and Calculation of Effect Sizes

Two researchers independently coded all 73 studies by extracting variables such as source (conference or journal), media of presentation (pen and paper, computer only or mixed), subject domain (science, technology, engineering, and mathematics), country (the US and "other countries"), design features which included randomization or no randomization, control of prior differences or none, test format (open-ended, multiple choice, mixed and unknown), media of test presentation (pen and paper or computer), and video types (duration: 0-15 minutes, 16-30 minutes, and 31-60 minutes and audiovisual format: video-based or narrated PowerPoint), pre-class instructional features (video only, video with non-assessment activities and video with assessment), strategy for starting in-class sessions (pre-class learning evaluation or no assessment), and structure of in-class activities (collaboration or no collaboration). We also coded the sample sizes, means, standard deviations, and effect sizes or other statistics provided by the authors that will enable effect size calculation. The mean interrater agreement among coders is 98% on all the coded variables, which indicates high reliability, according to Howell (2007).

Cohen's *d* effect size was calculated for each study included in this meta-analysis to obtain a standardized estimate of the difference in learning outcomes between students who learned with FI and those who learned with TI. For a few studies that did not provide basic descriptive statistics (e.g., mean and standard deviation), the effect sizes were obtained through estimations from other statistics using conversion formulas. For instance, since such studies provided standard errors, the standard deviations were obtained by converting the standard error to standard deviation using the formula. A positive and significant effect size indicates that FI is beneficial for learning.

Data Analysis

While following standard guidelines for running a meta-analysis of studies (Authors, 2010, 2017; Cooper & Hedges, 1994; Lipsey & Wilson, 2001), data were analyzed using Comprehensive Meta-Analysis 3.0 (Borenstein et al., 2008). The weighted mean effect sizes were aggregated to an overall weighted mean estimate of the effect of FI (i.e., $g+$). The significance of each weighted mean effect size was determined by its 95% confidence interval. When the lower confidence interval limit was greater than zero, the mean effect size indicated a statistically detectable result favoring FI. Since all the effect sizes were combined into a mean to determine if they all estimated the same population effect size, homogeneity of variance was examined using the *Q* and associated statistics generated by CMA. A p-value less than .05 means that the mean effect size is heterogeneous and does not estimate a common population mean.

Results

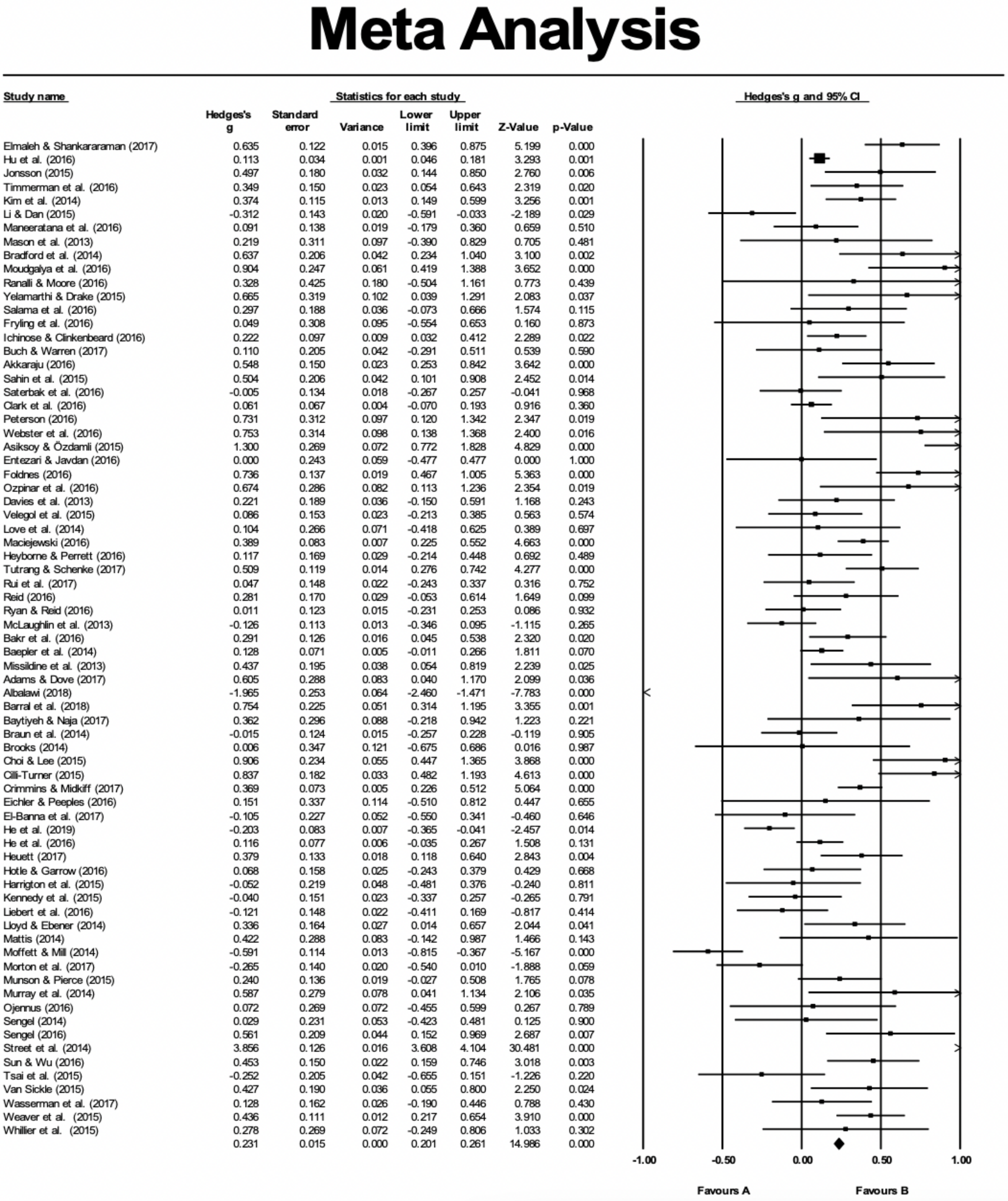
In all, 73 independent effect sizes were analyzed. Figure 2 shows a summary of statistics for individual studies, including their effect sizes (Hedges' *g*), standard errors, homogeneity of variance, *z* statistics value, significance levels (*p*-value), and confidence intervals. The positive effect size indicates that FI is beneficial for learning.

Effects of Flipped Instruction

The major purpose of this meta-analysis was to investigate the effect of FI on college students learning in STEM subject domains. Table 2 shows the overall results of the weighted mean of all independent effect sizes analyzed. Specifically, the table shows the number of participants (*N*), the weighted mean effect size ($g+$) and its standard error (SE), the 95% lower and upper confidence intervals (CI), the results of a test of homogeneity (*Q*) along with its degrees of freedom (*df*), and the percentage of variability that is due to true heterogeneity or between-studies variability (I^2). Overall, the results showed that FI is effective for learning STEM subjects (g = 0.23). The distribution was heterogenous $Q\ (72) = 1198.26$, $p < .001$, $I^2 = 93.99$. A total of 94% of the variance due to true heterogeneity was between-studies variance, while 6% of the variance was within-study variance based on sampling error. More variability was found among independent effect sizes than samples from a single population. Because of the high heterogeneity between studies, moderator analyses were further conducted.

