Does the Flipped Classroom Lead to Increased Gains on Learning Outcomes in ESL/EFL Contexts?

This research investigates whether the flipped classroom can lead students to increased gains on learning outcomes in 2 ESL/EFL contexts in Macau, China, and the US. A pretest posttest quasi-experimental mixed-methods design ($N = 64$) was used to determine any differences in student achievement that might be associated with the flipped approach (FA). The effectiveness of the FA on students’ achievement with grammar-student learning outcomes was evaluated with a pretest and posttest grammar test, along with students’ perceptions of their increased comfort and confidence using English grammar through a grammar survey. These data were triangulated with student focus groups and means of completed grammar assignments. The findings suggest that although both the control and experimental groups showed increased comfort in the self-report data, gains on actual achievement were significant only for the experimental groups. The researchers of this study make recommendations for a flipped curriculum and materials design for ESL/EFL teachers in any context globally.

Introduction

By now, everyone in education has probably heard of the flipped classroom. The flipped model is gaining attention among educators from all levels and fields around the world. According to the definition provided in The Flipped Learning Network (2014), a flipped classroom is:

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

In the flipped model, videos are used to deliver instruction to students outside of class, thus freeing class time for hands-on learning, individualized instruction, group collaboration, and creative projects in order to master learning objectives. Through these means, students receive extensive and intensive language input and are able to work at their own pace. Although the flipped classroom was originally a pedagogical approach used in the K-12 system, it has become widely applicable to the tertiary level. Therefore, we can say it is also an andragogical approach to teaching and learning. For language learning, the skills that children use are much different from those that adults use, so it is important to make this distinction but to show that the flipped approach is applicable to both contexts.

In response to the popularity of the flipped paradigm, a plethora of online sources have become available for teachers in many fields and disciplines to get started with adopting a flipped approach in their classes. Such online sources vary from websites and blogs to guidebooks geared toward teachers who may be unfamiliar with the concept of flipping. Likewise, several books have been written that are rich in content, such as Bergmann and Sams (2012, 2014), Bretzmann (2013), and Fulton (2014).

Despite the vast amount of materials available about the flipped approach, empirical evidence in support of the approach is seemingly lacking. Even less is available regarding the ESL/EFL classroom because much of what flipping entails is similar to what language teachers have been doing for decades. Besides Kate Petty (2015) and Troy Cockrum (2014), there are few sources that ESL/EFL teachers can turn to for hints and concrete lessons for using in a flipped classroom. These limited resources, while good starting points, offer no empirical support of the flipped model. Therefore, this is an area of the research that is missing. In particular, while research into student perceptions of the flipped model have been undertaken, few researchers have questioned the effects of the approach on student achievement.

The current investigation, therefore, aims to address the gap in the research in the area of achievement of students in flipped classrooms by addressing the following research questions, which are related to meeting student learning outcomes (SLOs) in grammar.

RQ 1: Do students in a flipped ESL/EFL classroom have a significant difference in gains on grammar SLOs compared to students in nonflipped ESL/EFL classes?
RQ 2: Do students in flipped ESL/EFL classes and nonflipped ESL/EFL classes experience changes in attitudes toward their grammar skills through time?
RQ 3: Do students in flipped ESL/EFL courses differ from students in nonflipped ESL/EFL classes in their perceived grammar skills?
RQ 4: Do ESL/EFL students in a flipped classroom have student buy-in with the flipped methodology?

How the Flipped ESL/EFL Classroom Differs From the Traditional Classroom

There are various arguments that the flipped language classroom is not that different from the traditional language classroom. In fact, since the 1980s with the advent of communicative language teaching (CLT), language classrooms have been dominated with approaches that allow students to receive a variety of input and then be given ample opportunities to produce output. For example, teachers often assign passages or texts to be read at home so that in class, students can engage in discussion on the topic of the reading assignments. To the researchers of this study and many like-minded researchers who have flipped their language classrooms, this is not an example of flipping. Let us explain why.

When a student takes a text and reads it at home, the teacher has no idea about the student’s ability to understand or comprehend the material. The teacher is unaware of the skills that the student may be lacking to comprehend the text and is thus unable to identify the areas in which the student needs assistance. As reading is process oriented, rather than product oriented, it is a skill that needs to be practiced inside the classroom, but not by traditional means of reading silently. Rather, it can and should be done through interactive and engaging activities in which students seek and provide details, fill in knowledge that is lacking, and infer and predict outcomes. Several reading activities include jigsaw reading,5 running dictations,6 and activities that require students to annotate, summarize, or respond to texts in class under the supervision of the teacher and often with the help of pair or group mates. Outside of the classroom, reading can be done on the learning-management system, as students in flipped classes are often required to post and read comments on certain topics or themes on discussion forums or through blogs.

Under the flipped approach, the aspect that is moved outside the classroom is the amount of teacher-talk. This can include the description of an upcoming assignment, details of a rubric, or scaffolding of a skill by showing the teacher doing an assignment through a talk-
aloud method and recording it. This can also include a recording of a brief lecture that the teacher thinks the students need to be aware of. Screencasts and videos are excellent platforms for this, as they can be rewound, fast-forwarded, or played as many times as necessary. Students can get the information they need at home to better prepare themselves for the upcoming lesson.

For the current study, the effects of the flipped approach on the learning of grammar are measured. Again, under the CLT model, grammar is not taught explicitly, especially at the university level. However, this is one fault that many researchers posit for CLT—that there is no room in the model for any explicit teaching (Savage, 2010). Yet as L2 learners in the K-12 system have been taught grammar as well as the four language skills explicitly, it is a sudden shock for them as university students to suddenly be required to fill in the lack of knowledge individually through independent study. Thus, it can certainly be argued that explicit teaching still has its place in the classroom, albeit in limited amounts. The flipped classroom allows students to receive that explicitly taught lesson, but without sacrificing classroom time, which is usually very limited and precious in the context of language learning. And as mentioned above, students can review those videos as much as they need until they finally feel that they have acquired the prerequisite knowledge they need to be successful.

