Spring 2019

English Language Learners and the Effectiveness of Flipped Learning in Middle-School Science

Stephen Elliott
California State University, Monterey Bay

English Language Learners and the Effectiveness of Flipped Learning in Middle-School Science

Stephen Elliott

Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Arts in Education

California State University, Monterey Bay

May 2019

English Language Learners and the Effectiveness of Flipped Learning

in Middle-School Science

Stephen Elliott

APPROVED BY THE GRADUATE ADVISORY COMMITTEE

____
Kerrie Chitwood, Ph.D.
Advisor and Program Coordinator, Master of Arts in Education

____
Dennis Kombe, Ph.D.
Advisor, Master of Arts in Education

Approved by the Dean of Graduate Studies 05/21/2019

Kris Roney, Ph.D. Associate Vice President
Academic Programs and Dean of Undergraduate & Graduate Studies

Abstract

English Language Learners (ELLs) who study in mainstream courses (e.g., science) alongside English Only (EO) peers may require teaching strategies which increase personal interactions with the instructor. The flipped classroom allows for more of this type of interaction by pushing the lecture portion out of the classroom as video homework. The purpose of this study was to see if flipped learning would have an effect on the academic achievement of ELLs in science. It was hypothesized that exposure to flipped learning would increase student post-test scores when compared to a control group. The study was a two group quantitative pre-test post-test quasi experimental design. The intervention consisted of a four-week period, during which the treatment group (n=24) received the flipped classroom, while the control group (n=20) received traditional lessons, which included lecture and practice (i.e., non-video) based homework. There was no significant difference between the post-test scores of the treatment and control; therefore, both methods can be seen as viable teaching strategies for ELLs. However, there were several limitations, including the length of the intervention, and sensitivity of the measure which may have influenced the results. Therefore, further research is recommended with this age group in order to determine if the flipped classroom is a viable strategy for teaching ELLs in science.

Keywords: flipped learning, science, middle school, ELL, technology in the classroom

Acknowledgements

To my beautiful, amazing wife. During this difficult year, the only argument we ever had was over whether or not I could really accomplish this. As I write these final words, I have proven you right. Thank you.

Table of Contents

Abstract…..………………………………………………………………………………iii

Literature Review.. 1

Methods.. 8

Research Question……………………………………………………………………9

Hypothesis 9

Research Design 9

Setting and Participants 10

Measures.. 11

Intervention.. 12

Procedures 13

Ethical Considerations... 14

Data Analysis... 16

Results 17

Discussion 18

References...24

Appendix A...28

Appendix B...32

English Language Learners and the Effectiveness of Flipped Learning in Middle-School Science

Literature Review

In the traditional learning model, the classroom has always been teacher-centered (Chuang, Weng, & Chen, 2018). In such classrooms, a teacher delivers didactic instruction (i.e., lecture) while students take notes. Homework assignments are given for practice, and students are told to complete them at home (González-Gómez, Jeong, Rodríguez, & Cañada-Cañada, 2016). Though this model may be effective for some, it does make a few important assumptions. First of all, the model assumes that hearing content only once in class will prepare students for the practice required by traditional homework. Second, this model assumes that every student is able to learn at the pace set by the instructor (Umansky et al., 2015). Finally, the traditional model of teaching assumes students do not need assistance from the instructor when completing homework. This traditional model ignores the needs of English Language Learners (ELLs). ELLs, who may have trouble retaining academic language, may need to work at a different pace, and may need direct, individualized feedback in order to attain mastery (Bergman, 2011). The flipped learning model is a new way of approaching classroom content instruction (Bergmann & Sams, 2012). Flipped learning turns instruction around through the use of technology, moving the lecture portion of class into videos for students to view at home. Through this model, individualized instruction and scaffolding of content to meet the diverse needs of ELLs is possible.

ELLs in California Content Classrooms

The numbers of students who identify as ELLs are increasing in California (Ed Data, n.d.). Currently, within the state, around twenty percent of the total number of students enrolled are identified as ELLs (Ed Data, n.d.). In addition, there are several pockets, including the diverse Central Coast, whose ELL enrollment averages around 40%, twice that of the rest of the state (Umansky et al., 2015). Given the need to develop English proficiency, some ELLs may be pulled from core content classrooms (Umansky et al., 2015). The current practice is to give students who struggle with English additional targeted class time for language. This means that core content, like science, may not be available to all ELLs. The result is a divided classroom, wherein students with previous exposure to core content are ready to move on, while other students in the same class may have had little to no exposure to the fundamentals (Umansky et al., 2015).

A report by Umansky and colleagues (2015) examined the issues of ELL access to core curriculum in the state of California. The authors noted that in science in particular, ELLs are nearly four times as likely as English only (EO) students to be pulled from science classrooms. This means that learning of scientific academic language (e.g., vocabulary, the scientific method), may be curtailed until after students reclassify as fluent English proficient (Umansky et al., 2015). It typically takes several years for ELLs to acquire conversational English, also known as basic interpersonal communication skills (BCIS). However, BCIS is not the language of the secondary classroom. More advanced ability, known as cognitive academic language proficiency (CALP) may take many more years to develop (Perez & Holmes, 2010). As a result, many students may appear to have mastered the basic vocabulary of science because of their BCIS skills, but are in reality underprepared for the academic rigor of core content (Janzen, 2008).

Additionally, many instructors are not trained to meet the demands of teaching content to students who are classified as ELLs. Janzen (2008) collected data on non-English content teachers who teach ELLs in their classrooms and found that of the 41% of teachers who had

ELLs in their classrooms, only about 12% had more than eight hours of training on how to meet their needs. Furthermore, even though scaffolding content has been identified as a key feature of best practices for ELLs, many teachers are unable to provide proper scaffolding of content through traditional didactic instruction (Bergman, 2011). If a classroom is set up with the teacher as the only source for content information, there will only be one way in which students can access content. This leaves ELLs at a disadvantage in content classrooms.