Analysis of Moderator Effects on Flipped Instruction

Moderator effects were analyzed for ten variables to examine the conditions under which FI was more or less effective for learning STEM subjects. The analyses were organized into research design features (control of prior differences, randomization of participants, and test format), instructional features (pre-class instructional features, in-



Meta Analysis

Study name	Statistics for each study						
	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Elmaleh & Shankararaman (2017)	0.635	0.122	0.015	0.396	0.875	5.199	0.000
Hu et al. (2016)	0.113	0.034	0.001	0.046	0.181	3.293	0.001
Jonsson (2015)	0.497	0.180	0.032	0.144	0.850	2.760	0.006
Timmerman et al. (2016)	0.349	0.150	0.023	0.054	0.643	2.319	0.020
Kim et al. (2014)	0.374	0.115	0.013	0.149	0.599	3.256	0.001
Li & Dan (2015)	-0.312	0.143	0.020	-0.591	-0.033	-2.189	0.029
Maneeratana et al. (2016)	0.091	0.138	0.019	-0.179	0.360	0.659	0.510
Mason et al. (2013)	0.219	0.311	0.097	-0.390	0.829	0.705	0.481
Bradford et al. (2014)	0.637	0.206	0.042	0.234	1.040	3.100	0.002
Moudgalya et al. (2016)	0.904	0.247	0.061	0.419	1.388	3.652	0.000
Ranalli & Moore (2016)	0.328	0.425	0.180	-0.504	1.161	0.773	0.439
Yelamarthi & Drake (2015)	0.665	0.319	0.102	0.039	1.291	2.083	0.037
Salama et al. (2016)	0.297	0.188	0.036	-0.073	0.666	1.574	0.115
Fryling et al. (2016)	0.049	0.308	0.095	-0.554	0.653	0.160	0.873
Ichinose & Clinkenbeard (2016)	0.222	0.097	0.009	0.032	0.412	2.289	0.022
Buch & Warren (2017)	0.110	0.205	0.042	-0.291	0.511	0.539	0.590
Akkaraju (2016)	0.548	0.150	0.023	0.253	0.842	3.642	0.000
Sahin et al. (2015)	0.504	0.206	0.042	0.101	0.908	2.452	0.014
Saterbak et al. (2016)	-0.005	0.134	0.018	-0.267	0.257	-0.041	0.968
Clark et al. (2016)	0.061	0.067	0.004	-0.070	0.193	0.916	0.360
Peterson (2016)	0.731	0.312	0.097	0.120	1.342	2.347	0.019
Webster et al. (2016)	0.753	0.314	0.098	0.138	1.368	2.400	0.016
Asiksoy & Özdamli (2015)	1.300	0.269	0.072	0.772	1.828	4.829	0.000
Entezari & Javdan (2016)	0.000	0.243	0.059	-0.477	0.477	0.000	1.000
Foldnes (2016)	0.736	0.137	0.019	0.467	1.005	5.363	0.000
Ozpinar et al. (2016)	0.674	0.286	0.082	0.113	1.236	2.354	0.019
Davies et al. (2013)	0.221	0.189	0.036	-0.150	0.591	1.168	0.243
Velegol et al. (2015)	0.086	0.153	0.023	-0.213	0.385	0.563	0.574
Love et al. (2014)	0.104	0.266	0.071	-0.418	0.625	0.389	0.697
Maciejewski (2016)	0.389	0.083	0.007	0.225	0.552	4.663	0.000
Heyborne & Perrett (2016)	0.117	0.169	0.029	-0.214	0.448	0.692	0.489
Tutrang & Schenke (2017)	0.509	0.119	0.014	0.276	0.742	4.277	0.000
Rui et al. (2017)	0.047	0.148	0.022	-0.243	0.337	0.316	0.752
Reid (2016)	0.281	0.170	0.029	-0.053	0.614	1.649	0.099
Ryan & Reid (2016)	0.011	0.123	0.015	-0.231	0.253	0.086	0.932
McLaughlin et al. (2013)	-0.126	0.113	0.013	-0.346	0.095	-1.115	0.265
Bakr et al. (2016)	0.291	0.126	0.016	0.045	0.538	2.320	0.020
Baepler et al. (2014)	0.128	0.071	0.005	-0.011	0.266	1.811	0.070
Missildine et al. (2013)	0.437	0.195	0.038	0.054	0.819	2.239	0.025
Adams & Dove (2017)	0.605	0.288	0.083	0.040	1.170	2.099	0.036
Albalawi (2018)	-1.965	0.253	0.064	-2.460	-1.471	-7.783	0.000
Barral et al. (2018)	0.754	0.225	0.051	0.314	1.195	3.355	0.001
Baytiyeh & Naja (2017)	0.362	0.296	0.088	-0.218	0.942	1.223	0.221
Braun et al. (2014)	-0.015	0.124	0.015	-0.257	0.228	-0.119	0.905
Brooks (2014)	0.006	0.347	0.121	-0.675	0.686	0.016	0.987
Choi & Lee (2015)	0.906	0.234	0.055	0.447	1.365	3.868	0.000
Cilli-Turner (2015)	0.837	0.182	0.033	0.482	1.193	4.613	0.000
Crimmins & Midkiff (2017)	0.369	0.073	0.005	0.226	0.512	5.064	0.000
Eichler & Peeples (2016)	0.151	0.337	0.114	-0.510	0.812	0.447	0.655
El-Banna et al. (2017)	-0.105	0.227	0.052	-0.550	0.341	-0.460	0.646
He et al. (2019)	-0.203	0.083	0.007	-0.365	-0.041	-2.457	0.014
He et al. (2016)	0.116	0.077	0.006	-0.035	0.267	1.508	0.131
Heuett (2017)	0.379	0.133	0.018	0.118	0.640	2.843	0.004
Hotle & Garrow (2016)	0.068	0.158	0.025	-0.243	0.379	0.429	0.668
Harrigton et al. (2015)	-0.052	0.219	0.048	-0.481	0.376	-0.240	0.811
Kennedy et al. (2015)	-0.040	0.151	0.023	-0.337	0.257	-0.265	0.791
Liebert et al. (2016)	-0.121	0.148	0.022	-0.411	0.169	-0.817	0.414
Lloyd & Ebener (2014)	0.336	0.164	0.027	0.014	0.657	2.044	0.041
Mattis (2014)	0.422	0.288	0.083	-0.142	0.987	1.466	0.143
Moffett & Mill (2014)	-0.591	0.114	0.013	-0.815	-0.367	-5.167	0.000
Morton et al. (2017)	-0.265	0.140	0.020	-0.540	0.010	-1.888	0.059
Munson & Pierce (2015)	0.240	0.136	0.019	-0.027	0.508	1.765	0.078
Murray et al. (2014)	0.587	0.279	0.078	0.041	1.134	2.106	0.035
Ojennus (2016)	0.072	0.269	0.072	-0.455	0.599	0.267	0.789
Sengel (2014)	0.029	0.231	0.053	-0.423	0.481	0.125	0.900
Sengel (2016)	0.561	0.209	0.044	0.152	0.969	2.687	0.007
Street et al. (2014)	3.856	0.126	0.016	3.608	4.104	30.481	0.000
Sun & Wu (2016)	0.453	0.150	0.022	0.159	0.746	3.018	0.003
Tsai et al. (2015)	-0.252	0.205	0.042	-0.655	0.151	-1.226	0.220
Van Sickle (2015)	0.427	0.190	0.036	0.055	0.800	2.250	0.024
Wasserman et al. (2017)	0.128	0.162	0.026	-0.190	0.446	0.788	0.430
Weaver et al. (2015)	0.436	0.111	0.012	0.217	0.654	3.910	0.000
Whillier et al. (2015)	0.278	0.269	0.072	-0.249	0.806	1.033	0.302
	0.231	0.015	0.000	0.201	0.261	14.986	0.000

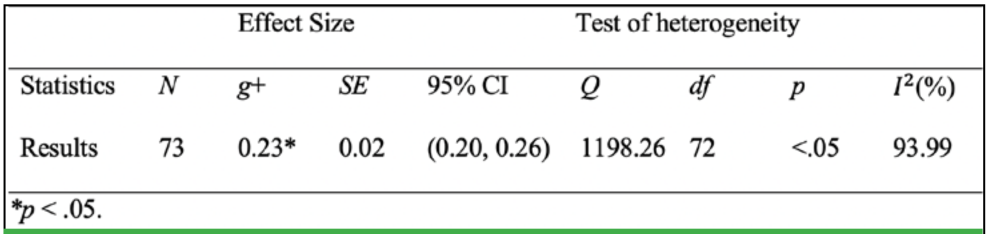
Figure 2. Statistics for Individual Studies

class starting, structure of in-class activities), and context (subject domain, country, and source). The results are presented in the following section, while their significance is discussed in the discussion section. Other variables coded for, such as video types and media of test presentation, were excluded from the analyses because most studies did not provide information on them. In contrast, the treatment context was excluded because all the study treatments were conducted in a classroom context.

