INVESTIGATING THE FLIPPED CLASSROOM IN UNDERGRADUATE EDUCATIONAL PSYCHOLOGY

A dissertation submitted to the
Kent State University College
of Education, Health, and Human Services
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

By

Tracy Arner

December 2020

A Dissertation written by

Tracy Arner

B.S., University of Phoenix, 2009

M.E.T., Boise State University, 2014

Ph.D., Kent State University, 2020

Approved by

__________________________, Director, Doctoral Dissertation Committee
Bradley J. Morris

__________________________, Member, Doctoral Dissertation Committee
Christopher Was

__________________________, Member, Doctoral Dissertation Committee
John Dunlosky

Accepted by

__________________________, Director, School of Lifespan Development and
Frank Sansosti Educational Sciences

__________________________, Dean, College of Education, Health and Human
James C. Hannon Service

ARNER, TRACY, DECEMBER 2020 Educational Psychology

INVESTIGATING THE FLIPPED CLASSROOM IN UNDERGRADUATE EDUCATIONAL PSYCHOLOGY (144 pp.)

Director of Dissertation: Bradley J. Morris, Ph.D.

Three experiments investigated the difference between the traditional and flipped instructional models on students' learning outcomes and reported engagement, preparation, and participation in class. Experiment 1 found increased engagement, preparation, and participation reported by students when in the flipped classroom compared to the traditional classroom. However, Experiments 2 and 3 did not show the same results. These results suggest that students' judgements of their engagement, preparation, and participation may not be calibrated when only experiencing one instructional model.

Students' learning in Experiment 1, measured with knowledge checks and exam scores was not statistically different between models. Experiments 2 and 3 used a pre-/post-/delayed posttest design to evaluate differences in students' learning between the flipped and traditional instructional models and found that there were no statistically significant differences between the two models at immediate or delayed posttest. These results suggest that the content and design of the learning activities may be a more important factor in students' learning and reported engagement, preparation, and participation than whether the class model is flipped or traditional.

ACKNOWLEDGEMENTS

I would like to thank my committee for their guidance and support throughout this adventure. John Dunlosky was incredibly helpful in pointing out other ways to think about so many things. Chris Was told me the important things that I needed to know about undergraduates and statistics. Jenny Roche encouraged me to think about my question, find my theoretical framework, and helped with my analyses and statistical understanding. I would especially like to thank my advisor, Brad Morris, who always let me follow my ideas down the rabbit hole but also reeled me back in with "there is some literature on that" when I strayed too far.

I would like to thank my lab mates, especially Grace Willer and Rachael Todaro. I could always rely on them to keep me from losing my mind or at least help me find it when I did. These two wonderful scientists and educators were instrumental in this work as colleagues and friends and I am forever grateful to them.

Last and most important, my undying love and gratitude belong to my two amazing children. Ashtin and Nathan encouraged me to follow this crazy path and guaranteed me that they would be fine while I did. They gave me their strength and love when I needed it and always believed that I would make it to the end even when I didn't. I will be forever grateful for their patience and attention every time I climbed up on my soap box ranting about education and all the things I want to do. Instead of ignoring me, they cheered me on and pushed me to the finish line. They will always be the candle that lights my way and the strength to proceed.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

LIST OF TABLES

CHAPTER I

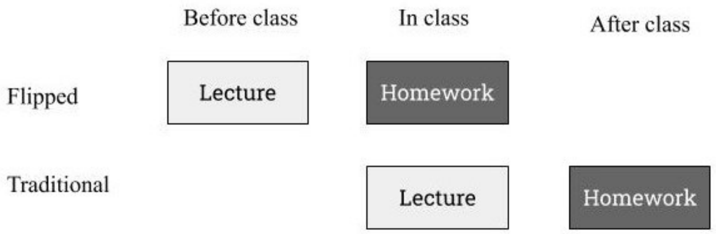
INTRODUCTION

The classroom model used in both K-12 and higher education settings has changed little over the last century. The traditional classroom is a “teaching as telling” model that frequently consists of an instructor introducing knowledge to students via lecture (Giannakos et al., 2018; Stephenson et al., 2008). This model generally includes little, if any, interaction with or between students (Burke & Fedorak, 2017; Gardiner, 1998; Pierce; 2013). Following its introduction, students practice using newly gained knowledge through ‘homework’ activities. Students in the traditional instruction model may be asked to read textbooks or articles prior to class but the instructor remains the primary source of knowledge (Gardiner, 1998).

The flipped classroom model inverts the in- and out-of-class activities so that initial instruction (i.e., introduction to new knowledge) takes place prior to the class meeting and practice and application of knowledge (e.g., homework problems) is now done in class (Bergmann & Sams, 2012; Mazur, 1997; O’Flaherty & Phillips, 2015; Strayer, 2012; Tucker, 2012). Evidence suggests that the benefit of the flipped model may be influenced by the structure and alignment of learning activities in and out of the classroom (Bergmann & Sams, 2012; Mason et al., 2013; Peterson, 2016) and students’ engagement (Burke & Fedorak, 2017; Lo & Hew, 2017). Figure 1 depicts the basic structure of the flipped and traditional classroom models.

Figure 1

Structure of basic flipped and traditional classroom models



Investigations of the flipped classroom have included student perceptions (Hao, 2016; McLaughlin, et al., 2013), teacher perceptions (Hao & Lee, 2016; Scott et al., 2015), and learning outcomes (Peterson, 2016; Pierce, 2013; Roach, 2014). However, there is little rigorous research investigating best practices for the instructional design of pre-class (i.e., initial instruction) and subsequent in-class activities that will support student engagement, learning, and retention (Njie-Carr et al., 2017; O'Flaherty & Phillips, 2015).

Three experiments investigated differences in learning between flipped and traditional instructional models on students' engagement, immediate learning outcomes, and long-term retention. Experiment 1 addresses a gap in the literature using a switching replications (Trochim, 1986) design to evaluate within- and between-subject differences in students' engagement, preparation, participation, and learning outcomes in both flipped and traditional formats. Students in the flipped model were introduced to new concepts via reading and question-embedded interactive videos. Students in the traditional model were introduced to new concepts via the same reading and in-class lecture. Students' learning, engagement, preparation, and participation are evaluated the same way in both models.

The second experiment addresses three gaps in the literature. First, it evaluated the impact of evidence-based practices (i.e., distributed practice, practice testing, and self-explanation) on students' learning in the flipped and traditional formats. Second, Experiment 2 evaluates differences in students' engagement, preparation, and participation-between instructional models, with and without assigned instructional supports. Third, this experiment evaluated the durability of learning using a 30-day delayed posttest.

Experiment 3 addresses a gap in the literature by investigating the benefit of guided interaction with peers (Chi, 2009; Wiley & Chi, 2014) on immediate learning outcomes and retention in the flipped instructional model.

The more active nature of the flipped classroom and the access to the instructor during the application of knowledge, should lead to improved learning outcomes for students. Three experiments sought to examine this benefit by comparing students' learning outcomes in flipped and traditional classrooms.

Practical and Theoretical Implications

The 'flipped classroom' instructional model has gained traction in recent years (Akcayir & Akcayir, 2018; Barbour & Schuessler, 2019; Karabulut-Ilgu et al., 2018) despite limited evidence supporting specific design features (DeLozier & Rhodes, 2016; Hamdan et al., 2013; Yarbro et al., 2014). However, deviations from a solely lecture-based, direct instruction model have been implemented previously under different labels such as inverted (Lage et al., 2000; Strayer, 2012), hybrid (Foldnes, 2016; Prunuske et al., 2012) or blended learning (Olapiriyakul, & Scher, 2006). The flipped instructional model falls under the larger umbrella of blended instruction as one of several instructional models including some level of technology use (e.g., computer labs, classroom computer, mobile device; Blended Learning Network, 2020).

While relatively new in K-12 education, a version of the flipped classroom was implemented in higher education in the early 1990's when Eric Mazur began inverting students' in- and out-of-class activities and incorporated the “peer instruction model” in his undergraduate physics class (Crouch & Mazur, 2001; Mazur, 1997; O’Flaherty & Phillips, 2015). Since its inception, the flipped classroom has been implemented in a variety of ways and as of yet, no single definition or instructional design exists to direct effective implementation (Lee et al., 2017; McLaughlin et al., 2016; Tucker, 2012). Yet, educators are beginning to change the way they structure their classes despite the paucity of rigorous research investigating design of the flipped classroom. (DeLozier & Rhodes, 2016; Hamdan et al., 2013; Karabulut-Ilgu et al., 2018; McNally et al., 2016; O’Flaherty & Phillips, 2015). As instructors migrate to any version of the flipped instructional model, it is important to determine what features influence students’ engagement, learning, and retention (Barbour & Schuessler, 2019; Karabulut-Ilgu et al, 2018; Lee et al., 2017). These three experiments sought to address this gap in the literature.

Below I discuss evaluation of students’ engagement in the learning process through motivation and cognitive processing frameworks, respectively. Then, I will discuss the evidence-based practices used in this work, the instructional design of the traditional and flipped classrooms, and how the aforementioned cognitive processing framework was applied to them.

Student Engagement

Students’ engagement in the educational process may be evaluated through two frameworks. First, students’ interactions with instructors, learning materials, and peers demonstrate engagement with the learning process which may be evaluated as a measure of students’ motivation to learn (Chi, 2009; 2018; Pintrich & De Groot; 1990; Zimmerman, 1990). Second, students’ engagement may also be evaluated based on behaviors that result in different levels of cognitive processing (Chi, 2018). Students’ level of processing (e.g., passive vs. active; Chi, 2009) affects comprehension of material and ability to transfer it (Chi & Wiley, 2014).

Engagement, measured through the lens of motivation, is a function of students' choices and beliefs (i.e., the learning strategies they choose). The levels of cognitive processing that students engage in are a function of learning activities provided in the instructional model (i.e., listening to a lecture compared to taking a practice test). Each of these frameworks and how they were applied in this work are discussed below.

Engagement as Motivation

Students who invest time in learning activities (e.g., reading before class, doing homework) and/or choose learning strategies to help them understand and retain content are actively engaging with content suggesting that they are motivated to learn (Blumenfeld et al., 2006; Pintrich & DeGroot, 1990). Additionally, students who set learning goals, plan study time, monitor performance, and reflect on the success or failure selected strategies are exhibiting motivated behaviors to learn. Students' motivational beliefs, strategy selection, and self-regulated behaviors may be predictive of academic achievement (Pintrich & DeGroot, 1990; Rotgans & Schmidt, 2012; Zimmerman & Schunk, 2012). The next section discusses student engagement as levels of cognitive processing.

Engagement as Cognitive Processing

Instructor-designed activities can elicit different levels of cognitive processing categorized as passive, active, constructive, interactive based on the learning behaviors students engage in (ICAP Model; Chi & Wiley, 2014; Chi et al., 2018). These overt (i.e., visible) behaviors that students exhibit may lead to increased levels of cognitive processing and improved learning outcomes (Chi, 2009; Menekse et al., 2013). The ICAP Model is discussed below and its application in this work is described in Chapter II.

Passive

Passive learning activities require little, if any, cognitive processing by the student which can negatively impact learning (Chi & Wylie, 2014; Chi, 2018). A traditional, classroom where students listen to an instructor, may be a passive learning environment. Students who are only listening, are not engaging in any overt behaviors (e.g., note taking, concept mapping) that may lead to knowledge integration (Chi & Wylie, 2014). Simply viewing a lecture, without prompts to synthesize or integrate new information, may result in increased mind wandering and loss of attention resulting in decreased learning (Peper & Mayer, 1986; Wammes & Smilek, 2017). In other words, attention alone may not be enough to encode new information (Wammes & Smilek, 2017; Wammes et al., 2019).

Active

Learning becomes active when students engage in any overt behaviors that may increase active cognitive processing such as activating prior knowledge or encoding new knowledge (Chi, 2018). Taking notes, copying presentation slides, or copying problem examples verbatim are examples of active behaviors students may engage in that lead to improved learning (Chi, 2009; Chi & Wylie, 2014).

Constructive

Students engaging with material and generating new ideas are demonstrating constructive cognitive processing (Chi, 2009). The construction of new knowledge through making inferences, generating self-explanations, creating concept maps, and similar activities has been shown to have greater benefit to learning than active engagement with material alone (Chi, 2018). When students engage in inference making or explaining a concept to themselves, they are updating and revising schema which may lead to improved transfer (Chi & Wylie, 2014).

Interactive

Interactive learning takes place when a student engages in a sequentially constructed or co-constructed dialog with a person or interacts with a technology tool (Chi & Wylie, 2014). During interactive learning, a speaker is constructing knowledge via generating explanations with the added benefit of input from peers which may include corrective feedback, a new line of inquiry, or aid in assimilation of new material (Chi, 2009; Roscoe & Chi, 2008). The dialogic nature of interactive cognitive processing requires that all parties (human or technology) engage in both explanation or question generation and response or feedback. Interaction that is not reciprocal will not benefit learning in the same way (Chi & Wylie, 2014; Chi et al., 2018; Menekse et al., 2013). The instructional models used in this work are discussed below followed by the a description of how the ICAP framework (levels of cognitive processing) was applied to each of them.

Instructional Models

The two instructional models compared in this work are the traditional, lecture-based model and the flipped instructional model (referred to as traditional and flipped, hereafter). The design of both models may vary by instructor, but both generally include an 'instruction phase' where the students are introduced to new concepts and an 'application phase' where students apply or practice new knowledge (Arner, 2020). The instructional models used in this work are defined below.

Traditional Model

The traditional model introduces students to new concepts through direct instruction (i.e., lecture). The instructor is the focus in the classroom providing information to students who are receivers of new knowledge. Students are assigned a learning activity to complete outside of class to apply or practice the information introduced by the instructor during the lecture (Crouch & Mazur, 2001; McLaughlin et al., 2013; O’Flaherty & Phillips, 2015). In this work, students in the traditional section were asked to read an article or chapter prior to class and were assigned a handout as homework after they attended lectures each week. The levels of cognitive processing included in the design of the traditional instructional model are described below.

Cognitive Processing in the Traditional Model

Students may engage in active, passive, or constructive processing, or some combination, in the traditional lecture model (Chi, 2009; Chi & Wylie, 2014). The traditional classroom in this work included the instructor as the giver of knowledge and students as largely passive receivers (Baepler et al., 2014; Bishop & Verleger, 2013; Lee et al., 2017; Nouri, 2016; Mazur, 1997). Students had the option to take notes and those who did engaged in some active processing of lecture content while writing. Students who elaborated or inferred while taking notes were engaging in constructive processing (Chi, 2009). The instructor occasionally posed questions as a method of refreshing attention which also elicits active or constructive processing depending on the question type (i.e., “Remember when we talked about. . . vs. “What did we learn about. . . ?“). Weekly knowledge checks with multiple-choice and short answer questions required active and constructive processing, respectively (Chi & Wiley, 2014).

Students in the traditional model were asked to complete two activities outside of the classroom. First, students had assigned reading (e.g., article or book chapter) prior to each lecture with no other explicit directions which may have been passive (i.e., skimming or reading without making connections). Those who did make inferences or connections to prior knowledge were engaging in active or constructive processing (Wylie & Chi, 2014). Second, students were assigned handouts with questions designed to elicit self-explanations which requires constructive processing (Chi, 2009).

Flipped Model

The flipped classroom is not necessarily a new instructional model but there is not a single design (Bergmann & Sams, 2012; DeLozier & Rhodes, 2017; Karabulut-Ilgu et al., 2017). However, two features are consistent across models. First, the flipped classroom model inverts learning activities so that new content is now introduced prior to a corresponding class meeting. In this work, students were introduced to new content through a combination of readings (same as traditional) and a variety of videos (e.g., question-embedded, content expert, voiceover lecture; discussed further in Chapter III). Second, in-class time is spent applying new knowledge in an interactive setting involving collaboration with peers or the instructor (Bergmann & Sams, 2012; Hamdan et al., 2013; Karabulut-Ilgu et al., 2017; Zainuddin & Halili, 2016). The in-class activity in this work included small-group, (i.e., four to six students) facilitated discussion using the interactive peer learning model (Mazur, 1997) and projected questions (Experiment 3). As in the traditional classroom, the cognitive processing that students engage in can vary with the assigned activities, as discussed below.

Cognitive Processing in the Flipped Model

Similar to the traditional classroom, students may engage in passive, active, and constructive cognitive processing in the flipped model. However, the flipped classroom also includes interactive learning activities which may lead to more durable learning (Chi, 2009; Chi & Wylie, 2014; Menekse & Chi, 2018). Students in the flipped model engaged in the same types of cognitive processing as students in the traditional model outside of class because they had the same assignments. Students in the flipped model also completed a pre-class activity (i.e., replacement for lecture in traditional classroom) that required constructive and interactive cognitive processing. The use of question-embedded videos in Experiments 1 and 2, required students to engage in constructive processing while responding to questions. Students also interacted with the video itself (i.e., stop, rewind, re-watch; these data are beyond the scope of this study) which does have benefit to learning (Biard et al., 2018). In Experiment 3, students were given lecture videos to watch before class that did not have questions embedded, but did include reflection prompts. Students in both sections were able to choose how they engaged with the ‘lecture’ as they were not explicitly instructed to take notes, infer, or generate explanations. Therefore, students may have engaged in passive, active or constructive processing during this activity.

Evidence-based Strategies

The use of evidence-based strategies can improve student learning outcomes regardless of the instructional model used or how new material is introduced to students. An abundance of research supports the benefit of distributed practice, practice testing, self-explanation (Dunlosky et al., 2013) and collaborative learning (Chi, 2009). The instructional design of both traditional and flipped classrooms in this work included distributed practice, practice testing, and self-explanation. However, there are some differences in the implementation in each model that may affect learning outcomes (e.g., level of cognitive processing; Chi, 2009). Additionally, the flipped classroom also incorporated interactive peer learning during in-class learning activities (i.e., table group discussions).

Distributed Practice

Distributing practice requires multiple instances of cognitively processing content with time between each practice session (Cepeda et al., 2006). The benefit of distributing practice has been demonstrated across content areas and grade levels (Dunlosky et al., 2013).

The first instance of cognitive processing occurred when students in both models were asked to read material (e.g., textbook, article) prior to class. The second instance of cognitive processing in the traditional model occurred when content was revisited, introducing some amount of space between interactions (Foertsch et al., 2002). Students in the flipped classroom are also introducing space between reading, watching video lectures, and in class discussion.

The way students engage with the learning activity may influence retention regardless of model. If they are passively engaged with text or video lecture they may not garner the benefit of distributed practice in either instructional model. However, the design of the learning activities in the flipped classroom (e.g., question embedded videos) may increase students' level of engagement (e.g., active and constructive) thereby increasing the benefit of distributing practice (Chi & Wylie, 2014; Foertsch et al., 2002).

Practice Testing

Practice testing is one of the most widely investigated strategies for improving learning outcomes (Rawson & Dunlosky, 2012) and is beneficial in two ways. First, multiple attempts at effortful retrieval in a low- or no-stakes setting improves learning and long-term retention (Bjork & Bjork, 2011; Little et al., 2012). Second, practice testing may also guide students' restudy efforts when used as a mechanism for metacognitive monitoring (Dunlosky et al., 2013; Rawson & Dunlosky, 2012). Students can evaluate their learning by whether or not their retrieval attempt was successful which can suggest areas of focus for future study (Dunlosky & Rawson, 2015;

McDaniel & Little, 2019). Practice tests were incorporated into both instructional models for all three experiments as knowledge checks (Experiments 1 and 2) and reading quizzes (Experiment 3). The timing and type of practice test was altered in Experiment 3 to maintain temporal alignment of students' interactions with content and is discussed further in Chapter III.

Self-explanation

Self-explanations occur when students study new material and then generate an explanation of the just-learned information (Chi et al., 1994; Hausmann & VanLehn, 2007). This requires students to actively engage in cognitive processing new information during which they may elaborate on presented content, make connections between new and prior knowledge, or make inferences (Chi et al., 1989). Self-explanations can benefit students' learning in by identifying gaps in their knowledge and through the generation effect (Chiu & Chi, 2014; Hausmann & VanLehn, 2007) which occurs when constructing all or part of a response to a question or cue (Bertsch et al., 2007). Self-explanations require more effortful processing which can improve students' comprehension of new material and support transfer to novel contexts (Dunlosky et al., 2013; Hausmann & VanLehn, 2007; Rittle-Johnson, 2006).

Students in both classrooms generated self-explanations on 10-question handouts submitted for a grade each week. They were identical in both models for Experiments 1 and 2. In Experiment 3, the use of handouts was manipulated as described in Chapter III.

Interactive Peer Learning in the Flipped Classroom

The benefit of constructing knowledge in active and interactive learning activities in traditional and technology-enhanced classrooms is well documented (Chi, 2009; Freeman et al., 2014; Prince, 2004). Peer-to-peer learning activities provide opportunities for learners to share their understanding of content, using evidence-based explanations in dyads, triads, or small groups (Chi, 2009; Menekse & Chi, 2018). However, introducing collaboration in the classroom doesn’t always lead to improved learning outcomes. Students with minimal prior knowledge or experience in collaborating may benefit from structured activities to guide students’ discourse (Crouch & Mazur, 2001; Menekse & Chi, 2018). Interactive peer learning was incorporated into the flipped classroom in two ways. First, students responded individually to projected multiple-choice questions. Students then discussed the question with their table group, explaining and providing to their peers. Following discussion, students responded individually a second time to ensure understanding. The second method of incorporating interactive peer learning was instantiated with handout questions projected for students to discuss with table groups. After students discussed, the instructor reviewed the answers with input from students to ensure understanding. In both versions, conversations were timed and monitored by the instructor.

Several questions remain unanswered regarding design of the flipped classroom model. Three experiments investigated differences in students’ engagement, preparation and participation and the influence of evidence-based strategies embedded in the design of both traditional and flipped classrooms. The goals of each experiment are discussed below.

Experiment Rationale

The present work investigated differences in students’ learning, engagement, preparation and participation in the flipped and traditional models with three experiments conducted in undergraduate Educational Psychology. The first experiment compared students’ learning outcomes and reported engagement, preparation, and participation using a switching replications design to compare within- and between-subject outcomes and reduce teacher effects. It was hypothesized that students would have improved learning outcomes and higher engagement, preparation, and participation when in the flipped classroom compared to the traditional classroom. A limitation of Experiment 1 was that students were only given a pretest to evaluate prior knowledge and group equivalence at the start of the semester. Learning was solely evaluated based on assessments during the semester.

Experiment 2 addressed this limitation by adding a pre-/posttest design to evaluate learning in each model. The goals of Experiment 2 were to investigate (1) differences in learning between models over and above the benefit of incorporating instructional supports (e.g., practice testing, self-explanations); (2) the durability of learning in both classrooms and; (3) change in students' engagement, preparation, and participation throughout the semester. It was hypothesized that students in the flipped classroom would outperform students in the traditional classroom at the immediate posttest and after a 30-day delay. It was also hypothesized that students would perform better with the assigned instructional supports than when left to their own devices to select learning strategies. I hypothesized that students in the flipped classroom would report higher levels of engagement, preparation and participation than students in the traditional classroom. The primary limitation of Experiment 2 was that students were experiencing slightly different content in a different temporal arrangement in the flipped and traditional models. For example, while the main ideas were covered in both models, emphases may have been placed in different areas and the videos may or may not have been watched by students in the traditional section depending on time constraints.

Experiment 3 addressed the limitations of Experiment 2 by ensuring that students in both sections received the same content and that it was temporally aligned (i.e., same content the same order; see Table 13 in Chapter III). The goals of Experiment 3 were to further elucidate the features of the flipped model (pre-and in-class learning activities) that may lead differences in learning outcomes compared to the traditional model. Specifically, Experiment 3 compared video lectures and interactive learning groups in the flipped classroom to in-class lectures in the traditional classroom to evaluate differences when students received the same content in the same order. I hypothesized that students in the traditional classroom would outperform students in the flipped classroom on reading quizzes but that students in the flipped classroom would outperform students in the traditional classroom on posttests and the delayed posttest. This hypothesis stems from the more effortful retrieval introduced by the longer gap between study and retrieval in the flipped classroom compared to recent instruction in the traditional classroom. I also hypothesized that students in the flipped classroom would report higher levels of preparation, participation, and engagement than students in the traditional classroom.

The next chapter includes a review of literature pertaining to student engagement and cognitive processing that support student learning, evidence-based practices, the definition and design of each classroom, and how levels of cognitive processing may occur in each of them.

CHAPTER II

LITERATURE REVIEW

Students' engagement with the learning process is discussed below, first as a teacher-driven construct described as levels of cognitive processing using the ICAP framework (Chi, 2009). Then, student engagement as a student-driven motivational construct is discussed. Following student engagement, four evidence-based practices and the instructional models (flipped and traditional) used in this work are discussed. After discussing the models, the levels of cognitive processing (i.e., ICAP framework; Chi, 2009) and how they may be implemented in each model is described. Finally, some limitations of the flipped classroom research are introduced followed by how this work seeks to address them.