Relevant Research on Teaching ELLs

The challenges of teaching language along with core content are not new. One foundational concept which has informed instruction for diverse learners comes from Vygotsky (1978). Vygotsky's Zone of Proximal Development (ZPD) rejects the Piagetian model, which describes learning as directly related to brain development (i.e., phases of brain development determining student ability) (Piaget, 1928). Instead, Vygotsky's research indicated that students may be able to work even beyond their own ability, if working alongside another person with higher ability (e.g., an advanced peer or teacher). Around the time that English translations of Vygotsky's work became available, Bandura's Social Learning Theory also claimed that learning is a social process, and that all learners benefit from working together with content (1977).

Another model for differentiation and scaffolding is Bloom's mastery learning (Bloom, 1968). Bloom postulated that each student's individual needs translate to a timeline needed for mastery, as opposed to whether or not mastery was achievable for all. In other words, students who fall behind in content may only need more time to practice until they reach the same level of mastery as their peers. From this idea, instructors have sought out and developed strategies which maximize student productive time in the classroom. Additionally, there is a need for strategies in which students with shorter timelines to mastery may move on without being frustrated by repeated exposure to content they have already mastered. Flipped learning is a strategy which can meet those needs, by setting individual timelines for student growth, while differentiating to meet individual needs (Bergmann & Sams, 2012).

The Flipped Learning Model

Flipped learning as a teaching strategy is relatively new. The first mention of it comes from Bergmann and Sams (2012), a pair of high school chemistry teachers in Colorado. The idea of flipped learning has gained popularity in the past few years, with studies on the effectiveness of the model being published worldwide (Cabi, 2018; Gonzalez-Gomez et al., 2016; Lin & Hwang, 2018).

In a flipped classroom, the lecture is treated as homework, usually in the form of videos from online resources (e.g., YouTube, Khan Academy) or video screencasts created by the instructor (Bergmann & Sams, 2012). The purpose of students hearing their lecture at home is to free up time in the class for practice and completion of activities related to content. Students can achieve more when doing these things alongside the support of the instructor and their peers (Bandura, 1977; Bloom, 1968; Lin & Hwang, 2018).

The greatest benefit for this strategy goes to students who have been underserved by the traditional model, like ELLs (Altemueller & Lindquist, 2017). Furthermore, Altemueller and Lindquist (2017) found that while the flipped classroom benefitted the academic gains of all students, the greatest gains were made by students classified as low achievers. Often, these are students who merely need more specific feedback in order to perform at required levels of mastery (Janzen, 2008). The format of flipped learning gives instructors more flexibility, allowing diverse needs to be met.

Affordances and Challenges of Flipped Learning

According to Bergmann and Sams (2012), the key element of this model has been misunderstood by many. Some educators attribute the videos that students watch at home to the success of the model. Contrary to this belief, Bergmann and Sams (2012) argued that the intended focus is what happens in the classroom, and not how the content is presented at home. When combined with engaging, hands-on learning opportunities, the majority of studies on flipped learning have found improved academic performance and an increase in positive attitudes toward learning in the classroom (Barral, Ardi-Pastores, & Simmons, 2018; Bhagat, Chang, & Chang, 2016; Lin & Hwang, 2018).

Bergmann and Sams (2012) insist that the success of their strategy depends greatly on how well activities in class provide students with a hands-on approach to learning. Therefore, for a flipped classroom to be truly effective, the activities presented in the classroom must follow a robust, interactive, and engaging format, such as inquiry learning or project based learning (Bergman, 2011). The work in the classroom is meant to put the video lecture in context, thereby reinforcing its necessity.

Despite the use of video lecture as a tool for completion of classroom work, student reluctance to complete homework is an important challenge of flipped learning. According to a study by González-Gómez and colleagues, (2016) some students engaged in a flipped classroom model admitted to not watching all of the video lectures, or sometimes even neglecting the homework altogether. Additionally, instructors who merely assign videos without checking comprehension have no way of knowing whether or not students watched the whole video, or whether they understood the content (Lee, 2018). Furthermore, many students may not have access to internet at home. This could be an equity problem, assuming that the student does not have smartphone access, or access to computer facilities at the school. However, this technology gap is disappearing through one-to-one device programs and moves toward implementation of the California statewide "No Child Left Offline" initiative (California Department of Education, n.d.).

Despite the best intentions of any instructor, the flipped model of learning cannot be effective if students begin classwork unprepared. Students are expected to complete their homework outside of the classroom, and the intervention will be compromised if they do not (Chuang et al., 2018). For this reason, student monitoring software, like EdPuzzle, can be beneficial. EdPuzzle allows teachers to embed comprehension questions into videos and also collects data on student viewing patterns such as time spent watching (EdPuzzle, n.d). By embedding formative assessments within the flipped class videos, there is an added assurance that students will come to class better prepared for the rigors of in-class instruction. These measures could also include short comprehension quizzes given to students before starting class activities (Leo & Puzio, 2016). Extra time with the videos can be given to students who did not master the content, while students who are ready to move on can work on in-class collaborative projects without having to review material that they are already comfortable with (Bergmann & Sams, 2012). This strategy will ensure that all students will have access to content before beginning classwork.