Instructional Features

The instructional features of FI, including the pre-class and in-class features, were considered essential variables that could have moderated the effect of FI since they are a critical part of the design and implementation of FI. While coding studies for this meta-analysis, it was observed that studies had different ways of designing and implementing FI. For the pre-class session, some authors provided students with some tasks to do alongside watching videos in preparation for the in-class session, while others did not. The tasks given to students were categorized into assessment tasks (e.g., quizzes and solving exercises) or non-assessment tasks (e.g., online discussions and preparing questions for class). As shown in Table 3, the results showed that pre-class instructional features were a significant moderator variable with a significant between-levels difference $QB(2) = 44.64$ ($p < .001$). Giving students videos with an assessment task produced the highest effect for FI ($g = 0.38$). In contrast, smaller effects were associated with studies that had videos with non-assessment activities (($g = 0.18$) and videos only ($g = 0.16$). The post hoc analysis further demonstrated that giving students videos with assessment tasks was associated with a larger weighted mean effect size.

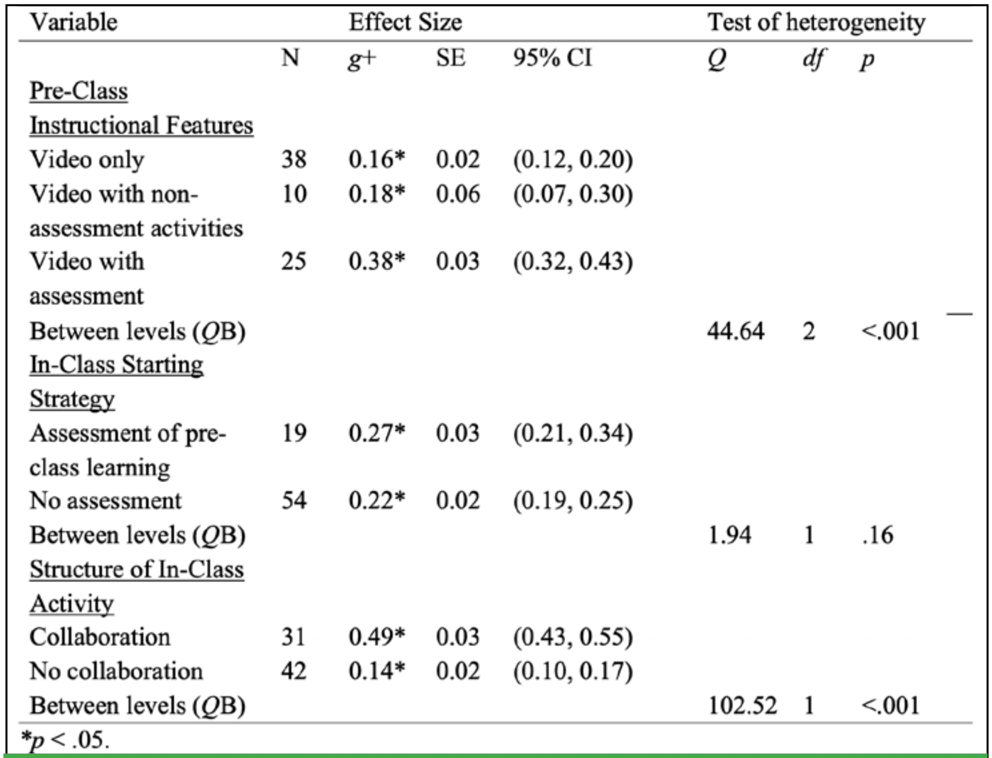
For the in-class instructional features, the instructor's strategy for starting in-class sessions, which includes whether there was an assessment of pre-class learning (using quizzes or question and answers) or not, and the structure of in-class activities, which included the use of collaboration among students (e.g., group discussions and group problem solving) or not were analyzed. The results showed that the strategy for starting in-class sessions was not a significant moderator of the effect of FI QB(1) = 1.93, (p = .16). Nevertheless, the structure of in-class activities was a significant moderator with a significant between-levels difference QB(1) = 102.52, (p < .001). Studies where collaboration was used showed a large effect for FI (g = 0.49) and were significantly different from studies that did not involve collaboration (g = 0.14).



		Effect Size			Test of heterogeneity			
Statistics	*N*	*g*+	*SE*	95% CI	*Q*	*df*	*p*	I^2(%)
Results	73	0.23*	0.02	(0.20, 0.26)	1198.26	72	<.05	93.99

*$p < .05$.

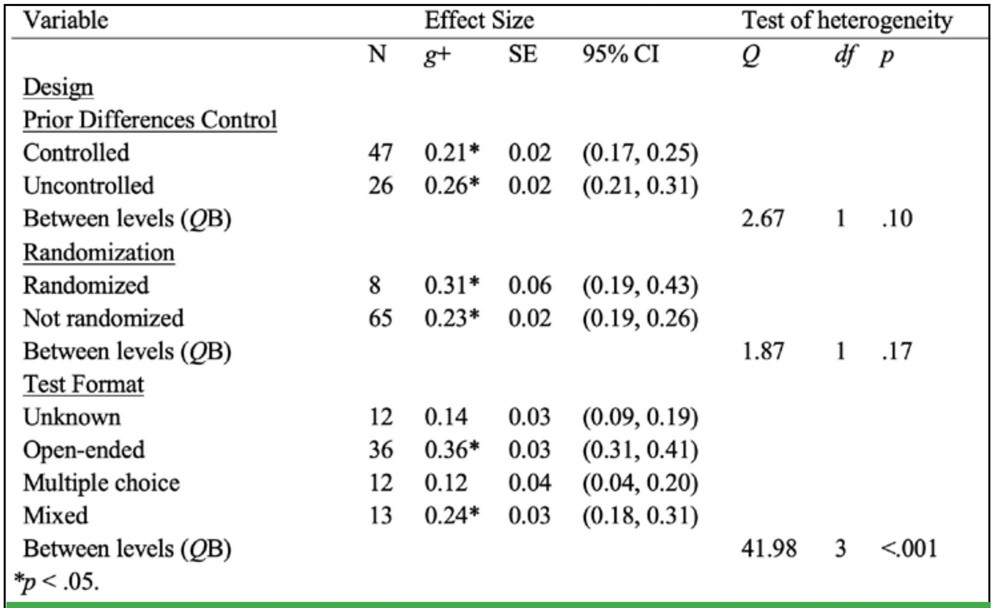
Table 2. Overall Weighted Mean Effect Size for all Studies



Variable	Effect Size				Test of heterogeneity		
	N	*g*+	SE	95% CI	*Q*	*df*	*p*
Pre-Class Instructional Features							
Video only	38	0.16*	0.02	(0.12, 0.20)			
Video with non-assessment activities	10	0.18*	0.06	(0.07, 0.30)			
Video with assessment	25	0.38*	0.03	(0.32, 0.43)			
Between levels (*Q*B)					44.64	2	<.001
In-Class Starting Strategy							
Assessment of pre-class learning	19	0.27*	0.03	(0.21, 0.34)			
No assessment	54	0.22*	0.02	(0.19, 0.25)			
Between levels (*Q*B)					1.94	1	.16
Structure of In-Class Activity							
Collaboration	31	0.49*	0.03	(0.43, 0.55)			
No collaboration	42	0.14*	0.02	(0.10, 0.17)			
Between levels (*Q*B)					102.52	1	<.001

*$p < .05$.

