

FLIP TO WIN: IMPROVING ARGUMENTATIVE WRITING IN CHINESE EFL CLASSROOMS

YU WANG¹

SERIAZNITA BINTI HAJI MAT SAID²

^{1,2}FACULTY OF SOCIAL SCIENCES AND HUMANITIES, UNIVERSITI TEKNOLOGI MALAYSIA,
SKUDAI, JOHOR, MALAYSIA

Abstract

Flipped classroom approaches have shown positive effects on writing, yet empirical evidence regarding argumentative writing remains scarce. This study investigates their impact on Chinese EFL undergraduates' argumentative writing skills using an exploratory sequential design. The qualitative phase involved focus-group interviews with 18 students, leading to the development of an assessment rubric for evaluating argumentative texts. In the subsequent quantitative phase, 83 written texts were collected at four time points and analyzed statistically. Finally, qualitative and quantitative findings were integrated to address the research question. Results reveal improvements in both genre knowledge, identified in qualitative analysis, and writing performance, evidenced by quantitative data. The findings suggest that combining sufficient pre-class preparation with interactive in-class activities significantly enhances argumentative writing. This study provides theoretical insights and pedagogical guidance for instructors seeking to improve argumentative writing instruction through the flipped classroom model.

Keywords: argumentative writing, flipped classroom, Chinese EFL undergraduates, an assessment rubric, exploratory sequential design

1. INTRODUCTION

Writing is considered a productive skill through which writers organize and express their main thoughts and opinions. Compared to descriptive and narrative writing, argumentative writing, a significant genre in most English proficiency tests, aims to persuade readers or audience to accept the writer's viewpoints through critical and well-organized arguments (Aziz, 2021; Yang, 2022). In normal writing classrooms, traditional ways are common to instruct students' writing, which is lecture-centered, hindering students' engagement and increasing their learning difficulties (Omar, 2020; Fathi & Rahimi, 2020; Sinaga & Feranie, 2017). Most students struggle with writing, due to the insufficient instructional time in the traditional teaching methods, which causes a weak writing foundation and insufficient writing practice.

As most Chinese EFL students and teachers regard it as a significant yet challenging task, an appropriate teaching innovation needs to be explored to address the current issue (Yang, 2022; Peltzer et al., 2024). Flipped classroom, one of models of the blended learning, reverses the traditional learning environment by delivering instructional content outside of the face-to-face classroom sessions, frequently through digital or online media. Consequently, in-class time is dedicated to interactive, student-centered learning activities (Bergmann & Sams, 2012). It encourages students to preview before classes, and positively participate in class activities. According to the previous research of the integration of the flipped classroom in language teaching, it is admitted that the flipped classroom stimulates students' academic performance and personalized learning (Clark & Kaw, 2020). However, most prior studies pay attention to both teachers' and students' satisfaction towards the flipped classroom (Shooli et al., 2022), while little empirical research has examined its effectiveness in enhancing English argumentative writing among EFL learners (Roohani and Rad, 2022). Considering the research background and research gap outlined above, this study aims to identify an effective pedagogical method to prompt students' English argumentative writing. Therefore, the flipped classroom was chosen to teach English argumentative writing to Chinese EFL undergraduates, and its effectiveness was examined through an exploratory sequential design.

1.1 Research Objectives

To identify the effectiveness of the flipped classroom on English argumentative writing skills of Chinese EFL undergraduates.

1.2 Research Questions

What are the effectiveness of the flipped classroom on English argumentative writing skills of Chinese EFL undergraduates?

1.3 Theoretical Framework

In terms of the research question and objective in the present study, the theoretical framework in Figure 1 involves the theory of second language acquisition (SLA theory), Bloom's revised taxonomy, and constructivism theory. As shown in the triangular framework of Figure 1, the second language acquisition theory and Bloom's revised taxonomy complement the constructivism theory, and the second language acquisition theory is related to the Bloom's revised taxonomy.

Constructivism is generative learning, because students can seek knowledge actively through cooperative learning and problem-solving activities, which is applied in English teaching to motivate students' active construction and student-centered teaching (Zulela & Rachmadtullah, 2019). Based on the prior study of Al-Huneidi and Schreurs (2013), the constructivism theory is closely related to blended learning, as the constructivism theory prompts the student-centered learning, and blended learning in turn facilitates the practical application of constructivist principles. Admittedly, the constructivism theory is significant in directing this study of implementing the flipped classroom in argumentative writing.

SLA theory complements constructivism theory in the current study, as Krashen (1982; 1984) thought abundant input as a key foundation in language acquisition, and learning process can be achieved when learners comprehend the input. It should align with the input hypothesis of Krashen (1982), which emphasizes the concept of 'i+1', referring to language input that is just slightly beyond the learners' current proficiency level. Therefore, both the quality and quantity of the input are crucial in language acquisition (Yang, 2018). However, traditional teaching methods often fail to provide sufficient comprehensive input in the instruction of argumentative writing, which is primarily an output-based skill. The implementation of the flipped classroom in the present study is guided by the constructivism theory and SLA theory, since the preparation before classes facilitates access to comprehensive input necessary for effective learning.

Another theory is Bloom's revised taxonomy, which is applied to the educational training and learning process widely. It consists of six hierarchical levels, and they are remembering, understanding, applying, analyzing, evaluating, and creating (Anderson & Krathwohl, 2001; Krathwohl, 2002). The first three levels are classified as low-order cognitive skills, while the latter three are considered as high-order thinking skills. In the context of the flipped classroom, this approach facilitates high-order thinking by enabling students to acquire lower-order skills, such as remembering and understanding through preparation before classes. Meanwhile, high order skills, including analyzing and evaluating, can be developed by interactive activities and in-class engagement (Zou & Xie, 2018; Fisher et al., 2021).

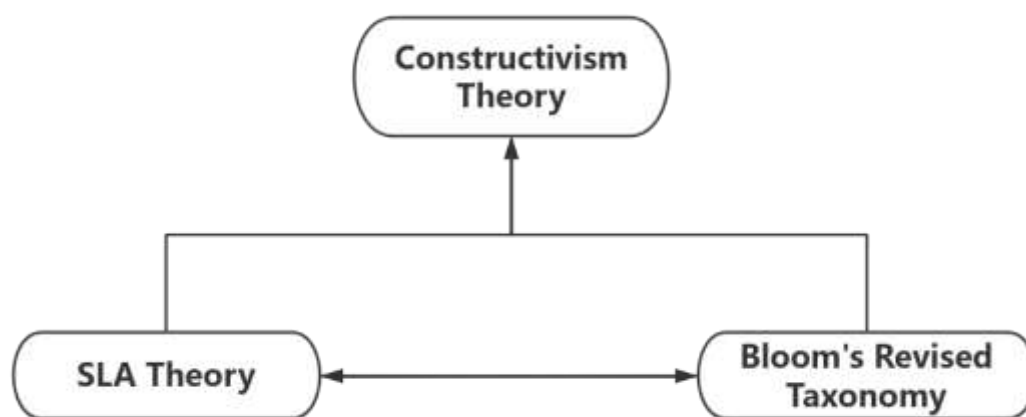


FIGURE 1 Theoretical Framework

2. LITERATURE REVIEW

This section reviews two key topics relevant to the current study: argumentative writing in EFL contexts and the flipped classroom in teaching argumentative writing.

2.1 Argumentative Writing in EFL Contexts

Writing is generally thought more challenging than the other three language skills, consisting of speaking, reading, and listening. Among various genres in writing, argumentative writing needs high-order thinking skills and critical thinking skills (Liao & Liao, 2020; Chang et al., 2024; Aziz, 2021). In order to persuade audience or readers to agree with the writers' viewpoints, logical opinions and solid arguments should be organized ((Latifi, Noroozi, & Talaei, 2021; Latifi, Noroozi, Hatami, & Biemans, 2021). In other words, well-organized arguments can direct students to make thoughtful decisions.