One of the biggest challenges for teaching language and almost any subject is individualized instruction. Students in almost every context are coming in with different knowledge bases and levels of motivation that affect any classroom. Though motivation was not a question in this study, researchers agree that motivation is tied into formative assessment (Cauley & McMillan, 2010). Thus the researchers in this study seek to define the flipped classroom as a classroom in which students receive formal input through teacher-recorded video lessons and then further engage in a high level of low-stakes formative assessment outside of the classroom. The flipped model in this study was characterized by such spontaneous feedback from teachers and students outside of the classroom through formal discussion posts. Students were encouraged to point out flaws or make other positive comments on their grammar discussion posts just as students do naturally inside of the classroom, though such grammatical conscious-ness raising can be more surprising for students when it is in a “public space” online for everyone to see and remember. All teachers hope for such asynchronous learning for students when they leave the classroom—that they are learning in an independent way; flipping helps to foster such learning habits by placing the responsibility on them not only inside the classroom but at home in a way that is more scaffolded.
The difference between low-stakes quizzes conducted outside of versus inside the classroom is that students have access to their teachers’ input at a moment’s notice while reviewing for the quiz. Many teachers allow students chances to make corrections on quizzes that students took in class. The flipped methodology can incorporate just that by allowing students to retake their online quizzes two times. Critics might say that students can easily cheat on quizzes online, though most learning management systems (LMSs) now have multiple techniques to lower the feasibility of students’ copying their answers. Again, the concept is that students are engaging in extra practice outside of the classroom while simultaneously receiving formal feedback through computer-generated scores or discussion posts.

**Review of the Literature**

*Flipped Research Outside the Humanities*

Given the increasing interest in the effectiveness of the flipped classroom, Bormann (2014) undertook an investigation into 30 peer-reviewed journal articles that explored the effects of the flipped classroom on student preparedness and gains in achievement. Few of Bormann’s findings produced studies that were of a quantitative nature; rather, more were derived from mixed-methods or qualitative studies. Bormann was able to categorize his findings according to three major themes:

1. Students who prefer the flipped model over a traditional model;
2. Student perceptions of engagement in flipped learning; and
3. Student achievement of learning outcomes.

A majority of the empirical studies that Bormann (2014) found reported student preferences for the flipped model over a more traditional model of teaching. Some of the findings from these studies were that the flipped classroom reaches all learners, the flipped environment promotes student empowerment, and the flipped model leads to student preparedness, addresses critical-thinking skills, and promotes computer literacy skills.

Included in an extensive literature review on students’ satisfaction or perceptions of the flipped approach, Doman and Webb (2014, 2015) found that most studies on student perceptions about the flipped approach were overwhelmingly positive. Papadopoulos, Santiago-Roman, and Portela (2010), after administering a survey to students, found the approach to be worthwhile, lessons to be useful.
and interactive, and the flipped classes to be preferable to traditional lecture-only classes. These findings are comparable to those of Pedroni and Meyer (2006), who collected data and found that students’ satisfaction with the software used to run the online part of the class increased from 2.7 to 2.9 based on a 1-5 Likert scale. Finally, Zappe, Leicht, Messner, Litzinger, and Lee (2009) found that students preferred only about half of the classes flipped so as also to allow a reasonable amount of time for traditional lectures. By using video lectures, Zappe et al. (2009) found that the online component allowed for increased teacher-student interaction and that students were willing to use out-of-class time to watch the videotaped lessons, with many of them watching sections of the videos multiple times. A majority of their students thought that the flipped classrooms were helpful, and most agreed that the additional time spent in class working on problem-solving activities greatly enhanced their understanding of the concepts.

Another major theme presented in the literature that Bormann (2014) described was that the flipped approach helped to promote student engagement. Because most knowledge in the flipped model is delivered by video clips that students watch at home before attending class, class time can be spent more effectively, allowing students to synthesize the information they learned, to ask questions as they come up, and to work out problems collaboratively. As cited in Bormann (2014), 73.6% of the students in Davies, Dean, and Ball’s (2013) study, 75% in Gaughan’s (2014) study, 79% in Murphree’s (2014) study, and 80% in Willey and Gardner’s (2013) study regularly completed flipped tasks outside the classroom, which was shown to be a measure of engagement.

Finally, a handful of isolated case studies reported increases in student outcomes from flipped learning environments. Several studies reported no significant difference between the achievements of students in the flipped versus the nonflipped classes. Though the studies reported no significant differences in student achievements, there was still strong support for the flipped model as students in the flipped classes did report higher perceptions of a better learning environment than did their fellow students in the nonflipped classes. Farah (2014), however, did report the success of the flipped classroom on the writing performance of high school learners in the Emirates in an EFL context. By observing significant differences in writing attainment of students in the flipped class over the nonflipped class, Farah provided data in support of the flipped approach as a means for increasing student achievement.
History of Flipped ESL/EFL Classes

When one is learning a new language, a range of methods and approaches are possible, and the flipped classroom has been one alternative for the past 10 years. As we know, achieving the optimal learning experience is a challenge in the classroom where time and resources are limited, despite our best efforts to introduce communicative, interactive, and authentic tasks. In addition, students often come in with mixed language abilities, different goals, and various learning styles that often cannot be addressed in the confines of the classroom. To provide more effective learning, tools and resources that can be used not only inside the classroom but also outside the classroom come in handy.