Student Engagement

Student engagement in the learning process as a construct that is controlled by the instructor, implemented through instructional design is discussed below. Intentional instructional design may be used to elicit specific levels of cognitive processing as described in the ICAP (Interactive-Constructive-Active-Passive) framework developed by Chi (2009). Next, engagement as a student-driven construct is discussed. This construct includes students' motivation to engage with content and select and use beneficial learning strategies.

Engagement as Cognitive Processing

The levels of cognitive processing that students engage in can lead to improved learning outcomes (Chi, 2009) and instructors can design learning activities that encourage specific levels of cognitive engagement (Chi & Wiley, 2014; Chi et al., 2018; Prince, 2004). One concern when designing instructional activities is that students may not engage with content in a way that is beneficial to learning. Therefore, instructors should consider the behaviors that students will have to engage in during the learning activity that are more likely to lead to desired learning outcomes (Chi & Wylie, 2014; Chi et al., 2018; Menekse et al., 2013). This framework defines four levels of cognitive processing that students may engage in during learning, each of which subsumes the prior, thus becoming more beneficial to learning (Chi, 2009). The use of intentional design by instructors within the ICAP framework may lead to increased cognitive processing and improved learning for students (Chi & Wylie, 2014). In other words, instructors may design activities that require higher levels of processing such as explain the text to your neighbor rather than read your textbook. Below I discuss the evidence relating to each level of cognitive processing.

Passive vs. Active

The first two levels of processing are often included in discussions of student engagement; passive compared with active learning (Jensen et al., 2015; Menekse et al., 2013; Prince, 2004). The latter is frequently associated with improved learning outcomes when compared with the former (Chi, 2009; Chi & Wylie, 2014; Menekse et al., 2013). The framework suggests that comprehension occurs when the learner is actively generating connections between new information and existing knowledge (Wittrock, 1989).

Researchers have found that students who take notes, rather than viewing only during an instructional activity, are actively engaging with content and have improved learning outcomes on assessments of recognition and recall (Peper & Mayer, 1986). One study demonstrated this benefit by moving the lecture portion of class to an online environment where students watched videos prior to attending an in-person class. Students reported actively watching lectures at home and a preference to video-lecture over a reading assignment. Those who watched the lectures performed better on assessments of lower-order cognitive skills (e.g., remembering, understanding) and there was no difference between groups on assessments of higher-order cognitive skills (e.g., analyzing). This suggests that active engagement with the lecture improves remembering but did not improve higher level understanding (Prunuske et al., 2012).

In sum, activities requiring students to engage with the content in a way that facilitates connection making, will have a greater benefit to learning than activities that do not (Chi, 2009; Wittrock, 1989). Research suggests that constructive or interactive activities are necessary for students to be able to generalize and transfer new knowledge (discussed below; Chi, 2009; Menekse et al., 2013).

Active vs. Constructive

The third level of cognitive processing is constructive and, as the name implies, includes learning activities in which students are constructing new knowledge. This level of cognitive processing may be associated with discovery or exploratory learning as well as learning activities that encourage students to infer or elaborate upon new knowledge provided during instruction (Chi, 2009; Chi, 2018). The benefit of actively engaging with content by completing active learning tasks such as finding solutions (Trafton & Trickett, 2001) and constructing self-explanations afterward (Chi, 2009; Chi & Wylie, 2014) has a greater benefit to learning than active engagement alone (e.g., note taking; Chi, 2009; Peper & Mayer, 1986). In sum, activities that encourage constructive cognitive processing are more beneficial to student learning than those requiring active processing (Chi, 2009; Chi & Wylie, 2014; Menekse et al., 2013).

Constructive vs. Interactive

The last level of cognitive processing is interactive. Interactive cognitive processing also involves students engaging in inference, elaboration, and connection making as with constructive processing. However, students who are engaging in interactive processing are doing so in dyads, triads, or small groups so that they are engaging in dialogues with peers, instructors, or technology (Chi, 2009). Interactions with others benefit learning by introducing missing knowledge, expanding connections to prior knowledge, and providing a way for misconceptions to be cleared up during the exchanges (Chi & Menekse, 2015; Menekse et al., 2013). The interactive level of cognitive processing subsumes the three prior levels, just as each level subsumes the levels below it. Students' engagement at each level of cognitive processing influences retention and durability of learning (Chi, 2009; Chi & Wylie, 2014). However, it is important to note that not all activities that are designed with one of these intended levels of processing will necessarily lead to improved learning outcomes when compared to the level(s) below it (Chi, 2009; Chi et al., 2018; Klahr & Nigam, 2004).

Active and passive classroom learning activities have been compared in numerous content areas and instructional models; however, the ICAP framework has not been applied to the flipped classroom model or used to compare the flipped and traditional classrooms.

Engagement as Motivation

Motivation has long been considered as a predictor for academic achievement (Pintrich & DeGroot, 1990; Rotgans & Schmidt, 2012; Zimmerman & Schunk, 2009). Successful students are also self-regulated learners (Pintrich & DeGroot, 1990; Zimmerman, 2002). A self-regulated learner is one who adopts a three-phase model that includes goal setting, performance, and reflection when approaching learning activities (Zimmerman, 2002). The goal setting phase involves learners analyzing the task and choosing strategies to accomplish the goal they set. The performance phase includes using the selected strategies and monitoring their performance to ensure they are using the chosen strategies and to determine whether or not they are in fact making progress toward their goal(s). Finally, a self-regulated learner will engage in reflection on their performance to evaluate task success, strategy choice, and decide on adaptations for future goal setting and performance phases (Pintrich & Degroot, 1990; Zimmerman, 2002). Engaging in these phases of self-regulation has been linked to academic success in several studies yet the origin of students' becoming self-regulated learners is less clear (Rotgans & Schmidt, 2012; Zimmerman, 2002). Students may develop a self-regulated process for learning independently (proactive) or they may develop the process after prompts from an educator to do something like tracking goals and progress (reactive; Zimmerman, 2002). In order for students to develop and maintain the self-regulatory process, they must be motivated to do so. That motivation may be derived from their goal orientation, beliefs about their competence, the value of the task, or the causes and expectancy of outcomes (Pintrich & DeGroot, 1990; Zimmerman & Schunk, 2009).

Pintrich and DeGroot (1990) investigated the intricate relationship between self-regulated learning, motivation and academic achievement through the development of a self-report questionnaire on which individuals determine whether items are *very true of me* to *not at all true of me* on a seven-point scale. They found that intrinsic motivation, self-efficacy, strategy use (e.g., distributed practice, studying with peers), and self-regulation were all positively correlated with academic achievement while test anxiety was negatively correlated with academic achievement (Pintrich & DeGroot, 1990). Pintrich and colleagues (1993) further developed the instrument and went on to evaluate the internal consistency, reliability, and predictive validity of

the final, 81-item MSLQ thus suggesting that it can accurately measure the relation between students' reported self-regulation, motivation, and academic achievement (Pintrich, 1991).

Rotgans and Schmidt (2012) investigated whether or not students' reported behaviors on the MSLQ were aligned with their observed classroom behaviors related to academic achievement. Achievement related classroom behaviors were rated by the teacher based on students' participation in discussions, engagement with content and persistence in self-directed learning, and presentation quality during each class period. Prior academic achievement was measured with a standardized college entrance exam and current academic achievement was measured with exams occurring in each of five classes, every four weeks, for a total of 20 tests per student. Rotgans and Schmidt (2012) collapsed the motivation and learning strategy subscales so that each was a single factor. Results indicated that students' motivation was not directly associated with achievement-related classroom behaviors nor was it related to academic achievement. However, their motivation score was strongly related to the use of learning strategies. The use of learning strategies had a positive relation to students' achievement-related classroom behaviors (observed by the teacher). The achievement-related classroom behaviors were a strong predictor of academic achievement. Not surprisingly, they found that prior academic achievement was a good predictor of both academic achievement and achievement-related classroom behaviors. These results align with those of Pintrich and DeGroot (1990) that cognitive strategy use (self-regulation) was a better predictor of achievement than students' motivation.

Students' motivation may be a starting point to improve learning outcomes. Research suggests that students' engagement in classroom learning activities, such as discussion participation and learning strategy use are related to student achievement. However, students' engagement in classroom learning activities (e.g., participation in discussions) may be mediated by their learning strategy selection and use (e.g., distributing practice to prepare for class). In other words, students may not be motivated to participate in classroom learning activities if they have not chosen strategies that help them prepare to do so (Pintrich & DeGroot, 1990; Rotgans & Schmidt, 2012). As such, instructors may better support students' learning through including evidence-based learning strategies and incorporating active, constructive, and interactive levels of cognitive processing in the design of course activities (Blumenfeld et al., 2004; Chi & Wiley, 2014; Rotgans & Schmidt, 2012; Zimmerman, 2002).

For the purpose of this work, students were asked to rate their engagement with learning materials, preparation for class, and participation in class learning activities. To my knowledge, students' ratings of their engagement in the learning process have not been used when comparing the flipped and traditional classrooms.

Evidence-based Strategies

An abundance of evidence supports several strategies that are beneficial to learning across grade levels and content areas independent of the instructional design of the classroom (Dunlosky et al., 2013). The evidence supporting distributed practice, practice testing, self-explanation, and interactive learning used in this work is discussed below.

Distributed Practice

Distributed practice occurs when students engage with content in multiple learning events with space between them (Dunlosky et al., 2013). Research on distributed practice has shown benefits across grade levels and content areas (Cepeda, Pashler et al., 2006; Donovan & Radosevich, 1999; Dunlosky et al., 2013; Vlach & Sandhofer, 2012). The benefit of distributing practice on learning has been demonstrated in laboratory settings using word pair learning (Cepeda et al., 2009, Dunlosky et al., 2013), verbal recall tasks (Cepeda et al., 2006) and motor skills (Donovan & Radosevich, 1999). More importantly for the purpose of this work, numerous studies have reported the benefit of distributing practice in applied research in math learning (Barzagar Nazari & Ebersbach, 2018; Schutte et al., 2015), science learning (Svihla et al., 2018; Vlach & Sandhofer, 2012), and reading skills (i.e., phonics; Seabrook et al., 2005).

Benefit of Distributed Practice

The benefits of distributed practice, while ubiquitous, may vary depending on the task complexity, the interstudy lag, and the time delay until the assessment (i.e., retention interval; Cepeda et al., 2009; Donovan & Radosevich, 1999; Dunlosky et al., 2013; Seabrook et al., 2005). Analyses looking at effect size of distributing practice while comparing task complexity found that tasks with lower cognitive requirements, but higher motor requirements had a large beneficial effect from introducing space. However, as the cognitive requirements for the task increased, the benefit of distributing practice decreased but still demonstrated a positive effect compared to massed practice (Cepeda et al., 2006; Donovan & Radosevich, 1999).

The temporal arrangement of study sessions has varied from minutes to days in both basic and applied settings (Barzagar Nazari & Ebersbach, 2018; Cepeda et al., 2006; Donovan & Radosevich, 1999; Schutte et al., 2015). Temporal arrangement includes the interstudy interval (AKA interstudy gap, lag), the time between each study session and the retention interval, the length of time between the last study session and the final test (Cepeda et al., 2006; Donovan & Radosevich, 1999). Research on the benefit of distributed practice in laboratory settings has explored a variety of interstudy intervals and has found that longer interstudy intervals are more beneficial to retention than shorter intervals at test. However, this effect is also influenced by the type of learning task (i.e., motor skill, verbal recall, etc., Cepeda et al., 2006; Donovan & Radosevich, 1999) and age of the participant (Seabrook et al., 2005). Research suggests that the length of the retention interval will influence the overall benefit gained by distributing practice in addition to the interstudy interval (Carpenter et al., 2012; Cepeda et al., 2006; Rohrer, 2015). One study found that students who were given a review of history knowledge 16 weeks after the learning period performed better than those given a review one week after the learning period at the delayed posttest, nine months later. Similarly, Bahrick (1979) found that students who learned translations with a 30-day gap performed better at posttest, 30 days after the last learning session. However, this may not always be feasible for classroom applications and instructors may have to consider methods to increase the instructional time and number of study sessions that students receive (Carpenter et al., 2012; Rohrer, 2015).

Distributed Practice in the Classroom

Practically speaking, it is important to determine the most beneficial interstudy gap for the desired retention interval (test-delay) when considering the design of instruction. However, much of the research on distributed practice is limited to laboratory settings that only include two learning sessions and a test session (Carpenter et al., 2012; Svihla et al., 2018). This makes it difficult to generalize the effect to classroom where ideally, students will retain information for a much longer term (i.e., be able to recall it) and develop understanding (i.e. transfer to novel situations; Svihla et al., 2018) thus instruction could benefit from a predetermined balance between review and introduction of new material (Carpenter et al., 2012; Rohrer, 2015). Given the potential limitations of class time, scheduling (i.e., class sessions per week/quarter/semester) and curricula, it becomes necessary to consider what methods might be used to distribute practice both in and out of the classroom (Carpenter et al., 2012; Rohrer, 2015; Schutte et al., 2015). Therefore, teachers may need to instantiate distributed practice through assigned work and due dates because students may not space their study voluntarily, even when provided evidence of the benefit of doing so (Barzagar Nazari & Ebersbach, 2018; Bjork et al., 2013;

Dunlosky et al., 2013). Furthermore, some evidence suggests that when students do revisit concepts voluntarily, they may not do it in such a way that it is beneficial for learning and retention (Svihla et al., 2018). One mechanism for distributing practice in this work is in the design of the pre-class, in-class, and after-class learning activities.

Practice Testing

Practice testing is an implementation of effortful retrieval in which the learner attempts to recall previously encoded knowledge (Karpicke, 2012). While the idea of testing is generally met with disdain, the use of retrieval practice has been found to be a successful mechanism for learning over decades of research (Dunlosky et al., 2013; Rawson & Dunlosky, 2012). There are two important differences between practice tests and those which students dislike. The first is that they are no- or low-stakes. Second, they can be completed in class such as a formative assessment or independent of class (by students), such as using flashcards.

Benefits of Practice Testing

The benefit of practice testing has been demonstrated with multiple question formats (i.e., multiple-choice, short-answer; Agarwal et al., 2008), multiple test conditions (i.e., open book, closed book; Little et al., 2012), and with a variety of age groups (i.e., K-12, undergraduate, graduate, and adult learners; Pyc et al., 2014), and across content types and knowledge levels (i.e., general education, advanced medical students; Dunlosky et al., 2013). Benefits from practice testing may be direct, derived from the taking of the test or indirect, derived from what occurs as a result of taking the test. Examples include identifying knowledge gaps, providing feedback, and focused restudying (Dunlosky et al., 2013; Dunlosky & Rawson, 2015).

Direct benefit. The direct benefit of practice testing (i.e., the testing effect) has been widely researched, with multiple types of tests and many experiments have shown that recall tests (e.g., generated response; short-answer; Greving & Richter, 2018; Hinze & Wiley, 2011) have a greater benefit than recognition type tests (e.g., multiple choice, selected response; Dunlosky et al., 2013). These results align with the effortful retrieval hypothesis suggesting that the direct benefit of practice testing may stem from the effortful retrieval of previously learned content (Pyc & Rawson, 2008). In other words, this theory suggests that exerting more effort to pull information out of long term memory to generate a response has greater benefit to retention after a delay than if students select one of several given options or have a cue to help them retrieve a specific target (Karpicke, 2012; Rawson & Dunlosky, 2012). The direct effects of practice testing in laboratory settings have repeatedly demonstrated that there is a greater benefit of recall than recognition testing, especially when compared to restudying even when students are unsuccessful at retrieving the target (Agarwal et al., 2008; Butler & Roediger, 2008; Little et al., 2012).

Indirect benefit. Learners may also benefit indirectly from taking a practice test, in addition to the direct benefit, such as when they are given feedback that clarifies misconceptions and fills in missing knowledge (Dunlosky et al., 2013; McDaniel & Little, 2019; Roediger et al., 2011). These indirect benefits of practice testing are an important consideration for practice testing in educational contexts as they can help teachers identify concepts to reteach and guide students' allocation of study time (McDaniel & Little, 2019; Pyc et al., 2014).

The identification of knowledge gaps and failed retrieval attempts may also lead to improved metacognitive monitoring. Research has demonstrated that students who engage in repeated practice testing develop more accurate metacognitive judgements of learning at the final test. Of particular importance here, is that the metacognitive judgements of learning by students who engaged in practice testing were significantly more accurate than those of students who engaged in rereading (Agarwal et al., 2008; McDaniel & Little, 2019). Another indirect benefit is that students who know that they will be completing frequent practice tests in class are incentivized to study more, in addition to being better able to target content to be restudied (Pyc et al., 2014). More frequent no- or low- stakes tests that students feel prepared for and are successful at may also decrease test anxiety during higher stakes tests (McDaniel & Little, 2019). In sum, providing students frequent low stakes practice tests may encourage them to study more, increase their accuracy about what they have already learned, provide guidance on where to focus study time and may reduce the anxiety associated with testing as they become accustomed to them and are more confident in their learning (McDaniel & Little, 2019; Roediger, et al., 2011). Despite the known benefits of practice testing, students persist in the use of less effective strategies (e.g., restudying) even after being shown the benefit of practice testing (Karpicke, 2012; McDaniel & Little, 2019; Miyatsu et al., 2018; Rawson & Dunlosky, 2012) suggesting that implementing practice tests may fall to the instructors in educational contexts (Dunlosky & Rawson, 2015). As such, it may fall to the instructor to instantiate practice testing in the classroom as was done in this work.

Self-explanation

Self-explanation involves a learner explaining new knowledge to themselves as though they were explaining it to another based on what they have previously learned (Hausmann & VanLehn, 2007). Explanations of the thinking process or the steps of solving a problem may be generated automatically or students may be encouraged or trained to do so (Chi et al., 1994; Hausmann & VanLehn, 2007). Research suggests that students who self-explain or self-instruct during the acquisition of knowledge or problem solving are better able to encode and retain new knowledge and skills (Chi et al., 1989).

Benefit of Self-explanation

Research supports the benefits of self-explanations while learning with multiple age levels and in several domains including anatomy (Chi et al., 1994), math (Rittle-Johnson, 2006), medical knowledge (Larsen et al., 2013), and physics (Chi et al., 1989). Additionally, the benefit of self-explanation has been evaluated in combination with different instructional strategies such as direct- and discovery- instruction (Rittle-Johnson, 2006) and via text-based instruction (Chi et al., 1989). Self-explanation has also been evaluated in combination with other study strategies such as practice testing (Butler et al., 2013), note-taking (Peper & Mayer, 1986; Trafton & Trickett, 2001) and with the use of multimedia learning (Roy & Chi, 2005; Wylie & Chi, 2014). One example of the benefit of self-explanation is through an investigation of students' learning Newton's laws of physics from a textbook in which the example problems were insufficient to explain the concept. Students completed a knowledge acquisition phase developing sufficient background knowledge to support acquisition and comprehension of the target knowledge (Newton's laws of motion). During the learning phase, participants were given a test that included declarative (facts from textbook), qualitative (making inferences from textbook information) and quantitative (procedural such as calculating force) questions as well as being asked to complete a verbal explanation protocol. Students were given opportunities to relearn until each chapter test was error free. The target chapter learning phase proceeded the same way except students were given declarative and qualitative questions prior to studying worked examples. After studying worked examples, students completed isomorphic problems (same structure, different values) and far transfer problems while providing 'talk-aloud' protocols (Chi et al., 1989). Results suggest that students who were categorized as 'Good', based on assessment outcome, produced more unique idea units, than students at the opposite end of the spectrum, classified as ‘Poor’. Good students also produced a higher proportion of explanations compared to ‘Poor’ students who produced more monitoring and other statement types. The quality of statements produced by the ‘Good’ students was higher than those produced by the ‘Poor’ students. These results suggest that students who generate sufficient high-quality explanations while studying worked examples had improved learning of procedural skills at posttest (Chi et al., 1989). The benefit of self-explanation on procedural learning was demonstrated in this study; however, similar benefits have been found with understanding declarative knowledge and integrating new information with prior knowledge. In addition to the benefit self-explanation has on knowledge acquisition, some research suggests that it may improve reasoning skills (Chamberland & Mamede, 2015, Kong, 2015) and it may lead to generalization or improved transfer of procedural knowledge to novel contexts, especially when combined with initial direct instruction (Chi et al., 1989; Rittle-Johnson, 2006).

Rittle-Johnson (2006) investigated the benefit of self-explanation on procedural learning, procedural transfer, and conceptual understanding with third through fifth grade students learning mathematical equivalence using A+ (2 + 3 + 4 = 2 + __) and +C (2 + 3 + 4 = __ + 4) problem types. Students experienced a pretest, a learning phase, and immediate and delayed posttests. Students in the explanation condition were shown correct and incorrect solutions from children at another school and asked to describe *how* the students solved the problems and *why* the answers were right or wrong. The benefit of self-explanation was demonstrated at both immediate and delayed posttests for procedural learning and at the delayed posttest for conceptual learning. Furthermore, students in the self-explanation groups (invention + and instruction +) performed better on procedural transfer questions (i.e., mathematical equivalence problems in a different format) than students who did not engage in self-explanation. Students in both groups struggled with the +C problem type, where the first addend after the equals sign was the blank; however, this struggle was decreased for those in the self-explanation group. Students in the direct instruction plus self-explanation group outperformed all other groups (Rittle-Johnson, 2006). These results further reinforce the benefit of producing explanations during learning on long-term retention and the increased performance on transfer tasks by students suggests that the constructive processing required to generate self-explanations may be a key advantage of this strategy. However, it is important to note that students learned the procedure but were less successful at acquiring the conceptual knowledge in contrast to prior studies that evaluated acquisition of conceptual knowledge from studying text passages (Renkl & Eitel, 2019; Rittle-Johnson, 2006). The differing results suggest that the type of prompt, the type of content, and the instructional modality (e.g., direct instruction, discovery, multimedia) are important considerations when using self-explanation as a learning strategy (Renkl & Eitel, 2019; Rittle-Johnson & Loehr, 2017).

Interactive Learning

Interactive learning refers to an activity in which learners are conversing with one or more people (e.g., peer, tutor, teacher; Roscoe & Chi, 2008). In order for the conversation to be considered an interactive learning activity, all learners must have opportunities to engage in generative activities like self-explanation. While one learner is explaining, other members of the group engage in questioning, feedback, or elaboration. All members of the group should have roughly equal opportunities to explain, question, elaborate, and provide feedback (Chi, 2009). Dyads or groups in which one individual does all the talking, is not interactive learning as others are merely listening to a speaker while in a small group (Chi, 2009; Roscoe & Chi, 2008). This combination of activities can help students encode new knowledge, develop connections between new and existing knowledge, fill gaps in knowledge as well as correct misconceptions during the interaction (Chi & Wylie, 2014; Prunuske et al., 2012; Roschelle et al., 2010). Students may also experience interactive learning using technology (e.g., intelligent tutor, interactive video; Zhang et al., 2006), or physical interaction (e.g., building a model; Chi, 2009) and interactions may occur in or out of the classroom. Extant literature suggests that the benefit of interactive learning activities is greater than that of constructive, active, and passive learning activities (Chi, 2009; Chi & Wylie, 2014). One reason for this may be due to the addition of feedback and elaboration to the constructive processing that occurs in co-constructing explanations (Chi & Menekse, 2015). Below I discuss the benefit of interactive learning with technology and interactive learning with peers including a specific strategy (Interactive Peer Learning; Mazur, 1997) that were used in this work.

Interactive Learning with Technology

Interactive learning may also use technology applications such as intelligent tutoring systems (Chi, 2009; Pilegard & Fiorella, 2016), question response systems (e.g., “clickers”; Paiva et al., 2014), or interactive video applications (e.g., EdPuzzle™; Arner, 2020, Zhang et al., 2006). For the purpose of this work, I will discuss the use of interactive video and question response systems below.

Interactive video. The use of video is not new to teaching and learning (Fiorella & Mayer, 2018; Hung et al., 2018). Instructional video has increased in popularity based on evidence demonstrating a benefit to learning when compared to reading alone (Bergmann & Sams, 2012) and due to the increased constraints on teachers’ classroom instructional time (Fox, 2018; Hung et al., 2018). The implementation of video-based instruction may be passive (i.e., view only; Fadde & Sullivan, 2013), active (i.e., note-taking occurs; Peper & Mayer, 1986) or interactive (i.e., learner manipulates some aspect; Fiorella & Mayer, 2018) and may be used in a variety of instructional models (e.g., fully online, flipped; Hung et al., 2018). Videos may be considered interactive in two ways. The first is through student control such as replay or accessing clickable resources (Zhang et al., 2006). Student controlled interaction may be influenced by cues in the video (i.e., "stop here to reflect"; Mayer, 2017) but ultimately, it is up to the student to act. A frequently mentioned benefit to student control of instructional video is the opportunity to rewind portions that they do not understand or when they find themselves disengaged (Bergmann & Sams, 2012; Foertsch et al., 2002). The second type of video interaction is instructor-controlled using design features such as segmentation (e.g., video stops, question pops up, learner answers, video resumes; Hung et al., 2018). Instructor-controlled interactive video aligns with the ICAP framework used in this work and is discussed further below (Chi, 2009).

