INVESTIGATING CRITICAL THINKING IN THE ARGUMENTATIVE
WRITING OF ENGLISH MAJORS AT A MAINLAND CHINESE
UNIVERSITY: IMPLICATIONS FOR POLICY CHANGES

LIU, FULAN

PhD

THE HONG KONG INSTITUTE OF EDUCATION

2014

LIU, FULAN   PhD   2014   HKIEd

Investigating critical thinking in the argumentative writing of undergraduate English majors at a mainland Chinese university: Implications for policy changes

by

LIU, FULAN

A Thesis Submitted to

The Hong Kong Institute of Education

in Partial Fulfillment of the Requirement for

the Degree of Doctor of Philosophy

August 2014

STATEMENT OF ORIGINALITY

I, LIU, Fulan, hereby declare that I am the sole author of the thesis and the material presented in this thesis is my original work except those indicated in the acknowledgement. I further declare that I have followed the Institute's policies and regulations on Academic Honesty, Copy Right and Plagiarism in writing the thesis and no material in this thesis has been published or submitted for a degree in this or other universities.

_____

LIU, Fulan

August 2014

THESIS EXAMINATION PANEL APPROVAL

Members of the Thesis Examination Panel approved the thesis of LIU Fulan defended on 31/07/2014.

Principal Supervisor
Dr. STAPLETON, Paul
Head
Department of ELE
The Hong Kong Institute of Education

External Examiner
Prof. LEE, Icy
Professor
Faculty of Education
The Chinese University of Hong Kong
Dr. YOU, Xiaoye
Associate professor
English Department
The Pennsylvania State University

Associate Supervisor
Prof. BENSON, Phillip
Professor
Department of ELE
The Hong Kong Institute of Education
Dr. MA, Qing
Assistant professor
Department of LML
The Hong Kong Institute of Education

Internal Examiner
Dr. WANG, Lixun
Associate professor
Department of LML
The Hong Kong Institute of Education

Approved on behalf on the Thesis Examination Panel:

Chair, Thesis Examination Panel
Prof. Wong, Wen Chung
Associate Vice President
The Hong Kong Institute of Education

ABSTRACT

Investigating critical thinking in the argumentative writing of undergraduate English majors at a mainland Chinese university: Implications for policy changes

by LIU, Fulan

for the degree of Doctor of Philosophy

The Hong Kong Institute of Education

Argumentation is generally perceived as a tool for critical and analytical thinking (Crammond, 1998; Walton, 2007). Empirical studies suggest that students' argumentative writing is an effective vehicle for promoting student learning and critical thinking (Kuhn & Crowell, 2011; Stapleton, 2001). The present research project, comprising three successive studies, investigated critical thinking abilities in the argumentative writing of undergraduate English majors in mainland China.

Study 1 examined students' performance in and perceptions of written argumentation. It further explored what factors might have influenced students' argumentative writing and critical thinking. The findings of Study 1 indicated that both typical classroom instruction and writing prompts, among other factors, did little to enhance students' critical thinking in terms of acknowledging and refuting alternative viewpoints (counterargumentation) in their written argumentation.

Study 2 investigated the effect of an instructional intervention in counterargumentation on students' argumentative writing and critical thinking. Using a modified Toulmin model of argumentation (1958), the intervention aimed to improve students' argumentative strategies, especially counterargumentation skills. A pretest-posttest design was used on experimental and control groups with 125 participants at a Chinese university. The control group received instruction in argumentative writing (which typically ignores counterargumentation), while the experimental group received instruction which included counterargumentation. The results of the study demonstrated the efficacy of explicit classroom instruction in counterargumentation. The inclusion of counterarguments and rebuttals was found to be significantly positively correlated with the overall quality of an argumentative essay, and the posttest score of the experimental group was significantly higher than that of the control group. Additionally, the experimental group displayed significantly improved critical thinking ability.

To extend the inquiry into the area of assessment, a third study was devised to investigate how the writing prompt might be having an impact on students' critical thinking in their argumentative writing. Study 3 consisted of two phases. In Phase 1, the prompts from three high-stakes tests, TOEFL, IELTS and TEM4, were investigated for two elements: rhetorical function and object of enquiry. Results revealed that both elements converged around a narrow set of functions and content. In Phase 2, control and experimental groups comprising 129 undergraduates in China wrote essays on a prompt deemed "conventional" by the findings of Phase 1, and an exploratory prompt respectively. Various differences between the two sets of essays

were noted related to standardized indexes of writing quality, as well as other rhetorical and linguistic features including: use of metadiscourse, essay organization, and use of certain lexical items. The results suggested that conventional prompts tended to produce formulaic responses while prompts engaging problem-solving could stimulate high-order thinking.

The findings of the three studies may have important implications for writing assessment as well as argumentative writing pedagogy in China and beyond. It is proposed that counterargumentation be considered in the writing prompts and rubrics of high-stakes English tests, and included in classroom instruction on argumentative writing. It is also contended that a wider range of prompts may broaden the scope of written language and forms of critical reasoning to the benefit of students.

ACKNOWLEGEMENTS

I am particularly grateful to my three supervisors: principal supervisor, Dr. Paul Stapleton and associate supervisors, Prof. Phillip Benson and Dr. Ma Qing. Dr. Stapleton's kind support, great patience and constant guidance got me through this research project. Without him, I would not have completed the project. Prof. Benson and Dr. Ma gave valuable advice at critical moments during the three years of study. Thank you! I am lucky to have you as my supervisors.

I am indebted to the teacher and student participants at Jiangxi Normal University where I collected most of my data. My colleagues and students actively engaged in the project in various ways: Ms. Fu Dan helped me with the instructional intervention; Ms. Chen Yanru and Ms. Hu Bo undertook the hard work of data processing and coding. I really appreciate their work.

Special thanks go to my thesis examiners Prof. Icy Lee, Dr. You Xiaoye and Dr. Wang Lixun. I would like to thank them for reading and commenting on my thesis. They are researchers and educators that I admire.

I obtained help from many people when writing the thesis. Academic staff at the Hong Kong Institute of Education helped me in research with their expertise, and supported me with their friendship! The Dean, Prof. Lo Sing kai, and other staff of the Graduate School at HKIED provided assistance in my study. Especially, Prof. Ann Johns helped me strengthen the introduction chapter of the thesis. Many friends and

fellow PhD students here in Hong Kong and back in mainland China helped to ease the tension during stressful times. I would like to take this opportunity to express my heartfelt gratitude to them.

Finally I would like to thank my family. Their unconditional love and selfless support provided me with the strength to complete the PhD project, and to carry on.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

CT   critical thinking

EAP   English for academic purposes

EFL   English as foreign language

ESL   English as second language

IELTS   International English Language Testing System

L1   English as first language

L2   English as second language

NACFLT   National Advisory Commission on Foreign Language Teaching in

            Higher Education

TEM   Test for English Majors

TOEFL   Test of English as a Foreign Language

WSU   Washington State University

WTO   World Trade Organization

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Motivation for the Research

Before I started my PhD research in Hong Kong, I had been lecturing on English writing at a mainland Chinese university for more than 10 years. For undergraduate English majors at that provincial public university, the long-standing English composition textbook was the same one that was being used in dozens of key universities in China. Like other composition teachers, I prepared my teaching materials based on that textbook, focusing on academic writing. Exposition and argumentation were the two types of writing dominating most of the classroom teaching, which has been the practice of most teachers in order to help students in the high-stakes exams they are to take: Test for English Majors Band 4 (hereafter TEM4), Test for English Majors Band 8 (hereafter TEM8), and others. Until the point of time that I commenced my doctoral studies, I found my teaching fulfilling. Because much of my teaching focused on written argumentation, I was drawn to this topic and began reading some of the books and journal articles in the area. During the reading, it promptly came to me that good argumentative form always includes a consideration of alternatives views. However, I realized that neither in my teaching nor in writing tests were alternative views appropriately accommodated. In other words, the classroom instruction and writing prompts of TEM4 & 8 tests did not encourage students to consider views that are alternative or opposite to the writer's own point of view, and without such a consideration, rebuttals were also absent.

Furthermore, the writing prompts used in high-stakes tests seemed stereotypical, i.e., the types of writing required of students and the spheres of discussion tended to be confined to a narrow range. As a consequence, students' responses tended to be quite uniform. I noticed that it was common among students to use writing templates, which, with changes in content words and phrases, could be easily molded into a response for the writing tasks on most high-stakes tests. Noticeably, these templates rarely acknowledge alternative views and respond to them (the process is termed counterargumentation) in argumentative essays. However, as I came to realize after extensive reading of the literature in this area, counterargumentation has been widely acknowledged as an essential element in written argumentation. Moreover, argumentative writing is closely related to argumentative reasoning, which is generally perceived as an important vehicle for critical thinking. This recognition motivated me to undertake further research in the area. Thus, I decided to carry out an investigation into undergraduate English majors' critical thinking as demonstrated in their argumentative essays at the mainland Chinese university mentioned above.

1.2 The Background and Key Concepts

This section outlines the background for the research project and explains the rationale behind the research while introducing key terms and concepts.

1.2.1 Argumentation, Counterargumentation and the Toulmin Model

Argumentation[1] refers to the process of constructing and supporting a claim (one's position about an issue) by providing data - evidence and grounds for the claim (Palmer, 2012; Toulmin, 1958). Although argumentation has a long history dating back to at least the ancient Greeks when Aristotle classified means of persuasion into ethos, pathos and logos, it is particularly since British philosopher Stephen Toulmin introduced his audience-based scheme of argumentation (1958), which included consideration for those who do not agree with the writer (Ramage, Bean & Johnson, 2010), that the study of argumentation has entered a new epoch of informal logic[23].

In essence, Toulmin (1958) perceived argumentation as a verbal exchange procedure between an arguer and a critical listener, and the latter has the right to challenge during the process. The original Toulmin model of argumentation (1958) is composed of six elements (see Figure 1.1 for the Toumin scheme). The first four are claims, data, warrants and backing: data are evidence that support the claim; warrants are similar to assumptions in that they authorize the inferences that arrive with the data; and backing is support for the warrant. The two other elements are qualifiers, which place limits on the strength of the initial claim, and rebuttals, which acknowledge that

---

[1] In this thesis, until it is specifically pointed out otherwise, argumentation refers to written argumentation.

[2] Informal logic is the term used as opposed to formal logic which refers to the traditional way of logic reasoning, e.g., by syllogism. Stephen Toulmin (1958) maintained that people do not argue with formal logic in their ordinary communication. He proposed a new argumentation model which is considered as starting an age of informal logic in the study of argumentation.