In what was first known as “blended” learning in 2000, the traditional language classroom was supplemented with self-study e-learning materials. Computer-assisted language learning (CALL) came with the arrival of the Internet, which contained a variety of self-study websites. Independent learning centers began to appear all over the world as warm environments where students could study languages alone in comfortable and safe locations. Televisions, video recorders, newspapers, magazines, DVDs, cassette recorders, and language textbooks were often located throughout these types of independent learning centers so that students would have a world of language input right at their fingertips. Learning management systems, such as Blackboards, WebCT, Moodle, Desire2Learn, and Canvas, popped up everywhere—from primary schools to universities—where teachers could place learning materials online for students to access 24/7 from any location.

CALL has gradually given way to the newer terms technology-enhanced language learning (TELL) and mobile-assisted language learning (MALL), as the computer is not the only device through which languages can be learned. Mobile devices such as smartphones and iPads are now common, and in many contexts, every student has at least one of these gadgets.

Today, many tools are used to enhance students’ learning. Word-processing software, such as Google Documents, is used to engage students in collaborative writing. Wikis and blogs are used to help develop students’ writing skills. Facebook, Twitter, and messaging forums are used to create discussions on topics of interest and to allow students to feel part of a greater learning community. Online books with quizzes built in help students to practice their comprehension skills. An abundance of authentic podcasts or videos (such as TED talks) give students access to input for listening practice and engage
them in higher-order thinking skills on controversial or current topics. And this list goes on and on.

Flipping the language classroom is one branch of TELL and MALL. The flipped classroom provides a digital solution to the lack of time that obstructs us in the classroom. More learning can be done outside the classroom. With an overwhelming amount of resources available today, it is up to teachers to find ways to integrate these technologies into the learning experience. Flipping the class is one solution.

As student achievement of grammar skills is the topic of investigation for this study, the researchers wanted to observe any correlation between students’ perception of their grammar skills to that of their actual performance in grammar (through tests). This study will seek to fill the gap in the literature about flipping in language courses and whether or not students actually perform better in meeting course outcomes.

**Methodology**

**Participants and Context**

This investigation of the flipped model in ESL/EFL contexts is a case study that is part of a larger two-year experiment in Macau and the US. Research was undertaken at a two-year community college in the US and the English Language Centre (ELC) at a university in Macau, China. All incoming students in Macau are required to take a yearlong English language course at the intermediate level as part of their requirements for General Education (GE). Most students complete their English studies during their first year, though some students take two years to complete this requirement if their proficiency is below the intermediate level upon entering the university. In the US, students who identify as ESL learners and who wish to take credit-bearing English courses are tested and placed into one of two developmental composition courses before entering English 100—necessary for advancement to transfer-level course work.

The participants in this study are considered high-intermediate EFL learners in Macau, and high-intermediate ESL learners in the US. Some of the students in the US are Generation 1.5. Additionally, some students in the US are parents or working professionals who returned to school at a later age.

Although the contexts of learning for ESL students are very different from those of EFL students, the researchers found the proficiencies and the needs of the students in both contexts to be similar. Figure 1 provides more information about the participants in this study.
<table>
<thead>
<tr>
<th>Program type</th>
<th>Class type</th>
<th>Ages</th>
<th>Nationalities</th>
</tr>
</thead>
<tbody>
<tr>
<td>US</td>
<td>Two-year community college</td>
<td>Intermediate academic writing</td>
<td>17-65 years old</td>
</tr>
<tr>
<td>Macau</td>
<td>University, first-year undergraduates</td>
<td>Intermediate EAP</td>
<td>18-20 years old</td>
</tr>
</tbody>
</table>

Figure 1. Participants.

The researchers for this study and one additional teacher taught four sections of classes for this study. A total of 64 students participated in this experiment. There was one experimental class in Macau and one experimental class in the US, forming a total of 39 students. There was also one control class in Macau and one control class in the US, totaling 25 students in both contexts.

Instructional Design

Both teaching contexts have a focus on critical-thinking skills and essay writing, and both researchers agreed that there was a need for more explicit grammar instruction in their courses to help students complete their writing assignments with better accuracy. Students in both contexts are assessed on grammar as part of a formal writing rubric, but teachers were not necessarily teaching explicit grammar skills—lower-level thinking skills—as the courses are more focused on higher-order thinking skills such as content and organization and the synthesis and evaluation of sources. Formal course-content objectives shared in both classes are detailed in Table 1.

Although students in both contexts may take some kind of developmental grammar courses, it is not a requirement in either context, though grammar is an essential ingredient toward passing their required composition courses. This is where flipping sought to fill the gap. With limited time in the class schedule for explicit grammar instruction, teachers in the flipped classroom had their students get such lectures online while the nonflipped students got mini-grammar lectures in class.

To address the question of whether the flipped model is viable in different contexts and more specifically Macau and US, Doman, Webb, and Pusey (2015) found that students leave the classroom with similar attitudes regarding technology. More specifically, instrumen-
Table 1
Common Intended Student Learning Outcomes in Macau and the US

<table>
<thead>
<tr>
<th>Critical reading and thinking</th>
<th>Writing process</th>
<th>Research skills</th>
<th>Grammatical/structure principles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifying an author’s purpose and main idea</td>
<td>Prewriting and planning</td>
<td>Searching library databases</td>
<td>Sentence variety and boundaries</td>
</tr>
<tr>
<td>Distinguishing major ideas from minor ideas and details</td>
<td>Drafting</td>
<td>Identifying reliable and relevant source material</td>
<td>Structure of simple, compound, and complex sentences</td>
</tr>
<tr>
<td>Responding to readings</td>
<td>Revising</td>
<td>Using reliable sources</td>
<td>Verb tense and form</td>
</tr>
<tr>
<td>Proofreading/editing</td>
<td></td>
<td></td>
<td>Comma usage</td>
</tr>
</tbody>
</table>

tality, comfort, digital literacy, and anxiety attitudinal ratings in both contexts during posttest time became more favorable. Thus the researchers of this study found it appropriate to conduct research of the flipped model across borders.