Table 3. Weighted Mean Effect Sizes for Instructional Features



Variable	Effect Size				Test of heterogeneity		
	N	*g*+	SE	95% CI	*Q*	*df*	*p*
Design							
Prior Differences Control							
Controlled	47	0.21*	0.02	(0.17, 0.25)			
Uncontrolled	26	0.26*	0.02	(0.21, 0.31)			
Between levels (*Q*B)					2.67	1	.10
Randomization							
Randomized	8	0.31*	0.06	(0.19, 0.43)			
Not randomized	65	0.23*	0.02	(0.19, 0.26)			
Between levels (*Q*B)					1.87	1	.17
Test Format							
Unknown	12	0.14	0.03	(0.09, 0.19)			
Open-ended	36	0.36*	0.03	(0.31, 0.41)			
Multiple choice	12	0.12	0.04	(0.04, 0.20)			
Mixed	13	0.24*	0.03	(0.18, 0.31)			
Between levels (*Q*B)					41.98	3	<.001

*$p < .05$.

Table 4. Weighted Mean Effect Sizes for Design Features

Design Features

Because design is critical to the degree of effectiveness of FI (Strayer, 2012), control of prior differences, randomization of participants, and test format, were analyzed to see how they might have influenced the effect of FI. Table 4 presents the results of the analyses. Controlling for students' prior knowledge differences and randomization of participants when assigning them to study groups was an essential factor to consider in judging the effectiveness of FI as it was part of the study methodology (Peterson, 2015). However, it was not a statistically significant moderator. The test format of the studies was a significant moderator of the effect of FI $Q_B(3) = 41.98$ ($p < .001$). Studies with open-ended tests showed the highest effect for FI ($g = 0.36$), followed by studies with a mix of open-ended and multiple-choice tests ($g = .24$). In contrast, studies with multiple-choice tests showed the least effect ($g = .12$). Post hoc analysis revealed that studies that had open-ended test formats yield high weighted mean effect size compared to studies in other categories of the test format.

Context

STEM subject domains, country of study, and source of papers were analyzed as moderator variables under context. Since this meta-analysis specifically investigated the effect of FI on STEM subject domains, it was important to examine whether there were differential effects of FI in different subject domains. Results are presented in Table 5. Results showed that the subject domain was a significant moderator of the overall effect of FI $Q_B(3) = 14.24$, $p < .001$. Higher effects were found in the mathematics domain ($g = 0.29$) and science domain ($g = 0.28$) while smaller effects were found in engineering ($g = 0.18$) and technology ($g = 0.15$). Post hoc analysis revealed that studies in mathematics and science domains were associated with a larger weighted mean effect size and were significantly different from studies in other STEM subject domains.

Another interesting moderator variable analyzed is the country where studies were conducted. We were specifically interested in exploring whether the effect of FI in studies conducted in the US might vary from that of studies carried out in other countries. Results showed that the between-levels difference was statistically significant, $QB(1) = 10.14$, $p < .001$, while studies conducted in the U.S. produced a larger weighted mean effect size ($g = .27$) than studies conducted outside the U.S. ($g = .17$). Post hoc analysis showed that the results in studies conducted in the US were significantly different from studies conducted outside the U.S. The source of papers, whether a study was published as a journal or conference paper, was not a significant moderator of the effect of FI.

Variable	Effect Size				Test of heterogeneity		
	N	$g+$	SE	95% CI	Q	df	p
Subject Domain							
Science	32	0.28*	0.02	(0.23, 0.32)			
Technology	10	0.15	0.03	(0.09, 0.21)			
Engineering	14	0.18*	0.04	(0.10, 0.26)			
Mathematics	17	0.29*	0.04	(0.21, 0.36)			
Between levels (QB)					14.25	3	<.001
Country							
US	48	0.27*	0.02	(0.23, 0.31)			
Others	25	0.17*	0.02	(0.13, 0.22)			
Between levels (QB)					10.14	1	<.001
Source							
Conference	13	0.19*	0.03	(0.13, 0.24)			
Journal	60	0.25*	0.02	(0.21, 0.29)			
Between levels (QB)					3.35	1	.07

*$p < .05$.

Table 5. **Weighted Mean Effect Sizes for Study Context**

Examining Publication Bias

One of the criticisms meta-analyses and systematic reviews have faced the potential for publication bias. Indeed, it is well-documented that non-significant studies may not be published as frequently as studies that produced statistically significant findings (Borenstein et al., 2009; Rosenthal & DiMatteo, 2001). Significant results are often favored "while nonsignificant results are relegated to file drawers" (Rosenthal & DiMatteo, 2001, p. 66). In the present study, Comprehensive Meta-Analysis was used to compute two approaches to examine the presence of potential publication bias that could threaten the validity of its overall result. The first was the funnel plot which estimates the effects of the individual studies against the standard error. The funnel plot, as presented in Figure 3, showed a nearly symmetrical distribution around the weighted mean effect. According to Song et al. (2002), symmetrical funnel plots indicate the absence of publication bias. Because the distribution of the obtained funnel plot was not perfectly symmetrically distributed, a second approach was further computed to cross-check the absence of publication bias in the study.

The second approach, "Classic Fail-safe N," was used to determine the number of null effect studies that would be needed to raise the p-value associated with the average effect above an arbitrary alpha level (set at $\alpha = .05$). This approach was used because it explains the robustness of significant results of meta-analysis studies. The results showed that it would take 4,359 additional studies to this meta-analysis to diminish the statistical significance of the results of this study. Obtaining 4,359 more studies to include in this meta-analysis is unrealistic and signifies the unlikeness that the effect size of this study can be reduced to zero. These results suggest an absence of publication bias in this study.

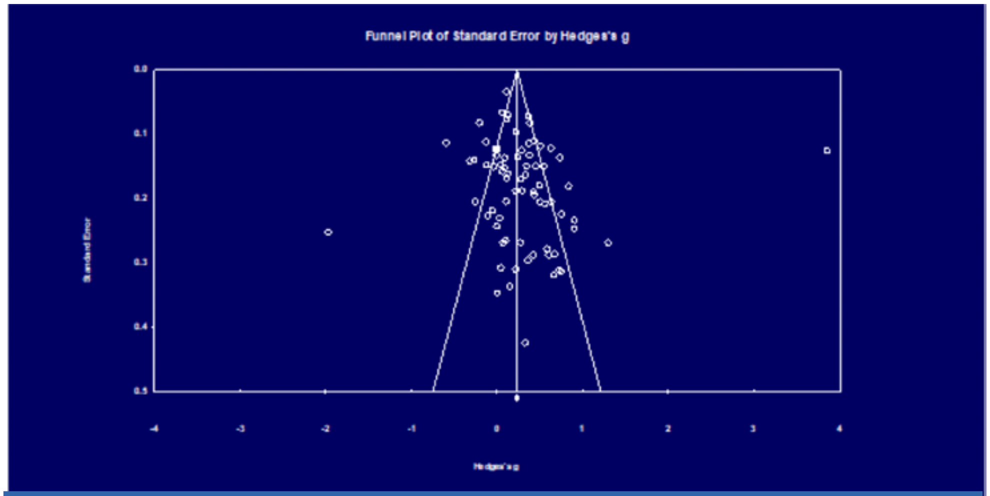


Figure 3. **Funnel Plot**

Discussion

This meta-analysis examined a sample of 73 studies that met the inclusion criteria for this study. Overall, the analysis revealed a statistically detectable positive effect of FI on learning STEM subjects. As no published meta-analysis investigated the effect of FI specifically on STEM subject domains, the findings of this meta-analysis demonstrate that FI is beneficial for learning in STEM subject domains at the college level. Moreover, the results of this meta-analysis align with the hypothesis and research findings that constructivist learning strategies like FI increase learning in STEM subject domains (Freeman et al., 2014; Margulieux et al., 2015). The rationale behind FI is that learning increases when students take active responsibility for learning and constructing knowledge (Kummer & Godoy, 2015). This meta-analysis informs classroom practice and contributes to theories of learning in that FI is one of the ways constructivist learning could be implemented, and the different instructional features implemented in the studies such as problem-solving, discussion, hands-on activities, simulations, and case studies offer practitioners a range of options on how to adapt FI in the classroom.