Another important consideration, especially when working in K-12, is the potential for cognitive overload. A study by Slemmons and colleagues (2018), found that the length of video shown as part of a flipped classroom had a significant effect on student ability to assimilate and use content to bolster achievement on assessments. Considering these findings, best practice would require assigning shorter videos, less frequently. Cognitive overload is especially of concern with ELLs and students who struggle with content (Perez & Holmes, 2010). These students have an extra load of work, as they must learn language along with content. Chunking information into smaller pieces gives all students more time to process new information (Bergman, 2011).

Flipped Learning as a Solution for ELLs

According to Bergman (2011), ELLs learn best in environments that include group work and hands-on activities. This environment is possible through flipped learning (Bergmann & Sams, 2012). Language learning in a flipped environment has already been studied in several contexts (Chuang et al., 2018; Lee, 2018; Lin & Hwang, 2018). For example, Lin and Hwang (2018) found that Taiwanese university students were able to use social media and videos in flipped learning to improve English oral performance over a control group in a typical lecture-based setting. Likewise, during Lee's 2018 study, English students in South Korea also mentioned the effectiveness of video instruction as a positive feature of flipped learning, as it allowed them to pause, replay, and take more notes as needed. Additionally, those same students felt that the structure of the flipped classroom model allowed them to ask questions of the instructors without feeling embarrassed (Lee, 2018).

Few studies have, however, addressed the effectiveness of flipped learning on ELLs in core content (i.e., science, history, and math). Studies have found that at-risk populations (e.g., ELLs) benefit the most from flipped learning (Altemueller & Lindquist, 2017; Lee, 2018). The reason for this is because the strategy allows ELLs to have extra needed time with the online curricular content (Gonzalez et al., 2016). Additionally, during class time, flipped learning provides increased access to instructor feedback and peer assistance (Lee, 2018). The student centered focus also removes fear of peer disapproval when asking for help, a common struggle for ELLs (Perez & Holmes, 2010). In core content like science, academic skills like question asking and seeking instructor feedback are essential to student success (Bergman, 2011).

Setting a Learning Environment for Middle School ELLs in Science

The importance of research on effective strategies for teaching ELLs in science should not be minimized. ELLs are at risk of falling behind their EO peers in all subjects, including science (National Council of Teachers of English, n.d.). This is because of the wide range of required knowledge and advanced demands on vocabulary in a typical science course. ELLs may also have gaps in science content knowledge, due to reduced time given for science instruction in elementary school (NSTA, n.d.). In order for science instruction to be implemented with fidelity for ELLs, new ways of thinking about teaching styles and classroom resources have to be considered. Resources which could fill the gaps left by the ELL transitions into mainstream classrooms are needed. Online curated resources like Stemscopes have been integrated into many school curricula (Stemscopes, n.d.). Stemscopes is a science instruction platform, which is directly aligned to the state-adopted Next Generation Science Standards (NGSS), and addresses each standard in an engaging format. Standards are cited and cross referenced, so that direct connections to any NGSS-aligned curriculum can be made (Stemscopes, n.d.). Some of the features which make products like Stemscopes an important resource for support of ELLs are the collection of high quality videos with available captions, lesson handouts in English and Spanish, as well as scaffolded reading options for students with limited academic vocabulary (Stemscopes, n.d.). These resources allow students to learn at their own pace, and therefore can be used to support the integration of a flipped class model.

Methods

The purpose of this study was to add to the growing literature on the subject of flipped learning. Specifically, this study explored if flipped learning was effective for ELLs in a middle school science classroom. Current research on flipped classrooms has been primarily conducted at the college and university level (Barral et al., 2018; Cabi, 2018; Chuang et al., 2018), although some studies have focused on middle and high school (Bhagat et al., 2016; Lee, 2018; Leo & Puzio, 2016). For ELLs, many have their first exposure to science in middle school, having spent the majority of time in elementary school working on English language acquisition (Umansky et al., 2015). The potential for frustration and cognitive overload from new information and disciplinary thinking continues to be a real concern for ELLs (Slemmons et al., 2018). While it was known that flipped learning had a positive effect in the context of language learning (Chuang et al., 2016; Lee, 2018; Lin & Hwang, 2018), there was a lack of research on the effectiveness of flipped learning specifically for ELLs who learn core content in mainstream classrooms.

Research Question

Does the flipped classroom model improve the academic achievement of ELLs in a seventh grade science classroom?

Hypothesis

Based on studies performed on the flipped classroom in the context of language learning (Chuang et al., 2016; Lin & Hwang, 2018), and in the context of middle and high school (Lee, 2018; Leo & Puzio, 2016; Slemmons et al., 2018), it was believed that the flipped learning model would have a positive effect on the academic achievement of ELLs in a seventh grade science class.

Research Design

This study utilized a two group, quasi-experimental, pretest/posttest research design. This design was chosen for its strong internal validity, and for the fact that many of the research articles reviewed in preparation for this study also followed this design (Altemueller & Lindquist, 2017; Bhagat et al., 2016; Leo & Puzio, 2016). True experimental design was not possible, due to the selected students being already assigned to their respective class sections. By nature of flipped learning, students had to receive the intervention as a whole class. The two groups were a treatment group that received the intervention and a control group that did not. Before the intervention, both groups took a pretest and then after the intervention both groups took the posttest.

Independent variable. The independent variable was the flipped classroom model (Bergmann & Sams, 2012). The flipped classroom model combined high quality video instruction outside the classroom with collaborative, engaging group practice and productive group work (e.g., group presentations, experiments, and research projects) inside the classroom (Leo & Puzio, 2016). The flipped classroom model upended the traditional classroom model of lecture during class with homework given as practice (Chuang et al., 2016). The lecture became homework in the form of videos provided by the instructor. Removing the lecture increased time in class given to students to practice what they learned. Instead of lecture, students worked in groups and partners, while the instructor was able to monitor and give feedback more quickly (Bergmann & Sams, 2012).