Although English argumentative writing plays a vital role in English proficiency test, such as CET (College English Test), TEM (Test for English Majors), IELTS (International English Language Testing System), as well as TOEFL (Test of English as a Foreign Language), it still faces various challenges in the learning process. Language barriers or difficulties is one issue in completing argumentative essays for EFL learners, due to the inadequate vocabulary and grammar, frequent syntactic errors, and even difficulties in constructing complicated sentences and selecting phrases (Lan, 2024; Liao & Liao, 2022; Aziz & Said, 2019; Liu, 2019; Sundari & Febriyanti, 2021). However, in Yang's opinion (2022), the improvement and better organization of argumentative writing for EFL learners can not be limited only in language level. According to the comparison of eight EFL undergraduates' English proficiency levels in Knouzi's study (2023), low-level proficiency students focused on language-related difficulties, while students with higher English proficiency level emphasized discourse synthesis process, such as selecting related information and organizing the structure. Accordingly, previous studies have discussed several factors that impede students' acquisition of argumentative writing. One such factor is the traditional education system that tends to be exam-oriented, and lecturer-centered (e.g. Aziz et al., 2023; Aziz & Said, 2020).

The education system in East Asia is acknowledged as exam-based education, in which students have less chances to participate in problem-solving activities, so it is tough for EFL learners to construct well-organized arguments with logical thoughts (Aziz et al., 2023; Zainuddin & Rafik-Galea, 2016). In the lecture-centered classroom, students are supposed to follow teachers' instructions in order to pass exams, and they are constricted in the definite answers during their learning process with few chances to solve problems and explain reasons (Dong et al., 2019; Aziz & Said, 2019; 2020). Additionally, rote memorization and the product approach are two common ways in the conventional class, according to the prior study by Zhang & Zhang (2021) and Qin (2020). Both of them are against the purpose of argumentative writing, impeding students' ability to make judgments. Therefore, deficient practice and inappropriate guidance in conventional classes keep students from developing well-organized argumentative essays.

Furthermore, according to Toulmin's model of argumentation, a well-organized argument includes three basic elements: claim, data, and warrant, and three secondary elements: qualifier, backing, and rebuttal. Previous analyses of argumentative writing, suggest that the arguments' organization is often neither adequate nor reasonable in terms of the six elements. While claims and data are present in argumentative essays written by Chinese EFL learners, the remaining elements are seldom observed (Zhang, 2018; Qin & Karabacak, 2010).

Although pedagogical methods are rarely employed (Yang, 2022), it is essential to apply innovative strategies and models to improve argumentative writing (Aziz & Said, 2020). Previous studies have shown that the mind-mapping approach can enhance students' writing performance in English argumentative writing. Additionally, genre-based pedagogy has been implemented to support the organization of arguments and overall writing improvement, according to Zhang and Zhang (2021). In this study, the flipped classroom approach incorporating a variety of activities focused on argumentative writing is adopted in the teaching of argumentative writing to address the challenges faced by both teachers and students. Hence, this research investigates the effectiveness of this approach in enhancing argumentative writing skills among Chinese EFL undergraduates.

2.2 Flipped Classroom in Teaching Argumentative Writing

As presented in Figure 2, flipped classroom model, belonging to the classification of blended learning, happens between face-to-face practice that is teacher-centered in campus and online delivery of learning instructions outside classes (Staker & Horn, 2012). In general, the previous studies achieved the implementation of the flipped classroom by delivering instructional content outside of the face-to-face classroom sessions, and promoting interactive, student-centered learning activities during in-class time (e.g. Roohani & Rad, 2022; Fathi & Rahimi, 2020; Shafiee Rad et al., 2021). Based on the definition of the flipped classroom, it is more possible to achieve asynchronous and synchronous learning.

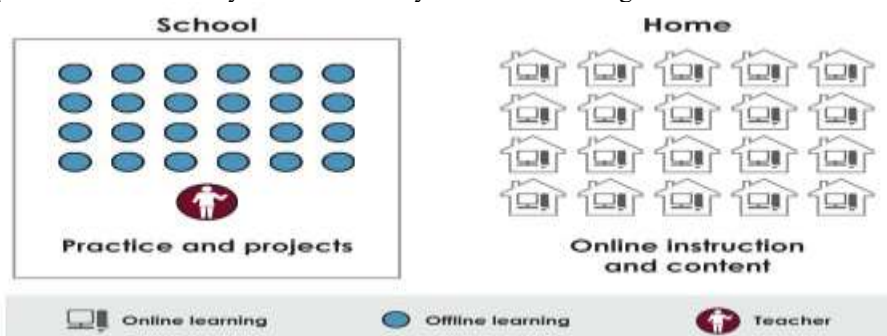


FIGURE 2 Flipped Classroom Model (Staker & Horn, 2012)

Admittedly, flipped classroom is one innovative product in linguistic pedagogy (Luo, 2021), and it combines the benefits of online and face-to-face learning. First of all, the flipped classroom emphasized the mastery of

basic learning contents before classes, which is a foundation of the in-class success. Before classes, the flipped classroom makes use of social media tools, so students can access learning sources without the limitation of time and space. Besides, the flipped classroom is also committed to enriching and designing in-class interactive activities between the teacher and students (Shooli et al., 2022). This kind of novel pedagogical method contributes to active learning in the student-centered environment (Shooli et al., 2022; Roohani & Rad, 2022), which is more effective than teacher-centered classroom. Moreover, the full preparation before classes improved the ability to solve problems, contributing to the development of critical and creative thinking (Zou & Xie, 2018; Asep et al., 2023). The application of the flipped classroom is related to the theoretical framework, as students can construct knowledge about argumentative writing actively, and prompt high-order thinking skills based on the preparation before classes and the active learning in classes. The prior studies have emphasized the features and advantages of the implementation of the flipped classroom, and most of them used a quasi-experiment design by comparing the results between the control and experimental groups. The flipped classroom was admitted as an effective approach to improve students' writing quality (Soltanpour & Valizadeh, 2018). Furthermore, in terms of the overall writing performance and fluency, EFL students in the flipped classroom are better than those in the non-flipped classroom (Fathi & Rahimi, 2022). However, they have few differences on complexity and accuracy in their writing. Additionally, a genre-based pedagogy is implemented to support the development of argumentation and writing skills, with changes in students' initial argumentation knowledge through questionnaires and interviews (Zhang & Zhang, 2021). The data analysis is guided by Tardy's genre knowledge model (Tardy, 2009), which comprises four dimensions: formal knowledge, process knowledge, rhetorical knowledge, and subject-matter knowledge. The findings indicate that the conventional writing approach is not effective in improving students' ability in argumentative writing, while students in the experimental group show notable progress under the genre-based pedagogy, particularly in their understanding of discourse structure and the language features of the argumentative genre. Consequently, it contributes to a more systematic classification and evaluation of students' argumentative writing skills in the current study.

2. METHODOLOGY

This part begins with an overview of the research design, followed by a description of research instruments and participants. Then, the data collection and analysis is elaborated.

3.1 Research Design

The exploratory sequential design, one of major mixed methods designs, was used in the current study, and it is a three-phase design (Creswell & Plano Clark, 2011; 2018). As shown in Figure 3, this research design starts with the qualitative data collection and analysis, and then an instrument or tool of the following quantitative phase is developed. The last phase involves the quantitative data collection and analysis, with the purpose to test or generalize the qualitative findings in the initial phase (Creswell, 2024).

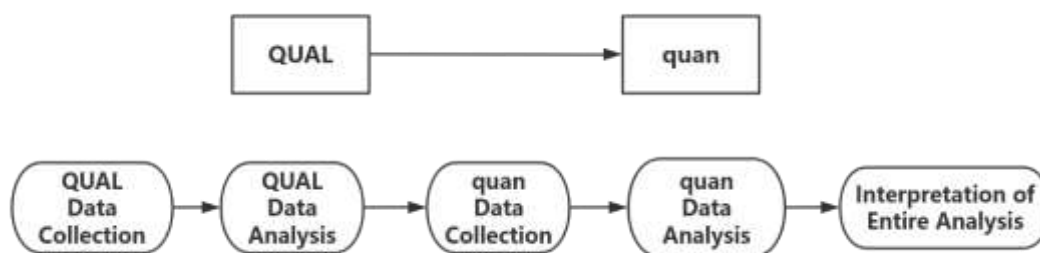


FIGURE 3 Exploratory Sequential Design (Creswell, 2009)

In terms of the implementation of the flipped classroom for argumentative writing of Chinese EFL undergraduates, this study employs an exploratory sequential design, which starts with a qualitative phase including focus-group interviews with 18 students. Then, a quantitative instrument, a targeted evaluation method is developed, grounded in the qualitative findings. In the following quantitative phase, 83 students' written texts were analyzed through four different time points. There are 83 written texts collected each time, resulting 332 texts in total. At last, both qualitative and quantitative data are interpreted to comprehensively achieve the research objective: To identify the effectiveness of the flipped classroom on English argumentative writing skills of Chinese EFL undergraduates.