Moderator variables were analyzed to examine the situations under which FI had more excellent learning effects for STEM subject domains. Design of studies and methods for implementing FI have been associated with the mixed results found in some studies (Peterson, 2015; Strayer, 2012). The moderator variables that were majorly related to study design and implementation were analyzed. Three moderator variables were examined under the plan, and instructional features were insignificant (strategy for starting in-class sessions, randomization of participants, and control for prior knowledge differences). For the critical moderators, the results of the analyses confirmed some of the issues raised in the literature concerning FI to a reasonable extent. For instance, the pre-class instructional feature analysis results address one of the concerns about implementing FI. One of the challenges associated with FI is the failure of some students to watch pre-class videos. Some authors have suggested that videos must be accompanied by a task to address this challenge. The findings showed that FI was more effective when students had a task, mainly an assessment task, that evaluated their understanding of the video content or engaged them with the video. This finding is critical for designing and implementing FI, particularly in STEM domains, where students often encounter complex concepts/topics. FI might be more profitable when students have some assessment tasks to engage them with the pre-class videos. Likewise, the results of the analyses for test format indicate that the format of tests used to evaluate students' learning outcomes may need to be considered when implementing FI. The highest weighted mean effect associated with the open-ended test format, followed by a mix of open-ended and multiple-choice tests, suggests that giving students open-ended tests that are not only proxies for transfer-type questions but also allow students to demonstrate learning in detail might serve better in approximating the learning gains of FI than other test formats.

One of the strengths of FI lies in the use of constructivist learning methods that support students to construct knowledge and learn content deeply. In the studies included in this meta-analysis, the classroom activities were majorly problem-solving, discussion, hands-on activities, simulations, and case studies, which are common in flipped in-class learning section (Akçayır & Akçayır, 2018; DeLozier & Rhodes, 2017; Rahman et al., 2014). The structure of activities could not only determine the kinds of cognitive processes they elicit from students but also students learning and behaviors (DeLozier & Rhodes, 2017). As the structure of the activities was analyzed, the findings revealed that collaboration between/among students yielded greater learning effects. This result differs from the conclusions of another meta-analysis where both individual and collaborative active learning activities were equally effective (Shi et al., 2020). However, it aligns with the literature that supports collaborative learning in flipped classrooms to increase students' engagement and learning (DeLozier & Rhodes, 2017; Galway et al., 2014).
For study context, country of studies and STEM subject domains were significant moderators of the effect of FI. While it is puzzling that studies conducted in the U.S. showed higher effects for FI than studies from other countries, we encourage caution in interpreting the results. Future meta-analyses need to explore the impact of FI more robustly in different subject domains, educational levels, and study countries. More significant learning gains of FI were associated with math, followed by science, while lesser gains were associated with engineering and the least with technology. These findings align with those of a previous meta-analysis. Cheng et al. (2019) conducted moderator analyses for subject domains that included STEM and non-STEM domains and reported results for STEM subject domains, excluding the technology domain. The findings showed that the learning effects were higher for math, followed by science, while engineering was associated with the least effect. However, the previous meta-analysis was cautious in making conclusions on the results for engineering because it had only five studies representing the domain. The present meta-analysis had 14 studies representing engineering and produced fewer effects for engineering compared to math and science domains. FI has received significant attention in the engineering domain since 2012, with studies showing other gains of FI, including engagement and self-efficacy. However, most studies did not report descriptive statistics to compute effect sizes (Karabulut-Ilgu et al., 2018). Even though engineering and technology domains were associated with fewer effects, their effect size was statistically significant, suggesting that FI seems promising in all STEM subject domains.

Conclusion

It could be gleaned that FI is a virtuous approach to deepen learning and student engagement in college STEM subject domains, although previous research has reported mixed results. As discussed earlier in the previous meta-analysis section, the mixed results in research are likely attributable to some design issues related to the implementation of FI. While we found diverse modes of implementing FI, one similar element that cuts across studies was using videos for pre-class sessions. The quality of videos is part of design issues that have been raised because it is thought to have an impact on students' learning outcomes in the flipped classroom (Zainuddin & Halili, 2016). Video type was one of the variables coded but was not included in the moderator analyses because most studies did not report the video features needed for such analyses. Hence, we could neither objectively determine the quality of videos used in the studies nor evaluate how it might moderate the overall effect of FI. We recommend that future studies on FI report in detail the features of videos used to enable such evaluation in future meta-analyses and support teachers in translating research into practice.

Moreover, one thing identified in research on FI is that many studies focused on students' perceptions and satisfaction. In contrast, some studies did not report descriptive statistics to be included in the present meta-analysis. As some existing studies on students' perception (e.g., He et al., 2016; Roach, 2014; and Van Sickle, 2016) suggest that students positively perceive FI for reasons including an increased opportunity for interaction during in-class sessions, more studies that measure students' learning outcomes in STEM subject domains are recommended. As this meta-analysis only analyzed the effect of FI on college STEM subject domains, similar studies with K-12 students need to be analyzed to examine the robustness of FI across educational levels.

References

References marked with an asterisk indicate studies 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 & Development, 34*(1), 1-14.

*Adams, C., & Dove, A. (2018). Calculus students flipped out: The impact of flipped learning on calculus students' achievement and perceptions of learning. *Primus, 28*(6), 600-615.

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of educational research, 80*(2), 207-245.

Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research, 87*(3), 659-701.

Akçayır, G., & Akçayır, M. (2018). The flipped classroom: A review of its advantages and challenges. *Computers & Education, 126*(2018), 334-345.

*Akkaraju, S. (2016). The role of flipped learning in managing the cognitive load of a threshold concept in physiology. *Journal of Effective Teaching, 16*(3), 28-43.

* Albalawi, A. S. (2018). The effect of using flipped classroom in teaching calculus on students' achievement at University of Tabuk. *International Journal of Research in Education and Science (IJRES), 4*(1), 198-207.

Anderson, H. M., Moore, D. L., Anaya, G., & Bird, E. (2005). Student Learning Outcomes Assessment: A Component of Program Assessment. *American Journal of Pharmaceutical Education, 69*(2).

*Aşıksoy, G., & Özdamlı, F. (2016). Flipped classroom adapted to the ARCS model of motivation and applied to a Physics course. *Eurasia Journal of Mathematics, Science & Technology Education, 12*(6).

*Baepler, P., Walker, J. D., & Driessen, M. (2014). It's not about seat time: Blending, flipping, and efficiency in active learning classrooms. *Computers & Education, 78*, 227-236.

*Bakr, M. M., Massey, W. L., & Massa, H. M. (2016). Flipping a dental anatomy course: a retrospective study over four years. *Education Research International, 2016*, 1-9.

Bergmann, J., & Sams, A. (2009). Remixing chemistry class: Two Colorado teachers make vodcasts of their lectures to free up class time for hands-on activities. *Learning & Leading with Technology, 36*(4), 22-27.

Bergmann, J., & Sams, A. (2012). *Flip your classroom: reach every student in every class every day*. ISTE.

*Barral et al. (2018). Student learning in an accelerated introductory biology course is significantly enhanced by a flipped- learning environment. *CBE—Life Sciences Education, 17*(3), ar38.