Dependent variable. The dependent variable was the academic achievement of the students. Academic achievement was conceptually defined as the ability of a student to meet established criteria (York, Gibson, & Rankin, 2015). Academic achievement was operationally defined in the context of this study as the achievement by students on the Stemscopes assessment module provided by the school district, as seen in Appendix A (Stemscopes, n.d.).

Setting & Participants

The school for this study was a Title I middle school in Central California, with over 80 percent of students qualifying for free and reduced lunches (Ed Data, n.d.). The percentage of ELL enrolment had increased in the last year from 40.9 percent to 49.7 percent. (Ed Data, n.d.) Students were part of the researcher's seventh grade integrated science class. Of the five class sections assigned to the researcher, four were selected for the equivalent size of the enrolled ELLs in each section. The use of a purposeful convenience sample of the students enrolled in the class assured that an appropriate number of ELLs were studied within the confines of the researcher's available students. The sample was convenient because students were selected from the researcher's available students. The sample was purposeful because students were selected based on the researcher's assessment of their academic achievement levels, and their status as ELLs.

Treatment group. The treatment group consisted of 24 ELLs enrolled in two sections of a science seven course. The group was comprised of 14 males and 10 females. Of these students, 23 had a native language of Spanish, and one of Tagalog. The students were learning alongside 36 EO peers who also received the intervention.

Control group. The control group consisted of 20 ELL students enrolled in two sections of a science seven course. The group was comprised of 13 males and 7 females. Of these students, all 20 had a native language of Spanish. They were learning alongside 30 EO peers, and none of these students received the intervention.

Measures

To measure academic achievement, a pre and posttest were administered. This pretest can be seen in Appendix A. These items were taken from the district approved online resource Stemscopes (Stemscopes, n.d.). Stemscopes provided assessment tools which were aligned to the California adopted Next Generation Science Standards (NGSS), which were taught in science seven courses throughout the school (NGSS, n.d.). Pre and posttests followed a pattern of weekly quizzes, to which the students were already accustomed. Each test contained six items, four multiple choice and two constructed response. Students were given 20 minutes to take the test on their provided Chromebook. Students who forgot their Chromebooks, or whose Chromebooks were not working were given a paper version of the test.

Validity. Test items were developed by a team of experts to align specifically to state-adopted NGSS standards upon which the school's curriculum is based (NGSS, n.d.). A two year study by Rice University, the originators of Stemscopes, found that utilization of the program, including the assessment bank, was associated with a statistically significant gain in eighth grade state science assessment scores (Rice University, 2012). This means the measure has strong internal validity. Also, since the assessments measure learning which is correlated with higher State standardized test score, the measure also shows external validity, in that the results can be generalized to other contexts.

Reliability. The Stemscopes assessment was not premade (i.e., selected items are not in a predetermined order). Instead, test items were selected from a bank of premade items which were aligned to the NGSS standards. After this, the instructor determined the order. The testing items selected were reviewed by two site based experts. The first was an expert in science instruction with over ten years of teaching experience. The second was an expert in ELL support, who also had a background in science instruction, and also had over ten years of teaching experience. All of the selected items were scored using an answer key also provided by Stemscopes. The measure can be seen in Appendix A (Stemscopes, n.d.). The researcher scored all tests. To further ensure reliability of scoring, the researcher also instructed another content expert to independently score 20 percent of the constructed response questions, for inter-rater reliability.

Intervention

In the study, students were divided into two groups, a control and a treatment group, of two class sections each. Sections were chosen for control and treatment in order to make the numbers of ELLs in each class section approximately match. Groups were also divided according to the researcher's estimation of proficiency level in science, for maximum equivalent of student ability. A pretest was administered. After this, the treatment group began a flipped classroom model, based on the parameters outlined by Bergman and Sams (2012). A total of 15 videos spaced out over four weeks were given as homework. During the class time, pre-class lecture was not given, and students worked collaboratively on projects pertaining to the content (e.g., labs, posters, group presentations, building scientific models).

Videos were not given daily out of concerns for cognitive overload in the treatment group (Slemmons et al., 2018). This was also an attempt to address concerns that students would avoid doing their homework if the amount was deemed to be significantly more than earlier in the year (González-Gómez et al., 2016; Lee, 2018). The control group also worked on the same projects, but were not given the video homework. Instead, they had direct instruction in class, followed by minimal group time to begin projects, being given the remainder of the unfinished work as homework. After the four weeks of intervention concluded, all students took a post test, in order to gauge the effectiveness of the treatment over the control group.

Procedures

For the introduction of the study, the unit began with a pretest. Following researcher analysis of the data, the control and experimental groups began their respective course. The control group had a daily lecture, followed by guided practice and minimal homework, as was the customary practice during the first semester of the course. The treatment group began with video homework to prepare for the collaborative group work that typified the flipped classroom model (Bergmann & Sams, 2012). Embedded comprehension questions were placed in the videos through the application EdPuzzle (EdPuzzle, n.d.). This ensured that students were paying attention to the video, and repeating the sections they did not understand. This application also collected data on the viewing habits of the students (e.g., total time spent viewing, what sections were repeated).

At the beginning of each class, all students (treatment and control) took a short, multiple choice comprehension quiz as part of an already established daily routine. This quiz was another way to ensure that students in the treatment group were comprehending the homework, and the students in the control group were comprehending the lecture. If students in the treatment group could not answer the questions, they were sent to listening stations for further practice before beginning the group activities. Collaborative activities aligned to the content (e.g., labs, group presentations, and building of scientific models) were assigned to the whole class, with the instructor checking in on groups, monitoring progress, and answering questions as needed. At the end of the four week unit, all students took the post test, and data was analyzed to see if there was a statistically significant difference between the two groups.