3.2 Research Instruments

In the current study, the instruments include focus-group interviews, written texts from all participants, and assessment rubrics.

3.2.1 Focus-group Interviews

The purpose of interviews is to collect detailed and in-depth information and perspectives with interviewees through interactions (Patton, 2015; Brinkmann & Kvale, 2015). Being different from individual interviews, focus groups aim to create interaction data through participants' discussions. Guest et al. (2017) regard focus groups as interpersonal and interactive instruments, which may generate more viewpoints, which is also admitted as the essence of focus groups (Greenbaum, 2003; Krueger & Casey, 2015). During the interactions of focus groups, participants can get support when sharing sensitive themes and information (Kruger et al., 2019; Guest et al., 2017). Concerning the size of focus groups, it ranges from six to twelve, and discussions generally last for ninety minutes (Guest et al., 2017; Creswell & Poth, 2016). Four focus groups were initially planned, with students participating in groups of six to ensure fairness in the findings. However, based on data saturation, the final number of focus groups was reduced to three, as the fourth group did not yield any new or additional insights from its transcript.

In the present study, a checklist in Appendix A by Creswell (2015) was used to monitor the interview process, and two experienced lectures were invited to validate the focus-group interview protocol according to a list of criteria in Appendix B (Adam et al., 2018). The focus-group interview includes three questions, presented in Appendix C. The first and the second question are to collect students' viewpoints on the flipped classroom, involving pre-class and in-class periods. The last question focuses on the overall description of their writing experience in this study.

3.2.2 Written Texts

Previous studies about writing collect written texts as their research data (e.g. Qin, 2020; Aziz & Said, 2019). As the present study aims to examine the effectiveness of the flipped classroom on argumentative writing of Chinese EFL undergraduates, their written texts during the research serve as documents. Each student participant is required to finish four written texts according to four topics, and these written texts are collected across four different time points with 83 pieces at each time point. The four time points were arranged in the first week (before the implementation of the flipped classroom), the seventh and tenth weeks (during the implementation of the flipped classroom), and the twelfth week (after the implementation of the flipped classroom). The findings from written texts were integrated with the qualitative data, due to the phases of the exploratory sequential design.

3.2.3 Assessment Rubrics

In the exploratory sequential design, a quantitative assessment tool was developed to evaluate students' written texts. A rubric, defined as a document outlining evaluation criteria and clearly articulating performance expectations, serves this purpose (Reddy & Andrade, 2010). The rubric used in the current study, called as asTTle Writing Scoring Rubric (See Appendix D), is a primary research instrument. It is admitted that a well-designed rubric is recognized for helping students to identify their weaknesses and strengths, which supports more objective and consistent evaluation (Faieza, 2019). Additionally, rubrics contribute to better understanding of learning targets, the quality standard, and promote reliable assessment ((Reddy & Andrade, 2010). As a result, rubrics can facilitate both students' achievement, and the overall evaluation systems. In this study, two independent raters were invited to score students' written texts in order to ensure inter-rater reliability.

3.2.4 Participants

There are 83 students who are voluntary to be included in the current study through advertising. They were from the same research university, and all personal information was confidential and anonymous during the process of data collection and analysis. They are from 18 to 22 years old, including 43 female and 40 male students, who are involved in various programs in the research university, such as computer science, laws, and preschool education. All of them participated in this study for 14 weeks, and submitted their written texts. Therefore, all written texts among the four different time points were collected with students' permission during the research. Furthermore, eighteen students were chosen to participate in focus-group interviews due to the mix methods design in the current study.

3.2.5 Research Procedure

The researchers asked for permission of the research university, and recruited participants through advertisement. The study lasted for fourteen weeks, and in the first week, the overview and purposes of this study were introduced to all students, and they were asked to download the online application used in the flipped classroom. From the second to the twelfth week, all participants were involved in the implementation of the flipped classroom with various activities of argumentative writing. As seen in Figure 4, the implementation of the flipped classroom is closed related to the theoretical framework of this study, with the learning objective to improve Chinese EFL undergraduates' argumentative writing skills. The major part of the flipped classroom implementation is classified into two phases, each with tasks for both teachers and students.

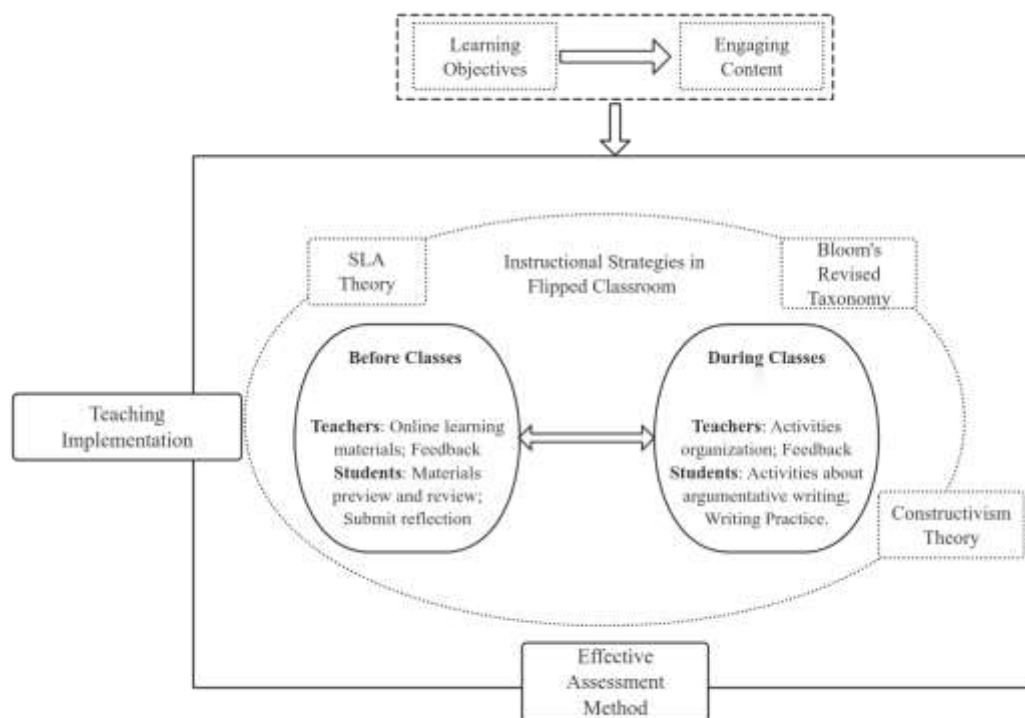


FIGURE 4 The Implementation of the Flipped Classroom

In the pre-class phrase of the flipped classroom, asynchronous learning is achieved through online application. Teachers offer learning materials and feedback, while students can preview and review the materials, as well as provide reflections. The process of preparation aligns with the SLA theory that suggests more comprehensive input before participating in in-class activities. The in-class phase focused on interactive activities, which were student-centered with active engagement in argumentative writing practice. The division of the flipped classroom follows Bloom's revised taxonomy, starting from lower-order thinking skills to higher-order cognitive processes. The overall implementation of the flipped classroom is supported by constructivism theory, emphasizing active learning through interaction and engagement. An assessment method was also included into the implementation of the flipped classroom, which was developed for the evaluation of the argumentative writing.

3.2.5 Data Collection and Analysis

The data in this study was collected through different instruments, and due to the exploratory sequential design, it began with focus-group interviews in the initial qualitative phrase. Then, a targeted evaluation method was developed in the subsequent quantitative phase for the analysis of written texts. Lastly, both the qualitative and quantitative data were interpreted for the research question and objective. The qualitative data from focus-group interviews was analyzed through thematic analysis. The codes and themes of the interview transcripts were generated in NVivo software. In terms of the quantitative phase, asTTle marking rubrics (See Appendix D) was used by two raters to score students' written texts. All statistics in this phase was analyzed in the SPSS. Due to the results in Shapiro-Wilk normality tests, it indicates that the statistical data of the written texts did not follow a normal distribution. As a consequence, the Friedman Test and Wilcoxon Signed-Rank Test were chosen in this study.