Grammar instruction started in week 3 in both the experimental and control classes. The four grammatical principles from Table 1 were used. The grammatical principles were part of the syllabus in the US context as well as a part of formal SLOs. The students in China did not have the same grammar expectations as a formal part of the syllabus, though they were asked to meet the same expectations as the US students through formal essays and intended grammar SLOs as mentioned in Table 1. Flipped teachers did not teach any of the grammar points in the class through formal instruction. They used 10 online video lessons related to the grammatical principles starting from week 3 and ending in week 13. The videos were an average of 10 minutes long and used a screencasted PowerPoint (PPT) lecture. The US researcher created five of the flipped videos and the Macau researcher created five. The Macau teacher used a video screencast in which students could see her face and the US teacher used a voice screencast. Students were asked to watch the video lesson, complete a follow-up online quiz, and post a sentence using the corresponding
grammatical principle through an online discussion tool. The flipped quizzes in both contexts were graded out of 10 possible points with 10 questions on each quiz. The discussion posts were pass/fail. In the US the flipped grammar homework counted for a total of 5% of the total course grade and in Macau 10%. The nonflipped students received identical grammar input from in-class PPT lectures and practice quizzes, though everything was conducted during class time. There was no formal assessment of the mini-grammar lessons in the control classes. Because of the time constraints in teaching in two different contexts, the teachers did not evaluate students’ improvement in grammar in terms of writing.

A grammar survey was given to both groups before they took a grammar test in week 2 and again in week 14 of the course. The survey was simple with only four items, which measured students’ comfort and confidence in their knowledge and use of English grammar. See Appendix A for a copy of the survey.

The pre- and posttest consisting of 32 items was given in weeks 2 and 14 of the course. The test consisted of 10 main items that were used for the instructional period of weeks 3-13.

Data Collection

Grammar Test

A 32-item test was given to both experimental and control groups during week 2 and week 14 in the study. A total of 25 students in the control group and 39 students in the experimental group completed the pretest and posttest. The test questions were designed to match the 10 grammar principles being taught during the semester. Three to four questions corresponded with each principle. Z-scores, kurtosis, and skewness were calculated, and none exceeded the absolute value of 1.96, suggesting that the data were within the acceptable range. Distributions were checked visually via histograms, in which data were found to be distributed normally. All data were used in the final analysis. A mixed factorial analysis of variance on the grammar test was run to evaluate three main hypotheses:

1. If there was significant difference between the control group and the experimental group;
2. If there were significant differences between the pretest and posttest conditions; and
3. If there was a significant interaction between the control and the experimental groups and the pretest/posttest conditions.
Grammar Survey

A four-item, 5-point Likert-scale grammar survey was also given out during the same weeks before the grammar test to get a more accurate picture of how students perceived their grammar skills.

The pretest alpha was .72 and the posttest alpha was .69; these values are within popular rules of thumb (i.e., around .70). Three of the four questions on the survey were positively worded, while one was negatively worded. A Cronbach’s alpha reliability score revealed that the negatively worded item on the survey (item number 3) lowered the total survey reliability. The question may have caused respondent confusion and therefore was deleted and not included in any of the data analysis. The overall grammar survey reliability score excluding item number 3 of .75 is considered reliable to get an understanding of students’ perceptions of their grammar skills.

Aggregate scores for the pretest and posttest grammar survey were obtained by getting the mean scores of the three items for each individual. Initially, a mixed-factorial analysis of variance was performed on the grammar survey data to answer the main hypothesis as stated above for the grammar test. However, the data were found to not be normally distributed and were measured on an ordinal scale. Because the requirements for a parametric test were not met, non-parametric tests were chosen for the data analysis instead (Hatch & Lazaraton, 1991). The Wilcoxin Signed-Rank Test was performed on the two groups (experimental and control) for each cultural context (Macau and US) through time (pretest and posttest).

Findings

Grammar Test

The ANOVA results in Table 2 and descriptive statistics in Table 3 reveal that all three hypotheses were statistically significant. Students in the experimental group scored higher compared to the students in the control group. Scores in the posttest were significantly higher compared to the pretest. Post hoc comparison of the four conditions, (a) pretest control group, (b) pretest experimental group, (c) posttest control group, and (d) posttest experimental group, revealed that the posttest experimental group had a significantly higher mean compared to the rest. There were no significant differences in the other pairwise comparisons. That means that the slight difference in starting points (means) between the control and experimental groups (.62 and .65) seen in Table 3 and Appendix B were not of significance, and both groups can be considered as starting at similar levels in terms of their grammar skills.
### Table 2
Analysis of Variance Results for the Full Sample

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>p</th>
<th>Partial η²</th>
<th>Observed power (1-β)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammar scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flipped/nonflipped</td>
<td>7.66(1, 62)**</td>
<td>.007</td>
<td>.11</td>
<td>.78</td>
</tr>
<tr>
<td>Pretest/posttest</td>
<td>7.06(1, 62)**</td>
<td>.010</td>
<td>.10</td>
<td>.74</td>
</tr>
<tr>
<td>Flipped/nonflipped*</td>
<td>6.69(1, 62)**</td>
<td>.012</td>
<td>.10</td>
<td>.72</td>
</tr>
</tbody>
</table>