[3] Argumentation as a field of study has originated and thrived in the West from ancient times. However, the East has its long tradition of argumentation as a social practice, which will be covered in detail in Chapters 2 & 5.

despite the careful construction of the argument, there may still be counterarguments. It is these two final elements, qualifiers and rebuttals, that are of particular interest for this study because they underscore the importance of entertaining alternative views.

```
┌─────────────────────────┐      ┌──────────────────┐      ┌─────────────────────────┐
│ Data                    │      │ Qualifier        │      │ Claim                   │
│ (reasons or evidence)   │ ──►  │ (limitation)     │ ──►  │ (statement of one's     │
│ e.g. Harry was born in  │      │ So, presumably   │      │ positions or opinions)  │
│ Bermuda                 │      │                  │      │ Harry is a British      │
│                         │      │                  │      │ subject                 │
└─────────────────────────┘      └──────────────────┘      └─────────────────────────┘
        │                                 │
        ▼                                 ▼
┌─────────────────────────────┐   ┌──────────────────────────────────┐
│ Warrant                     │   │ Rebuttal                         │
│ (linking data to claim)     │   │ (description of counter-examples)│
│ Since a man born in Bermuda │   │ Unless both his parents were     │
│ will generally be a British │   │ aliens                           │
│ subject                     │   │                                  │
└─────────────────────────────┘   └──────────────────────────────────┘
        │
        ▼
┌─────────────────────────────┐
│ Backing                     │
│ (support to back up the     │
│ warrant)                    │
│ The legal provisions have   │
│ it that…                    │
└─────────────────────────────┘
```

Figure 1.1 Toulmin's scheme of argumentation

In this thesis, the process of acknowledging alternative or opposing views and refuting them is termed counterargumentation. Counterargumentation involves considering counterarguments which are views brought up by those who disagree with the writer, and proposing rebuttals which are arguments used by the writer to refute the opposing views. It should be noted that Toulmin did not bring about the notion of counterargumentation directly. However, the notion has evolved from Toulmin's model of argumentation (1958) in general, and qualifiers and rebuttals in

particular. Actually the concept of counterargumentation originated from Toulmin's (1958) model and reached maturity in the theoretical work of Van Eemeren, Grootendorst, and Henkemans (1996) and Walton (1998, 2007). Van Eemeren et al. (1996), based on the Toulmin model, adopted a pragma-dialectical approach to argumentation that perceives the procedure of argumentation as providing arguments to justify one's standpoint, and to refute others' standpoints. Consonant with the Toulmin approach and the pragma-dialectical approach, Walton's (1998, 2007) dialog theory proposes that the dialectical conception has at its core raising critical questions, elaborating one's own position and responding to arguments with counterarguments and rebuttals. The notion of counterargumentation has also been developed in empirical studies. Texts that acknowledge and rebut opposing views are found to be more persuasive than those that did not (Crammond, 1998; Kuhn, 1991; Leitão, 2003; Nussbaum & Kardash, 2005; O'Keefe, 1999; Santos & Santos, 1999).

Studies on the argumentation schemata further underpin the importance of counterargumentation. Ferretti, MacArthur and Dowdy (2000) and Nussbaum and Schraw (2007) found that people make use of argumentation schemata when composing written argumentation. A complete argumentative schema comprises the production of arguments and counterarguments regarding an assertion, and the integration of all elements into a coherent piece (Piolat, Roussey, & Gomber, 1999). Under the argumentation schema hypothesis, a viewpoint-support schema is the basic schema while an argument-counterargument-rebuttal schema is termed an expanded or elaborate schema (Ferritti et al., 2000). Only when the two schemata work together is the argumentation sound and strong enough to be called good argumentation.

In this project, argumentation was analyzed in light of a modified Toulmin model of argumentation (1958). Four argumentative elements are identified: claim (an assertion in response to a contentious topic or problem), data (evidence to support a claim), counterarguments (the possible opposing views and evidence that can challenge the validity of a writer's claim), and rebuttals (statements and evidence in which the writer responds to a counterargument). Noticeably, these terminologies differentiate from the original elements in the Toulmin model (1958). In fact, the original terms of the Toulmin model have been changed to facilitate classroom instruction or to fit various research contexts (Gleason, 1999; Qin & Karabacak, 2010). Especially, the term *rebuttal* has grown into a meaning distant from Toulmin's use of the word.

In summary, counterargumentation has evolved from the Toulmin (1958) model of argumentation and developed in theoretical and empirical work. As a key notion in this research project, counterargumentation plays a fundamental role in students' argumentative writing and critical thinking development. The next two sections delineate how counterargumenation interplays with argumentative writing and critical thinking.

1.2.2 Counterargumentation and Persuasiveness of Argumentative Writing

There is no need to write arguments on a universally accepted fact; hence the very nature of argumentative essays is controversiality (Inch & Warnick, 2010). In argumentative writing, the writer presents arguments on an issue in order to persuade the reader to agree with a particular point of view (Chandrasegaran, 2008; Rothery,

1996; Schleppegrell, 2004). Within such a framework, persuasiveness could be defined as the extent to which a writer of an argumentative essay can convince her readers of a certain stance taken. Argumentative essays normally deal with a debatable issue. For a skilled writer of argumentative essays who takes a position on an issue and supports it, it is equally important to consider alternative views (Wolfe & Britt, 2008).

The importance of including counterarguments and rebuttals for making written argumentation persuasive has been underscored by much research (Kuhn, 1991; Leitão, 2003; Van Eemeren et al., 1996; Walton, 2007). Kuhn (1991), in the book *The Skills of Argument*, stated that one could not convince others if they were not aware of alternative views no matter how strongly they believed in their own views. Thus, Kuhn (1991) perceives counterargumentation as the core of persuasive argumentation and critical thinking. In empirical research, neglecting alternative views has been suggested as a general weakness in the argumentative writing of students at both the secondary level (e.g., Ferretti, Lewis & Andrews-Weckerly, 2009; McCann, 1989; Yeh, 1998) and the tertiary level (e.g., Nussbaum & Schraw, 2007; Wolfe & Britt, 2008).

The tendency of students to support only their preferred viewpoint while ignoring evidence against their own positions was first termed "myside bias" by Perkins and his colleagues (Perkins, 1985; Perkins, Farady & Bushey, 1991). Over the past two decades, myside bias has been investigated by researchers in the fields of critical thinking as well as argumentation and found to be a characteristic feature impeding

the persuasiveness of students' arguments (e.g., Baron, 1991, 1995; Goh, 2008; Santos & Santos, 1999; Toplak & Stanovich, 2003; Wolfe & Britt, 2008; Wolfe, Britt, & Butler, 2009).

A number of studies have demonstrated that the process of counterargumentation not only strengthens an arguer's position and helps achieve completeness in good reasoning, but also enables the writer to avoid myside bias; hence, incorporating counterarguments and refuting counterarguments are crucial for maximizing the extent of persuasiveness in argumentative writing (e.g., Leitão, 2003; Perkins et al., 1991; Stanovich & West, 2008; Toplak & Stanvich, 2003; Wolfe & Britt, 2008; Yeh, 1998; Zohar & Nemet, 2002). Highlighting this essential role of counterarguments was a meta-analysis of 107 studies on argumentative texts conducted by O'Keefe (1999) revealing that "two-sided" messages, where opposing arguments were acknowledged, were more persuasively efficacious than "one-sided" messages (those that ignore opposing arguments).

Counterargumentation not only makes an argumentative essay more persuasive, but is also evidence of the writer's critical thinking competence. The next section explains the connection between argumentative writing and critical thinking, and the significant role that counterargumentation plays in critical thinking development.

1.2.3 Counterargumentation, Argumentative Writing and Critical Thinking

Critical thinking and the writing of argumentative essays are closely connected

(Palmer, 2012; Ramage et al., 2010; Stapleton, 2001; Yeh, 1998). Critical thinking involves reasoning which is a sophisticated thinking process eventually communicated in language (oral or written arguments) (Palmer, 2012). As such, argumentation is considered a tool for critical and analytical thinking (Crammond, 1998; Walton, 2007). In general, the whole process of argumentative writing entails critical thinking. In an argumentative writing task, students are expected to analyze an issue, recognize multiple aspects of it, state and support a proposition (the main argument or thesis of the essay), anticipate counterarguments and respond to them (Yeh, 1998). These elements, especially acknowledging opposite views to one's own and responding to them, are key components of critical thinking (Brown & Keeley, 2012; Palmer, 2012).

Critical thinking has been pursued as a goal of education for over a hundred years (Ennis, 2003). It has been multiply defined and encompasses several facets, e.g., evaluating, analyzing and problem-solving. In essence, critical thinking entails two dimensions: abilities and disposition. Both dimensions embed a fairness of mind and viewing two or more sides of a case as essential elements (Perkins & Tishman, 2001). Critical thinkers tend to approach an issue with a balanced view that includes arguments for and against a position (Inch & Warnick, 2010). Critical thinkers also tend to be open-minded, which involves predicting an audience's needs, anticipating counterarguments and questioning their own assumptions (Ramage et al., 2010). In this sense, good skills in argumentation promote both dimensions of critical thinking. Empirical studies indicate that active engagement in written argumentation is a fruitful path to the development of critical thinking abilities and disposition (Goh,

2008). Thus, in this research, students' critical thinking is operationally defined as the process of working towards complete argumentation which addresses alternative viewpoints. It is further defined as an ability to respond to writing prompts that encourage widely ranging facets of higher-order thinking.

Given the significance of counterargumentation for both argumentative writing and critical thinking, it is desirable if students could acquire sufficient training in argumentative skills including counterargumentation. The next section describes the utilization of the Toulmin model (1958) in classroom instruction in written argumentation as scaffolding for students to develop counterargumentation and other argumentation skills.

1.2.4 The Application of the Modified Toulmin Model in Argumentative Writing Instruction

The Toulmin model of argumentation (1958) has been embraced by educators and scholars as a promising approach for generating and evaluating arguments, and has been employed in numerous writing classes and textbooks since the late 1970s. The informal logic approach that Toulmin adopted towards argumentation enables composition teachers to break the boundaries of syllogism when teaching argumentative skills. Furthermore, the Toulmin model (1958) facilitates the awareness of alternative views during the process of argumentation, which is regarded as fundamental for the pedagogy of writing as well as critical thinking. Since the Toulmin model was first introduced by Kneupper (1978), it has been

applied as an invention heuristic to help students produce arguments. The model's other function – to analyze argumentative texts – has also been explored by many scholars (Fulkerson, 1996b).

However, the application of the original Toulmin model is deemed too complicated for classroom instruction (Fulkerson, 1996b). Thus, modified versions have been employed in a considerable number of empirical studies. The literature reveals the Toulmin model has been modified in both simplified and elaborate ways. Simplified modifications, i.e., focusing on three or four of the argumentative elements while leaving out the others, were reported in studies by Bacha (2010), Burkhalter (1995) and Varghese and Abraham (1998). For instance, in their research targeting undergraduates, both Bacha (2010) and Varghese and Abraham (1998) utilized the modified Toulmin model in writing instruction and assessment focusing on three argumentative elements: claim, data and warrants in students' essays. In yet other studies, the Toulmin model has been adapted in elaborate ways, i.e., the *qualifier* and *rebuttal* element was elaborated into *opposition* and *response to opposition* or *counterarguments* and *rebuttals* (e.g., Crammond, 1998; Knudson, 1992; MaCann, 1989; Qin & Karabacak, 2010).

The application of the modified Toulmin model to the composition classroom can be explicit and implicit. Since the ability to write a good argumentative essay is not acquired, but rather needs intensive training (Graff, 2003), explicit instruction is considered critical to train students in essential argumentative skills (Graff, 2003; Voss & Means, 1991). Empirical studies (Gleason, 1999; Knudson, 1992; Nussbaum

& Kardash, 2005; Nussbaum & Schraw, 2007; Stapleton, 2001; Yeh, 1998) indicate salient efficacy when a modified Toulmin model is used in explicit instruction on argumentative writing.