4. RESULTS

In the exploratory sequential design, a qualitative phase is followed by a quantitative one, and it aims to integrate quantitative data and findings to qualitative findings (Creswell, 2024). In this study, data collected from the focus group interviews was analyzed using thematic analysis. Then, an evaluation method was designed for assessing argumentative texts for the quantitative analysis. The purpose of it is to systematize the effects of the flipped classroom on argumentative writing of Chinese EFL undergraduates.

4.1 Qualitative Results of Focus-group Interviews

The focus-group interview transcripts were analyzed by using the thematic analysis to examine the effectiveness of the flipped classroom on argumentative writing skills. Four themes were identified based on the study of Zhang and Zhang (2021), based on Tardy's genre knowledge. The four categories of the genre knowledge

are ‘rhetorical knowledge’, ‘formal knowledge’, ‘process knowledge’, and ‘subject-matter knowledge’. There is an overview of these themes and sub-themes, shown in Figure 5.

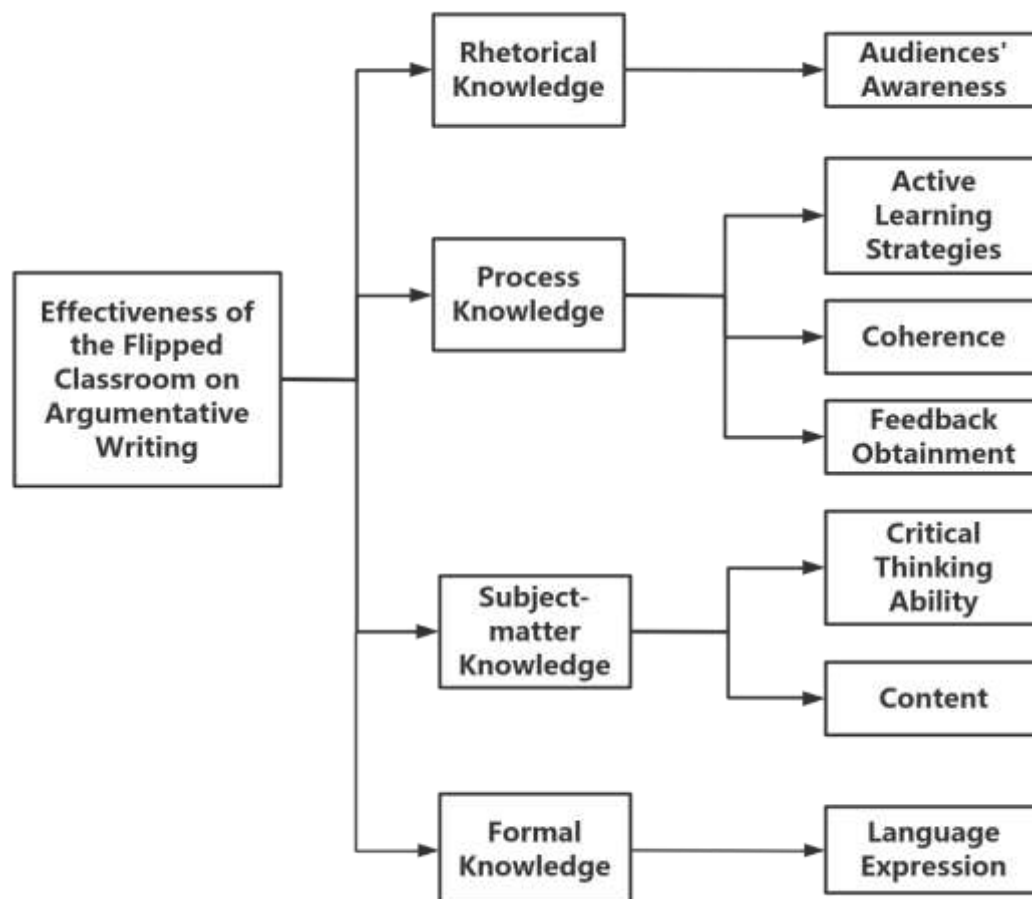


FIGURE 5 An Overview of Effectiveness on English Argumentative Writing of Chinese EFL Undergraduates

In terms of the rhetorical knowledge, it refers to the ability to communicate and persuade based on the targeted readers or audiences, purpose and context. In the current study, students in the second focus group had great consideration of audiences' perspectives in argumentative writing, called as audiences' awareness. The second crucial theme was process knowledge, referring to the understanding and steps involved in writing. It encompasses active learning strategies, coherence, and feedback obtainment, which were extracted from the transcripts in the three focus groups. The flipped classroom prompted students' active learning strategies, as it strengthened their engagement, established positive attitudes towards argumentative writing, facilitated positive learning ways, and increased learning enjoyment. Moreover, the feedback obtainment emphasized the importance of timely feedback from the teacher in students' learning process. Additionally, the coherence within the 'process knowledge' theme, was admitted as a fundamental part in the writing process, and it was the ability to structure and organize ideas logically and reasonably.

The third major theme was 'subject-matter knowledge', and it refers to the content and background information for well-organized arguments. As seen in Figure 4.1, it included the aspects of critical thinking abilities and content. Students in the three focus groups acknowledged their improvement in critical thinking skills, as the flipped classroom was helpful in developing critical thinking ability through sufficient preparation before classes and interactive activities in classes. Furthermore, the flipped classroom had positive effects on the content of argumentative writing, due to students' improvement in understanding topics and key points, as well as supporting and organizing arguments, according to students in the first and second focus groups. Finally, the last theme was 'formal knowledge', referring to the technical and structural aspects of writing, such as grammar and syntax. Students in the first and the second focus groups pointed out that students were not influenced by the flipped classroom on the language expression significantly, as they demonstrated a weak grasp of advanced vocabulary and expression, hindering their language quality in argumentative writing.

4.2 Quantitative Results of Written Texts

According to the exploratory sequential design in the current study, an assessment method for argumentative written texts was developed based on qualitative findings. Consequently, this study adopted asTTle Writing Scoring Rubrics (argue and persuade) (See Appendix D), which had been used in the previous study in argumentation and writing improvement (Zhang & Zhang, 2021). Compared with previous study that focused on the overall scores of students' written texts, this study aims to integrate both the qualitative and quantitative results. It analyzed the overall scores and six categories based on the asTTle Writing Scoring Rubrics, including content inclusion, coherence, audience awareness and purpose, language resources for achieving the purpose, vocabulary and grammar, and mechanics. In the quantitative phase, 332 written texts were collected in total, with 83 written texts at each time point. The four time points were named as T1, T2, T3, and T4 in the process of data analysis. The Friedman tests were applied to determine significant differences based on students' scores across the four time points. Additionally, pairwise comparisons were conducted using Wilcoxon signed-rank tests, adjusted by the Bonferroni correction for multiple comparisons. This quantitative analysis phase of written texts was integrated with the qualitative data comprehensively, aiming to examine the effectiveness of the flipped classroom on argumentative writing of Chinese EFL undergraduates.

4.2.1 Overall Grades

As shown in Table 4.1, the results of overall grades in the Friedman Test are presented. N stands for the number of written texts, which is 83, and the degree of freedom (df) is three, as four time points are being compared ($df = \text{the number of group} - 1 = 3$). The Chi-Square is large, indicating that at least one of the four time points differ from others significantly. The p-value suggests a statistically significant difference among T1, T2, T3, and T4.

TABLE 4.1 Statistics in Friedman Test of Overall Grades

Test Statistics ^a	
N	83
Chi-Square	178.358
df	3
Asymp. Sig.	.000

Considering the significant difference among the four time points found in the Friedman Test in Table 4.1, a Wilcoxon Signed-Rank Test was conducted to identify where differences occurred. As shown in Table 4.2, the results confirm statistically significant differences among T1-T3, T1-T4, T2-T3, T2-T4, T3-T4. Nevertheless, the difference between T1-T2 is not significant after applying the Bonferroni correction with an adjusted significance (Adj. Sig.) value of 0.105. The increases observed at later time points are stronger than that between T1 and T2, and a general increasing trend is confirmed from T1 to T4, indicating stronger transitions in later stages.