Segmentation is an example of instructor-controlled interaction that may include video features that require the student to perform an action to continue playing the video such as selecting a response target with a mouse-click or entering text response to a question (Mayer, 2017). Research suggests that segmenting video in this way may benefit learning by reducing cognitive load for learners as they may try to hold all of the information without time to process and encode it (Mayer, 2017; Mayer & Pilegard, 2014). Furthermore, Biard and colleagues (2018) evaluated the difference between continuous video (i.e., no pausing), learner-paced (i.e., learner stops the video when they think necessary), and system-paced (i.e., at the end of an instruction segment) in a medical procedure instructional video for novice learners. Results showed that roughly half of the students who had the option to pause, chose not to. In support of their hypothesis, students in the system-paced group demonstrated a much higher score on the measure of procedural learning than those in the self-paced or continuous groups but this benefit did not hold for conceptual recall. Not surprisingly, there was no difference between the continuous and self-paced groups. These results suggest that system-paced (i.e., instructor controlled) interactive videos may improve learning through preventing cognitive overload (Biard et al., 2018; Mayer & Pilegard, 2005). Segmenting videos and requiring student interaction may also benefit student learning due to increasing content salience at the pause point. A study evaluating the difference between unrestricted, learner-controlled video instruction and restricted, system-controlled question-embedded video, demonstrates this benefit (Vural, 2013). Learners who responded to questions cuing important concepts performed better at posttest than learners who were fully in control of video instruction (Vural, 2013). Two items of note in these analyses; first, while the analyses did control for pretest scores, they did not control for the increased amount of interaction with content required in the question embedded group and second, they did not control for the benefit of practice testing. While interactive video has been shown to be beneficial to learning, additional work is needed in this area.

The interactive videos used in this work include embedded questions that require students engage in constructive cognitive processing to generate a response before continuing. Students also received feedback when interacting with videos aligning with interactive learning in the ICAP framework.

Question response systems. Question response systems allow the instructor to create questions to which the learners will respond with a proprietary device (e.g., Turning Point™), a mobile device (e.g., text response), or a laptop (e.g., PollEverywhere™; Arner, 2020). The use of a question response system and question set may be the entirety of the interactive learning activity or they may be used to both formatively assess student understanding and/or to seed conversations for students in interactive learning groups discussed below (Deslauriers et al., 2011; Paiva et al., 2014; Roschelle et al., 2010). PollEverywhere™ was used as a question response system to guide students' table-group conversations about the topic in Experiments 1 & 2 as part of the interactive peer learning model described below.

Interactive Learning with Peers

Interactive learning activities with peers may take place in dyads or small groups during which learners engage in reciprocal dialogues about a topic of instruction or to achieve a common goal (Chi, 2009; Mazur, 1997; Nokes-Malach et al., 2019). According to the ICAP framework, interactive learning activities require all students in the group to take turns explaining a topic to their peer(s) as well as providing feedback or elaboration following the peer's explanation. Interactive learning must contain approximately equal proportions of explanation and feedback in order for it to be beneficial to all students' learning (Chi, 2009, Chi & Menekse, 2015; Roscoe & Chi, 2008). The goal of interactions may be problem solving (e.g., math; Roschelle et al., 2010), discuss or debate of formative assessment responses (interactive peer learning; Mazur, 1997), developing conceptual understanding (Crouch & Mazur, 2001; Menekse & Chi, 2018; Smith et al., 2011) or joint knowledge construction (Chi et al., 2008).

Interactive peer learning. Interactive peer learning is a method of implementing collaborative learning (Mazur, 1997). This method uses a question response system to guide discussion of formative assessment questions by small groups of students to improve learning. First, students independently respond to a formative assessment question using a question response system (e.g., "clickers"; Crouch & Mazur, 2001). Following the individual response, students discuss their responses with their peers. After discussion, students respond to the same formative assessment question independently to ensure that those who responded incorrectly have updated their understanding (Crouch & Mazur, 2001; Schell & Mazur, 2015). The percentage correct for the second response indicates to the instructor that the students are either understanding the content or that they need additional “just-in-time” instruction (i.e., immediate instructional response) to address misconceptions and/or clarify difficult concepts (Schell & Mazur, 2015). Instructors may set a criterion percent correct before moving on to the next concept which may be met on the first or second question response (Crouch & Mazur, 2001; Lasry et al., 2008). The peer instruction model of collaborative learning has been shown to improve student learning outcomes in physics (Lasry et al., 2008; Mazur, 1997), chemistry (Schell & Mazur, 2015), and biology (Smith et al., 2011). Comparisons in algebra-based and calculus-based physics courses revealed that students in the interactive peer learning model outperformed a historical sample of students who were taught using the traditional lecture model. Learning outcomes were evaluated using a pre-/post- design on standardized assessments of fundamental physics knowledge (e.g., ConcepTests, Mechanics Baseline Test, Force Concept Inventory; Crouch & Mazur, 2001). It is important to note here that the sample used for comparison was historical and the instructor may have improved their instructional practice over the years of comparison. Additionally, students in the traditional lecture course were not “expected” to prepare for class ahead of time whereas the intervention groups were (Crouch & Mazur, 2001; Mazur, 1997).

The evidence-based practices discussed here have been widely researched individually in numerous content areas and with a variety of student levels and in a variety of settings. To my knowledge, these strategies have not been explicitly used in work comparing the flipped and traditional classroom models. One goal of using the first three strategies in both classrooms was to evaluate whether or not implementing them in the flipped instructional model may increase their efficacy, leading to improved learning outcomes. For example, students may garner greater benefit of distributed practice in the flipped classroom when each practice session includes activities that elicit more active cognitive processing (Baepler et al., 2014; Jensen et al., 2015). The fourth strategy, interactive learning, was only included in the flipped classroom as it is one of the key differences between the flipped and traditional instructional models.

Instructional Models

The two instructional models compared in this work are the traditional, lecture-based model and the flipped instructional model. The components and evidence related to the traditional classroom followed by the flipped classroom are described below. Next, there is a description of how each of the four levels of cognitive processing in the ICAP framework (Interactive-Constructive-Active-Passive; Chi, 2009) may be applied in each.

Traditional Classroom

A common feature of the traditional classroom model is the introduction of new knowledge through an expert-provided lecture to a class of novice learners (Bergmann & Sams, 2012; Freeman et al., 2014; Gardiner, 1998). This method of instruction, sometimes referred to as the “sage on the stage” model, has a primary focus on the instructor and has been used since the inception of formal education. The role of the learner in the traditional classroom is to receive the knowledge being given out by the instructor therefore holding a passive role in their education (Freeman et al., 2014; Gardiner, 1998; Lage et al., 2000). The next section includes a description of the traditional classroom divided into pre-class, in-class, and after-class components.

Components of the Traditional Classroom

The instructional cycle in the traditional classroom frequently begins with a reading assignment as the pre-class learning activity (Miller et al., 2018; Strayer, 2012). The reading may be used as a pre-training method to help reduce the amount of cognitive effort students' need to exert during the lecture by increasing recognition or familiarity with terms to be covered in the lecture (Koedinger et al., 2013; Mayer, 2017; Smith & Ragan, 2004). The intent of instructors who assign reading as preparation for a lecture is to introduce students to new terms and concepts; however, the level of cognitive processing and whether or not students complete the pre-class reading influences their understanding (Chi, 2018; Mazur, 1997). Therefore, introduction of new knowledge still takes place in the classroom via an instructor's lecture in case students did not complete the assigned reading before class or did not have sufficient prior knowledge to understand it (Foertsch et al., 2002; Lage et al., 2000; Prunuske et al., 2012).

The in-class phase of the traditional instructional model is centered around the instructor who disseminates new knowledge through a lecture. Students generally take a passive role in learning in the traditional classroom by merely listening to the lecture (Bergmann & Sams, 2012; Jensen et al., 2015; Lage et al., 2000; Mazur, 1997).

The after-class learning activity is an opportunity for students to apply newly acquired knowledge (Bergmann & Sams, 2012; Mazur, 1997). The 'homework' that students complete after class often results in the realization that students do not fully grasp the new information and they do not have an instructor present to assist (Foertsch et al., 2002). While some instructors may reserve a portion of the in-class time for students to begin the after-class knowledge application activity, in this work the knowledge application phase occurs completely after class.

Traditional Classroom Evidence

The instructional method of “teaching as telling” or lecture-based instruction has become the de facto standard against which a variety of instructional methods have been compared (Freeman et al., 2014; Gardiner, 1998; Stephenson et al., 2008). The next section includes evidence related to student learning in comparisons of the traditional classroom with computer assisted instruction and discovery instruction, respectively. Then, there is a discussion of evidence related to students’ engagement in the traditional classroom.

Student learning in the traditional classroom. Student learning in the traditional classroom has been compared with computer assisted (AKA online, distance, and asynchronous) instruction in numerous content areas including medicine (Davis et al., 2007; Mohee et al., 2019), gerontology (Gallagher et al., 2005), research methods (Frederickson et al., 2005), and garment construction (Slocum & Beard, 2005) and human genetics (Stephenson et al., 2008). In some studies, the investigation comparing traditional instruction with computer assisted instruction was driven by cost or convenience and a determination of ‘not worse than’ traditional instruction on student learning was considered a positive result. For example, Davis and colleagues (2007) investigated the equality of learning outcomes in computer-based teaching and face-to-face lecture conditions as a method to accommodate the busy schedules of practicing physicians. They created equivalent online (recorded lecture with hyperlinks) and in person lectures and found that there was not a significant difference in learning for either condition. Similar results occurred in a study (Mohee et al., 2019) in the medical field with immediate posttest scores being roughly equivalent following web-based instruction. However, Mohee and colleagues (2019) found that students in the e-learning condition had significantly higher test scores than the traditional lecture at the two-week delay. Frederickson et al., (2005) used a switching replications design and found that there was not a significant difference in learning outcomes between the web-based and lecture-based instruction in research methods. However, they also evaluated students' satisfaction with both instructional conditions and found that students preferred the traditional over the web-based units. In sum, these results suggest that use of computer assisted instruction is at least as good as traditional lecture and in some cases may lead to improved retention (Davis et al., 2007; Frederickson et al, 2005; Mohee et al. 2019).

A second instructional method that has been compared with the traditional lecture method of instruction is discovery learning which is supported by proponents of constructivist learning theory (Klahr & Nigam, 2004; Mayer, 2004). Supporters of discovery learning suggest that students' construction of knowledge through 'discovering' underlying features or procedures will lead to longer retention and greater transfer (Klahr & Nigam, 2004; Rittle-Johnson, 2006). Conversely, supporters of direct instruction argue that discovery is too cognitively demanding for novice learners, they may not 'discover' the correct conceptual information or procedure, or learning may be superficial, preventing successful transfer (Glogger-Frey et al., 2017; Klahr & Nigam, 2004; Mayer, 2004).

Klahr and Nigam (2004) compared these two instructional models with children learning *control-of-variables strategy* where one group was given an apparatus to experiment with manipulating two variables (steepness and length) with the goal of developing unconfounded experiments. The direct instruction group had the same apparatus and goal but they observed the experimenter design both confounded and unconfounded experiments before trying it on their own. They found that the students who received direct instruction had a greater improvement than those in the discovery condition at immediate posttest. Similarly, Rittle-Johnson (2006) evaluated direct instruction and invention (discovering) in mathematical equivalence problems and found that students who received direct instruction on the correct procedure performed better on immediate and delayed posttests of procedural learning. The benefit of direct instruction was also bolstered by the addition of students' self-explaining the procedure. In fact, the addition of self-explanation was beneficial to learning for students in both types of instruction (Rittle-Johnson, 2006).

Student engagement in the traditional classroom. Research suggests students in traditional instruction settings may experience lapses in attention or engagement with the material due to mind wandering or technology-based multitasking (Peper & Mayer, 1986; Wammes & Smilek, 2017; Wammes et al., 2019). Additionally, detachment from the content can negatively impact students' learning outcomes (Chi & Wylie, 2014; Sana et al., 2013; Wood et al., 2012).

Wammes and Smilek (2017) investigated the students' reported mind wandering during class lectures and the impact that the break in attention had on learning. They used prompts in presentation slides to gather students' self-reported mind wandering. Students were asked to select from five possible responses ranging from 'completely on-task' (attending to the lecture) to 'completely mind wandering' (not attending to the lecture at all) with 'both on task and mind wandering' as the median response option. Results suggested that the mean degree of mind wandering remained fairly constant throughout the lecture with increased mind wandering at the beginning and the end of the lecture. Also, and more importantly, students reporting more mind wandering (i.e., not attending to the lecture) experienced decreased performance on quizzes taking place at the end of the lecture (Wammes & Smilek, 2017).

An additional concern with students' engagement with the lecture in the traditional classroom is the possible distraction that may occur with media use (i.e., phone or laptop; Sana et al., 2013; Wammes et al., 2019). It is becoming increasingly common for students to use technology in the classroom as a learning tool such as to take notes (Morehead et al., 2019; Sana et al., 2013). However, recent research has found that students may also use their technology for tasks unrelated to the lecture or learning such as checking social media, email, and other tasks of this nature (Wammes et al., 2019). Research suggests that the loss of attention to the lecture may have a negative impact on learning outcomes for the students who are multitasking as well as surrounding students (Sana et al., 2013; Wammes et al., 2019; Wood et al., 2012).

Mind wandering and media multitasking have been shown to have a negative influence on students' learning in the traditional lecture classroom. This may be due, in part, to the passive nature of the teacher-centered lecture (Braun et al., 2014; Chi & Wylie, 2014; Sana et al., 2013). A potential solution to prevent students from directing their attention away from the lecture is to create a more active learning environment where they are cognitively engaged in learning activities such as problem solving or discussion with peers (Jensen et al., 2015).

Summary of traditional classroom evidence. Evidence of student learning in the direct or lecture-based instruction model is mixed. When comparing lecture instruction with asynchronous, web-based instruction (i.e., video lecture, web-based module) results showed that web-based instruction was at least as good as lecture, if not better. Similarly, direct instruction was shown to have greater benefit to learning and transfer than students 'discovering' knowledge. Further benefit was found when direct instruction was combined with another evidence-based learning strategy. However, the potential for students to mind wander or multitask during a lecture does have a negative impact on learning outcomes. The combination of this evidence suggests that a more active classroom that includes direct instruction and guided constructive and interactive learning activities may lead to improved learning outcomes. The flipped classroom is an example of an instructional model that includes this type of activity and is discussed further below.

Flipped Classroom

The flipped instructional model falls under the larger umbrella of blended learning models (Cheng et al., 2019; Hamdan et al., 2013). Blended or hybrid instruction models generally replace some portion of previously in-class learning activities with technology-based instruction (Olapiriyakul & Scher, 2006). In the flipped classroom model, the online component is completed prior to the class meeting and is the introduction to new knowledge such that the ‘lecture’ and ‘homework’ of the traditional model are inverted (Bergmann & Sams, 2012; Lage et al., 2000; Karabulut-Ilgu et al., 2018; O’Flaherty & Phillips, 2015). A discussion of the flipped classroom divided into pre-class, in-class, and after-class components is below.

Components of the Flipped Classroom

The flipped classroom model does not have a single, required modality for pre-class activities. However, in the flipped classroom, the pre-class activity serves as the introduction to new knowledge that occurs during the in-class lecture in the traditional instructional model (Bergmann & Sams, 2012; Lage et al., 2000; Pierce & Fox, 2012; Schell & Mazur, 2015). New knowledge may be introduced through instructor-recorded video lectures, field-expert videos, individual, or collaborative reading activities (Bergmann & Sams, 2012; Bishop & Verleger, 2013; Miller et al., 2018; O’Flaherty & Phillips, 2015; Strayer, 2012). In contrast to the traditional classroom pre-training activity, instruction that would normally occur in the classroom is now presented prior to class.

The in-class portion of the flipped classroom model is now used for students to complete activities applying the knowledge learned in the pre-class activity. The primary difference between the flipped and traditional classroom models is this inversion of initial instruction and application of knowledge (DeLozier & Rhodes, 2017; Hamdan et al., 2013; Schell & Mazur, 2015; Zainuddin & Halili, 2016). The activities that students complete may be done individually (Foldnes, 2016) or in collaboration with peers (Crouch & Mazur, 2001; Freeman et al., 2014; Morton & Colbert-Getz, 2017). The instructor then takes a supporting role to answer students' questions and clarify observed misconceptions (Bishop & Verleger, 2012; Foertsch et al., 2002; Freeman et al., 2014; Schell & Mazur, 2015).

The after-class portion of the flipped instructional model may also vary in design. However, the purpose of any after-class activities is to extend classroom practice (Hwang & Lai, 2017; Morin et al., 2012). Students may complete unfinished problems from the in-class activity, be assigned additional problems, or choose to practice more based on feedback from peers or instructor (Schell & Mazur, 2015).

Flipped Classroom Evidence

The flipped classroom model has been evaluated using a variety of criteria with qualitative, quantitative, and case-study designs (Giannakos et al., 2017; O'Flaherty & Phillips, 2015; Zainuddin & Halili, 2016). Some of these criteria include student learning outcomes (Cheng et al., 2019; Lai & Hwang, 2016), students' perceptions of their learning (Gundlach et al., 2015), student perceptions of the format (Davies et al., 2013; Mason et al., 2013; Tune et al., 2016), and teacher's perceptions of the format (Hao & Lee, 2016). For the purpose of this work, evidence relating to students' learning outcomes and students' engagement and participation in the flipped classroom is included.

Student learning in the flipped classroom. The impact on student learning of the flipped classroom model has been evaluated with a variety of performance measures such as course grades, assessment scores and reduced failure or withdrawal rates as well as evaluations of students' and teachers' perceptions of the model. These evaluations have occurred in a variety of content areas including humanities (Kong, 2015), mathematics (Braun et al., 2014; Scott et al., 2016), statistics (Foldnes, 2016; Peterson, 2015; Winquist & Carlson, 2014), science (Baepler et al., 2014; Deslauriers et al., 2011; Prunuske et al., 2012), technology (Davies et al., 2013; Enfield, 2013), engineering (Foertsch et al., 2002; Morin et al., 2013;), and a variety of healthcare fields including nursing (Gilboy et al., 2015; Njie-Carr et al., 2017), pharmacology (McLaughlin et al., 2013; Mitroka et al., 2020; Pierce & Fox, 2012), and medical education (Morton & Colbert-Getz, 2017). Research on the flipped instructional model has found a positive impact on students' learning outcomes when compared with historical samples of the traditional instructional model of the same course using the same instructional content and assessments but with inverted structure (Baepler et al., 2014; Deslauriers et al., 2011; Hibbard et al., 2016; Morin et al., 2013). In other words, these studies found that moving the lecture to a pre-class activity and converting in-class activities to application or problem-solving had a positive impact on students' learning compared to prior semesters. Studies have also shown improved student learning in the flipped classroom when compared with traditional control classrooms (Bhagat et al., 2015; Davies et al., 2013; Peterson, 2016; Winquist & Carlson, 2014). Bhagat and colleagues (2015) compared the flipped instructional model with the traditional lecture on a trigonometry lesson with a pre-/post- quasi-experimental design over six weeks. They found that students in the flipped (experimental) condition performed significantly better on the posttest than the traditional (control) condition. In addition to the studies mentioned above, recent meta-analyses found trivial to small positive effects on learning (multiple content areas; Cheng et al., 2019), small positive effect (multiple content areas; van Alten et al., 2019) and a significant benefit to learning (health professions; Hew & Lo, 2018).

Research has also shown that student learning outcomes in the flipped classroom are not significantly different than those of students in the traditional classroom (Braun et al., 2014; Morin et al., 2013; Morton & Colbert-Getz, 2017; Scott et al., 2016). For example, results from a business course demonstrated little difference in scores between flipped and non-flipped classrooms taught by the same instructor (Findlay-Thompson & Mombourquette, 2014). Morton and Colbert-Getz (2017) found that post-graduate medical training conducted in flipped and traditional instructional models did not result in any significant difference in students' learning outcomes overall. However, analysis by question type, based on Bloom's taxonomy, revealed that students in the flipped condition did significantly better on questions at the *Analysis* level of the taxonomy. Similarly, a large study done at West Point Academy suggests that the flipped classroom may have some benefit in the short term but does not improve learning after a delay based on quiz and exam scores (Setren et al., 2019).

One difficulty with evaluating learning in the flipped classroom is the broad range of implementation methods used (Brewer & Movahedazarhouligh, 2018; Giannakos et al., 2018; O'Flaherty & Phillips, 2015). For example, one frequent method used to introduce content is through instructors sharing videos of themselves lecturing which students view prior to the class meeting (Abeysekera & Dawson, 2015). An advantage of a video-recorded lecture is that students may view it multiple times compared to a single instance of an in-person lecture (Burke & Fedorak, 2017). However, the use of a video-recorded lecture may not always improve student learning in the flipped classroom compared to an in-person lecture; an ineffective lecture that has been recorded is still ineffective (O'Flaherty & Phillips, 2015; Pierce, 2013; Stephenson et al., 2008; Walker et al., 2020). Instructors have also used video recorded by an expert, interactive video (e.g., questions embedded), cognitive/interactive tutors, interactive text, or some combination thereof (Cheng et al., 2019; Giannokos et al., 2018; Guo, 2019; Roach, 2014; Zainuddin & Halili, 2016). One reason simply switching to a pre-class introduction to content may not be effective is that students may struggle with comprehension and transfer when learning new concepts without guidance or scaffolding. Some literature suggests that lack of guidance or scaffolding bridging new knowledge to prior knowledge or knowledge application may be cognitively overloading, resulting in frustration or confusion (Kirschner et al., 2006; O'Reilly et al., 2019; van Alten et al., 2019). Additionally, research suggests that there should be cohesion between in- and out-of-class activities because it can impact students' understanding and ability to apply new knowledge (Hamdan et al., 2013; O'Flaherty & Phillips, 2015; Peterson, 2016; Strayer, 2012). In other words, simply inverting 'lecture' and 'homework' may not be the basis of any positive effect on learning outcomes. Students engaging in more active, constructive, and interactive cognitive processing may be the factor contributing to learning gains in the flipped classroom (Baepler et al., 2014; Bishop & Verleger, 2013; Chi & Wylie, 2014; Freeman et al., 2014; Jensen et al., 2015; Prunuske et al., 2012).

Student engagement in the flipped classroom. A common descriptor of the flipped classroom model is that it is more student-centric compared to the teacher-centric traditional classroom, suggesting that students have more control or are more actively involved or engaged in the learning process (Jenkins et al., 2017; O'Flaherty & Phillips, 2015). Student engagement in the flipped classroom may be evaluated through two lenses. First, students in the flipped classroom report increased *active* participation in both the pre-class and in-class learning activities. In other words, rather than the students' participation revolving around passively listening to a teacher, they are applying newly learned knowledge either individually or with peers (Bishop & Verleger, 2013; Brewer & Movahedazarhouligh, 2018; Foldnes, 2016; Gilboy et al., 2015). When working with peers, students may benefit from dialogue through adding to their new knowledge, correcting misconceptions, or getting feedback on knowledge gaps (Chi & Wylie, 2014; McLaughlin et al., 2013; Roscoe & Chi, 2008). Second, students in the flipped classroom are more cognitively engaged while working on projects or solving problems (Crouch & Mazur, 2001; Guo, 2019; Prunuske et al., 2012). While students are applying new knowledge, the instructor can observe problem-solving procedures, listen to discussions, answer questions, and address misconceptions expressed by students (Bergmann & Sams, 2012; Schell & Mazur, 2015). Students in the flipped classroom model report being more engaged in their learning, particularly when the in-class activities required higher order thinking or were challenging (Davies et al.,2013; Giannakos et al., 2018; Guo, 2019).

Summary of flipped classroom evidence. Evidence of learning and engagement in the flipped classroom is mixed. Results have shown increased learning, no difference, and a decrease in learning when comparing a traditional, historical or concurrent control, classroom. Students may experience improved learning outcomes and they may also report increased or decreased satisfaction. However, students in the flipped classroom model generally report increased engagement with learning activities and increased interactions with peers and the instructor when compared to the traditional lecture model (Guo, 2019; McLaughlin et al., 2013; Zainuddin & Halili, 2016). The variety of evidence may stem from the lack of a framework or system of design with which to develop learning activities in the flipped classroom. This work is comparing the flipped and traditional models using the ICAP framework as a guide. The levels of cognitive processing that may occur during learning activities in each model are described below.

Cognitive Processing in Instructional Models

Instructor planned activities are intended to improve student learning and retention (Smith & Ragan, 2004). However, the level of cognitive processing that students engage in will influence the benefit of activities on learning. In other words, the level of cognitive processing that students engage in is variable depending on the design of the activity, regardless of instructional model (Chi, 2018; Chi & Wylie, 2014). Below I describe levels of cognitive processing that may occur in each phase (pre-class, in-class, and after-class) in the traditional and flipped instructional models.