Whilst theoretical and empirical studies suggest the efficacy of the modified Toulmin model in cultivating students' argumentative writing and critical thinking (e.g. Fulkerson, 1996b; Van Eemeren et al., 1996), an extensive review of the literature reveals that even though a few recently published composition textbooks introduce the Toulmin model (e.g., Qu, 2005, 2013), the most popular ones ignore the Toulmin model and counterargumentation (e.g., Ding, Wu, Zhong, & Guo, 1994). This situation begs the question: how is argumentative writing taught at Chinese universities? The next section is a sketch about the pedagogy for, and the development of, undergraduate English majors' argumentative writing and critical thinking in China.

1.2.5 Pedagogy for Critical Thinking and Argumentative Writing of Undergraduate English Majors in Mainland China

Critical thinking has drawn considerable attention since American educator John Dewey (1910) made his appeal to the public about the significance of reasoning and thinking in education. However, despite the alarm given by Dewey, instruction in critical thinking is far from satisfactory in contemporary Western societies (Goh, 2008; Paul, Elder & Bartel, 2004). This deficiency has also been noted among Chinese EFL students at the tertiary level, either studying at home or abroad

(Atkinson, 1997; He, Yin, Huang, & Liu, 1999; Huang, 2010; Paton, 2005; Tian & Low, 2011).

In China, although the National Teaching Syllabus for English Majors (hereafter the Syllabus, NACFLT, 2000) lists developing students' "analytical thinking", and the capability of "approaching an issue from multiple perspectives" (p.12), as one of the chief pedagogical goals, a number of well-established Chinese scholars in the English as Foreign Language (hereafter EFL) circle have noted that university students majoring in English language seriously lack critical and analytical thinking competence (He et al., 1999; Huang, 1998, 2010; Wen & Liu, 2006; Wen, Wang, Wang, Zhao, & Liu, 2010). Ironically, however, these are the very scholars who took part in the writing of the Syllabus (NACFLT, 2000). They found the current situation dismaying despite the fact that they have listed the fostering of critical and analytical thinking as one of the primary goals of the education for English majors.

Undergraduate students majoring in foreign languages as a whole demonstrate weaker abilities in analysis, reasoning and evaluation than students in other academic disciplines (He et al., 1999; Huang, 1998). Huang (1998) asserted that "the critical thinking deficiencies impair the creativity, research and problem-solving abilities of people in the English language discipline."(p. 1) These warnings have drawn educators' and scholars' attention to the issue of critical thinking development of undergraduate English majors. In light of this, ongoing national research projects have been undertaken to address the issue. Wen et al. (2010), for example, conducted a nation-wide survey on university students' critical thinking development across

academic majors of liberal arts and found that, compared to students in other majors, English majors' critical thinking growth is slow. Nevertheless, these research projects were targeted more on revealing the gravity of the problem rather than tackling it. As such, the research on the critical thinking development of English majors at Chinese universities needs to be taken further.

Whilst the deficiency of critical thinking education for English majors has begun to be noted by researchers, pedagogy on argumentative writing remains underdeveloped in two senses: first, the quantity of studies on argumentative writing pedagogy is small. In particular, argumentation skills in argumentative essays of university undergraduates in mainland China is a field insufficiently explored; second, according to extant literature, the counterargumentation skills of university students' in their argumentative essays are underdeveloped. In one empirical study, it was found that counterarguments and rebuttals were virtually absent in argumentative essays produced by a group of undergraduate English majors at Chinese universities (Qin & Karaback, 2010). In other studies, there were also indications that this key notion of counterargumentation is lacking in Chinese students' written argumentation (Atkinson & Ramanathan, 1995; Tian, 2008).

This section sketches the current situations in the pedagogy of critical thinking and argumentative writing for undergraduate English majors in China. In an exam-oriented society such as the Chinese one, the impact of tests (especially high-stakes tests) on pedagogy can be substantial. In the next section, the writing prompts in two high-stakes tests are examined for their possible impact on students'

argumentative writing and critical thinking.

1.2.6 The Washback of High-stakes English Tests on Argumentative Writing Pedagogy

This section focuses on one prominent element of writing assessment - writing prompts. It deliberates the impact of writing prompts in high-stakes tests on the classroom instruction of argumentative writing and critical thinking.

The impact of tests on teaching and learning is referred to as washback (Alderson & Hamp-Lyons, 1996; Alderson & Wall, 1993; Weigle, 2002). It is believed that teachers are under certain pressure to adapt their teaching materials and classroom arrangement to cater for students' needs if they know that students are to take particular tests (Taylor, 2005). Washback can be positive and negative. Positive washback occurs when a testing procedure encourages teachers to adopt pedagogical practices that are based on the upgraded best thing in the field (Weigle, 2002), while negative washback results when a test's content and format is "based on a narrow definition of language ability," and hence is counterproductive in terms of teaching and learning (Taylor, 2005, p. 154; Weigle, 2002). The washback of high-stakes tests on students' argumentative writing has been taken into consideration in the research design of this project. In addition, the washback theory has been applied to the interpretation of students' performance and perceptions of argumentative writing. Illustrated below is how the writing prompts of two tests may have influenced English majors' argumentative writing.

The Two high-stakes exams that English majors are required to sit during their four-year study are TEM4 and TEM8. The TEM battery is aimed at evaluating the English language proficiency of Chinese university undergraduates majoring in English language. Since it was officially launched in 1992, it has gained considerable influence related to both academic qualifications and the job market (Jin & Fan, 2011). For instance, the TEM8 result has been used as a benchmark for an increasing number of companies and institutions to recruit employees. The TEM battery is now recognized as one of the most influential high-stakes English tests in China and is written by thousands of test takers each year (Cheng, 2008). In 2012, 205,004 university students took the TEM8 (TEM test center).

According to washback theory, the writing prompts of these two tests would have considerable impact on undergraduate English majors' argumentative writing development. In other words, classroom teachers as well as their students are acutely aware of the contents of this test and plan lessons and revision around it because of the potential future impact that the resulting score can have on a student's career and future. However, my initial investigation revealed that the writing prompts and rubrics in TEM evidently do not encourage counterargumentation, which is considered a key element of critical thinking in an argumentative essay. Taking the TEM8 writing prompt as an example, the prompt consists of a paragraph that introduces a topic followed by a goal instruction:

> *In the first part of your writing you should state your main*
> *argument, and in the second part you should support your*

argument with appropriate details. In the last part you should bring what you have written to a natural conclusion or make a summary.

This goal instruction requires test takers to make a claim (main argument) in their essay and support it with appropriate data. Notably, the writing prompt does not encourage students to consider alternative views, i.e., to counterargue. When the writing prompt of a high-stakes test completely ignores a fundamental argumentative skill – counterargumentation, it often has a subsequent washback effect on the teaching and learning. This research project is designed in part to uncover how the TEM battery, among other factors, impact on undergraduate English majors' argumentative writing and critical thinking. In return, the results of the research are expected to have important implications for the high-stakes tests.

In addition to their washback on pedagogy, writing prompts also exert a substantial effect on students' writing directly. As I stated in section 1.1, the writing prompts used in tests seem stereotypical in terms of topics and types of discourse, which may constrain students' argumentative writing and critical thinking. The next section provides a brief description of the prompt effects on students' writing.

1.2.7 Assessment for Argumentative Writing: the Writing Prompt Issue

This section discusses the effects of writing prompts on students' performance in argumentative writing via a brief review of pertinent studies. It further queries the

possibility of critical thinking cultivation through the use of prompts in writing assessment.

Key constituents of a writing prompt, particularly the object of enquiry and rhetorical function, can have a significant effect on students' writing performance in terms of rhetorical and linguistic features (Hinkel, 2002).

"Object of enquiry" refers to the subject matter in a writing prompt. Empirical studies have shown that different subject matter, i.e., personal or impersonal, specific or general, in writing prompts can elicit responses varying considerably in discourse (He & Shi, 2012; Hinkel, 1995, 2009; Lee & Anderson, 2007; Tedick, 1990; Yu, 2010). Yu (2010), for example, by examining the correlation between the object of enquiry and the lexical diversity in test-takers' essays, revealed that lexical diversity is higher in test-takers' response to impersonal topics. Hinkel (1995, 2009) found that students' language use when responding to particular writing prompts is dependent on the cultural and personal background, and topic knowledge. In light of these findings, Hinkel (2009) proposed that the exploration of new topics and prompts that are less proximate to test-takers' own experience and opinions may lead to richer linguistic variety. Since language use reflects one's higher order thinking, Hinkel's (2009) recommendation has given impetus to the design of this research project, specifically, the need to explore the possible influence of writing prompts with distinct objects of enquiry on students' written argumentation and critical thinking.

"Rhetorical function" is defined as the purpose of written or spoken discourse, e.g., explanation, recommendation or hortation (strongly encouraging or persuading other

people to accept an argument). As indicated by empirical research, different rhetorical tasks in writing prompts may produce various linguistic features (e.g., Hinkel, 2002; Reid, 1990; Spaan, 1993). However, there is a tendency for the rhetorical functions of writing prompts on high-stakes tests to converge. In one study pertaining to the present research, Moore and Morton (2005) compared the writing prompts used as university assignments and IELTS writing tasks. They found that compared to the wide range of university assignments, the rhetorical function of the IELTS prompts mainly involved evaluation and hortation. The object of enquiry in the IELTS prompts was also confined to a narrow range. Given the washback of prompts on argumentative writing pedagogy, more divergence in the rhetorical function of writing prompts may allow more space for students' argumentative writing and critical thinking development. Similar to Hinkel's (2009) study, Moore and Morton's (2005) study inspired me to investigate the effect of writing prompts.

In view of the salient effect of writing prompts on students' argumentative writing development, and given the close interaction between critical thinking and written argumentation, one question to be raised at the end of this section is: can critical thinking education be encouraged via writing assessment? The literature shows that cultivating students' critical thinking via argumentative writing has been explored by scholars. For instance, White (1993) contended that the principal concern of writing assessment is "to assess actively the active thinking of students when they write" because in essence writing is both "the means and the expression of critical thinking and problem solving" (p. 106). His idea of cultivating higher-order thinking via writing is upheld by other researchers and educators (e.g., Alderson & Wall, 1993;

Bean, 2011; Eckstein & Noah, 1993; Shirkhani & Fahim, 2011). It is worth noting here that higher order thinking does not equate exactly with critical thinking, but critical thinking is evidently a crucial dimension of higher order thinking. If critical thinking skills such as problem solving and analyzing are appropriately represented in writing prompts, which is the primary factor of assessment for writing, teachers and learners are more likely to be engaged in practice to this end. However, to date, few empirical studies have been conducted to explore the idea of enhancing critical thinking competence through writing assessment. This further manifests the research niches of this study which are discussed in the next section.

1.3 The Research Gap

A review of literature reveals that compared with studies on overseas Chinese students' critical thinking (e.g., Atkinson, 1997; Paton, 2005; Tian & Low, 2011; Turner, 2006), those concerning domestic Chinese students come later in time and seem fewer in number. Even fewer papers address critical thinking in the argumentative writing of mainland Chinese university students. In general, the research niche of this study lies in two aspects. First, in the L2 context, more research is needed with regard to the effect of explicit instructions in essential argumentative skills on fostering critical thinking. Second, in the Chinese context, there is a dearth of empirical studies on the cultivation of critical thinking in undergraduate English majors via their argumentative writing. Below is a specific description of the research gaps by comparing what has been done and what needs to be done in this field.