Notes. *p<.05, **p<.01

### Table 3
Descriptive Statistics for Grammar Test Scores

<table>
<thead>
<tr>
<th>Control group</th>
<th>Experimental group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest</td>
<td>Posttest</td>
</tr>
<tr>
<td>M</td>
<td>SD</td>
</tr>
</tbody>
</table>

### Grammar Survey
A number closer to 1 on the grammar survey means that students were more confident with their grammar skills, and a number closer to 5 means that students were less confident with their grammar skills. A Wilcoxon Signed-Rank Test showed that the in-class grammar lessons in the control group (in both US and Macau) did elicit a statistically significant change in students’ confidence throughout the semester ($Z = 2.60$, $P = 0.009$). As shown in Table 4, the median score reduced from 3.00 to 2.67. Similarly, the flipped grammar lessons in the experimental group also showed a statistically significant change in students’ confidence at posttest time ($Z = 2.92$, $P = 0.004$). The median score had no change at 2.67 though the mean score did show a small decrease from 2.85 to 2.54. The mean in the grammar survey was significantly lower for the posttest condition compared to the pretest condition in both groups.

The researchers wanted to see if the effect on the experimental group was larger than on the control group; the Pearson correlation
Table 4
(Wilcoxon) Results Experimental and Control (Combining US and Macau)

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Pretest median score</th>
<th>Posttest median score</th>
<th>Pretest mean score</th>
<th>Posttest mean score</th>
<th>Z</th>
<th>P-value (2-tail)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammar survey control</td>
<td>3.00</td>
<td>2.67</td>
<td>3.14</td>
<td>2.74</td>
<td>2.60</td>
<td>0.009</td>
</tr>
<tr>
<td>Grammar survey experimental</td>
<td>2.67</td>
<td>2.67</td>
<td>2.85</td>
<td>2.54</td>
<td>2.92</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Table 5
Correlations Experimental and Control (Combining US and Macau)

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Pearson correlation</th>
<th>Spearman correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammar survey control</td>
<td>.589**</td>
<td>.616**</td>
</tr>
<tr>
<td>Grammar survey experimental</td>
<td>.424**</td>
<td>.423**</td>
</tr>
</tbody>
</table>

Note. **Correlation is significant at the 0.01 level (2-tailed).

Coefficients revealed that the smaller value of ($R = .424$) in the experimental group compared with the larger value of ($R = .589$) in the control group means there was a larger effect on the control group than on the experimental group (see Table 5).

Focus Groups

Student-led focus groups were run in the experimental classes in both the US and Macau. Six students were in the US focus group and three in the Macau focus group. Students were asked to volunteer for the optional discussion, which took place outside of class time in week 14 of the semester. The teachers were not present during the focus group in order to prevent any bias, and one student was chosen to read the questions found in Appendix C and to lead the discussion. The student leader also answered the questions. Both focus groups were video recorded and transcribed. The researchers blind coded each of the transcriptions first and then recoded using a descriptive code, modeling the coding manual provided by Saldana (2009). Three themes were found to be common, and the following three descriptive
codes were identified: Grammatical Consciousness Raising (GCM), Learner Autonomy (LA), and Videos Led to Improvement of Grammar (VIG). Finally, researchers shared their descriptive coded transcriptions, agreed upon each occurrence, and counted the total number of occurrences. The focus group data are presented in Figure 2.

<table>
<thead>
<tr>
<th>Theme</th>
<th>Count/Percentage</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical Consciousness Raising</td>
<td>44/15%</td>
<td>At first, we may score 50, but after the 10 quizzes, our grammar will be improved to 80. I think this check our improvement (Student 7).</td>
</tr>
<tr>
<td>Development of Learner Autonomy</td>
<td>30/10.23%</td>
<td>English grammar ... that was the one thing I have to work on for myself (Student 3).</td>
</tr>
<tr>
<td>Videos Led to Improvement</td>
<td>20/6.82%</td>
<td>I forgot grammar we learned in high school. From the videos we can strengthen our grammatical knowledge (Student 9).</td>
</tr>
<tr>
<td>TOTAL THEMES</td>
<td>94/32.05%</td>
<td></td>
</tr>
<tr>
<td>TOTAL UTTERANCES</td>
<td>293/100%</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. Themes from focus groups (N = 9).

The term *grammatical consciousness-raising* (Sharwood-Smith, 1981) refers to having learners become aware of grammatical patterns by the explicit teaching of forms. It is supposed that as students’ consciousness is raised about grammatical patterns, the awareness will lead them toward using the patterns correctly, although this theory has been debated in SLA circles. Students in the flipped classrooms mentioned that their consciousness was raised by the video instruction, quizzes, and the production of example sentences on the online forum. This theme accounted for 15% of the total number of utterances.

Students also thought that autonomy was promoted through the grammar study. With the videos assigned as part of homework, students had to take responsibility themselves to ensure that they watched the videos and completed the quizzes and forum posts before the weekly deadlines.

Finally, students remarked that the grammar videos led them to increased improvement in using English grammar. As they could
rewind, fast-forward, or watch the videos multiple times, students thought that the self-paced nature of the videos helped them to learn.

**Discussion**

**RQ 1: Do students in a flipped classroom have a significant difference in gains on grammar SLOs compared to students in nonflipped classes?**

All of the students in the control and experimental classes scored low on the grammar pretest (62.68 and 65.85) and were at a similar level when they entered the class. It is clear that in contexts of both Macau and the US students needed help with their grammar skills, and an intervention was necessary. The results of the posttest remained the same for the nonflipped classes, revealing that the in-class mini–grammar lessons were not successful in raising students’ grammar skills based on the test scores. However, in the flipped classes, students had gains on grammar achievement, and in the posttest their mean scores increased from 65.85 points to 71.77 points (as shown in Table 3). In answer to the first research question, the flipped classroom did have a significant difference in gains on the intended grammar SLOs compared to the nonflipped classes.