Fidelity. In order to ensure fidelity, a team of two independent observers observed 20% of the class during intervention to make sure that conditions of the intervention were being met, and that there was no crossover between intervention and control (see fidelity checklist in Appendix B). The control observer was checking for a class-wide lecture before activities, and lack of productive group work within the context of the class. In contrast, the intervention observer was checking for productive group work, with no lecture component, and instructor feedback on an individual basis, as opposed to class-wide. This ensured that the intervention was implemented with 100% fidelity.

Ethical Considerations

Though the treatment was simple in design, there were a few ethical considerations to be addressed. Equal access was a consideration for this type of study. The video section of the assignments required access to technology and to the internet. Since the school district where this study took place had a one-to-one Chromebook policy, students had access to the technology. As a solution to internet access problems, the students had the option of renting a free Wi-Fi hotspot from the school to be used as long as they have need, up to the entire school year. Any student who was found to be without internet access at home was directed to these resources. Additionally, there were listening stations set up for review of materials just in case a student was somehow still unable to access the content.

One other possible ethical consideration is that either the control or the treatment group may have shown significant improvement in academic achievement over the other. If this was the case, the most effective strategy was to be adopted for all students enrolled in the course immediately following the final analysis of data.

Validity threats. One potential threat to validity for this study was the problem of participant attrition. The population of the district where this study took place was dynamic, and some students dropped out of the study in the middle, or were absent for a portion of it. This is the main reason why the study took place over a period of four weeks, and not a longer period of time. The shorter length of the study also addressed problems of maturation of subjects as a threat to validity (McMillan, 2016). Selection bias could have also played a role in this study. Students were selected as a whole class, as opposed to true random assignment, as was necessary for true experimental design. Pretests made sure that selected groups were not statistically different.

The diffusion effect was another real concern, as students were known to talk to each other about the daily lessons. For this reason, consecutive classes in the mornings and afternoons were selected, so that the lunch period would serve as a buffer, minimizing students' chances to interact and share information about class. This did not completely remove the possibility of the diffusion effect, so class assignments were also designed to be as similar to each other as the format allowed, to lessen the chance of students noticing the difference.

One further threat to validity which needed to be addressed is the fact that fidelity of intervention lay to some degree with students as well as the instructor. For flipped learning, students could not receive the effect of the intervention if they did not complete their homework. Lee (2018) found that at the beginning of a similar intervention, students in South Korea were not watching the videos at home, thereby reducing the effectiveness of the intervention. Other researchers, such as González-Gómez and colleagues, (2016) also found similar results. For this reason, each student took a short, multiple choice comprehension quiz at the beginning of the day, to make sure that students comprehended the content covered in the videos and were able to complete the in-class projects without needing a lecture. Listening stations were set up around the classroom so that students who did not pass the quiz could repeat the videos and then join their groups after they had caught up. This strategy was proposed by Lee (2018) for their flipped classroom investigation.

Data Analyses

All data were entered into the Statistical Package for the Social Sciences® (SPSS®) for Windows, version 24.0.0 (SPSS, 2016). No names or identifying information were included in the data analysis. Before analyses were conducted all data were cleaned to ensure no outliers were present (Dimitrov, 2012). After cleaning the data, independent and paired samples t-tests were conducted to determine the significant difference in between pre-test and post-test performance on the Stemscopes test in the control group or the treatment group. Further, before interpreting the analytical output, Levene's Homogeneity of Variance was examined to see if the assumption of equivalence has been violated (Levene, 1960). If Levene’s Homogeneity of Variance was not violated (i.e., the variances were equal across groups), data would be interpreted for the assumption of equivalence; however, if the variances were not equal across groups the corrected output would be used for interpretation.

Results

Two independent samples t-test were conducted on the whole sample ($n = 44$) for both the pre and post assessment scores. Results for the pre-test were: Levene's Homogeneity of Variance was not violated ($p > .05$), meaning the variance between groups was not statistically different and no correction was needed and the t-test showed non-significant differences between the mean scores on the pre-tests between the two groups $t\ (42) = .206$, $p > .05$. This means that the two groups could be compared without further adjustments (see Table 1). Results for the post-test were: Levene's Homogeneity of Variance was not violated ($p > .05$), meaning the variance between groups was not statistically different and no correction was need and the t-test showed non-significant differences between the mean scores on the post-tests between the two groups t (42) = -.190, p> .05. This means that there was no statistically significant difference between the post-test scores of the treatment and control. (see Table 1). Though both groups increased their scores, and the control group had a larger increase than the intervention group, the difference between their scores was not statistically significant enough to declare one teaching style as superior over the other.

Table 1

Results of Independent Samples T-Tests

	Mean	SD
Pre Test		
Treatment	3.92	2.02
Control	3.80	1.67
Post Test		
Treatment	4.04	1.98
Control	4.15	1.75

Note. SD = Standard Deviation.

After determining the differences between pre and post assessment scores between groups, two paired t-tests were run for both groups (i.e., treatment and control) to determine if participants mean scores from pre to post were significantly different within each group (See Table 2). Results for each group were as follows: treatment group, t (23) = -.293, p> .05; control group, t (19) = -1.095, p> .05. This means that both control and treatment group scores showed a change in mean scores from pre to post-test. Additionally, the negative t-value for each group indicates an increase in scores from pre to post assessment. However, the increase in scores from pre to post- test were not significant. This means that the hypothesis is rejected, as the flipped classroom did not indicate an increase in the academic achievement of the ELLs in the treatment group. Furthermore, the difference between the two group's post-test scores was not statistically significant, indicating that either teaching methodology may be viable for use in a middle-school science classroom.