With the underdeveloped critical thinking of university undergraduates studying social sciences having been noted by educators, some research projects and empirical studies have been conducted to address the issue. For instance, both Su (2011) and Wen and Liu (2006) surveyed critical thinking in undergraduate English majors' argumentative writing. In addition, a few scholars have stressed the connection between argumentative writing and critical thinking in books on argumentative writing (e.g., Zhang, 2011), while Qu (2005, 2013) has introduced the Toulmin model in his books. Nevertheless, three research gaps exist as follows. First, none of the aforementioned researchers or scholars have explored the role of counterargumentation in enhancing students' argumentative writing performance and critical thinking abilities. Second, there appear to be no empirical studies exploring the effect of explicit instruction in counterargumentation on students' argumentative thinking and critical thinking development. Third, the literature shows no empirical studies associating the cultivation of critical thinking with the argumentative writing prompts. This study was designed to fill these gaps.

1.4 Research Objectives and Questions

This research project investigated Chinese university students' critical thinking ability as demonstrated in their argumentative essays. An overarching hypothesis in this research project is that the argumentative writing and critical thinking abilities of Chinese university students can be enhanced through an instructional intervention in counterargumentation. It is further hypothesized that diversity in the objects of enquiry and rhetorical functions of a writing prompt will bring about variety in

students' responses and improvement in critical thinking skills.

The project comprises three studies, the second and third of which emerged from the results of the study that immediately preceded them. The first study examined students' performance in and perceptions of argumentative writing, and how some factors (identified via a questionnaire and interviews with students and teachers, and observations) impacted students' performance and perceptions. Three research questions were asked in Study 1:

RQ1: How do a group of mainland Chinese undergraduates respond to a typical argumentative writing prompt? In particular, do they produce any counterarguments and rebuttals?

RQ2: What are these students' perceptions of argumentative writing?

RQ3: What factors contribute to students' performance in and perceptions of argumentative writing? How are counterargumentation and critical thinking represented in these factors?

As explained in Chapter 4, Findings, in answer to these three questions, counterargumentation was found to be lacking in students' essays. Questionnaire and interview data revealed that counterargumentation was also substantially ignored in their perceptions of argumentative writing. The data showed that the top four factors having an influence on students' argumentative writing were: classroom instruction, high-stakes writing prompts, composition textbooks and the Syllabus. Investigation into these factors indicated that although critical thinking is highlighted in the

Syllabus, counterargumentation – one of the hallmarks of critical thinking – was underrepresented in the textbook, the high-stakes test writing prompts and their rubrics, and classroom instructions.

Informed by the findings of the first study, a second study was devised to explore the effect of explicit instruction in counterargumentation on students' argumentative writing performance and critical thinking ability, using an adapted Toulmin's model of argumentation. Thus, four research questions were asked in Study 2:

RQ4: How does an intervention in counterargumentation affect students' written argumentation performance in terms of quantity of counterarguments and rebuttals, and overall quality of an argumentative essay?

RQ5: Is there a correlation between the extent of counterargumentation and the overall score of an essay?

RQ6: What are the students' perceptions of argumentative writing after the intervention?

RQ7: How does the intervention impact students' critical thinking?

As explained in the Findings chapter, the results of Study 2 revealed that the intervention was efficacious in helping students generate counterarguments and rebuttals in their essays. The inclusion of counterarguments and rebuttals was found to be significantly positively correlated with the overall quality of an essay. Notably, after the intervention, students' critical thinking ability was enhanced in their argumentative essays. The findings of Studies 1 & 2 inspired a third study that

highlighted the role of writing assessment in the critical thinking development of students.

The third study, Study 3, was designed to extend the exploration of the critical thinking issue. Study 1 indicated that both typical classroom instruction and writing prompts did little to enhance students' critical thinking in terms of counterargumentation in their written argumentation. Accordingly, Study 2 focused on the effect of the instructional intervention; after completion, the results of Study 2 suggested the possibility of a pedagogical solution to the problem. However, it was realized that in the context of the Chinese educational system, the curriculum is very much influenced by high-stakes examinations; thus, any study that ignores the impact of external assessment by both the state and international bodies would fail to fully address a critically important dimension. Therefore, to extend the inquiry into the assessment aspect of the issue of critical thinking development, a third study was devised to investigate how one element of writing assessment, that is, the writing prompt, might be having an impact on students' critical thinking in their argumentative writing. As the literature reveals, writing prompts have significant washback on classroom instructions as well as student's writing performance (Moore & Morton, 2005; Weigle, 2002). Two of the key properties of a writing prompt have been singled out in this study following the literature (Moore & Morton, 2005), namely, the rhetorical function and object of enquiry. The findings in the Moore and Morton (2005) study and my preliminary observations were that the rhetorical function and object of enquiry of writing prompts in high-stakes tests tended to be confined to narrow ranges. Writing prompts with convergent rhetorical function and

object of enquiry were termed conventional prompts in this study. A hypothesis was thus proposed: conventional writing prompts in high-stakes tests constrain students' critical thinking in their writing. To test this hypothesis, Study 3 was designed to compare the effects of two writing prompts with distinct rhetorical functions and objects of enquiry on students' critical thinking abilities. In order to do this, it was deemed necessary to first investigate the nature of prompts in high-stakes exams. Thus, two research questions were asked.

RQ8: What patterns do the writing prompts in three high-stakes tests, namely, IELTS, TOEFL and TEM4, display in terms of their rhetorical function and object of enquiry?

RQ9: In what ways do the essays of students responding to a "conventional" writing prompt differ from those of an exploratory, non-conventional prompt?

1.5 Significance of the Research Project

Few empirical studies so far have investigated mainland Chinese students' counterargumentation skills in their argumentative essays. Even fewer studies have taken a combined look at undergraduate English majors' critical thinking and argumentation skill development in the Chinese context. The significance of this project arose out of the need to tackle these issues, with the fundamental concern being how increased strength of argumentation could turn students into more skillful writers and better critical thinkers. In addition, Study 3 also explored the role of writing prompts in cultivating students' critical thinking.

You (2010b) has used a metaphorical expression to depict Chinese people's English writing experience: "writing in the devil's tongue." This reflects two things: to Chinese people, English writing is hazardous; to outsiders, it is a mystery. The research design of this project served to clarify some unclear aspects concerning mainland Chinese undergraduate students' argumentative writing. Because most of the previous studies have been text analysis of argumentative essays by Chinese EFL learners (Lu, 2011), how these students perceive written argumentation remains unknown. RQ2 of this research was aimed at eliciting data in this regard from questionnaires and interviews. RQ3 further explored the influence that the teaching and testing environments exert on students' performances in written argumentation. The major anticipated outcome of the pedagogical intervention in Study 2 was that students would be able to write more persuasive and convincing arguments by acknowledging opposing viewpoints to their own and going on to refute them. Since realizing alternative viewpoints is a hallmark of critical thinking (Palmer, 2012), the training of being open-minded to alternative and opposite views would eventually enable undergraduate English majors to be more competent critical thinkers.

As such, the impact of Study 2 is two-fold: at the theoretical level and the pedagogical level. First, given the prevalence of argumentative writing for curricular ends and for test purposes (Zhu, 2001), and due to the considerable challenge it poses both to student writers (Ferretti, Andrews - Weckerly, & Lewis, 2007; Zhu, 2001), especially EFL student writers and to the classroom pedagogy (Gleason, 1999), this research is expected to add evidence to the existing scholarship on argumentative writing and provide guidance to the teaching and learning of argumentative essays in

Chinese tertiary writing classes. Secondly, this study endeavors to address the issue of Chinese students' critical thinking deficiency by increasing the strength of their written argumentation, i.e., the inclusion of counterarguments and rebuttals.

In the Chinese context, the importance of argumentative writing can be perceived by examining the high-stakes English tests either at the international level or at the national level. In the high-stakes English tests taken by an increasing number of Chinese people, e.g. IELTS, TOFEL or GRE, or TEM 4 & 8, CET 4 & 6 (College English Tests) and the EPT (English Proficiency Test for Chinese civil servants), the writing tasks consist of solely or mainly the writing of an argumentative essay (Xu, 2006; Zhou, 2004). Argumentative writing is also considered by students and teachers to be the foundation of academic writing (Cheng, 2008). Therefore, argumentation has attracted more attention than other genres of writing at the tertiary level. Although composition teachers dedicate a great amount of their writing instruction to developing written argumentation skills, studies indicate that EFL writers encounter various difficulties when performing argumentative writing in English (Zhu, 2001). It is anticipated that the current study will shed light on classroom pedagogy, so that students will be able to strengthen persuasiveness in their essays and benefit from it.

It is stated in the Syllabus that the fostering of critical thinking competence is the major educational objective for undergraduate English majors. But in reality, Chinese university students have been considered lacking critical thinking skills and dispositions (e.g., He et al., 1999; Wen et al., 2010), and undergraduate English majors are found to be weaker than undergraduates of other academic majors in such

basic critical thinking abilities as analysis, reasoning and evaluation (Wen et al., 2010). Notwithstanding the official requirement for critical thinking instruction, writing teachers generally are not qualified in this matter, and therefore do not know how to incorporate critical thinking development into their curriculum. Consequently, Chinese students may be confused about how to be critical in an appropriate way when writing argumentative essays. This study, while unable to facilitate full-scale training of critical thinking abilities, addresses an important aspect of critical thinking, i.e., that sound reasoning involves considering the opposite view or multiple views on a certain issue.

As such, the impact of this study has potential to go beyond classroom pedagogy to the assessment of argumentative writing. The outcomes of Study 2 are expected to raise awareness of effective prompts and appropriate scoring for argumentative essays, i.e., the making of rubrics incorporating counterarguments and rebuttals, given the salient washback that tests would assert on the teaching and learning of argumentative writing. The findings of the study may also generate reflection on the current instructional material on argumentative writing. The anticipation is, when classroom pedagogy, instructional material and assessment all accommodate counterargumentation, they can jointly contribute towards the critical thinking development of undergraduate English majors. In this sense, the English writing class can become a key arena for critical thinking education.

The intervention design in Study 2 is a (partial) replication of previous research (e.g., Gleason, 1999; Nussbaum & Schraw, 2007). There are two justifications for the

replication. On the one hand, replication in L2 writing research is important. Porte & Richards (2012) argue that key quantitative studies need to be replicated to have their robustness and generalizability tested. This is a basic requirement of scientific inquiry. Scholars in the L2 writing field (e.g., Ferris, 2005; Leki, 1992) have also made calls for replication studies.

On the other hand, this appears to be a strong replication. The findings from Study 2 suggest that Chinese undergraduates' critical thinking can be enhanced through an instructional intervention in counterargumentation. Plus, the overall quality of students' argumentative writing is improved by adding counterarguments and rebuttals. More importantly, the findings may have implications for the pedagogy and evaluation of argumentative writing in China. These two justifications for the replication further underscore the significance of this research project.

Study 3, despite being exploratory in nature, was quite revealing. It drew further attention to the important role writing assessment plays in the cultivation of critical thinking for students. The results indicate that high-stakes language tests may confine test-takers in their writing and thinking, due to the narrow range of rhetorical functions and objects of enquiry of writing prompts. In particular, the result that students responding to an experimental prompt demonstrate better abilities in higher order thinking may bear significant educational significance. It is thus contended that a wider range of prompts that takes better consideration of the many specific purposes encountered in academic contexts may broaden the scope of written language and forms of reasoning to the benefit of students.