The findings here were similar to those of Farah (2014), who found that the flipped classroom contributed to students’ scores on an English writing test. As in Farah, the findings from RQ 1 show that students in the flipped classes had ample time to learn the grammar points given to them and to ensure that they were ready for the larger grammar test at the end of the semester. These findings are also consistent with those of Davies et al. (2013); Mason, Shuman, and Cook (2013); McLaughlin et al. (2013); Murphree (2014); Strayer (2012); Tune, Sturek, and Basile (2013); and Wilson (2013), who found that the flipped classroom led to increases in student outcomes.

**RQ 2: Do students in flipped classes and nonflipped classes experience changes in attitudes toward their grammar skills through time?**

In answer to the second research question, students in the flipped classes and nonflipped classes experienced positive changes in attitudes toward their grammar skills through time, and it was statistically significant in both groups ($p = .004; p = .009$, with $p < 0.05$). Both classes at posttest time had the same median score of 2.67 and thus became slightly more confident in their English grammar skills. The in-class grammar lessons in the control class may have made the students feel more confident at the end of the semester just as the out-of-class grammar lessons did in the experimental class. However, regardless of the treatment condition (control or experimental), students reported
lower grammar survey results for the posttest, showing that they felt more confident in their abilities to use English grammar. These findings are consistent with earlier studies by the researchers on student attitudes (Doman & Webb, 2015). Equally, these findings are similar to those in the literature reported by Papadopoulos et al. (2010), Pedroni and Meyer (2006), and Zappe et al. (2009), who found that the flipped classroom produced students with positive attitudes toward their own progression.

**RQ 3: Do students in flipped courses differ from students in non-flipped classes in their perceived grammar skills?**

In answer to the third research question, students in the flipped courses did not differ from students in the nonflipped courses in their perceived grammar skills. The experimental class did experience a positive statistically significant change in attitudes but the effect was smaller than in the control group (experimental group, $R = .424$; control group, $R = .589$). The improvements of students in the flipped classes on their grammar test scores were statistically significant and that aligned with their perceptions. Students in the control classes also experienced a positive statistically significant change in attitudes (as exposed through the mini–grammar lessons in class), and the effect was slightly stronger ($R = .589$). One might assume that because the students in the control group thought they improved their grammar skills like the flipped class that they would also raise their grammar scores on the test, though this was not the case.

**RQ 4: Do ESL students in a flipped classroom have student buy-in with the flipped methodology?**

Past research in the field of ESL and flipping has shown that students have positive attitudes toward the flipped methodology and report feeling more enthusiastic about the class compared to more traditionally taught classes (Webb, Doman, & Pusey, 2014). However, how many of the students actually complete the flipped materials?

According to the literature, it is not uncommon for students across all disciplines to experience midsemester slumps (Bateman, 1990; Bolton, 2003; Duffy & Jones, 1995). There are obvious cycles at various points during the semester when subtle changes occur in the students’ enthusiasm for learning (Bolton, 2003). After the initial excitement of the beginning of the new academic year, students start to lose interest, become apathetic, stagnate, or show signs of listlessness around the midpoint of the semester (Bolton, 2003).

Results from the 10 grammar quizzes and discussion items given to students in both flipped classes in Macau and in the US show
that students had in fact begun to slow down and took less interest in completing the grammar assignments around the midpoint of the semester (especially weeks 7 and 8). See Appendix D, which shows this trend in grammar-assignment completion over the course of the 10 weeks of the experiment. The two researchers addressed this concern by holding one-on-one consultations and assigning reflective essays to make students more aware of the ebbing energies that seemed to be occurring. A resurgence of interest began to occur toward the final weeks of the semester, which is in line with the literature in this field (Bateman, 1990; Bolton, 2003; Duffy & Jones, 1995).

As shown in Table 6, the Macau students completed more of the assignments than the US students. On average between both contexts, 77.75% of all the flipped grammar assignments were completed. This figure is in line with the literature, which shows that generally 73-80% of students complete the outside-the-classroom flipped tasks regularly (Davies et al., 2013; Gaughan, 2014; Murphree, 2014; Willey & Gardner, 2013). The figure in Appendix D shows that in both contexts students completed more of the assignments at the beginning of the class, and then there was a slight decrease during the middle of the semester during weeks 6, 7, and 8. At the end of the semester, the average number of completed assignments rose again. The themes presented in the focus group and the average number of completed assignment show that students had more buy-in during the beginning and the end of the semester.

Table 6
Total Percentage of Flipped Grammar Assessments Completed

<table>
<thead>
<tr>
<th>Type of Assessment</th>
<th>US</th>
<th>Macau</th>
<th>Averaged totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Online quiz</td>
<td>79%</td>
<td>85%</td>
<td>82%</td>
</tr>
<tr>
<td>Online discussion</td>
<td>66%</td>
<td>81%</td>
<td>73.5%</td>
</tr>
<tr>
<td>Total of all flipped</td>
<td>72.5%</td>
<td>83%</td>
<td>77.75%</td>
</tr>
<tr>
<td>grammar assessments</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The findings for RQ 4 are consistent with the literature about student buy-in for the flipped classroom. Davies et al. (2013), Gaughan (2014), Murphree (2014), and Willey and Gardner (2013) all found that students generally completed the flipped tasks (more than 70%) and thus were more engaged with the class materials.