*Baytiyeh, H., & Naja, M. K. (2017). Students' perceptions of the flipped classroom model in an engineering course: A case study. *European Journal of Engineering Education, 42*(6), 1048–1061.

*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.

Bhattacharjee, J. (2015). Constructivist approach to learning–An effective approach of teaching learning. *International Research Journal of Interdisciplinary & Multidisciplinary Studies, 1*(4), 23-28.

Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. John Wily & Sons.

*Bradford, M., Muntean, C., & Pathak, P. (2014, October). An analysis of flip-classroom pedagogy in first year undergraduate mathematics for computing. In *2014 IEEE Frontiers in Education Conference (FIE) Proceedings* (pp. 1-5). IEEE.

*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.

*Brooks, A. W. (2014). Information literacy and the flipped classroom. *Communications in Information Literacy, 8*(2), 225–235.

*Buch, G. R., & Warren, C. B. (2017). The flipped classroom: Implementing technology to aid in college mathematics student's success. *Contemporary Issues in Education Research (CIER), 10*(2), 109-116.

Chen, K. S., Monrouxe, L., Lu, Y. H., Jenq, C. C., Chang, Y. J., Chang, Y. C., et al. (2018). Academic outcomes of flipped classroom learning: A meta-analysis. *Medical Education, 52*(9), 910–924.

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

Cheng, L., Ritzhaupt, A. D., & Antonenko, P. (2019). Effects of the flipped classroom instructional strategy on students' learning outcomes: A meta-analysis. *Educational Technology Research and Development, 67*(4), 793-824.

*Choi, J., & Lee, Y. (2015). To what extent does 'flipping' make lessons effective in a multimedia production class? *Innovations in Education and Teaching International, 55*(1), 3-12.

Cieliebak, M., & Frei, A. K. (2016, April). Influence of flipped classroom on technical skills and non-technical competencies of IT students. In 2016 *IEEE global engineering education conference (EDUCON)* (pp. 1012-1016). IEEE.

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

*Clark, R. M., Besterfield-Sacre, M., Budny, D., Bursic, K. M., Clark, W. W., Norman, B. A., ... & Patzer, J. F. (2016). Flipping engineering courses: A school wide initiative. *Advances in Engineering Education, 5*(3), 1-39.

Clark, R. M., Kaw, A., & Besterfield-Sacre, M. (2016). Comparing the effectiveness of blended, semi-flipped, and flipped formats in an engineering numerical methods course. *Advances in Engineering Education, 5*(3), n3.

Cooper H., & Hedges, L. V. (1994). *The handbook of research synthesis.* Russell Sage Foundation.

*Crimmins, M. T., & Midkiff, B. (2017). High structure active learning pedagogy for the teaching of organic chemistry: Assessing the impact on academic outcomes. *Journal of Chemical Education, 94*(4), 429–438.

*Davies, R. S., Dean, D. L., & Ball, N. (2013). Flipping the classroom and instructional technology integration in a college-level information systems spreadsheet course. *Educational Technology Research and Development, 61*(4), 563-580.

de Araujo, Z., Otten, S., & Birisci, S. (2017). Mathematics teachers' motivations for, conceptions of, and experiences with flipped instruction. *Teaching and Teacher Education, 62*(2017), 60-70.

DeLozier, S. J., & Rhodes, M. G. (2017). Flipped classrooms: a review of key ideas and recommendations for practice. *Educational psychology review, 29*(1), 141-151.

Dewey, J., Montessori, M., Strzemiński, W., Piaget, J., Vygotsky, L., von Foerster, H., ... & Morin, E. (1997). Constructivism (learning theory). *Journal of Social Sciences, Literature and Languages*, 9-16.

*Eichler, J. F., & Peeples, J. (2016). Flipped classroom modules for large enrollment general chemistry courses: A low barrier approach to increase active learning and improve student grades. *Chemistry Education Research & Practice, 17*(1), 197–208.

*El-Banna, M. M., Whitlow, M., & Mcnelis, A. M. (2017). Flipping around the classroom: Accelerated Bachelor of Science in nursing students' satisfaction and achievement. *Nurse Education Today, 56*, 41–46.

*Elmaleh, J., & Shankararaman, V. (2017, April). Improving student learning in an introductory programming course using flipped classroom and competency framework. In *2017 IEEE Global Engineering Education Conference (EDUCON)* (pp. 49-55). IEEE.

*Entezari, M., & Javdan, M. (2016). Active learning and flipped classroom, hand in hand approach to improve students learning in human anatomy and physiology. *International Journal of Higher Education, 5*(4), 222-231.

Feng, X., Chen, P., Liu, Y., & Song, Q. (2016, September). Using the mixed mode of flipped classroom and problem-based learning to promote college students' learning: An experimental study. In 2016 *International Conference on Educational innovation through technology (EITT)* (pp. 133-138). IEEE.

*Foldnes, N. (2016). The flipped classroom and cooperative learning: Evidence from a randomized experiment. *Active Learning in Higher Education, 17*(1), 39-49.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences, 111*(23), 8410-8415.

*Fryling, M., Yoder, R., & Breimer, E. (2016). Full flip, half flip and no flip: Evaluation of flipping an introductory programming course. *Information Systems Education Journal, 14*(5), 4-16.

Galway, L. P., Corbett, K. K., Takaro, T. K., Tairyan, K., & Frank, E. A. (2014). Novel integration of online and flipped classroom instructional models in public health higher education. *BMC Medical Education, 14*(1), 1-9.

Gillette, C., Rudolph, M., Kimble, C., Rockich-Winston, N., Smith, L., & Broedel-Zaugg, K. (2018). A meta-analysis of outcomes comparing flipped classroom and lecture. *American journal of pharmaceutical education, 82*(5), 433-440.

Gullayanon, R. (2014, December). Flipping an engineering mathematics classroom for a large undergraduate class. In *2014 International Conference on Teaching, Assessment, and Learning (TALE)* (pp. 409-412). IEEE.

Hao, Y. (2016). Exploring undergraduates' perspectives and flipped learning readiness in their flipped classrooms. *Computers in Human Behavior, 59*, 82-92.

*Harrington, S. A., Bosch, M. V., Schoofs, N., Beel-Bates, C., & Anderson, K. (2015). Quantitative outcomes for nursing students in a flipped classroom. *Nursing Education Perspectives, 36*(3), 179–181.

*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 & Instruction, 45*, 61–71.

*He, W., Holton, A., Gu, H., Warschauer, M., & Farkas, G. (2019). Differentiated Impact of Flipped Instruction: When would flipped instruction work or falter? *International Journal of Teaching and Learning in Higher Education, 31*(1), 32-49.

Hein, G. E. (1991, October 15-22). Constructivist learning theory. In *1991 International Committee of Museum Educators Conference.* Jerusalem, Israel.

*Heuett, W. J. (2017). Flipping the math classroom for non-math majors to enrich their learning experience. *Primus, 27*(10), 889–907.

Hew, K. F., & Lo, C. K. (2018). Flipped classroom improves student learning in health professions education: a meta-analysis. *BMC medical education, 18*(1), 1-12.

*Heyborne, W. H., & Perrett, J. J. (2016). To flip or not to flip? Analysis of a flipped classroom pedagogy in a general biology course. *Journal of College Science Teaching, 45*(4), 31-37.

*Hotle, S. L., & Garrow, L. A. (2016). Effects of the traditional and flipped classrooms on undergraduate student opinions and success. *Journal of Professional Issues in Engineering Education and Practice, 142*(1), 1–11.

Howell, D. C. (2007). *Statistical methods for psychology.* PWS-Kent Publishing Co.

Hu, R., Gao, H., Ye, Y., Ni, Z., Jiang, N., & Jiang, X. (2018). Effectiveness of flipped classrooms in Chinese baccalaureate nursing education: A meta-analysis of randomized controlled trials. *International Journal of Nursing Studies, 79*, 94-103.