TABLE 4.2 Grades of Six Categories

Pairwise Comparisons					
Sample 1- Sample 2	Test Sta- tistic	Std. Error	Std. Test Statis- tic	Sig.	Adj. Sig.a
T1-T2	-.476	.200	-2.375	.018	.105
T1-T3	-1.482	.200	-7.395	.000	.000
T1-T4	-2.428	.200	-12.114	.000	.000
T2-T3	-1.006	.200	-5.020	.000	.000
T2-T4	-1.952	.200	-9.739	.000	.000
T3-T4	-.946	.200	-4.719	.000	.000
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.					
Asymptotic significances (2-sided tests) are displayed. The significance level is .050.					
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.					

As shown in Table 4.3, it presents the results in Friedman tests of the six categories in the asTTle Writing Scoring Rubrics, including content inclusion, coherence, audience awareness and purpose, language resources for achieving the purpose, vocabulary and grammar, and mechanics. The Chi-Square values of the six categories indicate large differences among the four time points, but Chi-Square value of vocabulary and grammar is less than the other five categories. It means even though there are progressive improvements among the six categories, students make less progress in vocabulary and grammar. To further identify group differences, post-hoc pairwise comparison were conducted to determine which specific groups differ from one another.

TABLE 4.3 Statistics in Friedman Tests of the Six Categories

Test Statistics ^a						
	Content Inclusion	Coherence	Audience Awareness and Purpose	Language Resources for Achieving the Purpose	Vocabulary and Grammar	Mechanics
N	83					
Chi-Square	149.391	111.860	155.262	182.827	89.130	101.230
df	3					
Asymp. Sig.	.000					
a. Friedman Test						

Table 4.4 presents the results of six pairwise comparisons through conducting Wilcoxon Signed-Rank Test for content inclusion. The findings indicate that the comparison between T1 and T2 does not show a significant difference, because an adjusted p value is 0.554, indicating relative stability between these two time points. However, the remaining comparisons shown in Table 4.4 present highly significant differences, as p-value is less than 0.001, even after Bonferroni correction, which suggests substantial changes in scores of the content inclusion. Therefore, there is a consistent and statistically significant increasing trend in content inclusion over time, except for the non-significant change between T1 and T2.

TABLE 4.4, Wilcoxon Signed-Rank Test of Content Inclusion

Pairwise Comparisons					
T1-T2	-.337	.200	-1.683	.092	.554
T1-T3	-1.199	.200	-5.982	.000	.000
T1-T4	-2.054	.200	-10.251	.000	.000
T2-T3	-.861	.200	-4.299	.000	.000
T2-T4	-1.717	.200	-8.567	.000	.000
T3-T4	-.855	.200	-4.269	.000	.000
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.					
Asymptotic significances (2-sided tests) are displayed. The significance level is .050.					
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.					

In terms of coherence, the results of multiple comparisons determine the differences between the four time points. For the comparison between T1 and T2, the test statistic is 0.006, with a standard error of 0.200, and the standardized test statistic is 0.030. The p-value is 0.976, and the adjusted significance value (Adj. Sig.) is 1.000, indicating no significant difference between T1 and T2. By contrast, the other five comparisons show great differences, and the largest differences are observed between T1-T4 and T2-T4, as indicated by the large negative test statistics and very low p-values. Consequently, there is a clearly increasing trend in the coherence of students' written texts, with significant differences observed between the first time point and both the third and the fourth time points, although no meaningful change is detected between the first and the second time points.

TABLE 4.5 Wilcoxon Signed-Rank Test of Coherence

Pairwise Comparisons					
T1-T2	.006	.200	.030	.976	1.000
T1-T3	-.952	.200	-4.750	.000	.000
T1-T4	-1.620	.200	-8.086	.000	.000
T2-T3	-.946	.200	-4.719	.000	.000
T2-T4	-1.614	.200	-8.056	.000	.000
T3-T4	-.669	.200	-3.337	.001	.005

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.
Asymptotic significances (2-sided tests) are displayed. The significance level is .050.
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Considering the results of audience awareness and purpose from the Wilcoxon Signed-Rank Test, Table 4.6 presents that there is no significant difference between the first and the second time points, suggesting that the scores in the earlier time points of audience awareness and purpose are similar. In comparison, there are the largest differences observed between T1 and T4, as well as T2 and T4, indicating a dramatic increasing trend across the four time points. As a result, these results in Table 4.6 confirms a clear and significant progression over time considering students' performance in audience awareness and purpose.

TABLE 4.6 Wilcoxon Signed-Rank Test of Audience Awareness and Purpose

Pairwise Comparisons					
T1-T2	-.367	.200	-1.834	.067	.400
T1-T3	-1.488	.200	-7.425	.000	.000
T1-T4	-2.048	.200	-10.220	.000	.000
T2-T3	-1.120	.200	-5.591	.000	.000
T2-T4	-1.681	.200	-8.387	.000	.000
T3-T4	-.560	.200	-2.796	.005	.031
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.					
Asymptotic significances (2-sided tests) are displayed. The significance level is .050.					
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.					

Regarding language resources for achieving the purpose in students' written texts, as seen in Table 4.7, there are significant differences between T1 and T3, as well as between T1 and T4, which indicates considerable improvements in students' performance of language resources for achieving the purpose over time. Nevertheless, the difference between T3 and T4 is insignificant, suggesting a slower rate of improvement in the final stage.

TABLE 4.7 Wilcoxon Signed-Rank Test of Language Resources for Achieving the Purpose

Pairwise Comparisons					
T1-T2	-.428	.200	-2.134	.033	.197
T1-T3	-1.729	.200	-8.627	.000	.000
T1-T4	-2.181	.200	-10.882	.000	.000
T2-T3	-1.301	.200	-6.493	.000	.000
T2-T4	-1.753	.200	-8.748	.000	.000
T3-T4	-.452	.200	-2.255	.024	.145
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.					
Asymptotic significances (2-sided tests) are displayed. The significance level is .050.					
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.					

Table 4.8 presents students' performance in terms of vocabulary and grammar, and the findings reveal a progressive increase across the time, especially between T1 and T4. By contrast, the differences between T2 and T3, as well as between T3 and T4, are not statistically significant, with adjusted p-values of 0.326 and 0.105, respectively. It suggest a relatively stable performance between these intermediate time points. Hence, while Table 4.8 confirms the overall increasing trend in vocabulary and grammar, it also highlights a period of relative stability between T2 and T3, as well as between T3 and T4.

TABLE 4.8 Wilcoxon Signed-Rank Test of Vocabulary and Grammar

Pairwise Comparisons					
T1-T2	-.741	.200	-3.697	.000	.001
T1-T3	-1.127	.200	-5.621	.000	.000
T1-T4	-1.602	.200	-7.996	.000	.000

T2-T3	-.386	.200	-1.924	.054	.326
T2-T4	-.861	.200	-4.299	.000	.000
T3-T4	-.476	.200	-2.375	.018	.105
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same. Asymptotic significances (2-sided tests) are displayed. The significance level is .050.					
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.					

Lasting, students' performance in mechanics is shown in Table 4.9, presenting a clear statistically significant progression from the first time point to others. It confirms an overall increasing trend from T1 and T4, while the differences between T2 and T3 as well as T3 and T4 are insignificant, indicating a stable trend. These findings align with Friedman results, reinforcing the conclusion that students exhibit marked improvement from the early to the later stages, with performance becoming more gradual at higher levels.

TABLE 4.9 Wilcoxon Signed-Rank Test of Mechanics

Pairwise Comparisons					
T1-T2	-.892	.200	-4.449	.000	.000
T1-T3	-1.392	.200	-6.944	.000	.000
T1-T4	-1.500	.200	-7.485	.000	.000
T2-T3	-.500	.200	-2.495	.013	.076
T2-T4	-.608	.200	-3.036	.002	.014
T3-T4	-.108	.200	-.541	.588	1.000
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same. Asymptotic significances (2-sided tests) are displayed. The significance level is .050.					
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.					

5.DISCUSSION

This study aims to identify the effectiveness of the flipped classroom on argumentative writing of Chinese EFL undergraduates through conducting an exploratory sequential design by incorporating both qualitative and quantitative phases. This study began to collect qualitative data by focus-group interviews, which promoted the development of the subsequent quantitative phase, involving the design of an evaluation rubric for argumentative texts, and the quantitative assessment of students' written texts across the four time points. This study design ensures the data triangulation by using multiple instruments and methods to investigate the effectiveness of the flipped classroom on argumentative writing from different perspectives.