Table 2

Results of Paired T-Tests

	Mean	SD
Treatment Group		
Pre	3.92	2.02
Post	4.04	1.98
Control Group		
Pre	3.80	1.67
Post	4.15	1.75

Note. SD = Standard Deviation.

Discussion

Given the struggles of ELLs in content classrooms (Janzen, 2008), it is reasonable that more research on alternative instructional methods be made for ELLs in content areas. According to Umansky and colleagues (2015) ELLs are more likely than their EO peers to be behind grade level in science due to increased emphasis on English instruction at the elementary school level. The fact that a large number of ELLs have not been given time to develop science literacy in elementary school makes middle school critical in the educational career of an ELL. Therefore, more research is needed in a middle school context, where many ELLs will find their first academic exposure to science.

The purpose of this study was to see if the flipped learning model would be effective in raising the academic achievement of ELLs in a middle school science class. The flipped classroom reversed the roles of homework and class work, as compared to the traditional model

(Chuang, Weng, & Chen, 2018). Instead of giving a lecture in class explaining science content, the lecture was given as homework for students to watch at home. These videos replaced the lecture portion of the class entirely. The purpose of the flipped class room, according to the originators of the intervention was to increase time for student work in class (Bergmann & Sams, 2012). The strategy had been thoroughly examined with high school and university level students (Barral et al., 2018; Cabi, 2018; Chuang et al., 2018); however, less research has been conducted with middle school students.

In order to test the effectiveness of the flipped classroom, a two group, quasi-experimental, pretest/posttest research design was employed. The control group (n=20) received the traditional model of coursework, including minimal homework, a classroom lecture, and the remaining time for individual or partner work. The treatment group (n=23) received the flipped classroom strategy, which included videos to watch as homework, as well as productive group work for in-class time. Both groups were observed 20 percent of the time to ensure fidelity of intervention. A pre-test and post-test were administered to both groups. The results indicate non-statistically significant improvement for both groups; however a greater increase in academic achievement was observed in the control group. In comparing the pre and post-tests of the treatment, the mean score went from 3.92 to 4.04. Even though the mean of the control group started at a lower value of 3.80, their post test score increased to 4.15. This means that they had a greater change in post-test scores when compared to pre-test scores.

There are several reasons why the control group may have showed a greater increase in mean post-test scores over the flipped classroom. One potential reason was students in the 7th grade struggled to complete homework. Towards the beginning of the intervention, students were able to complete the homework, according to analytics collected by the educational application

Edpuzzle (Edpuzzle, n.d.). As the weeks went on, however, student engagement with videos dropped off dramatically, to the point that on the last week of the intervention, on average, only about one third of the students were actually completing the videos. This likely had a significant effect on student progress. Whereas the students in the control group were receiving teacher instruction every day, students in the treatment group would not receive it if they did not watch the videos. Both Lee (2018) and Gonzalez-Gomez and colleagues (2016) experienced this in similar studies. However, their results showed higher gain to the treatment group, as did the majority of the literature reviewed for this study (Altemueller, & Lindquist, 2017; Barral et al., 2018; Bhagat et al., 2016; Chuang et al., 2016 Leo & Puzio, 2016).

Consistent with the findings of this study, Cabi (2018) found that the control group outperformed the treatment group in their study. The study found that students were resistant to study the content outside of the classroom, and therefore often entered the classroom unprepared to complete the activities (Cabi, 2018). This information aligns with the observations in this study of student video engagement habits.

For ELLs, who have not had access to scientific skills, there may also be a lack of foundational skills needed for independent work (Umansky, 2015). Independent work and productive group work were integral to the intervention, and lack of appropriate training to work in groups would have had an effect on the final outcome. The videos used in the study were primarily content-based, and did not train students for the skill of working in a group. Though the students had a brief two-day training on productive group work, this may have been insufficient to help them adjust to the new style of class. Conversely, for the control group, there may have been an added benefit to the live modeling of academic language and skill that comes with traditional classes which include a lecture. Modeling of academic language is a key strategy for effective instruction of ELLs (Bergman, 2011). This may be another reason why the control group slightly outperformed the treatment.

It should be noted that the difference between the two groups is not statistically significant. The mean difference between pre and post-test for the intervention was -.293, while the mean difference between pre and post-test for the intervention was -1.095, which is not a significant difference. This means that both the flipped classroom and the traditional classroom are viable strategies for instruction, provided that they are implemented with the appropriate structure and scaffolds. In the literature reviewed, a majority of the studies found that both the treatment (i.e., flipped classroom) and control (i.e. traditional lecture) groups demonstrated an improvement on their post-test scores (Altemueler & Lindquist, 2017; Barral et al., 2018; Bhagat et al., 2016; Cabi, 2018; Chuang et al., 2016). This means that neither strategy should be fully abandoned, as they both had a positive effect on achievement. Therefore, the relative equivalency of the student post-test scores is evidence that either strategy would work in the classroom. This is important, because the study shows that the flipped classroom is viable, even though it may not produce results which are statistically significant over the traditional method.

Limitations and Recommendations for Future Studies

One important limitation of this study was due to the dynamic nature of the population of the students at the school site. Several students, in both treatment and control, moved out of state and out of the country during the course of this study. Other students moved in, and many lost a significant amount of time due to temporary relocation. During the time of the study, three students were also suspended for multiple days. Two students were in the treatment group, and one was in the control. All of the students suspended in the course sections used for this study were classified as ELLs. In order to address the problem of participant attrition, a four-week study was proposed. However, the problems of participant attrition were present nonetheless. Therefore, for future study, it is recommended that a longer period of time is utilized, perhaps a semester, or an academic year. Through a longer study, the trend of learning in a flipped classroom could be studied further.