*Hu, X., Liu, S., Xu, Z., & Xiao, G. (2016, September). Learning behavior analytics and learning effect evaluation for learners based on MOOCs. In *2016 International Conference Educational Innovation through Technology (EITT)* (pp. 1-5). IEEE.

Huber, E., & Werner, A. (2016, January). A review of the literature on flipping the STEM classroom: Preliminary findings. In *33rd International Conference of Innovation, Practice, and Research in the Use of Educational Technologies in Tertiary Education-ASCILITE 2016-show me the learning.*

Hughes, Y., & Lyons, N. (2017). Does the flipped classroom improve exam performance in medical education? A systematic review. *MedEdPublish, 6*(2), 1-18.

*Ichinose, C., & Clinkenbeard, J. (2016). Flipping college algebra: Effects on student engagement and achievement. *Learning Assistance Review, 21*(1), 115-129.

Jang, H. Y., & Kim, H. J. (2020). A Meta-Analysis of the Cognitive, Affective, and Interpersonal Outcomes of Flipped Classrooms in Higher Education. Education Sciences, 10(4), 115.

*Jonsson, H. (2015, October). Using flipped classroom, peer discussion, and just-in-time teaching to increase learning in a programming course. In *2015 IEEE Frontiers in Education Conference (FIE)* (pp. 1-9). IEEE.

Karabulut-Ilgu, A., Jaramillo Cherrez, N., & Jahren, C. T. (2018). A systematic review of research on the flipped learning method in engineering education. *British Journal of Educational Technology, 49*(3), 398-411.

*Kennedy, E., Beaudrie, B., Ernst, D. C., & St. Laurent, R. (2015). Inverted pedagogy in second semester calculus. *Problems, Resources, and Issues in Mathematics Undergraduate Studies, 25*(9–10), 892–906.

Khan, M., & Ibrahim, M. (2017, March). Flipped classroom in technology courses-impact on personal efficacy and perception based on learning style preferences. In *2017 IEEE Integrated STEM Education Conference (ISEC)* (pp. 135-142). IEEE.

*Kim, G. J., Patrick, E. E., Srivastava, R., & Law, M. E. (2014). Perspective on flipping circuits I. *IEEE Transactions on Education, 57*(3), 188-192.

Kozikoglu, I. (2019). Analysis of the studies concerning flipped learning model: A comparative meta-synthesis study. *International Journal of Instruction, 12*(1), 851-868.

Låg, T., & Sæle, R. G. (2019). Does the flipped classroom improve student learning and satisfaction? A systematic review and meta-analysis. *AERA open, 5*(3), 1-17.

Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *The Journal of Economic Education, 31*(1), 30-43.

*Li, S., & Dan, F. (2015, July). Influences from university students on the flipped classroom. In *2015 10th International Conference on Computer Science & Education (ICCSE)* (pp. 883-888). IEEE.

*Liebert, C. A., Lin, D. T., Mazer, L. M., Bereknyei, S., & Lau, J. N. (2016). Effectiveness of the surgery core clerkship flipped classroom: A prospective cohort trial. *American Journal of Surgery, 211*(2), 451–457.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical Meta-analysis*. SAGE publications, Inc.

*Lloyd, J., & Ebener, B. (2014). Flipping a non-majors biology class: Using video lectures, online resources, and a student response system to facilitate deeper learning. *Journal of the Idaho Academy of Science, 50*(1), 93-95.

*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.

*Maciejewski, W. (2016). Flipping the calculus classroom: an evaluative study. *Teaching Mathematics and its Applications: An International Journal of the IMA, 35*(4), 187-201.

*Maneeratana, K., Singhanart, T., & Singhatanadgid, P. (2016, December). A preliminary study on the utilization and effectiveness of a flipped classroom in Thailand. In *2016 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)* (pp. 345-352). IEEE.

Margulieux, L., McCracken, W. M. & Catrambone, R. (2015). Mixing in-class and online learning: Content meta-analysis of outcomes for hybrid, blended, and flipped courses. In 2014 *The computer-supported collaborative learning (CSCL) conference* (pp. 220–227). The International Society of the Learning Sciences.

*Mason, G. S., Shuman, T. R., & Cook, K. E. (2013). Comparing the effectiveness of an inverted classroom to a traditional classroom in an upper-division engineering course. *IEEE Transactions on Education, 56*(4), 430-435.

*McLaughlin, J. E., Griffin, L. M., Esserman, D. A., Davidson, C. A., Glatt, D. M., Roth, M. T., ... & Mumper, R. J. (2013). Pharmacy student engagement, performance, and perception in a flipped satellite classroom. *American Journal of Pharmaceutical Education, 77*(9), 1-8.

*Mattis, K. V. (2014). Flipped classroom versus traditional textbook instruction: Assessing accuracy and mental effort at different levels of mathematical complexity. *Technology, Knowledge and Learning, 20*(2), 231–248.

*Missildine, K., Fountain, R., Summers, L., & Gosselin, K. (2013). Flipping the classroom to improve student performance and satisfaction. *Journal of Nursing Education, 52*(10), 597-599.

*Moffett, J., & Mill, A. C. (2014). Evaluation of the flipped classroom approach in a veterinary professional skills course. *Advances in Medical Education and Practice, 5*, 415–425.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Group PRISMA. 2009. *Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. Annals of Internal Medicine, 151*(4), 264-269.

*Morton, D. A., & Colbert-Getz, J. M. (2017). Measuring the impact of the flipped anatomy classroom: The importance of categorizing an assessment by bloom's taxonomy. *Anatomical Sciences Education, 10*(2), 170–175.

*Moudgalya, K. M., Eranki, K. K., Ganguly, S., Mohanan, S., & Raja, D. (2016, March). Efficacy of a flipped method in an undergraduate class at IIT Bombay. In *2016 International Conference on Learning and Teaching in Computing and Engineering (LaTICE)* (pp. 186-193). IEEE.

*Munson, A., & Pierce, R. (2015). Flipping content to improve student examination performance in a pharmacogenomics course. *American Journal of Pharmaceutical Education, 79*(7), 1–7.

*Murray, L., McCallum, C., & Petrosino, C. (2014). Flipping the classroom experience: A comparison of online learning to traditional lecture. *Journal of Physical Therapy Education, 28*(3), 35–41.

Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of educational research, 76*(3), 413-448.

O'Flaherty, J., & Phillips, C. (2015). The use of flipped classrooms in higher education: A scoping review. *The Internet and Higher Education, 25*, 85-95.

*Ojennus, D. D. (2016). Assessment of learning gains in a flipped biochemistry classroom. *Biochemistry and Molecular Biology Education, 44*(1), 20–27.

Olakanmi, E. E. (2017). The effects of a flipped classroom model of instruction on students' performance and attitudes towards chemistry. *Journal of Science Education and Technology, 26*(1), 127-137.

*Özpinar, İ., Yenmez, A. A., & Gökçe, S. (2016). An application of flipped classroom method in the instructional technologies and material development course. *Journal of Education and Training Studies, 4*(12), 213-226.

*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.

Prashar, A. (2015). Assessing the flipped classroom in operations management: A pilot study. *Journal of Education for Business, 90*(3), 126-138.

Rahman, A. A., Aris, B., Mohamed, H., & Zaid, N. M. (2014, December). The influences of flipped classroom: A meta-analysis. In *2014 IEEE 6th Conference on Engineering Education (ICEED)* (pp. 24-28). IEEE.

*Ranalli, J., & Moore, J. (2016, October). Targeted flipped classroom technique applied to a challenging topic. In *Frontiers in Education Conference (FIE), 2016 IEEE* (pp. 1-4). IEEE.

*Reid, S. A. (2016). A flipped classroom redesign in general chemistry. *Chemistry Education Research and Practice, 17*(4), 914-922.