The flipped classroom enriches pre-class preparation and in-class interaction between the instructor and students to increase students' learning motivation and learning skills (Shooli et al., 2022; Chuang et al., 2018). The effectiveness of the flipped classroom has been acknowledged in previous studies, particularly in promoting active learning and feedback obtainment, which were included in the process knowledge in this study. The flipped classroom fostered student-centered learning with sufficient preparation before classes and interactive activities in classes, contributing to active learning and immediate feedback (Luo, 2021; Khodabandeh & Hemmati, 2022; Su Ping et al., 2019).

Furthermore, in this study, active learning strategies and feedback-rich flipped classroom can improve students' performance in content quality, critical thinking and coherence, which were confirmed by both qualitative and quantitative data. Critical and logical thinking has been promoted through problem-solving activities in the flipped classroom, according to Zou and Xie (2018). Students mastered the knowledge of genre and arguments' organization, so their scores in overall performance, content inclusion, and coherence have been improved across the four time points, which aligns with the qualitative findings from focus-group interviews. Regarding the language expression, categorized to formal knowledge in the qualitative findings, there was clear and consistent progress in this aspect. The results from the Friedman test and Wilcoxon Signed-Rank test indicated that students improve their ability in grammar and vocabulary, contributing to greater consistency in language use of their writing performance. Lastly, audiences' awareness in students' written texts was improved, which was acknowledged in both qualitative and quantitative data. This category was ignored in most previous studies, except for Zhang and Zhang's Study (2021). However, the findings in this study indicate obvious and significant differences between the first and the fourth time point, as well as the second and the fourth time point.

This study provides empirical evidence for the effectiveness of the flipped classroom in enhancing the argumentative writing skills of Chinese EFL undergraduates. In the qualitative phrase, data from focus-group interviews were categorized according to genre knowledge, including process knowledge, formal knowledge,

rhetorical knowledge, and subject-matter knowledge. Student participants admitted the positive impacts of the flipped classroom approach on promoting active learning strategies and enhancing feedback acquisition. Furthermore, key components of argumentative writing, such as content, coherence, audience awareness, language expression, as well as critical thinking ability, were found to be positively influenced by the flipped classroom, and these components were integrated into the genre knowledge framework adopted in this study. Subsequently, the effectiveness of the flipped classroom was examined through the further quantitative phase by using the evaluation rubric as a quantitative instrument. The results from the Friedman tests and Wilcoxon Signed-Rank tests revealed significant improvements in students' overall writing performance and in specific areas, including content inclusion, coherence, audience awareness and purpose, language resources for achieving the purpose, vocabulary and grammar, and mechanics, based on analysis of student writing across four time points.

6. CONCLUSION

The findings from the exploratory sequential design indicate that Chinese EFL undergraduates made substantial progress in argumentative writing with instructions in the flipped classroom. The positive effects of the flipped classroom are supported by prior study conducted by Roohani & Rad (2022) and Shooli et al. (2022), both of which highlighted its effectiveness in addressing the challenges of argumentative writing, and emphasized the advantages of the flipped classroom approach. However, given the limited empirical research on the effectiveness of the flipped classroom for developing argumentative writing skills among EFL learners, and instructors rarely adopted applicable pedagogical methods (Roohani & Rad, 2022; Liao & Liao, 2022), this study contributes to filling the research gap by investigating the effectiveness of the flipped classroom on Chinese EFL undergraduates' argumentative writing performance.

The existing findings of the current study have some theoretical and pedagogical implications. The theoretical implications emphasized the connection between this study's findings and the existing theories, with the theoretical framework underpinning the flipped classroom implementation in teaching argumentative writing. This study was grounded in three main theories, which were the Second Language Acquisition Theory (SLA Theory) (Krashen, 1982), Bloom's Revised Taxonomy Theory (Anderson & Krathwohl, 2001; Bloom, 1956), and constructivism theory. Among these theories, constructivism theory was recognized for its support of the learning process by building on prior knowledge, promoting active knowledge construction, and encouraging student-centered instruction. When integrated with blended learning and the flipped classroom model, constructivist principles confirm that engaging students in the process of constructing their own knowledge can enhance both their motivation and level of interaction (Subramaniam & Muniandy, 2019; Zulela & Rachmadtullah, 2019). The implementation of the flipped classroom approach facilitates an interactive environment through pre-class preparation and in-class interactive activities focusing on developing argumentative writing skills. According to the findings of this study, the flipped classroom approach fostered students' learning strategies and enhanced feedback obtainment, which aligned with the key principles of the constructivism theory. Regarding the SLA theory, which emphasizes the importance of comprehensive input (Krashen, 1982; 1984), this study was guided to provide ample and meaningful exposure to English argumentative writing through pre-class preparation and various interactive in-class activities. Finally, as high-order thinking abilities can be effectively stimulated in a flipped classroom (Zou & Xie, 2019), the overall design of the flipped classroom implementation was informed by Bloom's revised taxonomy. It enables Chinese EFL undergraduates to engage in deeper cognitive processes after establishing a solid foundation at the lower levels of learning.

Furthermore, the pedagogical implications in the study relate to the practical application of its findings in teaching argumentative writing to Chinese EFL undergraduates. First of all, this study aligns with the development of technology and pedagogical innovation, as it implemented the flipped classroom in teaching argumentative writing to facilitate both asynchronous and synchronous learning effectively. Secondly, in-class time in this study was dedicated to the development of high-order thinking skills due to the thorough pre-class preparation. These pedagogical strategies, integrated with the flipped classroom approach, offer valuable guidance for future instructional practices and research in similar educational contexts.

Similar to many studies, this study is not exempt from limitation. One limitation is that all student participants were from the same research university, despite their voluntary participation through study advertisements. While the effectiveness of the flipped classroom on argumentative writing was examined, it is suggested that participants can be invited from a broader range of universities across China to enhance the generalizability of the findings. Another limitation concerns the duration of the flipped classroom intervention. In this study, the implementation of the flipped classroom lasted eleven weeks, with one ninety-minute lesson per week. Although positive changes in students' written performance were observed across four time points, a longer intervention may yield more substantial and diverse outcomes. Consequently, a longitudinal research design can be adopted in future studies to more thoroughly examine impacts of the flipped classroom on writing performance among a wider population of Chinese EFL undergraduates.