Although the findings from the three studies bear significance on pedagogy and assessment in China, I believe that the findings may also carry a broader significance, i.e., beyond the Chinese context. Firstly, the population of Chinese students studying overseas has been increasing in the past few decades, and these results can shed light on the training of argumentative writing and critical thinking to Chinese students for international educational institutes. Second, the part of the results concerning the washback of high-stakes tests (IELTS and TOEFL) have the potential to lead test developers to reconsider their writing prompts, which has the possibility to eventually influence pedagogy to the benefit of English language learners beyond China.

This chapter outlines the framework for the three studies in this project. In the next chapter, a thorough review of the literature will substantiate the theoretical foundation and rationale for the research design.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter reviews the relevant literature, which provides evidence to substantiate the theoretical framework of the three studies in the research project. The main purpose of this project is to investigate the critical thinking ability in the argumentative writing of undergraduate English majors in a mainland Chinese university. Theoretical and empirical studies that underpin this research are reviewed in twelve sections briefly outlined below. Argument and argumentation are defined in section 2.2 and a succinct synopsis of the studies on argument and argumentation is given in section 2.3.

A primary foundation of the theoretical framework in this project is the Toulmin model of argumentation because it is the conceptual origin of counterargumentation – a central theme in this research and a hallmark of critical thinking in argumentative writing. The theoretical and empirical studies on the Toulmin model of argumentation and its use in argumentative writing pedagogy are outlined in sections 2.4 and 2.5. The role of counterargumentation in the persuasiveness of an argumentative essay is discussed in section 2.6.

The interplay between the argumentative writing and critical thinking skills is discussed in section 2.7. Research on these two sets of fundamental and interrelated

skills by L2 learners in general, and Chinese L2 learners in particular, is reviewed in sections 2.8 and 2.9 respectively. The influence of culture and culture-related factors on the critical thinking and argumentative writing of Chinese students is delineated in section 2.10.

Because Study 2 concerns the pedagogy of argumentative writing and Study 3 involves high-stakes language tests, studies on the washback effect of high-stakes language test on argumentative writing pedagogy are reviewed in section 2.11. The last section in this chapter, section 2.12 focuses on the scholarship on the writing prompt and its effect on argumentative writing and critical thinking performances.

2.2 Definitions of Argument and Argumentation

Argument and argumentation, two terms derived from the verb "argue", have been used interchangeably or without distinctive differences by some researchers (e.g., Dowdy, 1998; Kuhn, 1991) while being differentiated from each other by many other scholars (e.g., Toulmin, 1958; Stapleton, 2001; Van Eemeren et al., 1996). To these scholars, an argument is the "product" of arguing, and argumentation is the "process" of arguing. For instance, Toulmin, Rieke and Janik (1984) see an argument as "a train of reasoning", and define it as "the sequence of interlinked claims and reasons that establish the content and force of the position for which a particular speaker is arguing" (p.14). According to them, argumentation refers to the "whole activity of making claims, challenging them, backing them up by producing reasons, criticizing those reasons, [and] rebutting those criticisms" (Toulmin et al., 1984, p.14). Van

Eemeran et al. (1996) extend the definition of argumentation so that it "not only refers to the activity of advancing reasons but also to the shorter or longer stretches of discourse or text resulting from it"(p.5). From their perspective, argumentation is both the process and the product of arguing.

Based on Toulmin's (1958) model of argumentation, several composition textbooks (e.g., Inch & Warnick, 2010; Palmer, 2012; Ramage et al., 2010; Rottenberg & Winchell, 2009) provide simpler and more direct definitions for argument and argumentation. In summary, these authors agree that an argument consists of a claim, which is a statement of one's opinion on an issue, and supporting evidence for the claim; argumentation is the process of connecting individual arguments, counterarguments and rebuttals to construct the overall position. These definitions are widely acknowledged by researchers and educators; hence, they are adopted in this project.

In this project, argumentative writing refers to a piece of written argumentation, and university students' argumentative essays are usually a short essay on a controversial topic ranging from 200 to 400 in words. While such a short piece of writing would not normally be considered an "essay," the nature of the participants in the present research – freshmen and sophomores in mainland China writing time-limited responses in their second language (L2) – requires some allowances to be made regarding length.

2.3 The Evolution of Theories and Various Perspectives on Argument and Argumentation

Generally speaking, studies on argumentation can be categorized into formal and informal logic approaches; however, the Toulmin (1958) model of argumentation triggered the transition of the mainstream of argumentation research from the formal logic approach to the informal logic one (Van Eemeren, 2009).

2.3.1 The Formal Logic Approach to Argumentation

In essence, arguing is a way of logical reasoning, the study of which can be traced back to ancient Greek philosophers. Aristotle has been regarded as one of the earliest influential argumentation theorists. Theories of classical logic, dialectic and rhetoric developed by Aristotle have been the primary sources for modern theoretical thinking on argument and argumentation (Van Eemeren et al., 1996). At the centre of Aristotle's philosophy of argumentation is the assumption that existing opinions provide a basis for new opinions to be formed through reasoning or arguing. For this purpose, he divided arguments into two sorts: deductive syllogisms and inductive syllogisms. Aristotle also sorted out arguments based on their respective objectives; by such criteria, there are demonstrative, dialectical and rhetorical arguments (see Table 2.1). Aristotle's approach to argumentation laid the foundation for formal logic. The formal logic view of argumentation prevailed for centuries.

Table 2.1

Aristotle's three sorts of argument and their characteristics

ARGUMENTS	Demonstrative	Dialectical	Rhetorical
Objective	certainty	acceptability	persuasiveness
Status of the premises	evidently true	acceptable	Persuasive to audience
Inference	valid	valid	Persuasive to audience
Theory	logic	dialectic	rhetoric

(Extracted from Van Eemeren et al., 1996)

Before the 1960s, the conceptualizing of argument was restricted to the field of formal logic (Wenzel, 2006; Wood, 2009). Early concepts of arguments as logical constructions, i.e., syllogisms, suggest an image of argument-as-product, the dominant idea being that arguments were "little units that speakers and writers built into their discourses and that critics could take out for evaluation" (Wenzel, 2006, p.10). However, speakers or writers in real life seldom produce formal syllogisms, i.e., ordinary people do not argue with formal logic (Toulmin, 1958).

2.3.2 The Informal Logic Approach to Argumentation

Realizing this limitation, scholars brought forth alternative views on argument. In his seminal work *The Uses of Argument*, Toulmin (1958) introduced a model of argumentation counter to the traditional formal logic and the model was further developed and elaborated in *An Introduction to Reasoning* (Toulmin et al., 1984). In addition, Perelman & Olbrechts-Tyteca (1958) used the term "new rhetoric" to name

a framework for all forms of "nonanalytic thinking" in seeking a theory of argument complementary to formal logic. Informal logic is perceived as an approach to argumentation largely based on the works of Stephen Toulmin and Perelman & Olbrechts-Tyteca (Van Eemeren, 2009). Especially, Toulmin's (1958) *The Uses of Argument* has been regarded as a major precursor of informal logic which is defined by Blair and Johnson (1987) as a branch of logic whose task is to develop non-formal "standards, criteria, procedures for the analysis, interpretation, evaluation, criticism and construction of argumentation" (p. 148). Furthermore, informal logic has been partnered, and even equated with critical thinking (Johnson, 2000).

In another attempt to expand the limited, formal logical view on arguing, Brockriede (1975, 1980) and Wenzel (2006) developed three perspectives on argument: rhetorical process, dialectical procedure and logical product. The three perspectives are significant in that they construct argument in a more lively and vigorous way. The more recent theory of argumentation in the informal logic camp is the "pragma-dialectical" approach brought by Van Eemeren and Grootendorst (1988) who are later referred to as the Amsterdam school.

Another important source of the informal logic approach is dialog theory, introduced by Grice (1975) and developed by Walton (1989, 1998, 2007), which perceives argument as a dialog between two parties holding opposing views. Walton's dialogic theory and the Bayesian model are deemed complementary to the Toulmin model (Van Eemeren et al., 1996).

Billig (1987) contended that everyday thinking takes on the form of argumentation as people weigh the evidence for claims and counterclaims. This view is endorsed by Kuhn. In her book, *The Skills of Argument*, Kuhn (1991) distinguished between dialogic argument and rhetorical argument, defining the rhetorical argument as "an assertion with accompanying justification" (p. 12), and the dialogic argument "entails juxtaposition of two opposing assertions" (p. 12). Kuhn's actual intention was to demonstrate the distinction between the two kinds of argument, that is, both are reasoned arguments comprising a weighing process between opposing assertions (Kuhn, 1991).

In summary, the informal logic approach to argumentation is dominant in contemporary society. Although a number of theories in the informal logic camp have thrived since the 1960s, the Toulmin model of argumentation has been the most influential theory of argumentation. Other theories of the informal logic approach to argumentation are to varying degrees based on, or derived from, the Toulmin model. Therefore, the Toulmin model and its application are the cornerstones of this research, and are discussed in greater detail in the following sections.

2.4 Theoretical Studies on the Toulmin Model and Counterargumentation

As Hitchcock and Verheij (2006) point out, modern argumentation theories concerning defeasible reasoning, argumentation schemes and field-dependency conception are based on the Toulmin schemata to some extent. The focus of this section is how the concept of counterargumentation originates from Toulmin's model

of argumentation and becomes mature in the works of Van Eemeren et al (1996), Walton (1998; 2007) and Kuhn (1991). In this research, the process of acknowledging alternative views and refuting them is termed counterargumentation which consists of counterarguments and rebuttals.

Notwithstanding Toulmin's intention of writing *The uses of argument* being to question the usefulness of formal logic in advancing human epistemology (Toulmin, 1958), the model he proposed to represent the layout of arguments has had a great impact on the analysis of argument and argumentation (Van Eemeren et al., 1996). In terms of counterargumentation and refutation, although Toulmin did not fully develop these concepts, he laid the foundation for them (Hitchcock & Verheij, 2005; Van Eemeren et al., 1996). By formulating the dialectical structure of arguments, Toulmin highlighted the fact that arguing involved supporting views as well as considering attacks against them. To this end, Toulmin discussed the significance of counterargumentation in the procedure of arguing.

Basically, Toulmin perceived argument as a verbal exchange procedure between an arguer and a critical listener, and the audience has the right to challenge during the procedure (Van Eemeren et al., 1996). According to Toulmin, six steps (elements) are involved in producing an argument. The first step is to make an assertion or, in his terminology, a *claim*. The second step is appealing to *data* (facts, evidence) as support for the *claim*. Considering the possible challenge an arguer may face, the next step is giving *warrants* to justify that the *data* lead to the *claim* in an appropriate and legitimate way. The *warrants*, in other words, are the logical connection between the

claim and *data*. *Claim, data* and *warrant* are three basic elements for analyzing arguments. Toulmin (1958) illustrated the first three steps along with the second three steps of argumentation including *rebuttals* (condition of exception), *qualifier* (hedging for the conclusion of argument, e.g., "presumably") and *backing* (reinforcement for warrants) in his book. Rebuttals and qualifiers are necessary in arguing because conclusions of arguments can be either "tentative", or "subject to conditions, exceptions" (Toulmin, 1958). Proposing the elements of rebuttal and qualifier is one of the important contributions of the Toulmin model (Nussbaum, 2011). And it is the rebuttal and qualifier that evolves into the concept of counterargument (Palmer, 2012). In addition, Toulmin makes the case that the warrants, which justify the move from data to claim are dependent on the field of epistemology, that is, whether the authority of the warrants is accepted or not; in other words, the evaluation of the validity of an argument depends on different standards in different domains of knowledge.