Richardson, V. (2003). Constructivist pedagogy. *Teachers college record, 105*(9), 1623-1640.

Roach, T. (2014). Student perceptions toward flipped learning: New methods to increase interaction and active learning in economics. *International review of economics education, 17*, 74-84.

Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual review of psychology, 52*(1), 59-82.

*Rui, Z., Lian-rui, X., Rong-zheng, Y., Jing, Z., Xue-hong, W., & Chuan, Z. (2017). Friend or foe? Flipped classroom for undergraduate Electrocardiogram learning: A randomized controlled study. *BMC medical education, 17*(1), 1-9.

*Ryan, M. D., & Reid, S. A. (2015). Impact of the flipped classroom on student performance and retention: A parallel controlled study in general chemistry. *Journal of Chemical Education, 93*(1), 13-23.

*Sahin, A., Cavlazoglu, B., & Zeytuncu, Y. E. (2015). Flipping a college calculus course: A Case study. *Journal of Educational Technology & Society, 18*(3), 142-152.

*Salama, G., Scanlon, S., & Ahmed, B. (2017, April). An evaluation of the flipped classroom format in a first-year introductory engineering course. In *2017 IEEE Global Engineering Education Conference (EDUCON)* (pp. 367-374). IEEE.

*Saterbak, A., Volz, T., & Wettergreen, M. (2016). Implementing and assessing a flipped classroom model for First-Year Engineering design. *Advances in Engineering Education, 5*(3), n3.

*Sengel, E. (2014). Using the 'flipped classroom' to enhance physics achievement of the prospective teacher impact of flipped classroom model on physics course. *Journal of the Balkan Tribological Association, 20*(3), 488–497.

*Şengel, E. (2016). To flip or not to flip: Comparative case study in higher education in turkey. *Computers in Human Behavior, 64*, 547–555.

Shi, Y., Ma, Y., MacLeod, J., & Yang, H. H. (2020). College students' cognitive learning outcomes in flipped classroom instruction: a meta-analysis of the empirical literature. *Journal of Computers in Education, 7*(1), 79-103.

Smith, M. K., Vinson, E. L., Smith, J. A., Lewin, J. D., & Stetzer, M. R. (2014). A campus-wide study of STEM courses: new perspectives on teaching practices and perceptions. *CBE—Life Sciences Education, 13*(4), 624-635.

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

Strayer, J. F. (2012). How learning in an inverted classroom influences cooperation, innovation, and task orientation. *Learning Environments Research, 15*, 171–193.

*Street, S. E., Gilliland, K. O., McNeil, C., & Royal, K. (2014). The flipped classroom improved medical student performance and satisfaction in a pre-clinical physiology course. *Medical Science Educator, 25*(1), 35–43.

Strelan, P., Osborn, A., & Palmer, E. (2020). The flipped classroom: A meta-analysis of effects on student performance across disciplines and education levels. *Educational Research Review, 30*, 100314.

*Sun, J. C. Y., & Wu, Y. T. (2016). Analysis of learning achievement and teacher-student interactions in flipped and conventional classrooms. *International Review of Research in Open & Distance Learning, 17*(1), 79–99.

Tan, C., Yue, W. G., & Fu, Y. (2017). Effectiveness of flipped classrooms in nursing education: Systematic review and meta-analysis. *Chinese Nursing Research, 4*(4), 192-200.

Thai, N. T. T., De Wever, B., & Valcke, M. (2017). The impact of a flipped classroom design on learning performance in higher education: Looking for the best "blend" of lectures and guiding questions with feedback. *Computers & Education, 107*, 113-126.

*Timmerman, K., Raymer, M., Gallgher, J., & Doom, T. (2016, August). Educational methods for inverted-lecture computer science classrooms to overcome common barriers to stem student success. In 2016 *Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)* (pp. 1-4). IEEE.

*Tsai, C., Shen, P., & Lu, Y. (2015). The effects of problem-based learning with flipped classroom on elementary students' computing skills: A case study of the production of eBooks. *International Journal of Information and Communication Technology Education, 11*(2), 32–40.

*Tutrang, T. C., & Schenke, K. (2017). Can flipping the classroom work? Evidence from undergraduate chemistry. *International Journal of Teaching and Learning in Higher Education, 29*(3), 421-435.

*Van Sickle, J. (2015). Adventures in flipping college algebra. *Problems, Resources, and Issues in Mathematics Undergraduate Studies, 25*(8), 600–613.

Van Sickle, J. R. (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.

*Velegol, S. B., Zappe, S. E., & Mahoney, E. M. I. L. Y. (2015). The evolution of a flipped classroom: Evidence-based recommendations. *Advances in Engineering Education, 4*(3), n3.

*Wasserman, N. H., Quint, C., Norris, S. A., & Carr, T. (2017). Exploring flipped classroom instruction in calculus III. *International Journal of Science & Mathematics Education, 15*(3), 545–568.

*Weaver, G. C., & Sturtevant, H. G. (2015). Design, implementation, and evaluation of a flipped format general chemistry course. *Journal of Chemical Education, 92*(9), 1437–1448.

*Webster, D. R., Majerich, D. M., & Madden, A. G. (2016). Flippin'fluid mechanics--Comparison using two groups. *Advances in Engineering Education, 5*(3), n3.

*Whillier, S., & Lystad, R. P. (2015). No differences in grades or level of satisfaction in a flipped classroom for neuroanatomy. *The Journal of Chiropractic Education, 29*(2), 127–133.

*Yelamarthi, K., & Drake, E. (2015). A flipped first-year digital circuits course for engineering and technology students. *IEEE Transactions on Education, 58*(3), 179-186.

Zainuddin, Z., & Halili, S. H. (2016). Flipped classroom research and trends from different fields of study. *The International Review of Research in Open and Distributed Learning, 17*(3).

Chioma Ezeh is a doctoral candidate in the Language, Literacy, and Technology program at Washington State University – Pullman. Her research bridges multilingual education, literacy development, and learning technologies to leverage the linguistic and sociocultural practices of linguistically and culturally diverse learners for equitable education and policy. She is currently working on diverse teachers' and students' translation of language education policies into classroom literacy practices. Her work has been presented at 17 conferences and published in six journals and book chapters. As a graduate student, she has received 13 scholarship and fellowship awards in recognition of her scholarly engagement and service.

Olusola Adesope is a Boeing Distinguished Professor of STEM Education at Washington State University-Pullman. His research focuses on the use of systematic reviews/meta-analyses for advancing evidence-based practices, and cognitive and pedagogical underpinnings of learning with computer-based multimedia resources. His research is funded by the National Science Foundation. Dr. Adesope has published over 140 journal papers, book chapters and proceedings and presented about 110 papers in national and international conferences. He is an Associate Editor of the Journal of Educational Psychology and Senior Associate Editor for the Journal of Engineering Education and sits on the editorial boards of several top-tier journals.

Olasunkanmi Kehinde is a Washington State University doctoral student in Educational Psychology with a background in applied mathematics and statistics. His research interests include Assessment and Measurement, Psychometrics, Large Scale Assessment, Cognitive Diagnostic Models (CDMs), Multilevel Modeling, and Structural Equation Modeling in social, medical, and educational contexts. Currently, he is exploring learning progressions to contribute to the transformation of current large-scale assessments into learning progression-based assessments. His work has been disseminated through journal publications and conference presentations.

Emmanuel Jaiyeola holds a doctoral degree in cultural studies and social thoughts in education from Washington State University. Dr. Jaiyeola's research is domiciled in Cultural Studies, using social theories to interrogate societal power dynamics. This includes critical masculinity, African feminism, women's education, empowerment and entrepreneurship, gender studies, and inequality. He brings his expertise in critical studies into social justice and advocacy in higher education. As an instructor, he is interested in practical approaches and methodologies to make students learning more achievable in the era of developed technologies. He is a member of many educational professional bodies, including AERA and AESA.