REFERENCES

- Adams, A., Betz, A.R. & Dringenberg, E. (2018). Validation of an interview protocol to explore students' beliefs about intelligence. In ASEE annual conference & exposition proceedings, pp 1-12. <https://peer.asee.org/31221>
- Anderson, L.W., Krathwohl, D.R., Airasian, P.W., Bloom, B.S., Cruikshank, K.A., Pintrich, P. R. & Mayer, R.E. (2001). *A Taxonomy for Learning, Teaching and Assessing*, Addison Wesley Longman, Inc.
- Al-Huneidi, A., & Schreurs, J. (2011). Constructivism based blended learning in higher education. In *World Summit on Knowledge Society*. pp. 581-591. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Asep, S., Sarah, K. & Rajji, A. (2023). The Flipped Classroom: A Story from an EFL Classroom in Indonesia. *MEXTESOL Journal*, 47(2), p.n2. <https://www.researchgate.net/publication/370582442>
- Aziz, F. I.A. (2021). *The Persuasive Writing Features of Secondary School Students in Malaysia and Factors That Influence the Quality*. PHD Thesis, Universiti Teknologi Malaysia.
- Aziz, F.I.A. & Said, S.B.M. (2019). Developing a persuasive writing model for secondary school. *Educational Research for Policy and Practice*, 19, pp.143-158. <https://doi.org/10.1007/s10671-019-09253-6>
- Aziz, F. I.A., Salam, S. N. & Said. S.B.M. (2023). Improving English Persuasive Writing in Malaysia: A Recommendation. *Asian Journal of Research in Education and Social Sciences*, 5, 10. doi:<https://doi.org/10.55057/ajress.2023.5.2.7>
- Aziz, F. I.A. & Said. S.B.M. (2020). The Success Factors that Impact the Quality of Students' Persuasive Essays: Faktor-faktor Kejayaan yang Mempengaruhi Kualiti Karangan Persuasif Pelajar. *Journal of Management and Muamalah*, 10(1), pp. 30 - 45. Available at: <https://www.jmm.kuisjournal.com/index.php/jurnal/article/view/43> (Accessed: 8December2023).
- Aziz, F.I.A., Salam, S.N. & Said, S. B. M., (2023). INTEGRATING TOULMIN'S MODEL TO IMPROVE THE PERSUASIVE WRITING SKILLS OF SECONDARY SCHOOL STUDENTS IN MALAYSIA. In *31st MELTA INTERNATIONAL CONFERENCE*, pp. 124.
- Bergmann, J. & Sams, A. (2012). *Flip your classroom: Reach every student in every class every day*. International Society for Technology in Education.
- Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I. Cognitive Domain*. New York: Longman.
- Brinkmann, S. & Kvale, S. (2015). *InterViews: Learning the craft of qualitative research interviewing* (3rd ed.). Thousand Oaks, CA: Sage.
- Chang, J. L., Hung, H. T. & Yang, Y. T. C. (2024). Effects of an annotation-supported Socratic questioning approach on students' argumentative writing performance and critical thinking skills in flipped language classrooms. *Journal of Computer Assisted Learning*, 40(1), 37-48. <https://doi.org/10.1111/jcal.12856>
- Clark, R. M. & Kaw, A. (2020). Adaptive learning in a numerical methods course for engineers: Evaluation in blended and flipped classrooms. *Computer Applications in Engineering Education*, 28(1), 62-79. <https://doi.org/10.1002/cae.22175>
- Creswell, J.W. (2024). *Research design: Qualitative, quantitative, and mixed methods approaches*, 6th ed. Sage publications.
- Creswell, J. W. & Vicki L. Plano Clark. (2011). *Designing and Conducting Mixed Methods Research*, 2nd ed. Thousand Oaks: Sage.
- Creswell, J.W. & Plano Clark, V.L. (2018). *Designing and conducting mixed methods research*. 3rd ed. Sage publications.
- Creswell, J.W. (2015). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*, 5th ed. Pearson Education, Inc.
- Creswell, J.W. & Poth, C.N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.
- Dong, R., Fu, D., Zhou, X. & Wang, B. (2019). Writing instruction in China: challenges and efforts. *International Journal of Information and Education Technology*, 9 (10), pp.715-723.
- Faieza, C. (2019). Application of Rubrics in the Classroom: A Vital Tool for Improvement in Assessment, Feedback and Learning. *International Education Studies*, 12 (1), pp. 61-69. <https://doi.org/10.5539/ies.v12n1p61>
- Fathi, J. & Rahimi, M. (2020). Examining the impact of flipped classroom on writing complexity, accuracy, and fluency: A case of EFL students. *Computer Assisted Language Learning*, 35(7), pp.1668-1706. <https://doi.org/10.1080/09588221.2020.1825097>

- Fisher, R., Perényi, A. & Birdthistle, N. (2021). The positive relationship between flipped and blended learning and student engagement, performance and satisfaction. *Active Learning in Higher Education*, 22(2), pp.97-113. <https://doi.org/10.1177/1469787418801702>
- Greenbaum, T. (2003). The gold standard? why the focus group deserves to be the most respected of all qualitative research tools. *Quirk's Marketing Research Review*, 17, 22–27.
- Guest, G., Namey, E., Taylor, J., Eley, N. & McKenna, K. (2017). Comparing focus groups and individual interviews: findings from a randomized study. *International Journal of Social Research Methodology*, 20(6), pp.693-708. <https://doi.org/10.1080/13645579.2017.1281601>
- Khodabandeh, F. & Hemmati, F. (2022). Comparing the Effect of Peer Assessment in Blended, Flipped and Face-to-Face Classes: The Case of Argumentative Essays. *Iranian Journal of English for Academic Purposes*, 11(1), pp.1-12.
- Knouzi, I. (2023). The Cognitive Processes of ESL Writers Responding to an Integrated Argumentative Writing Task. *TESL Canada Journal*, 40(1), 61–80. <https://doi.org/10.18806/tesl.v40i1/1385>
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon Press. ISBN 0-08-028628-3
- Krashen, S.D., Terrell, T.D., Ehrman, M.E. & Herzog, M. (1984). A theoretical basis for teaching the receptive skills. *Foreign Language Annals*, 17(4), p.261.
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into practice*, 41(4), 212-218.
- Krueger, R. & Casey, M. (2015). *Focus groups: A practical guide for applied research* (5th ed.). Thousand Oaks, CA: Sage
- Kruger, L.J., Rodgers, R.F., Long, S.J. & Lowy, A.S. (2019). Individual interviews or focus groups? Interview format and women's self-disclosure. *International Journal of Social Research Methodology*, 22(3), pp.245-255. <https://doi.org/10.1080/13645579.2018.1518857>
- Lan, J. (2024). A Case Study on Writing Difficulties of Five Chinese English Majors. *Journal of Education and Learning*, 14 (2), pp.259-269.
- Latifi, S., Noroozi, O. & Talae, E. (2021). Peer feedback or peer feedforward? Enhancing students' argumentative peer learning processes and outcomes. *British Journal of Educational Technology*, 52(2), pp.768-784. <https://doi.org/10.1111/bjet.13054>
- Latifi, S., Noroozi, O., Hatami, J. & Biemans, H.J. (2021). How does online peer feedback improve argumentative essay writing and learning?. *Innovations in Education and Teaching International*, 58(2), pp.195-206. <https://doi.org/10.1080/14703297.2019.1687005>
- Liu, M. (2019). Effects of teacher feedback on the rewrites of Chinese undergraduates' English argumentative essays. *TESL Reporter*, 52(2).
- Liao, M. & Liao, Y. (2022). Improving Chinese College Students' Argumentative Writing: A Presentation-Assimilation-Discussion-Exercise Approach. *Frontiers in Psychology*, 13, p.874531. <https://doi.org/10.3389/fpsyg.2022.874531>
- Luo, Y. (2021). The Opportunities and Limitations of Blended Learning and the Flipped Classroom for Second Language Teaching. *BCP Social Sciences & Humanities*, 14, pp.190-194. DOI: 10.54691/bcpssh.v14i.193
- Omar, O. A. (2020). The Effectiveness of Blended Learning in Enhancing Saudi Students' Competence in Paragraph Writing. *English Language Teaching*, 13(9), 72-82. doi: 10.5539/elt.v13n9p72
- Peltzer, K., Lorca, A. L., Krause, U. M. & Busse, V. (2024). Effects of formative feedback on argumentative writing in English and cross-linguistic transfer to German. *Learning and Instruction*, 92, 101935. <https://doi.org/10.1016/j.learninstruc.2024.101935>
- Patton, M. Q. (2002). *Qualitative Evaluation and Research Methods* (3rd ed.). Thousand Oaks, CA: SAGE Publications.
- Qin, J. (2020). The Assessment of L2 Written English across the MENA Region. The Assessment of L2 Written English across the MENA Region, pp.225-240. L. McCallum and C. Coombe (eds.) https://doi.org/10.1007/978-3-030-53254-3_10
- Qin, J. & Karabacak, E. (2010). The analysis of Toulmin elements in Chinese EFL university argumentative writing. *System* 38 (3), 444–456. <https://doi.org/10.1016/j.system.2010.06.012>
- Roohani, A. & Rad, H.S. (2022). EFFECTIVENESS OF HYBRID-FLIPPED CLASSROOM IN IMPROVING EFL LEARNERS' ARGUMENTATIVE WRITING SKILL. *TEFLIN Journal: A Publication on the Teaching & Learning of English*, 33(2). DOI: 10.15639/teflinjournal.v33i2/349-366
- Shooli, E., Rahimi Esfahani, F. & Sepehri, M. (2022). Flipped classroom influence on the learner's outcomes: A study based on English writing courses in Iran. *Education Research International*, 2022, pp.1-9. <https://doi.org/10.1155/2022/1530290>
- Sinaga, P. & Feranie, S. (2017). Enhancing Critical Thinking Skills and Writing Skills through the Variation in Non-Traditional Writing Task. *International Journal of Instruction*, 10(2), 69-84. http://www.e-iji.net/dosyalar/iji_2017_2_5.pdf