Pedagogical Implications
The present study provides evidence for the relationship between students’ perceptions of their confidence and comfort levels
with English grammar after a treatment period of 10 weeks and their increased achievement on a post-treatment grammar test. Although both the control and experimental classes showed improvement on their grammar test scores, the measured gains were significant only for the experimental (flipped) classroom. The treatment of the flipped classroom significantly improved grammar test scores as well as the self-reported grammar proficiencies of the students.

An interesting finding from the study that should be taken into account is that the Macau students completed more flipped assignments on average than the US students did. Because students in Macau had a total class time of three hours per week and 14 weeks of instruction and the US students four hours per week and 17 weeks of instruction, this may account for the difference. As is common in EFL settings, students and teachers alike are looking for more class time and opportunities to practice. The same can be said for the US, though the students may have felt more overwhelmed by the flipped assignments than the students in Macau with less class time. Those looking to flip their classes should take into account the amount of student contact hours and carefully design flipped modules and videos that would replace traditional homework assignments without being an extra burden to the students.

Although further investigation on a larger scale is necessary with learners from various contexts, findings here suggest that the flipped approach is a viable methodology and can be an effective approach in ESL/EFL classrooms around the world. If revising an entire curriculum around a flipped approach, which involves teachers’ learning about new technologies and applying them in their classrooms, it is best that it be implemented slowly and with a great deal of training. Therefore, the researchers of this study make the following recommendations for teachers or administrators wishing to explore this option.

Teachers wishing to flip their classes should keep in mind the following:

1. Start by flipping only a small part of your class. Plan before the semester begins which aspects of the course will be flipped.
2. Flip modules that are most conducive to flipping. That is, identify modules in which online instruction would help to save class time for the application of skills gained after instruction.
3. Front-end your classes by preparing the instructional videos and online materials before the start of the semester.
4. Be willing to adapt your lessons depending on student responses and reactions.
5. Gather data from your students regarding their satisfaction with the flipped model, keeping in mind that it might take at least half of the semester before students begin to feel more comfortable with the additional use of technology.
6. As you and the students feel more comfortable with the flipped model, gradually increase the amount of flipped materials.

By following these simple instructions, teachers can experiment with a flipped classroom and see if this methodology suits their teaching styles.

**Conclusions**

The primary objective of this study was to compare the student learning outcomes (SLOs) of ESL/EFL students in Macau and the US who were exposed to a flipped classroom versus a traditional classroom. The goal was to assess any differences in achieving learning outcomes based on a pre- and post-grammar test. Data for the current study were gathered during the Spring 2015 semester at two institutes for higher education in Macau and the US. This study was part of a larger three-year study investigating students’ levels of satisfaction, perceptions, and achievements with the flipped approach to language teaching and learning.

Findings from this study suggest that the flipped classroom helped students to achieve their SLOs in grammar. Not only was achievement improved among flipped learners, but students also came to feel more comfortable and confident in their English grammar skills. Moreover, this teaching method boosted students’ buy-in. Students took charge and were more vested in their own learning. The weekly practice and mini-quizzes allowed students to check their knowledge of certain grammatical points and to practice those forms that they felt less confident with using. Although students in the nonflipped classes reported that they felt more confident with their English grammar skills, the results of the posttest grammar test did not reveal this to be correlated to their actual performances.

The outcomes of this study suggest that the flipped approach may be a viable methodology for ESL/EFL classes. This research has presented an innovative way of teaching grammar and addressing student learning outcomes.
Limitations

While our study provides compelling evidence in support of the flipped approach for helping students to make gains in meeting learning outcomes, caution should be used in generalizing the findings beyond the scope of the study. The learners in the current study were from a variety of backgrounds, but the number of participants ($N = 64$) is still too small to make broad generalizations.

Additionally, some researchers argue that in order to have a significant gain on achievement, a teaching strategy has to increase by 40% (Hattie, 2009), though it should be noted that there is no current literature in the flipped model to support this argument. The results from the current grammar study show that only a 6% score increase on the grammar test was possibly due to the flipped approach, which is well below the recommended increase, which means that further data are needed to support that the flipped approach does lead to an increase in achievement gains. On a similar note, students’ grammar in terms of writing achievement was not assessed because of slight differences in course requirements and perceived grading norms. Researchers in this study would like to address this in a future study in which teachers of multiple sections of classes calibrate and assess students’ grammar achievement in writing through a shared, formalized rubric.

A further limitation that should be noted is that the researchers did not account for students who may have received private grammar instruction in Macau or may have taken grammar courses outside of their required English classes in the US. This may have had an effect on how much grammar improvement they had throughout the semester, though it is unknown. No other courses were required in each context, though in the US students are encouraged to take a grammar review class.

Despite the limitations, the authors believe that the flipped classroom allows students the opportunity to achieve student learning outcomes through the self-paced nature of online videos that students can stop, rewind, and view again if necessary. This allows students more flexibility in learning and makes them more autonomous learners.

Acknowledgments

The researchers would like to thank the participants in this study as well as colleague Jose Rodriguez-Castillo, who helped instruct one of the courses in this study in 2015 at MiraCosta College. The researchers would also like to give a special thanks to Ivan Jacob Pesigan at the University of Macau for his insight and assistance with using SPSS. Finally, a warm thank-you goes to Kerry Jerome Pusey, who shared
with us insight into flipping the classroom and research methodology. Also, thanks to the University of Macau for a multiyear research grant that helped to make this research project possible.

Authors

Marie Webb teaches as a full-time nine-month instructor at the San Diego State University ALI and as an adjunct professor of language and linguistics at The University of San Diego’s MEd in TESOL, Literacy, & Culture Program. She also teaches writing as an associate faculty member in the American College of English/credit ESL program at MiraCosta College. More recently, she joined Anaheim University as a TESOL instructor at the David Nunan TESOL Institute, where she teaches TEYL and TESOL certificate programs. She has taught adult ESL in a wide range of settings in the US, Japan, China, and Korea.