Another limitation of the study was the small sample size. A total of 43 participants were spread out into two groups. With a sample of this size, it is difficult to generalize the effect to other contexts. In this case, one or two low test scores would have a significant effect on the final scores for the whole group, given the size. It is therefore further recommended that studies on a larger scale be carried out for results which can be generalized more readily.

A third limitation of the study was the length and sensitivity of the measure. The measure was purposefully short, to avoid the problem of cognitive overload (Slemmons et al., 2018). However, since the measure contained only six items, and had a total of ten points available, it is possible that a scale with more questions would have shown more precise results. Therefore, for future studies, it is recommended that a measure with greater sensitivity be developed.

One final limitation of the study which needs to be addressed is the lack of engagement of the students with the videos. This was an issue of concern mentioned in several studies reviewed (Cabi, 2018; Gonzalez-Gomez et al., 2016; Lee, 2018). Bergman and Sams (2012) also mentioned that the videos as a form of content instruction did not work for everyone. In fact, they mentioned that as their strategy evolved, videos slowly became one of many sets of tools for teaching content, also including textbooks and teacher mini-lectures, when needed (Bergmann & Sams, 2012). Therefore, a mixture of both the traditional and the flipped method may be most beneficial for all students, and could address the problem of lack of student engagement with the videos.

In conclusion, this study showed that both the traditional method, and the flipped classroom may be employed with middle school science classrooms. Although the traditional classroom showed greater improvement over the flipped classroom, the difference was not statistically significant. More studies should be conducted with middle school ELLs students to better understand the ways it can potentially prepare them for the future. Other instructional strategies should be tested as well, to ensure that the science knowledge gap is closed between ELLs and EOs in middle school.

References

Altemueller, L. & Lindquist, C. (2017). Flipped classroom instruction for inclusive learning. *British Journal of Special Education, 44*(3), 341-358. doi:10.1111/1467-8578.12177

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.

Barral, A. M., Ardi-Pastores, V. C., & Simmons, R. E. (2018). Student learning in an accelerated introductory biology course is significantly enhanced by a flipped-learning environment. *CBE—Life Sciences Education, 17*(3). doi:10.1187/cbe.17-07-0129

Bergman, D. (2011). Synergistic strategies: Science for ELLs is science for all. *Science Scope, 35*(3), 40-44. Retrieved from https://www.jstor.org/stable/43183155

Bergmann, J., & Sams, A. (2012). Flip your students learning. *Educational Leadership*, *70*(6), 16-20.

Bhagat, K. K., Chang, C. N., & Chang, C. Y. (2016). The impact of the flipped classroom on mathematics concept learning in high school. *Educational Technology & Society*, *19*(3), 134-142. Retrieved from https://www.jstor.org/stable/jeductechsoci.19.3.134

Bloom, B. S. (1968). Learning for mastery. *Instruction and Curriculum, 1*(2), 1-11. Retrieved from https://eric.ed.gov/?id=ED053419

Cabı, E. (2018). The impact of the flipped classroom model on students' academic achievement. *The International Review of Research in Open and Distributed Learning, 19*(3), 202-221. doi:10.19173/irrodl.v19i3.3482

California Department of Education. (n.d.). California Education Technology Blueprint. Retrieved from https://www.cde.ca.gov/eo/in/edtechbi.asp

Chuang, H., Weng, C., & Chen, C. (2016). Which students benefit most from a flipped classroom approach to language learning? *British Journal of Educational Technology, 49*(1), 56-68. doi:10.1111/bjet.12530

Dimitrov, D. M. (2012). Statistical methods for validation of assessment scale data in counseling and related fields. Alexandria, VA: American Counseling Association.

Ed Data. (n.d.). California Public Schools. Retrieved from https://www.ed-data.org/

Edpuzzle. (n.d.). Welcome to Edpuzzle. Retrieved from https://support.edpuzzle.com/hc/en-us/articles/360007260152-Welcome-to-Edpuzzle-

González-Gómez, D., Jeong, J. S., Rodríguez, D. A., & Cañada-Cañada, F. (2016). Performance and perception in the flipped learning model: An initial approach to evaluate the effectiveness of a new teaching methodology in a general science classroom. *Journal of Science Education and Technology, 25*(3), 450-459. doi:10.1007/s10956-016-9605-9

IBM Corp. Released 2012. IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp.

Janzen, J. (2008). Teaching English language learners in the content areas. *Review of Educational Research, 78*(4), 1010-1038. doi:10.3102/0034654308325580

Lee, M. (2018). Flipped classroom as an alternative future class model?: Implications of South Korea's social experiment. *Educational Technology Research and Development, 66*, 837-857. doi:10.1007/s11423-018-9587-9

Leo, J., & Puzio, K. (2016). Flipped instruction in a high school science classroom. *Journal of Science Education and Technology, 25*(5), 775-781. doi:10.1007/s10956-016-9634-4

Levene, H. (1960). *Contributions to probability and statistics; essays in honor of Harold Hotelling.* Stanford, CA: Stanford University Press.

Lin, C. J. & Hwang , G. J. (2018). A learning analytics approach to investigating factors affecting EFL students oral performance in a flipped classroom. *Educational Technology & Society*, *21(2), 205-219.* Retrieved from https://www.jstor.org/stable/26388398

McMillan, J. H. (2016). *Fundamentals of educational research* (7th ed.). Boston: Pearson.