The Toulmin model has been hailed by a number of scholars as a breakthrough in the study of argument in four ways: in establishing that arguments are open to exceptions; in unfolding the nature of argument as dialectical question-answer exchange. (e.g., data are often produced in response to the questions, "What have you got to go on?"); in pointing out the field-dependency feature of argument and in discussing the epistemological nature of the argument assessment (Fulkerson, 1996b; Nussbaum, 2011; Van Eemeren et al., 1996). Among the above-mentioned four dimensions of the significance of the Toulmin model, the first two dimensions have inspired the instruction on argumentative writing (Fulkerson, 1996a); the second two dimensions

have had a substantial influence on the education of critical thinking (Kuhn, 1991; Van Eemeren et al., 1996). As such, the Toulmin model of argumentation has significant importance for pedagogies of both argumentative writing and critical thinking.

While Toulmin put forward the concept of rebuttals, his model is deemed underdeveloped in counterargumentation and refutation (Van Eemeren, Grootendorst, Jackson, & Jacobs, 1993). The Amsterdam School, adopting a pragma-dialectical approach to argumentation, takes the Toulmin approach a step further. They contend that argumentation is needed only when two or more people hold differing or opposing views about a given subject, i.e. the occurrence of argumentation is embedded in the context of diverging views. Therefore, they regard the procedure of argumentation as putting forward a series of arguments to justify one's standpoint, and to refute others' standpoints (Van Eemeren et al., 1996).

Walton's (1998, 2007) dialogue theory is consonant with the Toulmin approach and the Amsterdam School in basic aspects (Dowdy, 1998). Walton (2007) holds that arguments are given in a dialogic interchange between two or more parties over some disagreement, conflict or discrepancy of opinion. According to the dialog theory, the dialectical conception has at its core raising critical questions, elaborating one's own standpoint and responding to arguments with counterarguments and refutation (Walton, 2007).

Kuhn (1991) uses the term "alternative theory" to refer to a view that might be held

by someone who disagrees with the writer. Novice writers/thinkers often fail to envision the possibility of alternatives to their views, but it is essential for skilled writers to generate a counterargument (alternative view) to their main argument (Crammond, 1998; Kuhn, 1991; O'Keefe, 1999; Wolfe & Britt, 2008). They (skilled writers) then go on to use the "cognitively complex argumentative skill" to rebut the "opposing line of reasoning" (Kuhn, 1991, p. 145). Rebutting counterargument is a crucial step towards integrating previous lines of argument and reinforcing the writer's proposition (Kuhn, 1991). Refutation basically involves criticizing the counterargument, arguing why it does not have force and hence restoring force to the main argument (Kuhn, 1991; Kuhn & Udell, 2007).

2.5 Empirical Studies on the Toulmin Model and the Instruction on Argumentative Writing

This section reviews empirical studies on the application of the Toulmin model to the argumentative writing pedagogy.

2.5.1 An Overview

Since the late 1970s, Toulmin's model has been embraced by argumentative writing instructors and researchers as a promising new approach for generating and evaluating arguments and it has been adapted to numerous writing classes and textbooks (Fulkerson, 1996b; Van Eemeren et al., 1996). In one of the first attempts, in the textbook *College Composition and Communication*, Kneupper (1978) used the

Toulmin model as an invention heuristic for producing argumentative reasoning. Following this lead, substantial inquiries have been made as to the modification and use of the model in the composition classrooms (e.g. Barnet & Bedau, 2011; Hairston, 1981; Palmer, 2012; Ramage et al, 2010; Rottenberg & Winchell, 2009; Simon, 2008; Spurgin, 1985; Vesterman, 2006). To date, two of the most popular books on argumentative writing, Annette Rottenberg's *Elements of Argument* and John Ramage and John Bean's *Writing Arguments* both base the construction of sound argumentation on the Toulmin approach. Fulkerson (1996b), who has made some reflections on the use of Toulmin model in argumentative teaching and learning, concludes that the heuristic feature of the Toulmin model might make it a "generative tool when used analytically" (p.58). By analyzing existing arguments using the Toulmin elements, Fulkerson argues, students may gain essential skills when constructing their own elaborated argumentation. However, Fulkerson (1996b) points out that the Toulmin model should be applied in a simplified way since the full model, with its complexity, is undesirable and infeasible in a composition classroom.

Broadly speaking, the significance of the Toulmin model in argumentative writing pedagogy is embedded in three aspects: firstly, the informal logic perspective enables educators and scholars to break the boundaries of syllogism when studying everyday argumentation (e.g., Voss, 2005); secondly, the six-part Toulmin model of argumentation provides a schematic tool to analyze the structure of argument and makes it possible for students to write sound and strong arguments; finally, the concept of rebuttal lays the foundation for future studies on alternative views in argumentation and critical thinking (e.g., Gleason, 1999; Knudson, 1992; O'Keefe,

1999; Stapleton, 2001; Yeh, 1998).

2.5.2 Modification of the Toulmin Model

In many composition and critical thinking textbooks, the Toulmin model of argumentation is utilized in a modified way (e.g., Hughes, Lavery & Doran, 2010; Palmer, 2012; Ramage et al., 2010). This is because the original Toulmin model is too sophisticated to be used directly in actual argumentative writing instruction and evaluation as well as critical thinking instruction, and the terms such as "qualifier" and "backing" are deemed difficult by teachers and students (Fulkerson, 1996b). For example, in the following claim and data a warrant is unnecessary because wanting to live a longer life is assumed:

> *Smoking is bad for you. Smokers are known to suffer more*
> *diseases than non-smokers and have a shorter lifespan.*

A substantial number of empirical studies have applied the Toulmin model to evaluate argumentative compositions or to facilitate teaching by modifying it. These empirical studies have involved students at different levels (from elementary to tertiary levels) and in various academic majors; accordingly the Toulmin model has been modified in various ways. Broadly speaking, the literature reveals the Toulmin model has been modified either in simplified or elaborative ways. Simplified modifications, i.e., focusing on three or four of the argumentative elements while leaving out the others, have been carried out by Bacha (2010), Burkhalter (1995), Connor (1990), and

Varghese and Abraham (1998). For instance, Bacha (2010) analyzed three argumentative elements: claim, data and warrants in EFL students' essays. In yet other studies, the Toulmin model has been adapted in an elaborative way, i.e., the *rebuttal* element was elaborated into *opposition* and *response to opposition* (e.g. Crammond, 1998; McCann, 1989; Knudson, 1992). These studies, however, represent only a small sampling of the many studies that have used the Toulmin framework.

Notably, after decades of scholars using and researching the Toulmin model, the original terminology has evolved and changed to help students understand fundamental argumentative elements or to fit varied research contexts (Gleason, 1999; Qin & Karabacak, 2010). Some terms have also been given different names while preserving the original meaning. For instance, *claim* has been termed proposition, opinion, conclusion, etc.; *data* has been named reasons, evidence, or arguments. Other terms like *qualifier* and *backing* have been used less often in empirical studies. With regards to the term *rebuttal*, the original terminology has usually been retained, but with evolved meanings. *Rebuttal*, together with *qualifier*, has been developed into *counterarguments* and *rebuttals*, or *opposition* and *response to opposition* as in Knudson (1992), to refer to the whole process of counterargumentation by acknowledging alternative or opposite views and refuting them.

2.5.3 Main Methods of Applying the Modified Toulmin Model

As for the methods of application of the modified Toulmin model, it has been applied in argumentative writing education and research in two ways: as scoring criteria and

in instruction. First, the model provides a tool for discourse analysis on students' argumentative texts. For instance, some researchers have used the adapted Toulmin elements as scoring criteria to measure the quality of students' argumentative essays (e.g., Ferris, 1994; Knudson, 1991, 1992, 1994; McCann, 1989; Qin & Karabacak, 2010). Second, the modified Toulmin model has also been widely employed in argumentative writing instruction in empirical research, i.e., researchers have used the adapted Toulmin model explicitly as an instructional intervention in the argumentative writing classrooms (e.g., Bacha, 2010; Ferretti et al., 2007; Gleason, 1999; Nussbaum & Kardash, 2005; Yeh, 1998). Besides being used for explicit instruction and training, the Toulmin model can be utilized in two other ways: implicit instruction (providing hints in writing prompts for certain argumentative elements such as counterarguments, e.g., Ferretti et al., 2000; Ferretti et al., 2009), immersion (using reading material to generate more convincing arguments, e.g., Genishi, 1992; Petraglia, 1995). The text below delineates the two most common applications of the Toulmin model.

2.5.3.1 The Toulmin Model as Scoring Criteria

One early modification and application of the Toulmin model as scoring criteria has been reported by McCann (1989). In a study on the difference in argumentative writing knowledge and the ability of students at three grade levels (6th, 9th and 12th), McCann measured the overall argumentative writing quality of students using scoring criteria based on a modified Toulmin model. The scoring criteria investigated six traits of students' written argumentation: claims, data, warrants, propositions,

opposition and response to opposition. McCann found that the 6th graders yielded fewer opposite views and responses to opposition; when the scripts were scored holistically, the youngest group was lower than that of the 9th and 12th graders. Notably, it was McCann in this study who first named the *opposition* and *response to the opposition* as two features of argumentation. Knudson (1992) adopted the Toulmin criteria as being representative of effective arguments when all elements were included. However, he admitted that the criteria might not be communicated clearly and explicitly to students, suggesting the limitations of the use of the modified Toulmin model as evaluation tool for students' essays. In another more recent study, Qin & Karabacak (2010) explored the relationship between the use of (adapted) Toulmin elements and the overall quality of the argumentative essays by a group of Chinese university students. They found that the use of secondary Toulmin elements (counterarguments and rebuttals) contribute favorably to the overall quality of argumentative essays.

2.5.3.2 The Toulmin Model in Explicit Instruction

Whilst Knudson's (1992) and McCann's (1989) studies indicated the efficacy of the modified Toulmin model as evaluating standards when scoring students' argumentative essays, a problem stood out; that is, it was difficult to explain these standards to students. In Knudson's study, students were presented with model arguments. However, they were still not clear about individual argumentative elements, especially how the elements could add up to form a good argumentative essay. A number of researchers have investigated the position of explicit instruction in

argumentative writing. In this section, the significance of explicit instruction in the Toulmin model is explained, and empirical studies involving the explicit teaching of the Toulmin model in class are reviewed.

The ability to write a good argumentative essay is not acquired, but needs intensive training (Graff, 2003). While the ability to argue orally seems to be an inborn human trait, written argumentation in the academic context needs scaffolding and practice (Graff, 2003). Explicit instruction is considered critical to train students in essential argumentative skills (Graff, 2003; Voss & Means, 1991). As the results of empirical research suggest, the more functional and possibly more efficacious way of applying the Toulmin model is to use it in explicit instructional activities in teaching students how to produce such elements as claim, data, counterarguments and rebuttals (e.g., Ferretti et al., 2007; Ferritti et al., 2009; Gleason, 1999; Nussbaum & Kardash, 2005; Nussbaum & Schraw 2007; Yeh, 1998). In most cases, these studies have also applied the modified Toulmin model in the assessment of student essays.