Dr. Evelyn Doman is associate professor of TESOL and coordinator of the MA TESOL program in the School of Education at the University of Guam. Dr. Doman received her doctorate in Applied Linguistics from Macquarie University, Sydney, Australia. Before her position in teacher training, she taught ESL/EFL at several universities in Korea, Japan, Macau, and the US. Her research interests include learner autonomy, processability theory, focus-on-form instruction, and teacher beliefs. She is particularly interested in the flipped language classroom, preservice teacher training, and ways to promote active learning in the classroom.

Notes

1 For example, see the Flipped Institute, 2015; Khan Academy, 2011; Knewton, 2011; The Flipped Learning Network, 2014.
2 For example, see Cybrary Man, 2015; Musallam, 2014; Petty, 2015.
3 These books are about flipping in general, and they are not specific to the field of ESL/EFL.
4 See researchers such as Davies, Dean, and Ball, 2013; Enfield, 2013; Gaughan, 2014; Lage, Platt, and Treglia, 2000; Murphree, 2014; Strayer, 2012; Willey and Gardner, 2013.
5 Jigsaw reading involves students working in mother groups to read assigned parts of a passage and then to get into break-out groups to share the information they read about in their parts and to learn what others in their groups read about in their assigned parts.
6 Running dictations involve students’ reading, summarizing, or looking for main ideas of a reading passages by taking turns with a partner and reading several sentences or passages posted on the walls around the classroom.
These include Butt, 2014; Davies et al., 2013; Enfield, 2013; Findlay-Thompson and Mombourquette, 2014; Lemmer, 2013; McLaughlin et al., 2013; Morin, Kecskemety, Harper, and Clingan, 2013; Murphree, 2014; Rowe, Frantz, and Bozalek, 2013; Strayer, 2012; Tune, Sturek, and Basile, 2013.

Davies et al., 2013; Enfield, 2013; Strayer, 2012.

Butt, 2014; McLaughlin et al., 2013.

Morin et al., 2013; Findlay-Thompson and Mombourquette, 2014; Rowe et al., 2013.

Enfield, 2013; Murphree, 2014; Rowe et al., 2013; Tune et al., 2013.

Rowe et al., 2013.

Students were enrolled in their engineering statistics classes.

Engineering students.

Davies et al., 2013; Gaughan, 2014; Murphree, 2014; Willey and Gardner, 2013.

Davies et al., 2013; Gaughan, 2014; Murphree, 2014; Willey and Gardner, 2013.

Davies et al., 2013; Mason, Shuman, and Cook, 2013; McLaughlin et al., 2013; Murphree, 2014; Strayer, 2012; Tune et al., 2013; Wilson, 2013.

Findlay-Thompson and Mombourquette, 2014; Morin et al., 2013.

As defined by Oudenhoven (2006), Generation 1.5 students are immigrant students who move to the US in their early preteen or teenage years, at the age of 12 or older, and enroll in school. They earn the label the “1.5 generation” because they bring with them characteristics from their home country but continue their assimilation and socialization in the new country, thus being “halfway” between the first generation and the second generation. Their identity is thus a combination of new and old culture and tradition, and they may thus experience what could be called a third culture. Because their native language is usually spoken at home, they often need to attend ESL classes in countries where English is the dominant language, such as the US.

The obvious differences are that ESL students are exposed to English outside the classroom, while EFL students generally get exposure to English only inside the classroom or through opportunities that they independently search out themselves.

Z-scores can also tell us how far a particular score is away from the mean, and kurtosis defines the distribution of scores around the mean. Skewness is asymmetry in a statistical distribution, in which the curve appears distorted or skewed either to the left or to the right. Skewness can be quantified to define the extent to which a distribution differs from a normal distribution.
References


Gaughan, J. E. (2014). The flipped classroom in world history. History Teacher, 47(2), 221-244.


**Appendix A**

**Flipped Grammar Survey**

Please answer the following questions HONESTLY and to the best of your ability. The survey should take you 5 minutes. Don’t be worried if some of the questions look similar.

1. I am comfortable with my English grammar skills.
   - Strongly Agree
   - Agree
   - Neutral
   - Disagree
   - Strongly Disagree

2. When asked about my English grammar skills I feel confident.
   - Strongly Agree
   - Agree
   - Neutral
   - Disagree
   - Strongly Disagree

3. I am not strong at English grammar.
   - Strongly Agree
   - Agree
   - Neutral
   - Disagree
   - Strongly Disagree

4. My knowledge of English grammar is acceptable to write papers.
   - Strongly Agree
   - Agree
   - Neutral
   - Disagree
   - Strongly Disagree
Appendix B
Plot of Control and Experimental Pre- and Post- Grammar Test Score Means

Appendix C
Focus Group Questions

1. How does your teacher use technology in this class?
2. Can you give specific examples about how technology has helped you to increase your English language skills?
3. What are some positive or negative experiences you have had with technology/online learning in this class?
4. How does your teacher help you to learn more about English grammar in this class?
5. How did you feel about your English grammar at the beginning of this class?
6. How did you feel about your English grammar at the end of the class? Is there anything you need more practice with?
7. How does your teacher encourage you to learn more about English grammar on your own?
8. What is your overall perception of the class? Is there anything that you particularly enjoyed/did not enjoy?
Appendix D
Percentage of Flipped Grammar Assignments Completed in Each Context (Macau and US)

Note. Flipped grammar assignments (FGA) includes the averages of the 10 weekly online grammar quizzes and 10 weekly online discussions after watching corresponding flipped video grammar lectures.