Cognitive Processing in the Traditional Classroom

The types of cognitive processing that students engage in during these activities may vary as well in spite of the intent of the instructor (Chi & Wiley, 2014; Chi et al, 2018). The pre-class activity in a traditional classroom often includes reading some material as pre-training for the in-class activity (e.g., textbook or article; Braun et al., 2014; Crouch & Mazur, 2001). While reading the assigned text, students may be passively engaging if they are simply reading the words and nothing else (Chi, 2009). Students who are highlighting what they identify as important content or taking summary notes on that content, are engaging in active cognitive processing. While we know that highlighting and summarizing are not as effective as other learning strategies, the active processing involved is superior to passive reading (Chi, 2009; Dunlosky et al., 2013). Students who are elaborating on what they are reading or generating self-explanations, while taking notes, are engaging in constructive cognitive processing (Chi, 2018).

The in-class learning activity, as discussed previously, generally includes an instructor disseminating knowledge with students potentially taking notes (Lage, et al., 2000; Morehead et al., 2019). Students may attempt to record the lecture verbatim, which is less active cognitive processing than if they summarize what is being said because their focus is on the words in the former rather than the meaning as in the latter (Chi & Wylie, 2014). Students who are making connections to prior knowledge (i.e., information from past classes or pre-class reading) while taking notes are engaging in active processing of new content whereas students who elaborate on what is being said in the lecture may be engaging in constructive cognitive processing (Chi, 2009; Chi, 2018). Instructors may also encourage active cognitive processing by introducing questions during lectures that students respond to via gesture (thumbs up/down) or 'clicker' or some other type of technology. Students' attempting to retrieve content or demonstrate understanding are engaging in at least active processing (Baepler et al., 2014; Levesque, 2011).

Students in the traditional model often have activities to complete after class, this is typically homework where students practice applying knowledge from the lecture (Braun et al., 2014). These activities may require these same levels of cognitive processing depending on how they are designed. However, one potential issue that may impact students' learning is how they *actually* engage with the content in the after-class learning activities. In other words, students' may not complete the before or after lecture activities or they may do so passively as described in the example above (Braun et al., 2014; Mazur, 1997). Additionally, students may struggle with the post-lecture assignment, fail to complete it, or develop misconceptions about the content (Bergmann & Sams, 2012; Mazur, 1997; O'Flaherty & Phillips, 2015; Prunuske et al., 2012). The flipped classroom model is intended to make learning more active and student centered by intentional design of pre- and in-class learning activities (Baepler et al., 2014; Berrett, 2012;

Burke & Fedorak, 2014; Hamdan et al., 2013; Roach, 2014; Schmidt & Ralph, 2015). The cognitive processing that may occur in the flipped classroom is described below.

Cognitive Processing in the Flipped Classroom

Instructors can design pre-class, in-class and after-class learning activities to elicit the desired levels of cognitive processing (Chi et al., 2018; Jovanović et al., 2017; Giannakos et al., 2018; Lai & Hwang, 2016).

The pre-class activity in the flipped classroom is the students' introduction to new knowledge. Similar to the traditional lecture classroom, students in the flipped classroom may engage in passive cognitive processing while viewing a video-based lecture introducing new concepts if the student listens or watches but doesn't do anything else (Chi, 2009; Lo & Hew, 2017; Mason, et al., 2013). Students may also engage in active processing when taking notes while watching the video lecture, replaying sections to ensure understanding or taking a moment to make connections to prior knowledge (Foertsch et al., 2002; Peper & Mayer, 1986). The ability for students to "self-pace" the lecture, pausing or replaying sections, is one advantage of the flipped classroom over the traditional classroom where the lecture is at the pace of the instructor (Chi et al., 2008; Foertsch et al., 2002; Zainuddin & Halili, 2016).

The structure of in-class time is the second key difference between the traditional lecture and flipped classrooms. The in-class activities that instructors design may require students to engage in more active, constructive, and interactive cognitive processing (Bergmann & Sams, 2012; Chi, 2018; Hamdan et al., 2013; Pierce & Fox, 2012). Students in the flipped classroom frequently complete activities in class similar to those assigned as homework in a traditional lecture format (Hamdan et al., 2013; Strayer, 2012). For example, students may engage in active cognitive processing while solving problems that require retrieval of a procedure that they were introduced to in the pre-class learning activity (Bishop & Verleger, 2013; Braun et al., 2014; Chi, 2009; Mazur, 1997; Pierce & Fox, 2012; Prunuske et al., 2012). The instructor may also provide learning activities that require students to engage in constructive cognitive processing. For example, instructors may ask students to solve problems beyond procedural execution (Peterson, 2016; Pierce, 2012; Wilson, 2013), apply knowledge to reason solutions (DesLauriers et al., 2011), or transfer recently learned knowledge to a novel, hands-on activity (Gundlach et al., 2015). Finally, students in the flipped classroom may engage in interactive cognitive processing when the instructor designs learning activities that require students to collaborate with a peer or small group to develop or construct knowledge (Chi et al, 2018; Menekse & Chi, 2018). For example, instructors can prompt students to answer questions (requires active processing) but then ask students to explain or defend their response to peers which requires interactive cognitive processing (Foldnes, 2016; Mazur, 1997; Roscoe & Chi, 2008). Similarly, students can engage in group problem solving that includes discussion of solution steps or strategies (McLaughlin et al., 2013; Peterson, 2016; Prunuske et al., 2012). Students may or may not be required to complete an after-class activity in the loosely defined flipped classroom model. However, for the purpose of this work, students were given an after-class activity requiring them to provide self-explanations to questions that were discussed in pre- and in-class activities during the week.

Current Investigation

The extant literature on implementations of the flipped classroom demonstrates mixed results and may not accurately reflect efficacy or attribution of benefit found in the flipped classroom model (O'Flaherty & Phillips, 2015). For example, some literature suggests that instructors can/should implement a quiz on pre-class material to encourage students to complete the pre-class activity (Giannokos et al., 2018; Guo, 2019; Scott et al., 2016). While this strategy may increase the fidelity of the flipped model, the research does not take into account that students’ learning may be benefitting from the testing effect (Pyc et al., 2014; Rawson & Dunlosky, 2012). Similarly, students who complete the pre-class, in-class, and after-class activities or assignments have three exposures with adequate space between to benefit from distributed practice suggesting that the benefit to learning may be due to the benefit of spacing (Dunlosky et al., 2013). In other words, the incorporation of known evidence-based practices could be what leads to improved learning outcomes in the flipped classroom when compared to the traditional classroom. Experiment 1 addressed this gap with a comparison of the traditional classroom with a flipped version (i.e., inverted pre-class and in-class learning activities) using two sections of Educational Psychology. A within- and between- subjects design was used to evaluate learning outcomes and compare students reported engagement, preparation, and satisfaction in each model. Experiment 2 addressed this gap by removing three evidence-based practices (distributed practice, self-explanation, and practice tests) from the flipped and traditional classrooms for half the semester. Students were assigned these instructional supports in the second half of the semester and their learning outcomes were compared between instructional models and with and without the evidence-based practices.

The cognitive processing that students engage in during learning activities is also a factor in the efficacy of the flipped instructional model. For example, students may be asked to watch a video with no other expectations, they may be required to take notes or answer questions during the video, or they may be asked to complete an interactive simulation (Cheng et al., 2019; O’Flaherty & Phillips, 2015; Zainuddin & Halili, 2016). Similarly, the activities that occur in the flipped classroom are more active, requiring students to solve problems independently or in collaboration with peers. Research on the flipped classroom model has shown that students report higher engagement and increased participation in flipped classrooms than in traditional lecture-based classrooms, suggesting that it is a more *active* learning environment (Giannakos et al., 2018; Nouri, 2016; McLaughlin et al., 2013; Prunuske et al., 2012). Furthermore, a key difference in the flipped and traditional classrooms is the interactive learning that takes place in class the flipped instructional model. The interactive learning activities included in the flipped model could be the more beneficial feature, yet this has not been clearly evaluated in the literature. Experiment 3 addressed this gap in the literature by comparing the flipped and traditional classrooms with identical content and temporal arrangement (i.e., all activities in the same order) but the flipped classroom included interactive learning and the traditional classroom did not. Students' reported engagement, preparation, and participation was also compared in Experiment 3.

CHAPTER III

METHOD, RESULTS, AND DISCUSSIONS

Experiment 1

The purpose of Experiment 1 was to compare the benefit of the flipped instructional model and the traditional, lecture-based model on student learning and their perceptions of their engagement, preparation, and participation during a full semester of undergraduate Educational Psychology. The method, procedure, results and discussion for Experiment 1 are described below.

Research Questions

(1) Is there a difference in student learning outcomes measured by knowledge check and exam scores between the flipped and traditional formats? and;

(2) Does instructional model influence students' reported engagement, preparation, and participation?

Hypotheses

1. Students' learning outcomes on knowledge checks and exam would be higher while in the flipped than when in the traditional format.
2. Students' reported engagement, preparation, and participation would be different between instructional formats.

Method

The design, participants, materials, measures, and procedure for Experiment 1 are described below.

Design

A switching replications design (Trochim, 1986) was used with two sections of undergraduate Educational Psychology in Spring of 2018. One instructor taught the first eight weeks of the semester using the flipped instructional model and the second eight weeks was taught using the lecture-based model. The second instructor did the reverse (See Table 1). The switching replications design was used to reduce teacher effects and extrapolate the benefit of the instructional model over and above any effect due to the instructor.

Table 1

Switching replications design used by two instructors of Educational Psychology

Instructor	First 7 weeks	Second 8 weeks
A	Flipped	Traditional
B	Traditional	Flipped

Participants

Fifty-four undergraduate students in a midsize Midwestern university completed the course and consented to share their data; 29 with instructor A and 24 with instructor B (76% Female; 81% Caucasian, 11% Black, 8% Other). Participants received one research credit for consenting.

Materials

The materials used in Experiment 1 are described below. The same materials were used in both instructional models except where noted.

Readings. Each week students in both sections were assigned readings from articles (practitioner-focused and empirical), chapters in the course textbook, and chapters or excerpts from other selected textbooks. Students in both sections were asked to read the assigned material prior to the designated class meeting (See Appendix A for sample course schedules). The readings were aligned to a weekly topic and both sections followed the same schedule with adjustments for non-school days.

Handouts. Students were provided with a handout at the beginning of each week with 10 short essay questions (See Appendix B for sample handout). The purpose of the handouts was two-fold. First, they were made available at the beginning of the unit so that students could use them to guide their reading and engage with the content making important concepts in the unit more salient. Second, students were told to generate self-explanations in response to the questions to encourage constructive processing of the most important concepts.

EDpuzzles. EDpuzzles are interactive videos in which the instructor has embedded questions that students must answer prior to continuing the video. Students accessed the EDpuzzles through the course learning management system (Blackboard™). The EDpuzzle assignments had a due date of 15 minutes before each corresponding class period to encourage students to complete them in preparation for class discussions. The EDpuzzle videos were purposefully selected to support or reinforce the concepts in the reading assigned for that class session. For example, in week four students read “Making Things Hard on Yourself, But in a Good Way: Creating Desirable Difficulties to Enhance Learning” (Bjork & Bjork, 2011) and the accompanying video was Dr. Bjork explaining ways teachers and students can introduce desirable difficulty to benefit learning. The instructor embedded questions immediately following several key points (number varied by length of the video) that required students to cognitively process the point that was just made to generate a response (Chi & Wiley, 2014).

Students in the traditional section watched the same videos during class without the embedded questions. See Appendix C for a complete list of videos.

Measures

All measures were course assignments and were assigned a point value for completion or a grade (Table 2). Students also received one research credit for consenting to data use. Student learning measures are described below followed by student engagement measures.

Table 2

Name, outcome, quantity and descriptions for all measures used in Experiment 1

Measure	Outcome	Quantity	Description
Pretest	Learning	1	50 multiple-choice questions to evaluate group equivalence
Knowledge Checks	Learning	15	5 multiple-choice and short answer questions
Exams	Learning	4	40 multiple-choice and 2 short essay questions
Participation Form	Student Engagement	30	5-point Likert scale rating *Engagement, Preparation, and Participation* from “Not at all” to “Very High” each class meeting

Student learning measures. Student learning in each instructional model was measured using the weekly knowledge checks and exam scores. Students’ prior knowledge was compared using a pretest on the first day of class. These measures are described further below.

Pretest. Students were given a pretest on the first day of class to assess existing knowledge of the content that would be taught throughout the semester. The pretest consisted of 50 multiple-choice questions and was taken in the course learning management system (Blackboard™). Students were told to do the best they could and that they would receive credit for completing the pretest.

Knowledge checks. At the end of the second class period each week, students took a knowledge check that covered the material for each week. The knowledge check consisted of four multiple choice questions with competitive responses (Little et al., 2012) and one short essay question. The short essay questions frequently asked students to compare and contrast concepts or theories (e.g., Piaget & Vygotsky, Classical & Operant Conditioning) or describe a model (e.g., working memory, concreteness fading) covered during the week. The knowledge checks acted as practice tests to induce retrieval practice in a low-stakes environment.

Exams. Four exams were given to both classes throughout the semester, two in each half (i.e., students took two exams in each instructional model). Exams included 40 multiple choice questions and two short essay questions similar to those found on the knowledge checks except the essay questions were more comprehensive. For example, students were asked to describe both classical and operant conditioning, provide an example of each, and how they might apply them in a classroom context. All exams covered only material introduced since the last exam.

Student engagement measures. Students' engagement was measured using self-reported ratings. They were asked to rate honestly their *Engagement* with the provided content; *Preparation* for class; *Participation* in class using a Google Form™ (see Appendix D) at the end of each class period. Each question used a 5-point Likert-scale from 0 - "Not at all" to 4 - "Very High". On the first day of class, students were told what was meant by each of the terms and that their responses about their engagement, preparation, and participation would not be used to assign their class participation grade. Students were told that *Engagement* referred to time spent reading/watching and thinking about all the material provided for each class meeting. Students were told that *Preparation* referred to completing all assigned tasks and being ready to participate in class. Students were told that *Participation* referred to how much they participated in each class meeting.

Procedure

The instructional models used in this work are compared using learning activities that were assigned or occurred pre-class, in-class, and after-class. Both models are discussed below and compared in Table 3.

Flipped instructional model. The procedure below was followed by both instructors whether teaching the flipped section in the first or second half of the semester.

Pre-class learning activity. Students in the flipped classroom were introduced to new concepts using one or two readings and a corresponding EDpuzzle™ video for each class meeting. Students were expected to complete these activities prior to the deadline, 15 minutes before class.

In-class learning activity. The first class session of each week began with the instructor reviewing the knowledge check from the prior week beginning with the first class after the first knowledge check. Students were called on to provide the correct answer to the displayed questions (one at a time) so that every student heard feedback with the correct answers and an explanation of why the other answer options were incorrect for the multiple choice questions. The instructor also reviewed the short answer essay questions to ensure that students were provided with the complete, correct answer and answered any questions students had about the prior week. The new instructional unit began with a short (8-15 minutes) lecture that included an overview of the most important concepts that students would be discussing during that class period (Mazur, 1997). After the mini-lecture, a multiple choice question, generated in Poll Everywhere™ (see Appendix E), was projected on a screen in the front of the room. Students responded independently using their phone (SMS response) or laptop (web-based response). When the class was at 90% or above correct response to the question, the instructor highlighted any key points and moved on to the next question. When students' responses were below 90% correct, they engaged in an interactive group (4-6 students per table) discussion for 5-8 minutes then responded to the same question, independently, a second time. Following the second response, the instructor asked students to explain their response reasoning to the entire class including why the correct answer was correct and why the incorrect answers were incorrect. The instructor guided the whole class discussions to ensure that all of the important points were made about each question and cleared up any misconceptions that students might have. Approximately 6-10 questions were covered in each class session. Students could choose to take notes and could use the handouts during class to guide their note-taking which could be done by hand or computer. The last two minutes of the first class session of the week was reserved for students to complete the participation form. The second class session of the week included the same mini-lecture and question-guided discussions but also included the weekly knowledge check and the participation form during the last 10 minutes of class.

After-class learning activity. Students were expected to complete and submit the 10-question handout (i.e., self-explanations) they were given at the beginning of the week. All submissions were via Blackboard™ and due by the end of the instructional week (i.e., Sunday, 11:59 pm).

Traditional instructional model. The traditional instructional model in this work is described using pre-class, in-class, and after-class learning activities. The procedure for each segment is described further below.

Pre-class learning activity. Students in the traditional model were expected to complete the same one to two readings to prepare for the in-class segment as students in the flipped classroom. See Appendix A for Course Schedule.

In-class learning activity. The first class session of each week began with the instructor reviewing the knowledge check from the prior week (beginning with the third class meeting). Afterward, the instructor provided direct instruction on new concepts supported by a projected slide show (e.g., Google Slides presentation). At the end of the lecture, the instructor showed the video that students in the flipped section engaged with during their EDpuzzle™ learning activity. The instructor in the traditional section occasionally asked the class a question and would call on one student to respond. Questions were not planned in advance and occurred at random points throughout the lectures. Students were allowed to raise their hand and ask questions when they didn’t understand a concept. Students could choose to take notes and could use the handouts during class to guide their note-taking which could be done by hand or computer. During the last two minutes of the first class meeting each week, students completed the participation form to rate their engagement, preparation, and participation for that class period. The second class session of each week included a lecture, the video, and the knowledge check (i.e., practice test) followed by the participation form during the last 10 minutes of class.

After-class learning activity. Students were expected to complete and submit the 10-question handout (i.e., self-explanations) they were given at the beginning of the week. All submissions were via Blackboard™ and due by the end of the instructional week (i.e., Sunday, 11:59 pm).

Table 3

Activities completed before, during, and after class by instructional model.

Condition	Pre-class Activities	In-class Activities	Post-class Activities
Flipped	Reading(s) EDpuzzle™ interactive video	Mini-lecture Interactive Peer Learning using multiple-choice questions to facilitate discussion (PollEverywhere™)	Complete and submit handout via Blackboard™
Traditional	Reading(s)	Lecture Supporting video* Reading quiz	Complete and submit handout via Blackboard™

**Note*. The supporting video is the same video that was used in the EDpuzzle for the flipped model

Results

The results of Experiment 1 are below. Learning measures are discussed first and student's reported engagement, preparation, and participation are discussed next.

Learning Measures

The first step in data analysis was to evaluate the descriptive statistics for knowledge checks and exams by condition. See Table 4 for knowledge checks and Table 5 for exams.

Table 4

Descriptive statistics for knowledge checks 1-15 by condition

	Flipped			Traditional		
Test	Mean(sd)	se	Median (iqr)	Mean(sd)	se	Median (iqr)
KC1	3.31(1.09)	0.223	3.25(1.12)	3.92(1.08)	0.204	4.00(1.62)
KC2	2.44(1.39)	0.284	2.50(3.00)	4.34(0.85)	0.157	5.00(1.00)
KC3	3.83(.929)	0.190	4.00(1.62)	4.21(1.21)	0.225	5.00(1.50)
KC4	4.60(.659)	0.135	5.00(0.63)	3.83(1.50)	0.279	4.00(1.50)
KC5	4.17(1.12)	0.229	4.25(1.00)	3.84(0.67)	0.124	4.00(1.00)
KC6	4.29(.736)	0.150	4.25(1.00)	3.72(0.83)	0.154	4.00(1.00)
KC7	3.77(1.01)	0.206	4.00(1.62)	4.59(0.98)	0.182	5.00(1.00)
KC8	3.19(1.45)	0.269	3.00(2.00)	4.15(1.25)	0.256	5.00(1.50)
KC9	4.14(1.16)	0.215	4.00(1.00)	4.25(0.90)	0.183	5.00(2.00)
KC10	3.83(1.29)	0.240	4.00(1.00)	3.25(1.02)	0.209	3.00(1.12)
KC11	4.45(1.32)	0.248	5.00(0.63)	4.21(0.61)	0.124	4.00(1.00)
KC12	2.93(1.68)	0.313	3.00(2.00)	3.71(1.29)	0.264	4.00(1.62)
KC13	3.88(1.44)	0.268	5.00(2.00)	4.00(1.15)	0.235	4.00(1.62)
KC14	3.79(1.52)	0.282	4.00(1.50)	3.55(1.35)	0.276	4.00(2.31)
KC15	4.00(1.44)	0.267	5.00(1.00)	4.29(1.08)	0.221	4.50(1.00)

Note. Instructional model switched at week 8. Weeks 1-7 flipped model (n=24) and traditional model (n=29). Weeks 8-15 were reversed.

Table 5

Descriptive statistics for exams 1 - 4. Dashed line indicates inversion of instructional model

	Flipped			Traditional		
Test	Mean(sd)	se	Median (iqr)	Mean(sd)	se	Median(iqr)
Exam 1	84.20(7.97)	1.63	84.5(12.5)	87.1(7.46)	1.38	88.00(9.00)
Exam 2	84.20(12.20)	2.50	83.50(17.2)	82.90(18.00)	3.34	86.00(11.00)
Exam 3	84.20(19.00)	3.54	90.00(10.00)	81.00(9.26)	1.89	83.00(10.50)
Exam 4	86.90(19.00)	3.52	92.00(10.00)	87.10(9.05)	1.85	88.00(12.00)

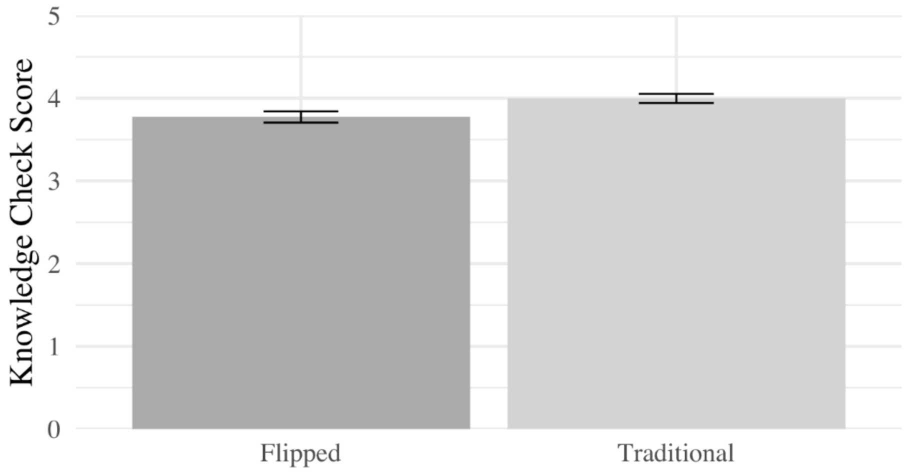
Note. The dotted line indicates inversion of instructional model between classes.

Two mixed random effects models were used to evaluate scores on knowledge checks (Model 1.1) and exams (Model 1.2). Scores for knowledge checks are reported as integers based on total number of points (partial and whole numbers) - treated as a continuous variable. Therefore, a linear, mixed, random effects model as used to evaluate scores with student and knowledge check were set as the random intercepts, with a fully maximal random effect structure implemented to meet model convergence. Results indicate a marginal effect for classroom type, such that students performed marginally higher on knowledge checks in the traditional classroom, relative to the flipped classroom (*ß* = 0.20, *SE* = 0.10, *p* = .051, R^2 = .74; see Figure 2).

Exams are reported as whole numbers with no partial points and treated as a count variable. Therefore, a generalized linear model with Poisson distribution was conducted with student and test set as the random intercepts. Results from the Poisson mixed radon effects model indicated that there was not a significant difference in exam scores as a function of instructional model (*ß* = -0.004, *SE* = 0.01, *p* = .76, R^2 = .72).

Figure 2

Knowledge check scores by condition



Students' Engagement

A repeated measures MANOVA was used to evaluate students' engagement, preparation, and participation as a function of instructional model. Results suggest that each of the student variables together accounts for differences in instructional model (Pillai = .04, $F(1,3) = 1.44$, $p < .001$). Univariate analyses indicated that students reported more preparation (mean difference = .08, $F(1,697) = 3.94$, $p = .05$) and participation (mean difference = .26, $F(1,697) = 29.92$, $p < .001$) in the flipped classroom.

Discussion

Experiment 1 sought to answer whether there was a difference in students' learning outcomes measured by exam and knowledge check scores when in flipped and traditional conditions and whether the instructional format would influence students' reported engagement, preparation, and participation. It was hypothesized that students would have better learning outcomes while in the flipped model due to the deliberate design of learning activities that required active processing and the use of interactive peer learning (Mazur, 1997). It was also hypothesized that students in the flipped classroom format would report higher levels of engagement, preparation, and participation than students in the traditional classroom due to the additional pre-class activity and the increased student discussion in the flipped classroom.