Gleason (1999), based on the findings of a series of three studies designed with the Toulmin scheme to address the difficulty that students encounter in argumentative writing, reports on the effectiveness of explicit instruction on argument structure. In that study, the two experimental groups, one receiving instructions on the basic version of argumentation (proposition, claim, data and warrants) and one the expanded version of argumentation (where acknowledging the opposition and response were added to the basic version), scored significantly higher than control groups who received no training in Toulmin elements in terms of basic argumentative

features. The expanded group significantly improved with regard to raising opposite views and refuting those views. The findings of other interventionist studies support Gleason's research results showing that instruction on elaborate argument structure (Chandrasegaran, 2008; Crowhurst, 1991; Reznitskaya, Anderson, & Kuo, 2007; Varghese & Abraham, 1998; Wolfe et al., 2009; Yeh, 1998) had a positive effect on the quality of students' argumentative essays.

2.5.4 The Toulmin model in Science Education

The application of the Toulmin model has gone beyond the composition course to subject matter classrooms in science education. In recent years, the Toulmin model has been increasingly employed in varied science classes to strengthen students' ability in argumentation.

A number of researchers have found that the Toulmin model is an effective tool in the teaching and learning of argumentation in a scientific context (e.g., Erduran, Simon & Osborne, 2004; Kaya, 2013; Osborne, Erduran & Simon, 2004; Zohar & Nemet, 2002). In Osborne et al.'s (2004) two-year study in a junior high school science class, the usefulness of the Toulmin model in evaluating the quality of argumentation was examined. In the first phase of the study, teachers' development in teaching argumentation was assessed, while in the second phase, the researchers evaluated the improvement in the quality of students' argumentation at the end of an interventional inquiry. In both phases, the Toulmin model was chosen as the foundation of the analytical framework and was found to be a feasible tool. Erduran et al. (2004) tested

two methodological approaches to applying the Toulmin model as a measurement of both the quality and quantity of argumentation in classroom discourse. Their study explored the extension of the use of the Toulmin model in tracing argumentation development for the whole class, thus indicating the usefulness of the Toulmin model for "collective reasoning behavior" (p. 932). In general, the findings in Osborne at al. (2004) and Erduran et al. (2004) suggested the efficacy of the Toulmin model in science teaching and learning. Similar findings occurred in a number of other studies (e.g., Venville & Dawson, 2010; Zohar & Nemet, 2002).

While researchers have noted the practicality of the Toulmin model in teaching argumentative skills to students and evaluating argumentation, some of them have also recognized the limitation of applying the model in science education. For instance, Sadler and Fowler (2006) found that it was difficult to differentiate the quality of students' argumentation using the Toulmin model in a genetic engineering context. According to some scholars (e.g., Sampson & Clark, 2008; Simon, 2008), the limitation of applying the model is due to its focus on the form rather than the content and quality of an argument.

2.5.5 Future Directions of the Toulmin Model for Various Classrooms

The efficacious use of the Toulmin model in composition and science classrooms has been suggested in many studies. Meanwhile, concerns and challenges have been raised regarding the future directions of the Toulmin model. Lunsford (2002) has cautioned that the application of the Toulmin model must be situated in the

instructional context, or be "heavily mediated by other writing instruction" (p. 160).

Lundsford outlined in detail a summer composition course where the adapted Toulmin model was employed by instructors to enhance students' argumentation. The results of that study proposed several ways of contextualizing the Toulmin model for better application, which mainly include aligning readers and texts (by naturalizing the model and using ample examples to illustrate) and mapping concepts onto texts (by interpreting or construing the concepts).

Some researchers (Driver, Newton & Osborne, 2000; Duschl, Ellenbogen & Erduran, 1999; Nielsen, 2013), however, challenge the use of the Toulmin model as an analytical tool for students' written argumentation. Driver et al. (2000) proposed that the model was insufficient for analyzing the "interactional aspects of argument" and the "linguistic and situational contexts" of argumentation (p. 294). In other words, the Toulmin model is effective for measuring argumentative elements rather than the context and nature of argumentative discourse. Some researchers (e.g., Erduran et al., 2004; Osborne et al., 2004) have aimed at improving the use of Toulmin Patterns. With their empirical study, Osborne et al. (2004) proposed a workable framework based on the Toulmin model to analyze the quality of the argumentation process in the classroom.

To conclude, the scholarship reviewed in this section is concerned with the application of the Toulmin model in the composition class as well as subject matter classes. The model can be used as an analytical tool for argumentation or be used implicitly or explicitly in argumentative writing instruction. It has been found useful

in many empirical studies, but with limitations. Future directions of the use of the model are discussed in some recent studies. Of salient significance to Study 2 of this project is that theoretical research suggests the necessity of explicit instruction in argumentative writing and empirical research indicates efficacy of explicit instruction in the Toulmin model in the writing classroom, which provides a theoretical basis for the instructional intervention of Study 2.

2.6 Persuasiveness of Argumentative Essays

2.6.1 Counterargumentation and Persuasiveness of Argumentative Essays

In the last section, the literature is reviewed pertaining to the application of the Toulmin model in the argumentative writing of students. But how does the inclusion of key argumentative elements such as counterarguments and rebuttals influence the quality of an argumentative essay? Since it is established that the primary purpose of an argumentative essay is to persuade (e.g., Ding et al., 1994; Palmer, 2012), the question actually is: how does the inclusion of key argumentative elements interrelate with the persuasiveness of an argumentative essay? In this section, relevant scholarship is reviewed.

The significance of including counterarguments and rebuttals for making written argumentation persuasive has been underpinned by much research. Walton (1989) listed two goals of persuasive argumentation: to support one's own position and to refute the opponent's argument by identifying its weaknesses. Kuhn (1991) holds that

at the core of competent argumentative reasoning is the handling of supporting elements, alternative views and counterevidence, while stating that writers' failure to "envision conditions that falsify their own theory" (p.117) is the main obstacle to effective argumentation and critical thinking. Indeed, neglecting alternative views has been identified as a common weakness in the argumentative writing of students at both the primary or secondary level (e.g., Ferretti et al., 2009; McCann, 1989; Yeh, 1998) and the tertiary level (e.g., Nussbaum & Schraw, 2007; Wolfe & Britt, 2008). This weakness in the argumentative essays of students was referred to as *myside bias* which will be discussed in detail in the next section.

2.6.2 Empirical Studies on the Persuasiveness of Argumentative Essays in the L1 Contexts

Many studies have demonstrated that the inclusion of counterarguments strengthens an arguer's position and helps achieve completeness in good reasoning; hence, incorporating counterarguments and refuting counterarguments are crucial for maximizing the extent of persuasiveness in argumentative writing (e.g., Crammond, 1998; Leitão, 2003; O'Keefe, 1999; Perkins et al., 1991; Stanovich & West, 2008; Toplak & Stanvich, 2003; Wolfe & Britt, 2008; Yeh, 1998; Zohar & Nemet, 2002). For instance, O'Keefe (1999) conducted a meta-analysis of 107 studies on argumentative texts. He found that two-sided messages (those acknowledging opposing views) were more effective in persuasiveness than one-sided messages. Moreover, O'Keefe (1999) found that refutational two-sided messages attempting to rebut opposing arguments tended to be more persuasive than nonrefutational

two-sided messages acknowledging opposite views, but not refuting them.

Three studies of particular interest to the present one had counterarguments as their central theme; two of these focused on the type of writing prompts given to student participants, while the other focused on classroom instruction. Nussbaum and Kardash (2005), who analyzed argumentative scripts from 77 undergraduates in the United States, found that by simply including a request for counterarguments (a "goal instruction") in their prompt, students significantly increased the number of counterarguments and rebuttals compared to those students who received a bare prompt. Similarly, Ferretti et al. (2000), in a study at the primary level, found that specific requests for counterarguments and rebuttals in essay prompts resulted in better persuasiveness and frequencies of these argumentative elements than in bare prompts where no goal instructions were stated. As for classroom treatment in counterarguments, Nussbaum and Schraw (2007) found that two treatments, 1) an organizer which graphically illustrated how to structure an argumentative essay and 2) explicit instruction in argumentative writing, had positive, but different effects on the production of counterarguments and rebuttals in undergraduate student writing.

Most recently, Wolfe et al. (2009) conducted a design of three studies on counterargumentation among university students. Results of the first study reveal that rebutting other-side information brings about better ratings of the quality of argumentation and impression of the writer. The second study was designed to uncover factors that decide the persuasiveness of an argument comprising claim and reasons. The results suggest that attitudes, beliefs and values shape people's

agreement with a claim and their judgment of the quality of reasons. In the third study, a tutorial was held to address three fundamental areas in argumentation schema: main claims (arguer's position), elaborative supporting evidence and presenting and refuting counterarguments. The findings of this study revealed that training in an elaborated argumentative schema can help students include other-side information, and thus enhance the persuasive effect of their essays. In addition, they found the overall quality of argumentative essays improved as a consequence of the tutorial.

2.6.3 Empirical Studies on the Persuasiveness of Argumentative Essays in the L2 Contexts

While the studies noted above have analyzed students writing largely in their native language (L1), the present study examines the argumentation ability of second language (L2) students. Of special interest to this inquiry is a study by Qin and Karabacak (2010), in which a text analysis was performed on students' argumentative essays to reveal the relationship between the use of the elements adapted from Toulmin and the overall quality of argumentative writing. The participants of that study were 133 second-year English majors from a Chinese university who received no instruction regarding the Toulmin model of argumentation. Two raters coded six Toulmin-like elements: *claim, data, counterargument claim, counterargument data, rebuttal claim* and *rebuttal data*. The raters then assessed the overall quality of the papers based on a 5-point scoring rubric developed from McCann (1989), and Nussbaum and Kardash (2005). This study found that undergraduate English majors had grasped the basic elements of argumentation, i.e., *claim* and *data*. However, the

use of the elements that consider alternative views, i.e., *counterargument, rebuttal,* was rather low in frequency. Another key finding, similar to the studies mentioned above, was that when counterarguments and rebuttals were included, they enhanced the overall quality of argumentative writing.

In this section, the relation between the inclusion of counterargumentation and the persuasiveness of argumentative essays is discussed; empirical studies in both L1 and L2 contexts are reviewed. The next section will discuss how critical thinking interacts with written argumentation, which is a fundamental component of the theoretical framework of this project.

2.7 Argumentative Writing and Critical Thinking

2.7.1 An Overview

The ultimate goal of this research project is to explore students' critical thinking development through their argumentative writing. To this end, it is essential to examine the interplay between argumentative writing and critical thinking via visiting relevant scholarship. In this section, critical thinking is first operationally defined. Then the connection between critical thinking and informal logic is discussed before an examination of the function of argumentative writing as a vehicle for critical thinking. In the end, the crucial role of counterargumentation in critical thinking is deliberated.