National Council of Teachers of English. (2008). English Language Learners. [PDF file]. Retrieved from http://www.ncte.org/library/NCTEFiles/Resources/PolicyResearch/ELLResearchBrief.pdf

Next Generation Science Standards. (2018). The Standards. Retrieved from https://www.nextgenscience.org/search-standards

NSTA. (n.d.). *Introduction*. Retrieved from https://www.nsta.org/about/positions/ngss.aspx

Perez, D., & Holmes, M. (2010). Ensuring Academic Literacy for Ell Students. *American Secondary Education,38*(2), 32-43. Retrieved from http://www.jstor.org/stable/41406160

Piaget, J (1928). La causalité chez l'enfant. *British Journal of Psychology*, *18.* 276–301. doi:10.1111/j.2044-8295.1928.tb00466.x.

Rice University. (2012). Evidence supporting the efficacy of the STEMscopes™ science curriculum: Two-year comparison of science scores in a medium-sized district in Texas. *Rice University Center for Digital Learning and Scholarship*.

Slemmons, K., Anyanwu, K., Hames, J., Grabski, D., Mlsna, J., Simkins, E., & Cook, P. (2018). The impact of video length on learning in a middle-level flipped science setting: Implications for diversity inclusion. *Journal of Science Education and Technology,27*, 469-479. doi:10.1007/s10956-018-9736-2

Stemscopes. (n.d.). Home. Retrieved from https://stemscopes.com/

Umansky, I. M., Reardon, S. F., Hakuta, K., Thompson, D. D., Estrada, P., Hayes, K…, Goldenberg, C. (2015). Policy brief [PDF file]. Retrieved from https://edpolicyinca.org/publications/policy-briefs

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* Cambridge, Mass.: Harvard University Press.

York, Travis T., Gibson, Charles, & Rankin, Susan. (2015). Defining and measuring academic success. *Practical Assessment, Research & Evaluation, 20*(5), 1-20. Retrieved from http://pareonline.net/getvn.asp?v=20&n=5

Appendix A

Pre-test and post-test items from Stemscopes

Mid-Quarter Quiz

* Required

1. **First Name** *

2. **Last Name** *

3. **Period** *

Mark only one oval.

- 1
- 3
- 4
- 5
- 6

4. **Student ID Number** *

5. **1. Matter is passed from the living part of an ecosystem to the non-living part when which of the following takes place?**

Mark only one oval.

- Plants absorb gasses
- Plants decompose
- Animals eat plants
- Animals drink water

6. **2. An example of a predator and a prey species for four separate ecosystems is provided. Which of the following would demonstrate an appropriate prediction for the changes to one of the populations of organisms due to a predator/prey interaction?**

Ecosystem	Predator Species	Prey Species
1	Dragonfly	Mosquito
2	Shark	Small fish
3	Mountain Lion	Deer
4	Snake	Mouse

Mark only one oval.

- ◯ When the population of sharks decreases because of fishing, the population of small fish will decrease.
- ◯ After a period of heavy rainfall and warmer temperatures, the population of mosquitoes will increase.
- ◯ During hunting season, when many deer are killed by hunters, the mountain lion population will decline.
- ◯ When abundant resources allow the population of mice to increase, the population of snakes will decrease.

7. **3. Refer to the tables provided. Which plants depend on and may compete for sandy soil in Zone 9?**

Soil requirements of East Texas Trees			
Species	**Texture**	**Water Retention**	**Calcium Content**
Black Hickory	Sandy	Low	High
Chinese Elm	Sandy	Low	Average
Live Oak	Loamy Sand	Average	Average
Loblolly Pine	Sandy Clay	High	Low
Post Oak	Loamy Sand	Average	Average
Sweetgum	Sandy	Low	High

East Texas Soil Types			
Climate Zone	**Elevation Above Sea Level**	**Texture**	**Calcium Content**
9	0-40 feet	Sandy	High
8	40-200 feet	Loamy Sand	Average
7	200-1000 feet	Sandy Clay	Low

8. **4. Describe how introducing a non-native species of a plant can reduce an area's biodiversity over time.**

9. **5. The chart below lists data on four different projects designed to restore a wetlands habitat destroyed by human activity and a recent hurricane.The project that will be chosen will have to be low in cost, benefit the greatest number of species, and be popular with the local citizens whose taxes will sponsor it. Based on the data and the criteria given, which project will be chosen?**

Project	Anticipated cost (millions of dollars)	% of Wildlife to Benefit	Community Support
1	13–16	15–25	High
2	25–30	50–75	Moderate
3	12–17	70–80	High
4	18–25	75–80	Low

Mark only one oval.

- Project 1
- Project 2
- Project 3
- Project 4

10. **6. Information on brown tree snakes is provided along with information on the island of Guam.Brown tree snakes were first brought to Guam in 1950.Which of the following would provide evidence that the change to this environment brought about by the introduction of the brown tree snake has had an effect on the populations of organisms that lived on Guam?**

Brown Tree Snakes
- Venomous and aggressive
- Live in trees
- Hunt at night
- Eat lizards and small birds

Island of Guam in 1950
- Home of growing US military base
- Jungle ecosystem
- No snakes
- Large lizard population

Mark only one oval.

- Many of the military bases on Guam were closed in the 1980s.
- Several other species of snakes were known to live on Guam before 1950.
- The population of lizards on Guam in 1960 was very small.
- The longest brown tree snake ever found on Guam measured 3 meters in length.

Appendix B

Fidelity Checklist

Control Group			
Week of observation	Date observed	Traditional Classroom (No group work, classwide lecture)	Initial

Treatment Group			
Week of observation	Date observed	Flipped Classroom Setup (Small group work with no lecture)	Initial