Students' Learning

Differences in individual knowledge check and exam scores between conditions could be a product of the differences in instructors. While the course schedules and the topics covered were the same, teaching styles and content delivery varied slightly. The knowledge checks were taken at the end of a class period and it is possible that the instructors provided different amounts of content that was almost immediately assessed. The traditional classroom lent itself to "spoon feeding" more content to students whereas the flipped classroom was more of a guided overview of the content with the questions students answered during table discussions serving to highlight key ideas (Moffett & Mill, 2014). Students in the flipped classroom may not adequately understand the content if it is not well-aligned to in-class activities and students may not adequately prepare for class in spite of reporting that they did (Peterson, 2016; Setren et al., 2019). Additionally, the course structure in this work included the same evidence-based practices for students in both conditions throughout the semester. It may be that these practices had more of an influence on student learning outcomes than the instructional model or that they may have compensated for deficiencies in one or both classrooms. The use of exams to compare learning within subjects in each instructional model may not have been an accurate measure of knowledge acquisition between models and these results are an indication of students' immediate learning outcomes but do not evaluate whether one model is better than the other for improving retention. These limitations are addressed in Experiment 2, comparing the flipped and traditional instructional models with and without the evidence-based instructional supports (i.e., knowledge checks, handouts) that were included in Experiment 1 using a pre-/post-/delayed post design.

Students' Engagement

The hypothesis that students would report higher levels of engagement, preparation and participation in the flipped classroom than in the traditional classroom was supported. Experiments 2 & 3 continue to investigate students' reported engagement, preparation, and participation.

Experiment 2

Experiment 2 addressed some of the limitations of Experiment 1 by using a pre-/post-/delayed posttest design and manipulating the use of instructional supports in the flipped and traditional classrooms. The goal of Experiment 2 was to investigate (1) the benefit of instructional model (flipped or traditional) on learning over and above the benefit of instructional supports (e.g., practice testing, self-explanations) used in both classrooms; (2) the durability of learning in the flipped and traditional models. Experiment 2 also continued the investigation of the influence of instructional model on students' reported engagement, preparation, and participation.

Research Questions

(1) Does the flipped classroom instructional model benefit learning over and above the use of instructional supports compared to the traditional instructional model?

(2) Does the flipped classroom instructional model lead to more durable learning than the traditional instructional model?

(3) Does students' reported engagement, preparation, and participation differ throughout the semester in the flipped and traditional classrooms?

Hypotheses

(1) Students in the flipped condition would have increased learning outcomes compared to the traditional condition, over and above the benefit of instructional supports at immediate and delayed posttest.

(2) Students' in the flipped classroom would have increased retention at delayed posttest compared to students in the traditional classroom.

(3) Students' reported engagement, preparation, and participation would be higher in the flipped condition than the traditional condition.

Method

The design, participants, materials, measures, and procedure for Experiment 2 are described below.

Design

The semester was divided in half for the purpose of measuring the change in learning outcomes in each condition, but instructors maintained the same instructional model for both halves of the semester. The manipulation for Experiment 2 was the removal of the instructional supports (e.g., knowledge checks, handouts) that were provided to students in Experiment 1 (See Table 8). In other words, students were left to choose their own study strategies for the first half of the semester and given instructional supports in the form of the handout assignments (i.e., self-explanation) and knowledge check assessments (i.e., practice tests) that were completed for a grade in the second half of the semester.

Table 6

Design for Experiment 2

Instructor	First 8 weeks	Second 8 weeks
A - Flipped	No Instructional Supports	With Instructional Supports
B - Traditional	No Instructional Supports	With Instructional Supports

Participants

Fifty-six undergraduate students at a midsize Midwestern university completed two sections of Educational Psychology in the Spring 2019 semester; 28 students with each instructor (75% Female; 89% Caucasian, 5% Black, 6% Other). Participants received one research credit for consenting to share their data.

Materials

The same materials were used for both instructional models except where noted, each is discussed below.

Readings. The readings used in Experiment 2 were the same as Experiment 1 (See Appendix A for course schedule).

Handouts. The handouts used in Experiment 2 were the same as those in Experiment 1. However, students were only given the handout assignments in the second eight weeks of the semester as they were an instructional support that was part of the manipulation. The timeline for receiving the handouts and submitting them and the instructions for completing them were the same as Experiment 1.

EDpuzzles. Edpuzzles were used the same in Experiment 2 as in Experiment 1.

Measures

The items used to measure student learning and student engagement are described below.

Student learning measures. Student learning was measured using a pre-/post-/delayed-post design to compare learning outcomes between instructional models as well as within-subjects for each half of the semester (i.e., with and without instructional supports). Students' learning was also measured in the second half of the semester using the knowledge checks (i.e., practice tests). Each of these is described below.

Pretests. Students were given a 25-question, multiple choice pretest on the first day of each half of the semester covering content that would be taught during each half of the semester. Students were told to do their best but that they would receive full credit for completing it.

Posttests. Students were given a 50-question, multiple choice posttest at the end of each half of the semester consisting of the 25 questions on the pretest for that semester half and 25 questions that were not previously seen on knowledge checks or the pretest. Students were told that the posttest counted as an exam and that they should prepare for it as they would any exam in any typical college class.

Delayed posttest. Students who completed the surveys at all three time points and both pre- and posttests were invited to take a delayed posttest 30 days after the semester ended and offered $10 if they did so. The delayed posttest consisted of the second 25 questions from each of the posttests given during the semester. In other words, the delayed posttest was made up of questions that had not been seen on the pretest but had been seen posttests 1 and 2.

Knowledge checks. The knowledge checks in Experiment 2 were the same as those in Experiment 1. Students were only given knowledge checks in the second half of the semester because they were considered a practice test, one of the instructional supports removed in the first half of the semester as part of the manipulation.

Student engagement measures. Students' reported their *Engagement, Preparation,* and *Participation* in Experiment 2 using the same form as in Experiment 1. (See Appendix D)

Procedure

The procedure for the flipped and traditional classrooms are described below.

Flipped instructional model. The flipped instructional model in this work is described in segments for pre-class, in-class, and after-class learning activities. The procedure for each segment is described further below. See Table 9 for a comparison.

Pre-class learning activity. The pre-class learning activity was consistent throughout the semester. Students in the flipped classroom were introduced to new concepts using one or two readings and a corresponding EDpuzzle™ video for each class meeting. Students were expected to complete these activities prior to the deadline, 15 minutes before class. See Appendix A for Course Schedule.

In-class learning activity. The instructor started each class meeting with an 8-15 minute lecture covering the most important points from the pre-class activity for that class period. Following the mini-lecture, the instructor used the same procedure as in Experiment 1 projecting multiple choice questions generated in Poll Everywhere™ to guide student discussion.

Beginning in week eight, students were given a knowledge check (i.e., practice test) at the end of the second class period each week. Subsequently, the first class period of week nine began with a review of the knowledge check and then the aforementioned interactive peer learning activity for the rest of the semester.

After-class learning activity. During the first eight weeks of the semester, students did not have any assigned after-class learning activity. They were free to choose any study method that they felt would prepare them for the posttest. During the second eight weeks, the students were assigned a handout (i.e. self-explanations) to complete and submit by the end of the week.

Traditional instructional model. The traditional instructional model in this work is also described in segments for pre-class, in-class, and after-class learning activities. Each segment is described further below.

Pre-class learning activity. Students in the traditional model were expected to complete the same one to two readings to prepare for the in-class segment as students in the flipped classroom. See Appendix A for Course Schedule.

In-class learning activity. The first eight weeks consisted of a lecture with a supporting presentation and the video viewed by students in the flipped classroom without questions. Occasionally the instructor asked questions to which one student was selected to respond.

During the second eight weeks, students were given a knowledge check (i.e., practice test) at the end of the second class period each week. Subsequently, the first class period of the following week began with a review of the knowledge check and then the aforementioned lecture and video viewing.

After-class learning activity. Students did not have any assigned after-class activity during the first eight weeks of the semester. They were free to choose any study method that they felt would prepare them for the posttest. Students were assigned a handout (i.e. self-explanations) to complete and submit by the end of the instructional week beginning in week 8. See Appendix B for an example handout.

Table 7

Activities completed by segment and instructional model during the first and second halves of the semester. ***Bold*** *indicates differences.*

Condition	Pre-class Activities	In-class Activities	Post-class Activities
Flipped (1)	Reading(s) Interactive video	Mini-lecture Interactive Peer Learning using multiple-choice questions to facilitate discussion (PollEverywhere™) discussion	No assignment
Traditional (1)	Reading(s)	Lecture Supporting video	No assignment
Flipped (2)	Reading(s) Interactive video	**Knowledge check review** Mini-lecture Interactive Peer Learning using multiple-choice questions to facilitate discussion (PollEverywhere™) **Weekly Knowledge Check**	**Complete and submit handout**
Traditional (2)	Reading(s)	**Knowledge check review** Lecture Supporting video **Weekly Knowledge Check**	**Complete and submit handout**

Results

The results of Experiment 2 are below. Learning measures are discussed first, followed by students' reported engagement.

Learning Measures

Descriptive statistics for knowledge checks and pre-/post-/ and delayed post are below in Tables 8 and 9, respectively.

Table 8

Descriptive statistics for knowledge checks 8 - 15 by condition

	Flipped (n=28)			Traditional (n=28)		
Test	Mean(sd)	se	Median (iqr)	Mean(sd)	se	Median(iqr)
KC8	3.71(0.99)	.187	3.75(2.00)	3.29(1.81)	.342	4.00(2.62)
KC9	3.56(0.88)	.169	3.88(1.00)	3.64(1.43)	.270	4.00(1.25)
KC10	3.42(0.95)	.179	3.50(1.00)	3.86(1.50)	.283	4.25(1.62)
KC11	4.67(0.62)	.118	5.00(.312)	4.52(0.83)	.159	5.00(1.00)
KC12	3.84(0.72)	.136	4.00(1.25)	3.93(1.37)	.259	4.00(1.50)
KC13	3.56(1.14)	.215	3.75(1.00)	3.82(0.95)	.179	4.00(2.00)
KC14	3.64(1.02)	.193	4.00(1.00)	3.04(1.73)	.327	3.50(2.25)
KC15	4.07(1.07)	.202	4.00(1.62)	4.32(1.19)	.225	5.00(1.00)

Note. Knowledge checks were only given in weeks 8-15

Table 9

Descriptive statistics for pretests, posttests, and delayed posttest by instructional model

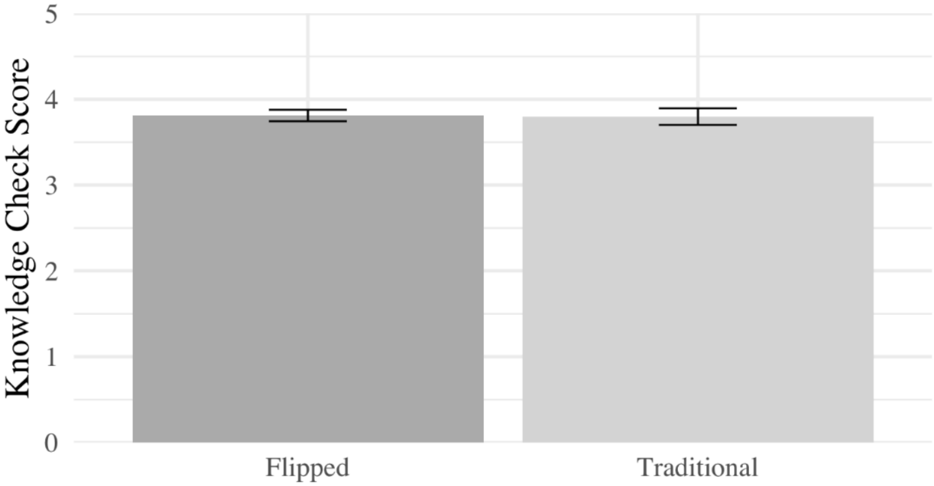
	Flipped (n=28)			Traditional (n=28)		
Test	Mean(sd)	se	Median (iqr)	Mean(sd)	se	Median(iqr)
Pretest 1	11.10(2.60)	.492	11.00(2.25)	10.90(2.87)	.542	11.00(3.25)
Pretest 2	14.30(3.05)	.577	14.50(4.25)	12.80(3.84)	.725	14.00(5.25)
Posttest 1	37.00(4.08)	.771	37.00(5.25)	36.50(6.88)	1.30	38.50(6.50)
Posttest 2	38.10(4.23)	.800	39.50(6.00)	36.30(4.94)	.933	36.50(7.00)
Delayed Post	31.80(7.62)	2.04	33.00(5.25)	29.90(9.58)	3.62	33.00(12.00)

Note. Results on the delayed posttest were calculated with Flipped (n=14) and Traditional (n=7) scores.

Two mixed random effects models were used to evaluate scores on knowledge checks (Model 2.1) and posttests (Model 2.2). Scores for knowledge checks are reported as integers based on total number of points (partial and whole numbers) and therefore were treated as a continuous variable. Because knowledge check scores are integers, a linear mixed random effects model was conducted with student and knowledge check set as the random intercepts. A fully maximal random effect structure was implemented to meet model convergence. Results indicate that there was not an effect of instructional model on students' knowledge check scores. (β = -0.01, SE = 0.17, p = .94, R^2 = .24; see Figure 3).

Figure 3

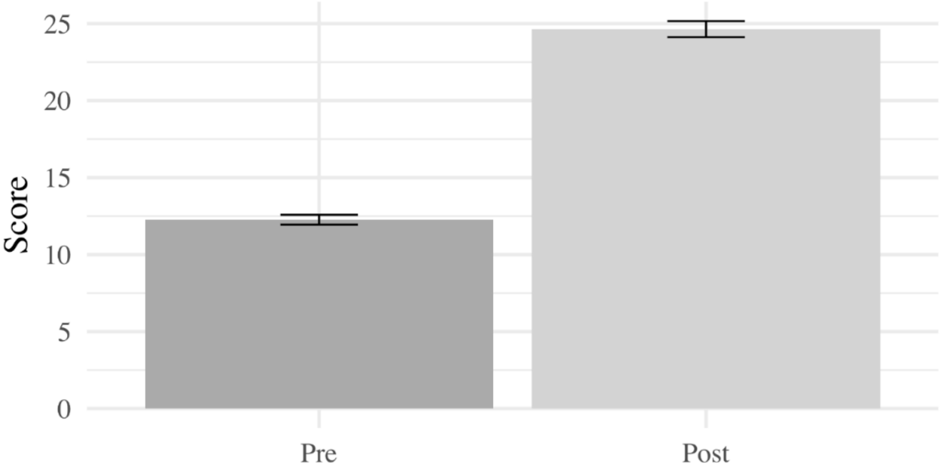
Knowledge check scores by condition



Posttests are reported as whole numbers with no partial points and treated as a count variable. Therefore, a generalized linear model with Poisson distribution was conducted with student and test as random effects and a condition by test type (Pre-Post) interaction was the fixed effect. Results from Poisson mixed random effects model indicated no significant difference between posttest scores as a function of classroom model (*ß* = -0.304, *SE* = 0.04, *p* = .350, R^2 = .79). However, pretest score was a significant predictor of posttest score in both classrooms (*ß* = -0.628, *SE* = 0.24, *p* = .008, R^2 = .79; See Figure 4).

Figure 4

Pretest and posttest scores



The durability of students' learning as a factor of instructional model was evaluated using scores on a 30-day delayed posttest. Scores were reported as whole numbers with no partial points and treated as a count variable. Therefore, a generalized linear model with a Poisson distribution was conducted with student as the random intercept. Results from the mixed random effect model indicated no significant difference in delayed posttest scores as a function of classroom model ($\beta = -0.067$, $SE = 0.12$, $p = .585$, $R^2 = .55$).

Students' Engagement

A repeated measures MANOVA was used to evaluate students' reported engagement, preparation, and participation in the flipped and traditional classrooms over the course of the semester. Results suggest that the student variables together do not account for a significant difference in students' reported engagement as a function of classroom model (Pillai = .08, $F(1,3) = 1.47$, $p = 0.233$).

Discussion

Experiment 2 evaluated whether or not there was an effect of model over and above any benefit of using evidence-based practices and whether or not students' in the flipped classroom would have more durable learning than those in the traditional classroom. Experiment 2 also investigated whether students' reported engagement, preparation, and participation would change over time in each instructional model.

Student Learning Outcomes

The influence of condition on learning was evaluated in the second half of the semester by comparing scores on individual knowledge checks and using a mixed model. No significant differences in learning between conditions were found. This experiment also investigated whether the interactive learning that took place in the flipped classroom would lead to increased retention over the passive traditional classroom. Results suggest that there was no difference in students' retention 30 days after the posttest. Therefore, the incorporation of interactive learning may not be a key feature that leads to improved retention. However, it is important to note that only 38% of the students took the delayed posttest making this result difficult to generalize. The lack of difference in student learning between instructional formats and with and without instructional supports may be a result of quality of instruction in both conditions. Another possibility is that students were sufficiently self-motivated to study beyond the requirements of the manipulation because pre-service teachers are required to pass the course.

Student Engagement

Experiment 2 also investigated the influence of model on students' reported engagement, preparation, and participation. Results indicated that there were no differences across conditions nor was there a difference due to the manipulation of adding instructional supports in the second half of the semester. Figure 10 shows that students reported preparation for class was stable throughout the semester suggesting that knowing that they would have quizzes did not change students' perceptions of their preparation. Interestingly, students in the traditional classroom reported higher preparation, though not significantly, than students in the flipped classroom who had more learning activities to complete before each class meeting. Similarly, students in the traditional classroom reported higher participation, though not significantly, than students in the flipped classroom in spite of opportunities to participate being limited. One possibility for the difference in preparation is that students in the flipped classroom decreased their reported preparation if they didn't complete all of the assigned learning activities whereas students in the traditional classroom only had one learning activity (reading) so they were more likely to report that they had done it. It is also possible that the reported ratings are inverse of the expected because students didn't experience both models, therefore they may have been unable to accurately calibrate their rating of how much preparation they did for each class meeting or how much they participated.

One limitation of both Experiments 1 and 2 is that the amount and structure of the learning activities and lecture content were not controlled. Instructors taught the same concepts using similar (not identical) presentation materials, however, the explanations and examples provided in class were ad hoc. Additionally, the instructors may not have covered all content equally (e.g., ran out of time to watch the video) due to time constraints and variability introduced by student questions. Therefore, comparisons between models may not be an accurate reflection of the model but more related to instructor or content delivery differences. Experiment 3 addresses this limitation.

Experiment 3

Experiment 3 expanded on Experiment 2 first by addressing the limitation related to the differences in the design (i.e., presentation slides) and delivery (i.e., lecture structure) of the content. The primary difference was that the content design and delivery was the same in both classrooms in Experiment 3. Additionally, the content was temporally aligned so that students in both classes received the same content in the same order (i.e., reading, lecture, video). Students in the traditional classroom received the same lecture, delivered live, as students in the flipped classroom viewed recorded. Students in the traditional classroom also watched the same supporting video following the in-person lecture during class as students in the flipped classroom (See Table 10). The goals of Experiment 3 were to compare learning outcomes between students in the flipped and traditional classrooms and to investigate students' reported engagement, preparation, and participation. As such, Experiment 3 included two changes to the content, learning activities and assessments in the flipped and traditional instructional models. First, Experiment 3 investigated the use of small group, interactive discussion facilitated by guiding questions in the flipped classroom that may lead to improved learning outcomes at posttest and delayed posttest compared to the traditional classroom model. Students were given a question prompt and asked to take turns generating explanations and providing elaborations, and feedback in their small group (Chi, 2009). Second, students in both sections were given quizzes in each class meeting after exposure to the same amount of content (see Table 13). The research questions that this study sought to answer are presented below.

Research Questions

(1) Does the flipped instructional model improve learning outcomes over the traditional lecture model when using the same instructional materials and temporal alignment?

(2) Does interactive peer learning using facilitated discussion in the flipped classroom benefit long-term retention compared to the traditional instructional model?

(3) Do students' reported engagement, preparation, and participation throughout the semester in the flipped and traditional classrooms when using the same instructional content?

Hypotheses

(1) Students in the traditional classroom will outperform students in the flipped classroom on reading quizzes in each class meeting.

(2) Students in the flipped classroom will outperform students in the traditional classroom on posttests and the delayed posttest.

(3) Students in the flipped classroom will report higher engagement, preparation, and participation than students in the traditional classroom.

Students in the traditional classroom were hypothesized to outperform students in the flipped classroom on the reading quizzes because of the recency of the introduction to new knowledge (Cowan, 2016). Students in the flipped classroom had a longer duration since the introduction to new knowledge thus increasing the retrieval difficulty (Dunlosky et al., 2013). The potential for increased forgetting and more effortful retrieval experienced by students in the flipped classroom was hypothesized to be beneficial to learning over the course of the semester thus resulting in improved performance at posttest and delayed posttest (Dunlosky & Rawson, 2015; Rawson & Dunlosky, 2012). Additionally, it was hypothesized that the interactive generation of explanations in the flipped classroom would lead to better performance and retention than the individual generation of explanations done by students in the traditional classroom (Chi, 2009).

Method

The design, participants, materials, measures, and procedure for Experiment 2 are described below.

Design

The design of Experiment 3 was the same as Experiment 2 without a mid-semester change in manipulation. See Table 10 for learning activity arrangement.

Table 10

Activities completed before, during, and after class by instructional model.

Condition	Pre-class Activities	In-class Activities	Post-class Activities
Flipped	Reading(s) Lecture via video Supporting video	Reading quiz Interactive Peer Learning	Complete and submit handout
Traditional	Reading(s)	Lecture Supporting video Reading quiz	Complete and submit handout

Participants

Ninety-three undergraduate students at a midsize Midwestern university completed two sections of Educational Psychology in the Fall 2019 semester; 48 students in the flipped classroom and 45 in the traditional (80% Female, 1% Nonbinary; 94% Caucasian, 5% Black, 1% Asian). Participants received one research credit for consenting to share their data.

Materials

The materials used in Experiment 3 are described below and were the same for both instructional models.

Readings. Students in both sections were assigned articles and book chapters to read prior to each class meeting (See Appendix K for Course Schedule).

Handouts. The handouts used in Experiment 3 were similar to those used in Experiments 1 & 2 (See Appendix B).

Lecture materials. The lecture materials in Experiment 3 were updated from Experiment 2 so that students saw the same content in both classes. The content used for the lecture was produced on a Google Slides presentation that was converted to a voiceover video lecture for students in the flipped classroom. The same presentation was used in a live lecture in the traditional classroom.

Supporting videos. The video schedule used in Experiment 3 was an updated version of the schedule from Experiments 1 & 2 and did not have embedded questions. (See Appendix G for video titles).

Measures

All measures were course assignments and assigned a point value for completion or a grade.

Student learning measures. Student learning in each model was measured using pre- and posttests for each half of the semester. Retention was evaluated using a 30-day delayed posttest. Student learning was also measured using the reading quizzes taken in class in both classrooms. The pretests, posttests, and delayed posttests were the same as those used in Experiment 2.

Reading quizzes. Knowledge checks in Experiments 1 & 2 were updated in Experiment 3. The reading quizzes administered each class meeting consisted of 10 multiple choice questions taken from the assigned readings.

Student engagement measures. Students' reported their *Engagement, Preparation and Participation* using the same form as in Experiments 1 and 2. (See Appendix D)

Procedure

The procedure for Experiment 3 was similar to Experiment 2. Both classes maintained the same instructional format throughout the semester and there was no mid-semester change in manipulation. Each model is described further below.

Flipped instructional model. The flipped instructional model in Experiment 3 had the same format as Experiment 2.

Pre-class learning activity. The pre-class learning activity structure was the same as Experiments 1 & 2 (See Table 10). The lecture presentation in Experiment 3 included new voice-over-presentations. The video lectures were updated as part of the manipulation comparing the flipped and traditional classrooms using the same materials. See Appendix F for course schedule and Appendix G for video titles.

In-class learning activity. The in-class learning activity was modified in two ways in Experiment 3. The learning activities were temporally aligned and the peer interaction was question-facilitated discussion. Class meetings started with the reading quiz taken on Blackboard™. Students had 10 minutes to complete it and were able to see their score and incorrect questions upon submission.

Next, the instructor reviewed the prior reading quiz calling on one student at a time to provide the correct answer and reasoning for all 10 questions clearing up misconceptions before moving on.

The remainder of the class period was spent on interactive peer learning through facilitated discussion. Students generated responses and explanations for five questions per class period for 5-8 minutes each. Afterward, the instructor called one or more students to share their responses until all of the main ideas were covered. When necessary, the instructor filled in missing information and clarified misunderstandings.

After-class learning activity. The after-class learning activity was the same as Experiments 1 & 2 except that the questions on the handouts were those discussed in class. The purpose for this change was to compare the influence of interactive response generation with independent generation in the traditional classroom.

Traditional instructional model. The traditional instructional model in Experiment 3 had the same structure as Experiments 1 & 2.