- Shafiee Rad, H., Roohani, A. & Rahimi Domakani, M. (2021). Flipping EFL learners' writing classroom through role-reversal and discussion-oriented models. *Language Learning & Technology*, 25(2), 158-177. <http://hdl.handle.net/10125/73438>
- Staker, H. & Horn, M.B. (2012). *Classifying K-12 blended learning*. Innosight institute.
- Soltanpour, F. & Valizadeh, M. (2018). A flipped writing classroom: Effects on EFL learners' argumentative essays. *Advances in Language and Literary Studies*, 9(1), pp.5-13. DOI: 10.7575/aiac.alls.v.9n.1p.5
- Su Ping, R.L., Verezub, E., Adi Badiozaman, I.F.B. & Chen, W.S. (2019). Tracing EFL students' flipped classroom journey in a writing class: Lessons from Malaysia. *Innovations in Education and Teaching International*, 57(3), pp.305-316. <https://doi.org/10.1080/14703297.2019.1574597>
- Subramaniam, S.R. & Muniandy, B. (2019). The effect of flipped classroom on students' engagement. *Technology, Knowledge and Learning*, 24(3), pp.355-372. <https://doi.org/10.1007/s10758-017-9343-y>
- Sundari, H. & Febriyanti, R.H. (2021). The analysis of Indonesian EFL argumentative writing using Toulmin's model: The structure and struggles from the learners. *Scope: Journal of English Language Teaching*, 5(2), pp.67-78. DOI: 10.30998/scope.v5i2.8544
- Tardy, C. (2009). *Building genre knowledge*. Parlor Press.
- Yang, J. (2018). On the Functions of Output in SLA from the Perspective of Constructivism. 3rd PIL International Conference on Education and Management Innovation. doi:10.5729/aer.2018.117.204
- Yang, R. (2022). An empirical study of claims and qualifiers in ESL students' argumentative writing based on Toulmin model. *Asian-Pacific Journal of Second and Foreign Language Education*, 7(1), pp.1-17. <https://doi.org/10.1186/s40862-022-00133-w>
- Yang, R. (2022). An empirical study on the scaffolding Chinese university students' English argumentative writing based on toulmin model. *Heliyon*, 8(12). <https://doi.org/10.1016/j.heliyon.2022.e12199>
- Zainuddin, S.Z. & Rafik-Galea, S. (2016). EFFECTS OF TRAINING IN THE USE OF TOULMIN'S MODEL ON ESL STUDENTS' ARGUMENTATIVE WRITING AND CRITICAL THINKING ABILITY. *Malaysian Journal of Languages and Linguistics (MJLL)*, 5(2), pp.114-133. DOI: 10.24200/mjll.vol5iss2pp114-133
- Zhang, Y. (2018). An Investigation into the Development of Structure and Evidence Use in Argumentative Writing. *Theory and Practice in Language Studies*, 8(11), 1441. <https://doi.org/10.17507/tpls.0811.08>
- Zhang, L. & Zhang, L.J. (2021). Fostering stance-taking as a sustainable goal in developing EFL students' academic writing skills: Exploring the effects of explicit instruction on academic writing skills and stance deployment. *Sustainability*, 13(8), p.4270. <https://doi.org/10.3390/su13084270>
- Zhang, T. & Zhang, L.J. (2021). Taking stock of a genre-based pedagogy: sustaining the development of EFL students' knowledge of the elements in argumentation and writing improvement. *Sustainability*, 13(21), p.11616. <https://doi.org/10.3390/su132111616>
- Zou, D. & Xie, H. (2018). Flipping an English writing class with technology-enhanced just-in-time teaching and peer instruction. *Interactive Learning Environments*, 27(8), pp.1127-1142. <https://doi.org/10.1080/10494820.2018.1495654>
- Zulela, M.S. & Rachmadtullah, R. (2019). Constructivism approach in learning to write narrative at elementary school. In 1st International Conference on Innovation in Education (ICoIE 2018), pp. 287-290. Atlantis Press.

APPENDIX A

A Checklist for Interviewing (Creswell, 2015 P221)

A Checklist for Interviewing	
Who will participate in your interviews?	Alice Zhang
What types of interviews are best to conduct?	Face-to-face individual interview
Is the setting for your interview comfortable and quiet?	Yes
If you are audiotaping, have you prepared and tested the equipment?	Yes
Did you obtain consent from the participants to participate in the interviews?	Yes
Did you listen more and take less during the interview?	Yes
Did you probe during the interview? (ask to clarify and elaborate)	Yes

Did you avoid leading questions and ask open-ended questions?	Yes
Did you keep participants focused and ask for concrete details?	Yes
Did you withhold judgments and refrain from debating with participants about their views?	Yes
Were you courteous and did you thank the participants after concluding the interview?	Yes

A Checklist for Interviewing	
Who will participate in your interviews?	Clytie Wang
What types of interviews are best to conduct?	Face-to-face individual interview
Is the setting for your interview comfortable and quiet?	Yes
If you are audiotaping, have you prepared and tested the equipment?	Yes
Did you obtain consent from the participants to participate in the interviews?	Yes
Did you listen more and take less during the interview?	Yes
Did you probe during the interview? (ask to clarify and elaborate)	Yes
Did you avoid leading questions and ask open-ended questions?	Yes
Did you keep participants focused and ask for concrete details?	Yes
Did you withhold judgments and refrain from debating with participants about their views?	Yes
Were you courteous and did you thank the participants after concluding the interview?	Yes

APPENDIX B Validation Form for Interview Protocols (Adams, et al., 2018)

No.	Criteria
1	The question reflects the research objectives.
2	The question satisfies the research objectives.
3	The question effectively gathers necessary data.
4	The question is clear and concise.

APPENDIX C Focus-group Interview Questions

What preparation do you do before classes learning argumentative writing? What do you think of it?
What are interactive and engagement activities during classes learning argumentative writing? What do you think of them?
How would you describe your experiences in the argumentative writing course?

APPENDIX D asTTle Writing Scoring Rubric (argue or persuade)

	Level 1	Level 2	Level 3
Content inclusion (20%)	Makes some arguments that related to the topic. The content can be tangential from the topic. (0-6)	Includes most argument elements (e.g., position statement, main points, evidence, restatement). Includes some useless information that does not contribute to argument. (7-12)	Includes all argument elements (e.g., position statement, main points, evidence, restatement). The content relates and contributes to the argument. (13-20)
Coherence (20%)	Only organized at sentence level. (0-6)	Attempt to structure content with grouping ideas across sentences. May use simple linking words (e.g., and, or, because). (7-12)	Effective ideas grouping and paragraphing. Use varied linking words or phrases (e.g., although, by the same token, nevertheless). (13-20)
Audience awareness and purpose (20%)	Writer recognizes that his/her opinion is needed in evidence. The writer uses language to state opinions with a personal perspective. (0-6)	Language use and writing style generally appropriate to audience. Writer states his/her position. Some attempt to influence the reader is evident. (7-12)	Language use and writing style appropriate and directed to audience (e.g., attempts to persuade reader). Clearly stated consistent position is evident. (13-20)
Language resources for achieving the purpose (20%)	Topic-related vocabulary present. Often speech-like in structure and uses a personal voice. (e.g., I reckon) (0-6)	Uses topic appropriate vocabulary. Attempts to use language to make arguments seem more objective (e.g., passive structures) and powerful (e.g., emotive language). (7-12)	May attempt to use persuasive language (e.g., emotive vocabulary) to influence readers or includes or refers to the reader (e.g., you would). Uses language to make arguments seem more objective (e.g., passive structures) and/or powerful (e.g., certainly, must, absolutely). (13-20)
Vocabulary and grammar (15%)	Limited vocabulary range; frequent errors of agreement, tense, number, articles, pronouns, prepositions and meaning confused or obscured. (0-4)	Appropriate vocabulary range; several errors of agreement, tense, number, articles, pronouns, prepositions but meaning seldom obscured. (5-9)	Sophisticated vocabulary range; few errors of agreement, tense, number, articles, pronouns, prepositions (10-15)
Mechanics (5%)	Frequent errors of spelling, punctuation, capitalization, paragraphing. (0-1)	Occasional errors of spelling, punctuation, capitalization, paragraphing. (2-3)	Few errors of spelling, punctuation, capitalization, paragraphing. (4-5)