Pre-class learning activity. The pre-class activity in the traditional classroom was the same as Experiments 1 & 2. (See Appendix F for Course Schedule)

In-class learning activity. The in-class learning activity in Experiment 3 was modified to adhere to the same temporal alignment and content presentation as the flipped classroom. The instructor started class with a review of the prior reading quiz, calling on one student at a time to provide the correct answer and reasoning for all 10 questions. The instructor clarified any misconceptions before moving on.

The lecture presentation was nearly identical as the flipped lecture video as the instructor followed the voiceover script closely. In other words, the content was nearly identical in the flipped and traditional classrooms including the length of the lecture. Following the lecture, the instructor played the supporting video watched by students in the flipped classroom.

During the last 10 minutes of each class meeting, students took the reading quiz on Blackboard™. They were able to see their score and which questions they got wrong upon submission. Afterward, students completed the participation form and were free to leave.

After-class learning activity. Students were given handouts after the second class meeting each week to complete and submit via Blackboard™ by Sunday at 11:59 pm. See Appendix B.

Results

The results of Experiment 3 are discussed below. Learning measures are discussed first followed by results of students' reported engagement, preparation, and participation.

Learning Measures

Descriptive statistics for reading quizzes and pre-/post-/delayed post are shown in Tables 11 and 12, respectively.

Table 11

Descriptive statistics for reading quizzes 1 - 23 by condition

	Flipped (n=48)			Traditional (n=45)		
Test	Mean(sd)	se	Median (iqr)	Mean(sd)	se	Median (iqr)
RQ1	5.88(1.55)	0.224	6.00(2.00)	6.40(2.00)	0.309	7.00(3.00)
RQ2	4.88(1.90)	0.271	5.00(3.00)	4.35(1.78)	0.272	5.00(3.00)
RQ3	5.33(1.66)	0.238	5.00(3.00)	7.60(1.36)	0.208	8.00(2.00)
RQ4	7.88(1.63)	0.233	8.00(2.00)	8.24(1.43)	0.218	8.00(1.50)
RQ5	5.69(2.03)	0.290	6.00(2.00)	6.24(2.26)	0.344	7.00(1.50)
RQ6	5.71(1.86)	0.266	6.00(3.00)	6.98(2.15)	0.329	7.00(2.50)
RQ7	5.67(1.93)	0.276	5.00(3.00)	6.63(2.10)	0.321	7.00(2.00)
RQ8	6.88(1.56)	0.223	7.00(2.00)	7.37(1.92)	0.292	8.00(3.00)
RQ9	5.38(1.73)	0.250	6.00(1.50)	6.54(1.47)	0.224	5.00(1.00)
RQ10	5.53(1.67)	0.239	6.00(3.00)	6.37(1.94)	0.296	7.00(3.00)
RQ11	5.00(2.04)	0.292	5.00(3.00)	5.74(2.06)	0.314	6.00(2.00)

RQ12	6.57(1.79)	0.256	7.00(2.00)	6.42(2.51)	0.383	7.00(2.50)
RQ13	6.82(1.87)	0.267	7.00(2.00)	8.21(1.74)	0.265	9.00(1.00)
RQ14	5.00(1.43)	0.204	5.00(2.00)	5.46(1.35)	0.206	6.00(1.00)
RQ15	5.61(1.62)	0.231	6.00(2.00)	6.33(1.90)	0.290	6.00(3.00)
RQ16	7.00(2.51)	0.358	8.00(3.00)	8.35(2.20)	0.336	9.00(1.5)
RQ17	8.45(1.49)	0.212	9.00(1.00)	8.26(1.89)	0.291	9.00(2.00)
RQ18	5.90(1.98)	0.283	6.00(2.00)	7.72(2.21)	0.337	8.00(2.00)
RQ19	5.82(1.55)	0.221	6.00(2.00)	7.05(1.93)	0.294	7.00(1.00)
RQ20	7.00(2.03)	0.293	7.00(2.00)	7.74(2.19)	0.337	8.00(2.00)
RQ21	6.42(1.11)	0.355	6.00(1.00)	8.57(1.85)	0.160	9.00(2.00)
RQ22	6.25(2.46)	0.355	6.50(3.00)	7.00(3.06)	0.467	8.00(2.00)
RQ23	6.49(2.44)	0.349	7.00(2.00)	8.31(1.14)	0.175	8.00(1.00)

Table 12

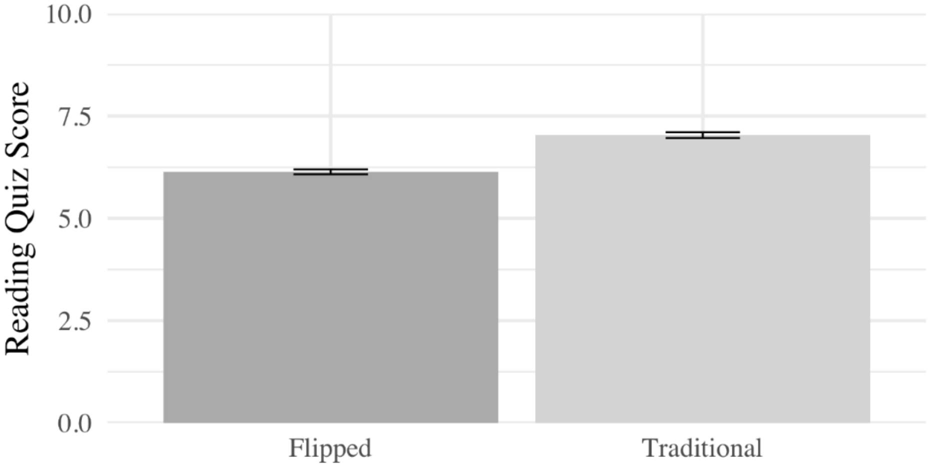
Descriptive statistics for pretests, posttests, and delayed posttest by instructional model

	Flipped (n=48)			Traditional (n=45)		
Test	Mean(sd)	se	Median (iqr)	Mean(sd)	se	Median(iqr)
Pretest 1	11.10(2.50)	.358	11.00(3.00)	10.60(3.27)	.498	11.00(4.00)
Pretest 2	13.10(4.29)	.613	14.00(4.00)	13.70(3.56)	.542	13.00(3.56)
Posttest 1	37.70(5.50)	.786	38.00(8.00)	38.60(3.88)	.592	39.00(5.50)
Posttest 2	37.00(4.93)	.704	38.00(8.00)	38.80(4.34)	.661	39.00(6.00)
Delayed Post	29.70(10.60)	2.44	32.00(15.50)	30.90(6.92)	2.44	31.50(7.25)

Two mixed random effects models were used to evaluate scores on reading quizzes (Model 3.1) and posttests (Model 3.2). Scores for reading quizzes are reported as integers based on total number of points (partial and whole numbers) and were therefore treated as a continuous variable. Because reading quiz scores are integers, a linear mixed random effects model was conducted with student and reading quiz set as the random intercepts. A fully maximal random effect structure was implemented to meet model convergence. Results indicate that there was a significant effect of instructional model on students' reading quiz scores favoring the traditional classroom. ($\beta = 0.90$, $SE = 0.21$, $p < .001$, $R^2 = .40$; see Figure 5).

Figure 5

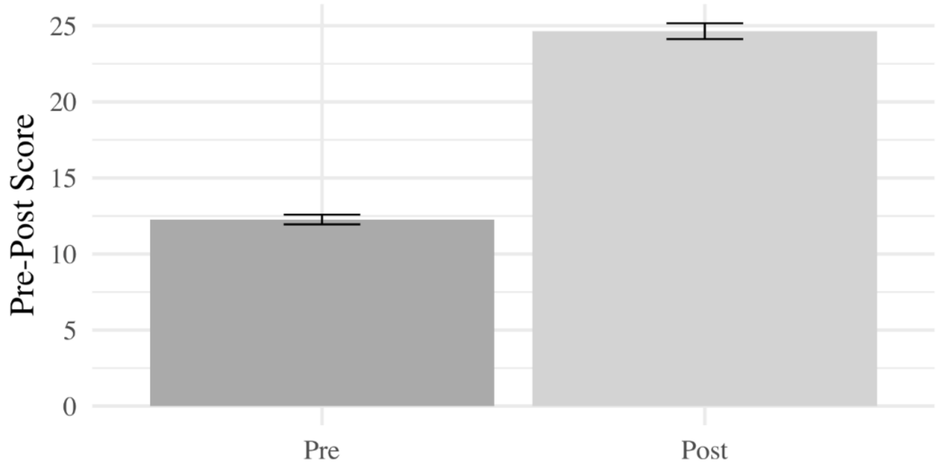
Reading quiz scores by condition



Posttests and delayed posttest scores are reported as whole numbers with no partial points and treated as a count variable. Therefore, a generalized linear model with Poisson distribution was conducted with student and test as random effects and a condition by test type (Pre-Post) interaction as the fixed effect. Results from the Poisson mixed random effects model indicated no significant difference between posttest scores as a function of classroom model ($\beta = 0.041$, $SE = 0.02$, $p = .097$, $R^2 = .80$). However, pretest score was a significant predictor of posttest score ($\beta = -0.671$, $SE = 0.24$, $p = .005$, $R^2 = .79$; See Figure 6). Results from the Poisson mixed random effects model for the delayed posttest indicated no significant difference between scores as a function of classroom model ($ß = 0.056$, $SE = 0.14$, $p = .695$, $R^2 = .66$).

Figure 6

Pretest and posttest scores



Students' Engagement

A repeated measures MANOVA was used to evaluate student engagement, preparation, and participation as a function of instructional model when using the same instructional content. Results suggest that each of the student variables together does not account for differences in students' reported engagement as a function of instructional model (Pillai = .03, $F(1, 3) = 1.007$, $p < 0.394$).

Discussion

The goal of Experiment 3 was to evaluate differences in students' learning and retention in the flipped and traditional instructional formats when using the same instructional materials and the same temporal arrangement (i.e., same content and assessments in the same order). Variations in content design and delivery was a limitation of Experiments 1 and 2 that was addressed in Experiment 3. This experiment also compared differences in students' reported engagement, preparation, and participation when using the same content and order. The time spent on learning activities in Experiments 1 & 2 was not controlled for. Experiment 3 was designed to address this limitation by attempting to equalize time spent on learning tasks for the flipped and traditional classes. The design of Experiment 3 was intended to parse out the influence on learning and student engagement of the flipped instructional model features, effortful retrieval with the reading quizzes and interactive cognitive processing with the interactive group discussions (Dunlosky & Rawson, 2015; Chi, 2009).

Students' learning outcomes were compared using reading quizzes and a pre-/post-/delayed post design. The hypothesis that students in the traditional classroom would perform better on reading quizzes was supported. Students in the traditional classroom did perform better on the reading quizzes than students in the flipped classroom. This outcome was hypothesized because the reading quizzes were given immediately following content delivery in the traditional classroom, thus the expectation was that new knowledge would have been active in students' short-term memory (Cowan, 2016). Students in the flipped classroom had a greater delay in exposure to material having engaged with initial instruction prior to class. The longer duration between the introduction of the information and assessment of knowledge in the flipped classroom was expected to allow more forgetting and increased retrieval difficulty for students in the flipped classroom potentially resulting in lower reading quiz scores.

The second hypothesis stems from the known benefits of effortful retrieval introduced by space between instruction and recall and the interaction with peers in class as both have been shown to benefit learning and retention (Chi, 2009; Dunlosky et al., 2013). Based on this evidence, it was expected that students in the flipped classroom would outperform students in the traditional classroom at posttests and the delayed posttest. However, this hypothesis was not supported. Results showed no difference between the two classes on any of the posttests. The similarity in scores between classes may stem from a lack of difference in instructional benefit between the models. Students in both classrooms were given the same content in the same order in an attempt to parse out the differences afforded by the effortful retrieval (Rawson & Dunlosky, 2012) and interaction with peers (Chi & Wiley, 2014) in the flipped classroom. The results of this experiment suggest that these features may not be what has led to improved learning outcomes in prior flipped classroom research. Another possible reason for the lack of difference between the classes is the importance of the educational psychology class used in this work as students are required to pass the course to complete their major. In other words, students' study behaviors independent of the assigned activities, could have minimized or negated the benefit of either classroom format.

The third hypothesis, that students' reported engagement, preparation, and participation would differ between the flipped and traditional classroom was not supported. There were no significant differences between their self-reported levels of engagement. Similar to Experiment 2, students' behaviors, independent of the design, influenced reported engagement. A second potential reason for the similarity in reporting is that students only experienced one instructional model so they didn't have a baseline to compare with as students did in Experiment 1.

Students in this experiment received nearly identical content and had no difference in learning outcomes. It's worth noting that the quality of the content students are provided and the levels of engagement elicited by assigned activities may be the more important features of instructional design. Instructors who are planning to use the flipped or traditional instructional formats should be intentional in their design choices, using evidence to inform decisions (Jensen et al., 2015; Peterson, 2016; Ryan & Reid, 2016).

CHAPTER IV

GENERAL DISCUSSION

The goals of this work were to investigate differences in student learning occurring in the flipped and traditional classrooms and to evaluate students' reported levels of engagement in the learning process. Analyses of learning measures in all three experiments showed no effect of the type of instructional model on students' learning outcomes. There was also no difference in learning outcomes when manipulating the assigned learning activities in both classrooms in Experiment 2. These results suggest that differences in learning between formats found in prior literature may stem from additional factors that were not included in this work.

Similarly, a difference in students' reported engagement, preparation, and participation was only seen when students experienced both instructional formats in Experiment 1. One possible reason there was not a discernable difference between conditions in Experiments 2 and 3 is that students did not have a baseline or alternative condition with which to compare their engagement levels. The reported levels of engagement, preparation and participation were near ceiling in both instructional formats in all three experiments suggesting that students believed that they were fully engaged, prepared, and participating in class regardless of format. Further work is needed to evaluate the accuracy of their judgements and the relation to learning outcomes.

The experiments in this work were not without limitations. First, the course used in these experiments is required for all education majors who must pass the course to continue with their program of study. This may have led students to study more, independent of the required course activities or instructional format. Students did not report on additional study activities nor were they instructed to limit their strategy use. Second, the content of the course is focused on how students learn and how to improve student learning, essentially telling students how to become more effective learners. Students using the evidence-based practices learned in their coursework may have masked the benefit of one instructional format over the other at both the immediate posttests and after the 30-day delay. Third, these data were collected over the course of the entire semester for all three experiments. The fluctuations that occur in students' daily lives may have influenced their study patterns and engagement behaviors throughout the course of the semester.

The results of this work add to the mix of results in the extant literature further suggesting that more research is needed to identify the feature(s) or learning activities included in the flipped classroom that may lead to improved learning outcomes over more traditional classrooms. One complication in research comparing the flipped and traditional classrooms is the lack of a single definition or framework of the flipped classroom (Abeysekera & Dawson, 2015; Bishop & Verleger, 2013; Giannokos et al., 2018). At the most basic level, the flipped classroom inverts the initial instruction (lecture) and the application of knowledge (i.e., problem solving or generating question responses) from the traditional classroom structure (Bergmann & Sams, 2012; Bishop & Verleger, 2013; Lage et al., 2000). The question remains as to whether any benefit from the flipped format is due to the order of learning activities (inverting in-class activities) or if there is some other factor or factors influencing learning outcomes. As such, one method for evaluating any differences in learning at the level of this basic inversion is suggested below. The suggested method is derived from the results of the present experiments in light of the main results of the current empirical literature.

The suggested experiment defines the flipped format as lecture occurring prior to class via video and application of knowledge occurring collaboratively in class. The traditional format is the inverse, a live lecture in class with application of knowledge occurring after class, independently. A switching replications design in two sections of a course taught by the same instructor to evaluate within and between subject learning outcomes in both formats is recommended.

Literature comparing the flipped and traditional instructional formats cites the passive nature of the traditional classroom as being a feature for which the flipped classroom is a solution to improve learning (Baepler et al., 2014; Freeman et al., 2014; Jensen et al., 2015). Collaborative problem-solving and question-response generation are cited in the literature as activities that make the flipped classroom an active learning environment (Bergmann & Sams, 2012; Schell & Mazur, 2015). The difference between *constructive* individual response generation in the traditional format and *interactive* group response generation in the flipped classroom may be a defining factor in learning differences (Chi, 2009; Menekse & Chi, 2018).

Evaluating the influence of instructional format on learning should minimize confounds and attempt to equalize the levels of processing elicited by initial instruction (Chi & Wiley, 2014) and include evaluation of students' attention to that instruction. Similarly, the knowledge application phase in flipped and traditional formats should compare interactive and independent generative activities, respectively. Below are recommendations to accomplish these goals.

The first recommendation is that the same instructor and content are used in both formats to minimize confounds introduced by differences in content design (i.e., presentation slides) or delivery (i.e., different examples or emphases). Also, the course would not be related to learning sciences or pedagogy to reduce confounds introduced by knowledge of or instruction on effective learning strategies. A voice over slide presentation introducing new information would be used for the asynchronous video lecture in the flipped classroom. The same presentation would be used in the synchronously delivered lecture in the traditional classroom. Questions would be embedded in both presentations requiring students to respond with a mobile device connected via login to the presentation. This feature serves to elicit the same levels of cognitive processing and attention to the in both formats. In addition to equalizing engagement and processing, response accuracy would be included in analyses comparing models and as a covariate in posttest comparisons.

The second suggestion is that students' attention and breaks in attention to instruction would be captured in both formats. Research suggests that mind wandering and/or media multitasking frequently occur in the traditional classroom and students disengage from the lecture (Sana et al., 2013; Wammes & Smilek, 2017; Wammes et al., 2019). Similarly, evidence suggests that there are also frequent occurrences of mind wandering or task unrelated thoughts while viewing video lectures that may negatively impact learning (Hollis & Was, 2016; Kane et al., 2017; Zhang et al., 2020). Attention data in the traditional classroom would be captured via video recordings and coded for students' observed behaviors (e.g., gazing out the window, looking at phone) and through logging their Internet IP address connections on computer and mobile device. In the case of video lectures, eye tracking during video viewing would be used to capture changes in attention at and between question probes (Zhang et al., 2020). Media multitasking would be captured using the same method as the traditional classroom. Evaluating these behaviors in both conditions can provide a better understanding of the influence of students' on- and off-task behavior during synchronous and asynchronous initial instruction. Furthermore, understanding how students engage and disengage with initial instruction when it includes active learning opportunities (i.e., question responses) and the relation to learning outcomes can provide insight into more effective instructional design.

The third suggestion is to eliminate additional learning tasks and strategies known to benefit students' learning. No pre-instruction reading would be assigned to prepare for traditional in-class lectures or flipped video lectures. Students would also be instructed to engage with the lecture but not to do any type of note taking during the lectures. These two oft used tasks may negate differences between the two instructional formats.

The knowledge application phase will occur 24 hours after the initial instruction in both formats. Students in the flipped format will work in small groups to solve problems and answer questions in class the day after initial instruction. Students in the traditional format will complete the same problems the day after the in-class instruction. Standardizing the time between initial instruction and knowledge application should reduce the confound of differences in time spent on learning tasks and the interstudy gap. Constraining the knowledge application this way should help identify differences between individual and interactive generation activities in the traditional and flipped formats, respectively.

During the second half of the switching replications design, students reverse conditions and learn new content with equivalent difficulty. Learning outcomes in both units would be measured with pre-, post-, and 30-day delayed posttests from the end of each unit of instruction. Each unit would last two weeks and include four class meetings to provide enough data points for comparison and reduce measurement error due to factors outside of the control of the experimenter such as students' schedules and time management conflicts.

In sum, the purpose of this design is to compare two elements of the flipped and traditional instructional formats, initial instruction and knowledge application, and the influence on students' learning outcomes. First, the use of identical, active-learning materials addresses the question of whether or not the primary difference that influences learning in the two formats is the amount of active processing elicited from students during initial instruction. Measures of students' attention during instruction are intended to evaluate whether or not the level of off-task behaviors (i.e., mind wandering and media multitasking) mediates learning outcome. Second, the impact of applying knowledge in collaboration with peers in the flipped classroom is compared with the application of knowledge independently. Removing any other learning tasks that may be assigned or students may choose to engage in can reduce the factors that influence learning outcomes to these specific design features of the flipped and traditional formats.

Conclusion

Recently there has been a push for classrooms to include more active learning opportunities due to the concern that lecture-based instruction is passive and may not be sufficient for learners to encode and store new knowledge (Baepler et al., 2014; Freeman et al., 2014; Prince, 2004; Prunuske et al., 2012). The call to make classrooms active and engaging learning environments is leading to an increase in implementations of the flipped classroom (Abeysekera & Dawson, 2015; Berrett, 2012; O'Flaherty & Phillips, 2015; Tucker, 2012). However, research comparing the flipped format and traditional formats has yet to provide clear evidence suggesting what design elements (e.g., voice over slides, interactive group work) are most beneficial for engaging students and improving learning outcomes. If the suggested research described above demonstrates equivalent learning outcomes in both instructional formats, this would suggest that the inversion of instruction and knowledge application may be less important than ensuring that students are engaged with the learning activities.

A large body of research identifies several evidence-based strategies known to improve learning (e.g., practice testing, interactive peer learning; Chi, 2009; Dunlosky et al., 2013) that may be implemented in any instructional format. A well-designed classroom (flipped or traditional) can and should include these strategies. The flipped classroom may lead to improved learning outcomes due to the incorporation of these practices inherent in the inverted design. Meaningful engagement with initial instruction is a feature that can influence learning and as such, myriad tools are available for instructors to include engagement prompts (e.g., formative assessment questions) during initial instruction to create a more active learning environment. A key difference in the flipped classroom occurs in the knowledge application phase where students often work collaboratively and have access to the instructor when they don't understand. Three features in this type of knowledge application phase are likely to lead to the benefit to learning seen in the flipped classroom. First, students who know that they will have to use initial instruction to solve problems with their peers are incentivized to adequately prepare prior to class with whatever mode of instruction is provided (Nokes-Malach et al., 2019; Rotgans & Schmidt, 2012). Conversely, students in the traditional classroom may not have the same incentive when listening to the instructor's lecture as initial instruction. Second, students are co-generating responses and extending their knowledge through elaborations with peers. The interactive learning activities provide the benefit of receiving feedback from peers as well as filling in an gaps in knowledge from elaborations from peers (Menekse & Chi, 2018; Roscoe & Chi, 2008; Nokes-Malach et al., 2019). The traditional classroom format may not include collaborative activities in class or during knowledge application (Gardiner, 1998; Lage et al., 2000). Students may or may not seek collaboration on their own, in spite of the evidence of their benefit (Dunlosky et al., 2013; Miyatsu et al., 2018). Finally, the students in the flipped classroom have access to the instructor when they do not understand new information and peers cannot help. Instructors can clarify misconceptions and students can immediately apply their new understanding of the concept while solving problems with the assistance of instructor or peers.

Students in the traditional format often complete knowledge application as homework, following initial instruction and if they are confused, they may develop misconceptions or simply give up (Bergmann & Sams, 2012).

Instruction that actively engages students and maintains attention is more beneficial than passive instruction that students may disengage from (Chi, 2009: Freeman et al., 2014; Jensen et al., 2015). The inverted design of the flipped classroom includes the aforementioned features that may lead to improved learning outcomes. However, it is worth noting that any format of instruction can intentionally incorporate opportunities for active learning and evidence-based strategies that will lead to improved student learning; the flipped classroom is an example of one.

APPENDICES

APPENDIX A

COURSE SCHEDULE EXPERIMENTS 1 AND 2

Appendix A

Course Schedule Experiments 1 and 2

	Date	Questions to Consider	Pre-Class Readings/Assignments	Class Activities
Week 1: Introduction to Educational Psychology	Tuesday: 1/16	What is Educational Psychology?	Mayer Chapter 1*	Review Syllabus Introductions
	Thursday: 1/18	What is Working Memory? What is Cognitive Load Theory?	Gathercole (2008) Cognitive Load Theory EdPuzzle #1	Interactive Learning Groups KC #1
Week 2: Cognitive Development Theories	Tuesday: 1/23	Is learning a social or individual process?	Newcombe (2013) EdPuzzle #2	Interactive Learning Groups
	Thursday: 1/25	Are we born with innate knowledge?	Newcombe (2013) EdPuzzle #3	Interactive Learning Groups KC #2
Week 3: Learning Theories	Tuesday: 1/30	What is classical conditioning?	Eggen & Kauchak Chapter 9 EdPuzzle #4	Interactive Learning Groups
	Thursday: 2/1	What is operant conditioning? Social Cognitive Theory revisited	Review Readings from weeks 1-3 EdPuzzle #5	Interactive Learning Groups KC #3
Week 4: Effective Learning Techniques	Tuesday: 2/6	Why don’t students like school?	Willingham Chapter 1 EdPuzzle #6	Interactive Learning Groups
	Thursday: 2/8	What are the best learning strategies? What are desirable difficulties?	Dunlosky (2013) Bjork & Bjork (2011) EdPuzzle #7	Interactive Learning Groups KC#4 Exam 1

Week 5: Background Knowledge	Tuesday: 2/13	Must we have the facts before we can attain the skill?	Willingham Chapter 2 EdPuzzle #8	Interactive Learning Groups
	Thursday: 2/15	How does background knowledge affect reading comprehension?	Hirsch (2003) EdPuzzle #9	Interactive Learning Groups KC #5
Week 6: Human Memory: Forgetting & Remembering	Tuesday: 2/20	How does human memory work? Is forgetting the key to remembering?	Willingham Chapter 3 EdPuzzle #10	Interactive Learning Groups
	Thursday: 2/22	What kinds of tests trigger productive retrieval processes?	Little et al. (2012) EdPuzzle #11	Interactive Learning Groups KC #6
Week 7: Effective Acquisition of Abstract Concepts	Tuesday: 2/27	Why is learning abstract concepts SO hard? Concrete v. abstract examples	Willingham Chapter 4 Kaminski et al. (2006) EdPuzzle #12	Interactive Learning Groups
	Thursday: 3/1	Concreteness Fading	Fyfe et al (2014) EdPuzzle #13	Interactive Learning Groups KC#7
Week 8: Practice, Practice, Practice	Tuesday: 3/6	Is drilling worth it?	Willingham Chapter 5 Baroody (2006) EdPuzzle #14	Lecture / Discussion
	Thursday: 3/8	What kinds of tests prove to be the best type of practice?	Agarwal et al. (2008)	Interactive Learning Groups KC#8 Exam 2
Week 9:	Tuesday: 3/13	How do we get students to think like experts?	Willingham Chapter 6	Lecture/ Discussion

	Thursday: 3/15	How do experts even become experts?	Ericsson (2008)	Lecture/ Discussion KC#9
Week 10: Learner Differences	Tuesday: 3/20	What are cognitive styles and cognitive abilities? Should multiple intelligences be taught in schools?	Willingham Chapter 7	Lecture/ Discussion
	Thursday: 3/22	Do learning styles exist?	Rohrer & Pashler (2012)	Lecture/ Discussion KC#10
Week 11: Beliefs About How We Learn	Tuesday: 4/3	What is intelligence? Is intelligence fixed or incremental?	Willingham Chapter 8	Lecture/ Discussion
	Thursday: 4/5	What do students believe about how they learn?	Dweck (2010)	Lecture/ Discussion KC #11
Week 12: Becoming a Better Teacher	Tuesday: 4/10	How can we sharpen our teaching skill?	Willingham Chapter 9	Lecture/ Discussion
	Thursday: 4/12	What are the characteristics of the effective teacher?	Hattie Chapter 3	Lecture/ Discussion KC #12 Exam 3
Week 13: Visible Learning	Tuesday: 4/17	How do teachers know if the techniques they are using are effective?	Eggen 14	Lecture/ Discussion
	Thursday: 4/19	How do we assess learning?	Eggen 15	Lecture/ Discussion KC#13
Week 14: Motivation, Attribution, & Self-Determination Theory	Tuesday: 4/24	What are intrinsic and extrinsic motivation? What is self-determination theory?	Eggen Chapters 10 & 11	Lecture/ Discussion
	Thursday: 4/26	What is productive attribution theory?	Kitsantas & Zimmerman (2006)	Lecture/ Discussion KC #14

Week 15: Classroom Management &	Tuesday: 5/1	How do we develop self-regulated learners?	Eggen Chapter 12	Lecture/ Discussion
	Thursday: 5/3	How much decoration is too much decoration in the classroom?	Fisher et al (2014)	Lecture/ Discussion KC #15 Exam 4

APPENDIX B

SAMPLE HANDOUT

Appendix B

Sample Handout

Handout 6

Hobiss 2017/Cognitive Load Theory - Paas, Renkl, Sweller (2003)

1. Discuss each of the four reasons that attention is an important area to understand for both researchers and educators.
2. Describe how attention can be the bottleneck in cognition and learning. Discuss the connection between attention and working memory. How does this connect to what we learned in Mayer chapter 1?
3. Given what we learned about attention and working memory, discuss the things that teachers should consider when developing their lessons. (Hint: Think back to last week)
4. Discuss the types of cognitive load including what teachers can and can't control. Include what happens when cognitive load is too high and ideas for minimizing the negative impact of cognitive load on student learning.
5. Discuss the recommendations for instructional design that can support students and reduce cognitive load. Include examples from your content area.

Willingham Ch. 4 and Fyfe et al., (2014)

1. Describe why acquisition of abstract concepts is so difficult for students. Give an example of abstract content from your content area.
2. Describe surface and deep structure in problems. How does comparing and contrasting help students understand problem structure? How can concreteness fading help students find the shallow and deep structures of problems?
3. Discuss the things teachers should consider when determining what type(s) of examples they use in lessons and activities.
4. Describe each stage of concreteness fading. Include examples of how you might present content from your subject area at each stage.
5. Students benefit from lots of examples, and Willingham suggests that concrete and abstract examples should be used. Discuss the recommendations Willingham makes including the pros and cons of both concrete and abstract examples. How does this relate to the Concreteness Fading model?

APPENDIX C

VIDEOS EXPERIMENTS 1 AND 2

Appendix C

Videos Experiments 1 and 2

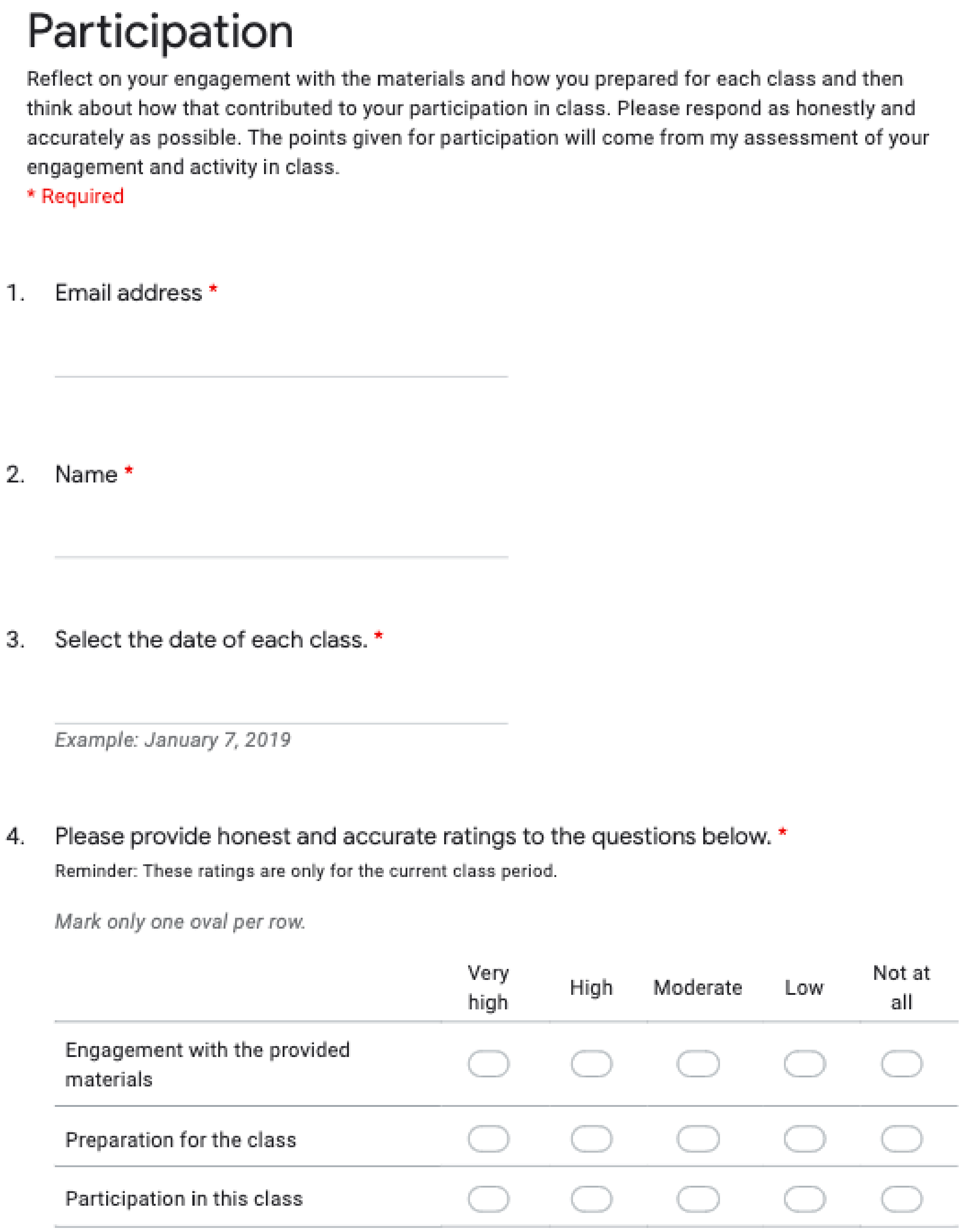
Week	Day One	Day Two
1	#1 Peter Doolitte: How your working memory makes sense TED Talk	
2	#2 CC: The Growth of Knowledge	#3 CC: Adolescence
3	#4 CC: How to Train a Brain	#5 CC: The Bobo Beatdown
4	#6 Learning to Learn: Conversation with John Dunlosky	#7 Using Desirable Difficulties to Enhance Learning by Bob Bjork
5	#8 Teaching Content is Teaching Reading by Daniel Willingham	#9 More Reading: Kelly Corrigan at TEDx Sonoma County
6	#10 CC: Remembering and Forgetting	#11 The Power of Forgetting by Bob Bjork
7	#12 Methodology of Singapore Math	#13 What is 'Transfer of Learning' and How Does it Help
8	#14 Is Testing the Friend or Foe of Education? Katherine Rawson TED Talk	#15 Testing Effect (test-enhanced learning)
9	#16 How to Think Like Dr House	#17 Peak by Anders Ericsson - by Brian Johnson
10	#18 Learning Styles Don't Exist by Daniel Willingham	#19 CC: Controversy of Intelligence
11	#20 Your Brain is Plastic	#21 The Power of Believing you Can Improve by Carol Dweck
12	#22 Is Teaching an Art or a Science by Daniel Willingham	#23 What makes a good teacher great? Azul Terronez TED Talk
13	#24 Real-time Assessment: Providing a Window into Student Learning	#25 Let's Teach for Mastery--not Test Scores- Sal Khan TED Talk
14	#26 CC: The Power of Motivation	#27 Promoting Motivation, Health, and Excellence by Ed Deci
15	#28 Teacher Advice: Classroom Management	#29 5 Mistakes New Teachers Make

APPENDIX D

PARTICIPATION FORM

Appendix D

Participation Form



Participation

Reflect on your engagement with the materials and how you prepared for each class and then think about how that contributed to your participation in class. Please respond as honestly and accurately as possible. The points given for participation will come from my assessment of your engagement and activity in class.

* Required

1. Email address *

2. Name *

3. Select the date of each class. *

Example: January 7, 2019

4. Please provide honest and accurate ratings to the questions below. *

Reminder: These ratings are only for the current class period.

Mark only one oval per row.

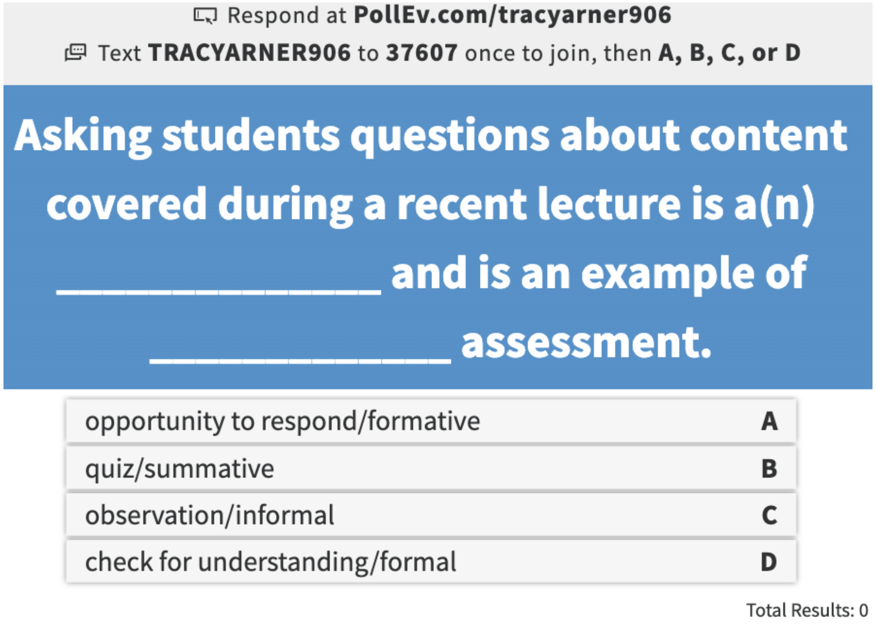
	Very high	High	Moderate	Low	Not at all
Engagement with the provided materials	○	○	○	○	○
Preparation for the class	○	○	○	○	○
Participation in this class	○	○	○	○	○

APPENDIX E

POLL EVERYWHERE EXAMPLE

Appendix E

Poll Everywhere Example



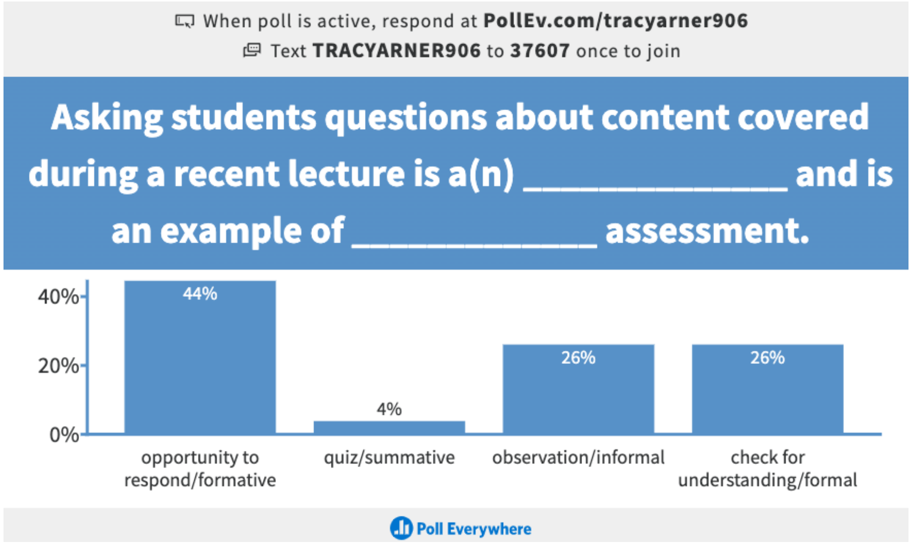
Respond at **PollEv.com/tracyarner906**

Text **TRACYARNER906** to **37607** once to join, then **A, B, C, or D**

Asking students questions about content covered during a recent lecture is a(n) _______________ and is an example of _______________ assessment.

opportunity to respond/formative	**A**
quiz/summative	**B**
observation/informal	**C**
check for understanding/formal	**D**

Total Results: 0



When poll is active, respond at **PollEv.com/tracyarner906**

Text **TRACYARNER906** to **37607** once to join

Asking students questions about content covered during a recent lecture is a(n) _______________ and is an example of _______________ assessment.

APPENDIX F

COURSE SCHEDULE EXPERIMENT 3

Appendix F

Course Schedule Experiment 3

Fall 2019 Course Schedule**		
Date	Questions to Consider	Before class readings
Week 1 - Introduction to Educational Psychology		
PRETEST # 1 MONDAY 8/26		
Monday 8/26	What is Educational Psychology?	Mayer Ch 1
Wednesday 8/29	Is learning a social or individual process?	Newcombe (2013)
SURVEY # 1 DUE August 30, 2019 at 11:59 pm		
Week 2 - Cognitive Development Theories		
Monday 9/2	LABOR DAY	
Wednesday 9/4	Are we born with innate knowledge?	Newcombe (2013)
Week 3 - Learning Theories		
Monday 9/9	What are classical and operant conditioning? Is learning social?	Eggen & Kauchak Ch 9
Wednesday 9/11	Why don’t students like school?	Willingham Ch 1
Week 4 - Effective Learning Techniques		
Monday 9/16	What are the best learning strategies?	Dunlosky (2013) Little et al., (2012)
Wednesday 9/18	When is difficulty desirable?	Bjork & Bjork (2011)
Week 5 - Background Knowledge		

Monday 9/23	Must we have the facts before we can attain the skill?	Willingham Ch 2
Wednesday 9/25	How does background knowledge affect reading comprehension?	Hirsch (2003)
Week 6 - Human Memory: Forgetting and Remembering		
Monday 9/30	How does human memory work? What is working memory?	Willingham Ch 3 Gathercole (2008)
Wednesday 10/2	What role does attention play? How much is too much?	Fisher et al., (2014) CLT website
Week 7 - Effective Acquisition of Abstract Concepts		
Monday 10/7	Why is learning abstract concepts SO hard? Concrete v. abstract examples	Willingham Ch 4
Wednesday 10/9	Concreteness Fading	Fyfe et al., (2014)
POSTTEST # 1 WEDNESDAY 10/9		
Week 8 - Practice, Practice, Practice		
PRETEST # 2 MONDAY 10/14		
Monday 10/14	Is drilling worth it?	Willingham Ch 5
Wednesday 10/16	Can tests be practice?	Rawson & Dunlosky (2012)
SURVEY # 2 DUE October 18, 2019 at 11:59 pm		
Week 9 - Becoming an Expert		
Monday 10/21	How do we get students to think like experts?	Willingham Ch 6
Wednesday 10/23	How do experts even become experts?	Ericsson (2008)
Week 10 - Learner Differences		
Monday 10/28	What are cognitive styles and multiple intelligences?	Willingham Ch 7

Wednesday 10/30	Should I teach to my students' learning styles?	Rohrer & Pashler (2012) Willingham Blog post
Week 11 - Beliefs About How We Learn		
Monday 11/4	What is intelligence? Is intelligence fixed or incremental?	Willingham Ch 8
Wednesday 11/6	What do students believe about how they learn?	Dweck (2010)
Week 12 - Special Populations		
Monday 11/11	VETERANS DAY	
Wednesday 11/13	How do we meet the needs of struggling students? What do I do when my students aren't from here?	Learning Disabilities Chapter Classwide Interventions for Students Educating English Language Learners Teaching English Language Learners-Tips from the Classroom
Week 13 - Motivation and Learning		
Monday 11/18	How can teachers motivate students to learn?	Eggen & Kauchak Ch 10 & 11
Wednesday 11/20	How do teachers develop self-regulated learners?	Eggen & Kauchak Ch 12
Week 14 - Becoming a Better Teacher		
Monday 11/25	How do we hone our teaching skills?	Willingham Ch 9
Wednesday 11/27	THANKSGIVING	
Week 15 - Visible Learning		
Monday 12/2	How do teachers know if the techniques they are using are effective?	Eggen & Kauchak Ch 14

Wednesday 12/4	How do we assess learning?	Eggen & Kauchak Ch 15
POSTTEST #2 WEDNESDAY 12/4		

APPENDIX G

VIDEO LIST EXPERIMENT 3

Appendix G

Video List Experiment 3

Week	Day One	Day Two
1	Peter Doolittle: How your working memory makes sense TED Talk	
2	CC: The Growth of Knowledge	CC: Adolescence
3	CC: How to Train a Brain	CC: The Bobo Beatdown
4	Learning to Learn: Conversation with John Dunlosky	Using Desirable Difficulties to Enhance Learning by Bob Bjork
5	Teaching Content is Teaching Reading by Daniel Willingham	More Reading: Kelly Corrigan at TEDx Sonoma County
6	CC: Remembering and Forgetting	Heavily Decorated Classrooms Disrupt Learning in Young Children
7	What is 'Transfer of Learning' and How Does it Help	Methodology of Singapore Math
8	Is Drilling Worth It?	Is Testing the Friend or Foe of Education? Katherine Rawson TED Talk
9	How to Think Like Dr House	Peak by Anders Ericsson - by Brian Johnson
10	CC: Controversy of Intelligence	Learning Styles Don't Exist by Daniel Willingham
11	Your Brain is Plastic	The Power of Believing you Can Improve by Carol Dweck
12	Rita Pierson-Every kid needs a champion	ELL Student Support Strategies
13	CC: The Power of Motivation	Self-regulation and Motivation
14	What makes a good teacher great? Azul Terronez TED Talk	
15	Real-time Assessment: Providing a Window into Student Learning	Let's Teach for Mastery--not Test Scores-Sal Khan TED Talk

REFERENCES

REFERENCES

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.

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.

Agarwal, P. K., Karpicke, J. D., Kang, S. H., Roediger III, H. L., & McDermott, K. B. (2008). Examining the testing effect with open-and closed-book tests. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 22*(7), 861-876.

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

Arner, T. (2020). Instructional design for the flipped classroom. In Walker, Z., Tan, D. & Ng Koh. (Eds.) *Flipped classrooms with diverse learners*. Singapore: Springer.

Arner, T., Aldosari, B., & Morris, B. J. (2017). Teaching students using evidence-based learning strategies through flipped classrooms. In R. Obeid, A. Schwartz, C. Shane-Simpson & P. J. Brooks (Eds.), *How we teach now: The GSTA guide to student centered teaching* (pp. 92-106). Retrieved from http://teachpsych.org/ebooks/

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.

Bahrick, H. P. (1979). Maintenance of knowledge: Questions about memory we forgot to ask. *Journal of Experimental Psychology: General,* 108(3), 296.

Barbour, C., & Schuessler, J. B. (2019). A preliminary framework to guide implementation of the flipped classroom method in nursing education. *Nurse Education in Practice*, 34, 36-42.

Barzagar Nazari, K., & Ebersbach, M. (2018). Distributed practice: Rarely realized in self-regulated mathematical learning. *Frontiers in Psychology*, 9, 2170.

Benassi, V. A., Overson, C. E., & Hakala, C. M. (2014). Applying science of learning in education: Infusing psychological science into the curriculum. Retrieved from the *Society for the Teaching of Psychology* website: http://teachpsych.org/ebooks/asle2014/index.php

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

Berrett, D. (2012, February 19). How 'flipping' the classroom can improve the traditional lecture. Retrieved from http://www.chronicle.com/article/How-Flipping-the-Classroom/130857

Bertsch, S., Pesta, B. J., Wiscott, R., & McDaniel, M. A. (2007). The generation effect: A meta-analytic review. *Memory & Cognition, 35*(2), 201-210.

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

Biard, N., Cojean, S., & Jamet, E. (2018). Effects of segmentation and pacing on procedural learning by video. *Computers in Human Behavior*, 89, 411-417.

Bishop, J. L., & Verleger, M. A. (2013, June). The flipped classroom: A survey of the research. In ASEE national conference proceedings, Atlanta, GA (Vol. 30, No. 9, pp. 1-18).

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, 64, 417-444.

Blended Learning Models. (n.d.). Retrieved April 8, 2020, from https://www.blendedlearning.org/models/

Blumenfeld, P. C., Kempler, T. M., & Krajcik, J. S. (2004). Motivation and cognitive engagement in learning environments. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 475–488). New York, NY: Cambridge University Press.

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

Brewer, R., & Movahedazarhouligh, S. (2018). Successful stories and conflicts: A literature review on the effectiveness of flipped learning in higher education. *Journal of Computer Assisted Learning, 34*(4), 409-416.

Budé, L., Imbos, T., van de Wiel, M. W., & Berger, M. P. (2011). The effect of distributed practice on students' conceptual understanding of statistics. *Higher Education, 62*(1), 69-79.

Burke, A. S., & Fedorek, B. (2017). Does "flipping" promote engagement?: A comparison of a traditional, online, and flipped class. *Active Learning in Higher Education, 18*(1), 11-24.

Butler, A. C., Karpicke, J. D., & Roediger III, H. L. (2007). The effect of type and timing of feedback on learning from multiple-choice tests. *Journal of Experimental Psychology: Applied, 13*(4), 273.

Butler, A. C., & Roediger, H. L. (2008). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Memory & Cognition, 36*(3), 604-616.

Carpenter, S. K., Cepeda, N. J., Rohrer, D., Kang, S. H., & Pashler, H. (2012). Using spacing to enhance diverse forms of learning: Review of recent research and implications for instruction. *Educational Psychology Review, 24*(3), 369-378.

Cepeda, N. J., Coburn, N., Rohrer, D., Wixted, J. T., Mozer, M. C., & Pashler, H. (2009). Optimizing distributed practice: Theoretical analysis and practical implications. *Experimental Psychology, 56*(4), 236-246.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*(3), 354.

Chamberland, M., & Mamede, S. (2015). Self-explanation, an instructional strategy to foster clinical reasoning in medical students. *Health Professions Education, 1*(1), 24-33.

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.

Chi, M. T. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science, 1*(1), 73-105. doi:10.1111/j.1756-8765.2008.01005.x

Chi, M. (2018). ICAP: How Students Engage to Learn. In proceedings of the *Cognitive Science Society*

Chi, M. T., Adams, J., Bogusch, E. B., Bruchok, C., Kang, S., Lancaster, M., ... & Wylie, R. (2018). Translating the ICAP theory of cognitive engagement into practice. *Cognitive Science, 42*(6), 1777-1832.

Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*(2), 145-182.

Chi, M. T., De Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*(3), 439-477.

Chi, M. T., & Menekse, M. (2015). Dialogue patterns in peer collaboration that promote learning. *Socializing intelligence through academic talk and dialogue*, 263-274.

Chi, M. T., Roy, M., & Hausmann, R. G. (2008). Observing tutorial dialogues collaboratively: Insights about human tutoring effectiveness from vicarious learning. *Cognitive Science, 32*(2), 301-341.

Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist, 49*(4), 219-243.

Chiu, J. L., & Chi, M. T. (2014). Supporting self-explanation in the classroom. *Applying science of learning in education: Infusing psychological science into the curriculum*, 91-103.

Clark, R. C., & Mayer, R. E. (2011). *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning* (3rd ed.). San Francisco, CA: Pfeiffer.

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review, 24*(4), 1158-1170.

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

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.

Davis, J., Chryssafidou, E., Zamora, J., Davies, D., Khan, K., & Coomarasamy, A. (2007). Computer-based teaching is as good as face to face lecture-based teaching of evidence based medicine: a randomised controlled trial. *BMC Medical Education, 7*(1), 23.

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.

Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved learning in a large-enrollment physics class. *Science, 332*(6031), 862-864.

Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology, 84*(5), 795.

Dunlosky, J., & Rawson, K. A. (2015). Practice tests, spaced practice, and successive relearning: Tips for classroom use and for guiding students' learning. *Scholarship of Teaching and Learning in Psychology, 1*(1), 72.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*(1), 4-58. doi:10.1177/1529100612453266

Ebert-May, D., Brewer, C., & Allred, S. (1997). Innovation in large lectures: Teaching for active learning. *Bioscience, 47(*9), 601-607.

Enfield, J. (2013). Looking at the impact of the flipped classroom model of instruction on undergraduate multimedia students at CSUN. *TechTrends, 57*(6), 14-27.

Ertmer, P. A., & Newby, T. J. (2013). Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly, 26*(2), 43-71.

Fadde, P., & Sullivan, P. (2013). Using interactive video to develop pre-service teachers' classroom awareness. *Contemporary Issues in Technology and Teacher Education, 13*(2), 156-174.

Fiorella, L., & Mayer, R. E. (2018). What works and doesn't work with instructional video. *Computers in Human Behavior*, 89, 465-470. doi:10.1016/j.chb.2018.07.015

Findlay-Thompson, S., & Mombourquette, P. (2014). Evaluation of a flipped classroom in an undergraduate business course. *Business Education & Accreditation, 6*(1), 63-71.

Foertsch, J., Moses, G., Strikwerda, J., & Litzkow, M. (2002). Reversing the lecture/homework paradigm using eTEACH® web-based streaming video software. *Journal of Engineering Education, 91*(3), 267-274.

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

Fox, T. A. (2018). Building a Foundation in Early Childhood Education: An Analysis of the Common CoreState Standards and Its Influence on Kindergarten (Doctoral dissertation, Southern Connecticut State University).

Frederickson, N., Reed, P., & Clifford, V. (2005). Evaluating web-supported learning versus lecture-based teaching: Quantitative and qualitative perspectives. *Higher Education, 50*(4), 645-664.

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.

Gallagher, J. E., Dobrosielski-Vergona, K. A., Wingard, R. G., & Williams, T. M. (2005). Web-based vs. traditional classroom instruction in gerontology: A pilot study. *American Dental Hygienists' Association, 79*(3), 7-7.

Gardiner, L. F. (1998). Why we must change: The research evidence. *Thought and Action, 14*(1), 71-88.

Giannakos, M. N., Krogstie, J., & Sampson, D. (2018). Putting flipped classroom into practice: A comprehensive review of empirical research. In *Digital Technologies: Sustainable innovations for improving teaching and learning* (pp. 27-44). Springer, Cham.

Gilboy, M. B., Heinerichs, S., & Pazzaglia, G. (2015). Enhancing student engagement using the flipped classroom. *Journal of Nutrition Education and Behavior, 47*(1), 109–114. doi:10.1016/j.jneb.2014.08.008

Glogger-Frey, I., Gaus, K., & Renkl, A. (2017). Learning from direct instruction: Best prepared by several self-regulated or guided invention activities?. *Learning and Instruction*, 51, 26-35.

Gluckman, M., Vlach, H. A., & Sandhofer, C. M. (2014). Spacing simultaneously promotes multiple forms of learning in children's science curriculum. *Applied Cognitive Psychology, 28*(2), 266-273.

Greving, S., & Richter, T. (2018). Examining the testing effect in university teaching: retrievability and question format matter. *Frontiers in Psychology*, 9, 2412.

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).

Guo, J. (2019). The use of an extended flipped classroom model in improving students' learning in an undergraduate course. *Journal of Computing in Higher Education, 31*(2), 362-390.

Halverson, L. R., Graham, C. R., Spring, K. J., Drysdale, J. S., & Henrie, C. R. (2014). A thematic analysis of the most highly cited scholarship in the first decade of blended learning research. *The Internet and Higher Education*, 20, 20-34.

Hamdan, N., McKnight, P., McKnight, K., & Arfstrom, K. M. (2013). The flipped learning model: A white paper based on the literature review titled a review of flipped learning. Flipped Learning Network/Pearson/George Mason University.

Hao, Y., & Lee, K. S. (2016). Teaching in flipped classrooms: Exploring pre-service teachers' concerns. *Computers in Human Behavior*, 57, 250-260.

Hausmann, R. G., & VanLehn, K. (2007). Explaining self-explaining: A contrast between content and generation. *Frontiers in Artificial Intelligence and Applications*, 158, 417.

Hew, K. F., & Lo, C. K. (2018). Flipped classroom improves student learning in health professions education: a meta-analysis. *BMC Medical Education, 18*(1), 38.

Hibbard, L., Sung, S., & Wells, B. (2016). Examining the effectiveness of a semi-self-paced flipped learning format in a college general chemistry sequence. *Journal of Chemical Education, 93*(1), 24-30.

Hinze, S. R., & Wiley, J. (2011). Testing the limits of testing effects using completion tests. *Memory, 19*(3), 290-304.

Hollis, R. B., & Was, C. A. (2016). Mind wandering, control failures, and social media distractions in online learning. *Learning and Instruction*, 42, 104-112.

Hung, I. C., Kinshuk, & Chen, N. S. (2018). Embodied interactive video lectures for improving learning comprehension and retention. *Computers & Education,* 117, 116-131.

Hwang, G. J., & Lai, C. L. (2017). Facilitating and bridging out-of-class and in-class learning: An interactive e-book-based flipped learning approach for math courses. *Journal of Educational Technology & Society, 20*(1), 184-197.

Jenkins, M., Bokosmaty, R., Brown, M., Browne, C., Gao, Q., Hanson, J., & Kupatadze, K. (2017). Enhancing the design and analysis of flipped learning strategies. *Teaching & Learning Inquiry, 5*(1), 1-12.

Jensen, J. L., Kummer, T. A., & Godoy, P. D. d. M. (2015). Improvements from a flipped classroom may simply be the fruits of active learning. *CBE—Life Sciences Education, 14*(1), ar5.

Jovanović, J., Gašević, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *The Internet and Higher Education, 33*(4), 74-85.

Kane, M. J., Smeekens, B. A., Von Bastian, C. C., Lurquin, J. H., Carruth, N. P., & Miyake, A. (2017). A combined experimental and individual-differences investigation into mind wandering during a video lecture. *Journal of Experimental Psychology: General, 146*(11), 1649.

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.

Karpicke, J. D. (2012). Retrieval-based learning: Active retrieval promotes meaningful learning. *Current Directions in Psychological Science, 21*(3), 157-163.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75-86.

Kitsantas, A., & Zimmerman, B. J. (2006). Enhancing self-regulation of practice: The influence of graphing and self-evaluative standards. *Metacognition and Learning, 1*(3), 201-212.

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science, 15*(10), 661-667.

Koedinger, K. R., Booth, J. L., & Klahr, D. (2013). Instructional complexity and the science to constrain it. *Science, 342*(6161), 935-937.

Kong, S. C. (2015). An experience of a three-year study on the development of critical thinking skills in flipped secondary classrooms with pedagogical and technological support. *Computers & Education*, 89, 16-31.

Kulik, J. A., & Kulik, C. L. C. (1988). Timing of feedback and verbal learning. *Review of Educational Research, 58*(1), 79-97.

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.

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, D. P., Butler, A. C., & Roediger III, H. L. (2013). Comparative effects of test-enhanced learning and self-explanation on long-term retention. *Medical Education, 47*(7), 674-682.

Lasry, N., Mazur, E., & Watkins, J. (2008). Peer instruction: From Harvard to the two-year college. *American Journal of Physics, 76*(11), 1066-1069.

Lee, J., Lim, C., & Kim, H. (2017). Development of an instructional design model for flipped learning in higher education. *Educational Technology Research and Development, 65*(2), 427-453.

Levesque, A. A. (2011). Using clickers to facilitate development of problem-solving skills. *CBE—Life Sciences Education, 10*(4), 406-417.

Little, J. L., Bjork, E. L., Bjork, R. A., & Angello, G. (2012). Multiple-choice tests exonerated, at least of some charges: Fostering test-induced learning and avoiding test-induced forgetting. *Psychological Science, 23*(11), 1337-1344.

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.

Mastropieri, M. A., Scruggs, T. E., Spencer, V., & Fontana, J. (2003). Promoting success in high school world history: Peer tutoring versus guided notes. *Learning Disabilities Research & Practice, 18*(1), 52-65.

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning?. *American Psychologist, 59*(1), 14.

Mayer, R. E. (2008). Applying the science of learning: Evidence-based principles for the design of multimedia instruction. *American Psychologist, 63*(8), 760-769. doi:10.1037/0003-066x.63.8.760

Mayer, R. E. (2017). Using multimedia for e-learning. *Journal of Computer Assisted Learning, 33*(5), 403-423.

Mayer, R. E., & Pilegard, C. (2005). Principles for managing essential processing in multimedia learning: Segmenting, pretraining, and modality principles. *The Cambridge handbook of multimedia learning*, 169-182.

Mazur, E. (1997, March). Peer instruction: getting students to think in class. In *AIP Conference Proceedings* (Vol. 399, No. 1, pp. 981-988). AIP.

McDaniel, M. A., & Little, J. L. (2019). Multiple Choice and Short-Answer Quizzing on Equal Footing in the Classroom. In J. Dunlosky & K. A. Rawson (Authors), *The Cambridge Handbook of Cognition and Education* (pp. 480-499). Cambridge University Press.

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), 196.

McLaughlin, J. E., White, P. J., Khanova, J., & Yuriev, E. (2016). Flipped classroom implementation: A case report of two higher education institutions in the United States and Australia. *Computers in the Schools, 33*(1), 24-37.

McNally, B., Chipperfield, J., Dorsett, P., Del Fabbro, L., Frommolt, V., Goetz, S., ... & Roiko, A. (2017). Flipped classroom experiences: student preferences and flip strategy in a higher education context. *Higher Education, 73*(2), 281-298.

Menekse, M., & Chi, M. T. (2018). The role of collaborative interactions versus individual construction on students' learning of engineering concepts. *European Journal of Engineering Education, 44*(5), 702-725 doi: 1-24.10.1080/03043797.2018.1538324

Menekse, M., Stump, G. S., Krause, S., & Chi, M. T. (2013). Differentiated overt learning activities for effective instruction in engineering classrooms. *Journal of Engineering Education, 102*(3), 346-374.

Miller, K., Lukoff, B., King, G., & Mazur, E. (2018). Use of a Social Annotation Platform for Pre-Class Reading Assignments in a Flipped Introductory Physics Class. *Frontiers in Education, 3*. doi:10.3389/feduc.2018.00008

Mitroka, J. G., Harrington, C., & DellaVecchia, M. J. (2020). A multiyear comparison of flipped-vs. lecture-based teaching on student success in a pharmaceutical science class. *Currents in Pharmacy Teaching and Learning, 12*(1), 84-87.

Miyatsu, T., Nguyen, K., & McDaniel, M. A. (2018). Five popular study strategies: Their pitfalls and optimal implementations. *Perspectives on Psychological Science, 13*(3), 390-407.

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.

Mohee, K., Haboubi, H., Protty, M., Srinivasan, C., Townend, W., & Weston, C. (2019). 45 Comparing impact of an e-learning package to lecture-based teaching in the management of Supraventricular Tachycardia(SVT): A randomized-controlled study. *Cardiac Rhythm Management*. doi:10.1136/heartjnl-2019-bcs.43

Morehead, K., Dunlosky, J., & Rawson, K. A. (2019). How much mightier is the pen than the keyboard for note-taking? A replication and extension of Mueller and Oppenheimer (2014). *Educational Psychology Review, 31*(3), 753-780.

Morin, B., Kecskemety, K. M., Harper, K. A., & Clingan, P. A. (2013). The inverted classroom in a first-year engineering course. In 120th ASEE Annual Conference & Exposition (1220 ed., Vol. 23, pp. 1-11). *American Society for Engineering Education.*

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.

Njie-Carr, V. P., Ludeman, E., Lee, M. C., Dordunoo, D., Trocky, N. M., & Jenkins, L. S. (2017). An integrative review of flipped classroom teaching models in nursing education. *Journal of Professional Nursing, 33*(2), 133-144.

Nokes-Malach, T. J., Zepeda, C. D., Richey, E., & Gadgil, S. (2019). Collaborative Learning: The Benefits and Costs. In J. Dunlosky & K. A. Rawson (Authors), *The Cambridge Handbook of Cognition and Education* (pp. 480-499). Cambridge University Press.

Nouri, J. (2016). The flipped classroom: for active, effective and increased learning–especially for low achievers. *International Journal of Educational Technology in Higher Education, 13*(1), 33.

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.

Olapiriyakul, K., & Scher, J. M. (2006). A guide to establishing hybrid learning courses: Employing information technology to create a new learning experience, and a case study. *The Internet and Higher Education, 9*(4), 287-301. doi: 10.1016/j.iheduc.2006.08.001

O'Reilly, T., Wang, Z., & Sabatini, J. (2019). How much knowledge is too little? When a lack of knowledge becomes a barrier to comprehension. *Psychological Science, 30*(9), 1344-1351.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38*(1), 1-4.

Paiva, F., Glenn, J., Mazidi, K., Talbot, R., Wylie, R., Chi, M. T., ... & Nielsen, R. D. (2014). Comprehension seeding: Comprehension through self explanation, enhanced discussion, and inquiry generation. In *International Conference on Intelligent Tutoring Systems* (pp. 283-293). Springer, Cham.

Peper, R. J., & Mayer, R. E. (1986). Generative effects of note-taking during science lectures. *Journal of Educational Psychology*, *78*(1), 34.

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. doi:10.1177/0098628315620063

Pierce, R. (2013). Student performance in a flipped class module. In Society for Information Technology & Teacher Education International Conference (pp. 942-954). *Association for the Advancement of Computing in Education* (AACE).

Pierce, R., & Fox, J. (2012). Vodcasts and active-learning exercises in a "flipped classroom" model of a renal pharmacotherapy module. *American Journal of Pharmaceutical Education, 76*(10).

Pilegard, C., & Fiorella, L. (2016). Helping students help themselves: Generative learning strategies improve middle school students' self-regulation in a cognitive tutor. *Computers in Human Behavior,* 65, 121-126.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), 33.

Pintrich, P. R. (1991). A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ).

Pintrich, P. R., Smith, D. A., Garcia, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement, 53*(3), 801-813.

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education, 93*(3), 223-231.

Prunuske, A. J., Batzli, J., Howell, E., & Miller, S. (2012). Using online lectures to make time for active learning. *Genetics, 192*(1), 67-72.

Pyc, M. A., Agarwal, P. K., & Roediger, H. L. III. (2014). Test-enhanced learning. In V. A. Benassi, C. E. Overson, & C. M. Hakala (Eds.), Applying science of learning in education: Infusing psychological science into the curriculum (p. 78–90). *Society for the Teaching of Psychology*

Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory?. *Journal of Memory and Language, 60*(4), 437-447.

Rawson, K. A., & Dunlosky, J. (2012). When is practice testing most effective for improving the durability and efficiency of student learning?. *Educational Psychology Review, 24*(3), 419-435.

Renkl, A., & Eitel, A. (2019). Self-Explaining: Learning about Principles and their Application. In J. Dunlosky & K. A. Rawson (Authors), *The Cambridge Handbook of Cognition and Education* (pp. 480-499). Cambridge University Press.

Rittle-Johnson, B. (2006). Promoting transfer: Effects of self-explanation and direct instruction. *Child Development, 77*(1), 1-15.

Rittle-Johnson, B., & Loehr, A. M. (2017). Eliciting explanations: Constraints on when self-explanation aids learning. *Psychonomic Bulletin & Review, 24*(5), 1501-1510.

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.

Roediger III, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*(3), 249-255.

Roediger III, H. L., Putnam, A. L., & Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. *Psychology of Learning and Motivation*. 55, 1-36.

Rohrer, D. (2015). Student instruction should be distributed over long time periods. *Educational Psychology Review, 27*(4), 635-643.

Roschelle, J., Rafanan, K., Bhanot, R., Estrella, G., Penuel, B., Nussbaum, M., & Claro, S. (2010). Scaffolding group explanation and feedback with handheld technology: Impact on students' mathematics learning. *Educational Technology Research and Development, 58*(4), 399-419.

Roscoe, R. D., & Chi, M. T. (2008). Tutor learning: The role of explaining and responding to questions. *Instructional Science, 36*(4), 321-350.

Rotgans, J. I., & Schmidt, H. G. (2012). The intricate relationship between motivation and achievement: Examining the mediating role of self-regulated learning and achievement-related classroom behaviors. *International Journal of Teaching and Learning in Higher Education, 24*(2), 197-208.

Roy, M., & Chi, M. T. (2005). The self-explanation principle in multimedia learning. *The Cambridge Handbook of Multimedia Learning*, 271-286.

Ryan, M. D., & Reid, S. A. (2016). 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.

Sana, F., Weston, T., & Cepeda, N. J. (2013). Laptop multitasking hinders classroom learning for both users and nearby peers. *Computers & Education*, 62, 24-31.

Schell, J., & Mazur, E. (2015). Flipping the chemistry classroom with peer instruction. In J. Gavier-Martinez & E. Serrano-Torregrosa (Eds.), *Chemistry Education: Best practices, opportunities and trends* (pp. 319-344). Weinheim, Germany: Wiley-VCH.

Schmidt, S. M. P., & Ralph, D. L. (2015). The flipped classroom: A twist on teaching. *Contemporary Issues in Education Research, 9*(200348548), 1–6.

Schutte, G. M., Duhon, G. J., Solomon, B. G., Poncy, B. C., Moore, K., & Story, B. (2015). A comparative analysis of massed vs. distributed practice on basic math fact fluency growth rates. *Journal of School Psychology*, 53(2), 149-159.

Schworm, S., & Renkl, A. (2002). Learning by solved example problems: Instructional explanations reduce self-explanation activity. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 24, No. 24).

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. doi: 10.1108/JARHE-04-2015-0024

Seabrook, R., Brown, G. D., & Solity, J. E. (2005). Distributed and massed practice: From laboratory to classroom. *Applied Cognitive Psychology*, 19(1), 107-122.

Setren, E., Greenberg, K., Moore, O., & Yankovich, M. (2019). Effects of the flipped classroom: Evidence from a randomized trial. *School Effectiveness and Inequality Initiative*.

Slocum, A. C., & Beard, C. A. (2005). Development of a CAI module and comparison of its effectiveness with traditional classroom instruction. *Clothing and Textiles Research Journal, 23*(4), 298-306.

Smith, P. L., & Ragan, T. J. (2005). *Instructional design*. Hoboken, NJ: J. Wiley & Sons.

Smith, M. K., Wood, W. B., Krauter, K., & Knight, J. K. (2011). Combining peer discussion with instructor explanation increases student learning from in-class concept questions. *CBE—Life Sciences Education, 10*(1), 55-63.

Stephenson, J. E., Brown, C., & Griffin, D. K. (2008). Electronic delivery of lectures in the university environment: An empirical comparison of three delivery styles. *Computers & Education, 50*(3), 640-651.

Strayer, J. F. (2012). How learning in an inverted classroom influences cooperation, innovation and task orientation. *Learning Environments Research, 15*(2), 171–193. doi:10.1007/s10984-012-9108-4

Svihla, V., Wester, M. J., & Linn, M. C. (2018). Distributed practice in classroom inquiry science learning. *Learning: Research and Practice, 4*(2), 180-202.

Trafton, J. G., & Trickett, S. B. (2001). Note-taking for self-explanation and problem solving. *Human-Computer Interaction, 16*(1), 1-38.

Trochim, W. M. (1986). *Advances in quasi-experimental design and analysis* (No. 31). Jossey-Bass Inc Pub.

Tucker, B. (2012). The flipped classroom. *Education Next,* 82-83.

Tune, J. D., Sturek, M., & Basile, D. P. (2013). Flipped classroom model improves graduate student performance in cardiovascular, respiratory, and renal physiology. *Advances in Physiology Education, 37*(4), 316-320.

van Alten, D. C., Phielix, C., Janssen, J., & Kester, L. (2019). Effects of flipping the classroom on learning outcomes and satisfaction: A meta-analysis. *Educational Research Review*, 28, 100281.

VanLehn, K. (2006). The behavior of tutoring systems. *International Journal of Artificial Intelligence in Education, 16*(3), 227-265.

Vlach, H. A., & Sandhofer, C. M. (2012). Distributing learning over time: The spacing effect in children's acquisition and generalization of science concepts. *Child Development, 83*(4), 1137-1144.

Vural, O. F. (2013). The Impact of a Question-Embedded Video-Based Learning Tool on E-Learning. *Educational Sciences: Theory and Practice, 13*(2), 1315-1323.

Walker, Z., Tan, D., Klimplová, L., & Bicen, H. (2020). An Introduction to Flipping the Classroom. In *Flipped Classrooms with Diverse Learners* (pp. 3-15). Springer, Singapore.

Wammes, J. D., Ralph, B. C., Mills, C., Bosch, N., Duncan, T. L., & Smilek, D. (2019). Disengagement during lectures: Media multitasking and mind wandering in university classrooms. *Computers & Education*, 132, 76-89.

Wammes, J. D., & Smilek, D. (2017). Examining the influence of lecture format on degree of mind wandering. *Journal of Applied Research in Memory and Cognition, 6*(2), 174-184.

Willingham, D. T. (2009). Why don't students like school?: A cognitive scientist answers questions about how the mind works and what it means for the classroom. Hoboken, NJ: John Wiley & Sons.

Winquist, J. R., & Carlson, K. A. (2014). Flipped statistics class results: Better performance than lecture over one year later. *Journal of Statistics Education, 22*(3). doi: 10.1080/10691898.2014.11889717

Wiseheart, M., Küpper-Tetzel, C. E., Weston, T., Kim, A. S. N., Kapler, I. V., & Foot-Seymour, V. (2019). Enhancing the Quality of Student Learning Using Distributed Practice. In *The Cambridge Handbook of Cognition and Education* (pp. 550–583). New York, NY: Cambridge University Press.

Wittrock, M. C. (1989). Generative processes of comprehension. *Educational Psychologist, 24*(4), 345-376. DOI: 10.1207/s15326985ep2404_2

Wood, E., Zivcakova, L., Gentile, P., Archer, K., De Pasquale, D., & Nosko, A. (2012). Examining the impact of off-task multi-tasking with technology on real-time classroom learning. *Computers & Education, 58*(1), 365-374.

Wylie, R., & Chi, M. T. (2014). 17 The Self-Explanation Principle in Multimedia Learning. *The Cambridge Handbook of Multimedia Learning,* 413.

Yarbro, J., Arfstrom, K. M., McKnight, K., & McKnight, P. (2014). Extension of a review of flipped learning. *Flipped Learning Network*/Pearson/George Mason University.

Zainuddin, Z., & Halili, S. H. (2016). Flipped classroom research and trends from different fields of study. *International Review of Research in Open and Distributed Learning*, 17(3), 313-340.

Zhang, H., Miller, K. F., Sun, X., & Cortina, K. S. (2020). Wandering eyes: Eye movements during mind wandering in video lectures. *Applied Cognitive Psychology, 34*(2), 449-464.

Zhang, D., Zhou, L., Briggs, R. O., & Nunamaker Jr, J. F. (2006). Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Information & Management, 43*(1), 15-27.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*(1), 3-17.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice, 41*(2), 64-70.

Zimmerman, B. J., & Schunk, D. H. (2012). Motivation: An Essential Dimension to Self-regulated Learning. In Motivation and Self-regulated Learning: Theory, Research, and Applications (pp. 1–30). New York: Routledge.