2.7.2 Definition of Critical Thinking

Since American educator John Dewey's (1910) thought-provoking discussion of "reflective thought", the conceptualization of critical thinking has evolved into a two dimensional construct – a combination of cognitive abilities and thinking disposition (Ennis, 2003; Facione, 1990; Halpern, 2003; Norris, 2003; Paul, 1993, 1995; Pithers & Soden, 2001). Earlier definitions of critical thinking focus on the cognitive factor of critical thinking, i.e., critical thinking is viewed as a skill or a set of skills, a mental procedure or merely correct assessing of statements (Baron, 1985; Ennis, 1962; McPeck, 1981). Over time, critical thinking theorists (e.g., Ennis, 1987) began to emphasize the intention of and initiative in seeking better judgment, which is termed critical thinking disposition (Facione, 1990; Halpern, 2003; Sears & Parsons, 1991; Siegel, 1988). Both dimensions of critical thinking embed a fairness of mind, objectivity and viewing two or more sides of a case as essential elements (Perkins & Tishman, 2001), by which critical thinking is operationally defined in this research project.

2.7.3 Critical Thinking and Informal Logic

The concept of argument as the core of thinking has in fact existed from antiquity. The major early philosophers – Plato, Socrates, Aristotle – were all centrally concerned with thinking, and all regarded the construction of reasoned arguments as the heart of thinking (Kuhn, 1991). In modern society, two key concepts: *critical thinking* and *informal logic,* have been developed. The connection between critical

thinking and informal logic, to a great extent, resembles the linkage between thinking and arguing (Billig, 1987). To be specific, the practice of treating the terms *critical thinking* and *informal logic* as coextensive started from educational reforms at Canadian and American universities in the early 1970s (Van Eemeren et al., 1996). In the then novel logic course for university students, the objectives of fostering critical thinking were achieved by teaching the analysis and evaluation of arguments (Van Eemeren et al., 1996). Therefore, the perspectives and methods of informal logic argumentation provide a means to seek the educational goal of critical thinking. Because of this, informal logic argumentation and critical thinking are conventionally deemed equivalent, even interchangeable (Billig, 1987; Fox, 1994; Inch & Warnick, 2010; Perkins et al., 1991; Van Eemeren et al., 1996; Walton, 1989). Another significant notion relevant to the argumentative writing and critical thinking connection is the notion of "writing to learn" (Atkinson, 1997). Language, written language in particular, is regarded as a "tool for intellectual exploration, an avenue for debate and dialectic" (Atkinson & Ramanathan, 1995, p.558). Since agonistic ideas are tested by using language, written argument is the key sign whether someone is a critical thinker or whether critical thinking has taken place (Atkinson, 1997). In this sense, Toulmin's contribution towards developing informal logic, especially his emphasis on the context of argumentation (hence rebuttals and qualifiers) can also be considered an important contribution to the development of critical thinking. The interrelation between argumentation and critical thinking provides a theoretical basis for the current research.

2.7.4 Argumentative Writing as a Vehicle for Critical Thinking

In fact, students' argumentative writing has long been viewed as a vehicle for promoting student learning and critical thinking (Goh, 2008; Kuhn & Crowell, 2011; Knipper & Duggan, 2006; Pandis, Ward, & Matthews, 2005). Writing allows students to reflect on and draw linkages among ideas. Students also learn to formulate and critique written arguments (Nussbaum, 2008). Kuhn and Crowell (2011) conducted an interventional study to inquire about the effect of the transfer of dialogic argumentation skills on middle school students' critical thinking development. In that multi-year intervention, the experimental group received treatment of electronically conducted dialogic reasoning on controversial topics, while the control group received normal writing training and discussion on the same topics. At the end of the intervention, both groups were asked to write essays on the topics. The findings suggest that the experimental group surpassed the control group in both reasoning skills and argument quality. Kuhn and Crowell's (2011) study corroborated the view that counterargumentation skills in argumentative writing enhance students' critical thinking competence. The next section further reviews the contribution of counterarguing and rebutting in argumentative writing to the improvement of students' critical thinking.

2.7.5 The Significance of Counterargumentation in Argumentative writing to Critical Thinking

Counterargumentation, an arguer's recognition of opposing views and refuting them,

has been deemed central to one's critical thinking abilities and disposition (Palmer, 2012; Perkins & Tishman, 2001; Walton, 1989). Palmer (2012) contends that acknowledging alternative views is the hallmark of critical thinking. In the words of Perkins and Tishman (2001), "thinking about the other side of the case is a perfect example of a good reasoning practice. It is a move one would ordinarily count as part of intelligent behavior" (p. 3). By applying the concept of argument as dialogue to critical thinking education, Walton (1989) maintained that an essential element of critical thinking ability is considering both sides of an argument.

Counterargumentation is also a necessary attribute of good argumentation (Kuhn, 1991; Perkins et al., 1991; Santos & Santos, 1999). Voss and Means (1991) contend that the evaluation of the soundness of argumentation is based on three criteria: (1) the acceptability of the reason; (2) the support for the claim; (3) the extent to which counterclaims are taken into account. The third criterion is particularly underscored by scholars. For instance, Walton (1989) maintains that skilled arguers understand the importance of addressing opposite views and undertake to refute these views. Kuhn (1991) holds that the handling of supporting elements, alternative views and counterevidence are crucial skills in argumentative reasoning. In addition, Kuhn (1991) perceives it as a main obstacle to critical thinking if an arguer fails to consider alternative or opposite views.

2.7.6 Critical Thinking and Myside Bias in Argumentative Writing

More than two decades ago, Perkins and his colleagues identified the tendency of

students to support only their preferred viewpoint while ignoring evidence against their own positions; the tendency was termed "myside bias" (Perkins, 1985; Perkins et al., 1991). Observations through empirical studies (e.g., Felton, 2004; Kuhn & Udell, 2003) revealed that unskilled student writers tended to pay more attention to their own views and supporting evidence but neglect opposite views and counterevidence. According to theorists (e.g., Perkins, Jay & Tishman, 1993; Stanovich & West, 1997), the myside bias phenomenon indicates a lack of open-mindedness; thus it is detrimental to one's critical thinking ability. Over the past two decades, myside bias has been investigated by researchers in the fields of critical thinking as well as argumentation and found to be a characteristic feature impeding the sound reasoning of students' arguments (e.g., Baron, 1991, 1995; Goh, 2008; Santos & Santos, 1999; Toplak & Stanovich, 2003; Wolfe & Britt, 2008; Wolfe et al., 2009). In an empirical study, Toplak and Stanovich (2003) examined the relation between the myside bias on an informal reasoning task and the amount of tertiary education. They found that the degree of myside bias decreased with year of education in university. The results suggest that school education can have positive effects on balanced and good reasoning which is in agreement with Kuhn's (1991) conjecture that unbiased argumentation may be a function of school education.

A number of empirical studies demonstrate that the inclusion of counterargumentation not only strengthens an arguer's position and helps achieve *completeness* in good reasoning, but also enables an arguer to avoid *myside bias*, hence, incorporating counterarguments and refuting counterarguments are crucial in argumentative writing as a tool of critical thinking (e.g., Goh, 2008; Kuhn & Crowell, 2011; Leitão, 2003;

O'Keefe, 1999; Perkin et al., 1991; Stanovich & West, 2008; Toplak & Stanovich, 2003; Wolfe & Britt, 2008; Wolfe et al., 2009; Yeh, 1998; Zohar & Nemet, 2002).

O'Keefe's (1999) findings that the "two-sided" message which considers both supportive arguments and opposing arguments is more persuasively efficacious than "one-sided" messages which ignore opposing arguments justifies the necessity of dealing with counterarguments. This position has also been endorsed by other researchers (e.g., Leitão, 2003; Stanovich & West, 2008) to prevent the dangers and pitfalls of excluding other-side arguments and resulting in myside bias (Wolfe & Britt, 2008) in argumentative writing.

This section is concerned with the interrelation between argumentative writing and critical thinking, and the significance of counterargumentation in argumentative essays to the critical thinking ability of students. Most empirical studies reviewed are conducted in the L1 (English as first language) context. The next section will be focused on relevant studies in the L2 (English as second or foreign language) context.

2.8 Argumentative Writing and Critical Thinking in the L2 Context

2.8.1 An Overview

Compared with the scholarship of argumentative writing in the L1 context, studies on L2 students' argumentative writing and critical thinking are fewer in number and with a shorter history. It could be said that most of the literature on critical thinking in L2

composition has been published since the 1990s. In some of the articles, researchers have raised concerns about the inclusion of critical thinking in the L2 academic writing class because they think that critical thinking is a product of Western culture, and thus is unteachable to learners of non-Western background (Atkinson, 1997; Atkinson & Ramanathan, 1995; Fox, 1994; Ramanathan & Kaplan, 1996a, 1996b). For instance, Atkinson (1997) cautions TESOL educators about teaching critical thinking to ESL students, arguing that critical thinking is primarily cultural thinking and that different sociocultural backgrounds of ESL students may make critical thinking unteachable in composition courses. This assumption that critical thinking is unteachable has generated extensive debate among researchers (e.g., Benesch, 1999; Durkin, 2008; Gieve, 1998; Raimes & Zamel, 1997; Stapleton, 2001, 2002a; Tian & Low, 2011). Researchers such as Brookfield (1987) and Benesch (1999) contend that students can be taught to be critical thinkers by encouraging them to be more aware of other people's assumptions. These conflicting views about critical thinking education arose because critical thinking is a concept adaptive to cultural contexts. The role of culture in argumentative writing and critical thinking (especially in the L2 context) will be covered at length in section 2.10.

2.8.2 Empirical Studies on the Critical Thinking in L2 Argumentative Writing

Since the current study concerns the acknowledging and refuting of counterarguments as a hallmark of critical thinking manifested in L2 students' argumentative essays, of particular relevance is one study - Stapleton's (2001) classroom treatments on the critical thinking in the writing of Japanese university students. In that treatment, half

of the participants were required to write argumentatively on a familiar topic, while the other half wrote on an unfamiliar one. The essays were then assessed against criteria consisting of four critical thinking elements: *argument*, *evidence*, *recognition of opposition* and *fallacies*. By referring to the Toulmin model and reviewing textbooks on writing and critical thinking, the author generated writing rubrics that covered the basic elements of critical thinking. Two raters then blindly scored the essays using the rubrics. The results indicated that content familiarity played a positive role in critical thinking, i.e., produces more arguments, evidence and refutations in the writing. Significantly, by making a case that participants in the experiment displayed critical thought, the author questioned Atkinson (1997) and other researchers' notion that Asians are generally deficient in critical thinking.

Chandrasegaran (2008) compared Singaporean students' argumentation performances in informal conditions (school forum postings by secondary school students) and formal conditions (an academic essay by a master student). The data yielded from this research demonstrated that Singaporean students' written argumentation, either in informal or formal contexts, contained definite stance assertions and certain forms of stance support, e.g., citing authority and appealing to values. What is worth noticing is that counterargumentation elements were found in forum postings, but not in the academic essays. Chandrasegaran theorized that secondary school students' acknowledging and refuting opposite views, though not instructed on such moves, were a type of "informal literacies" which was referred to by Moss (2000) as texts read by people in everyday informal settings other than the school settings. While the master student's failure to include counterargumentation elements in the essay affirms

Moss's (2000) remark that "informal literacies do not act as a powerful resource within schooled settings"(p. 62), suggesting that explicit instruction may be desirable for students to exploit their "informal literacies" and achieve a higher level of writing competence (Chandrasegaran, 2008).

Of special interest to the present study is the instructional approach that Bacha (2010) adopted in an experiment aimed to improve Arab university students' argumentative writing skills. The instructional model is based on Toulmin's (1958) argumentative elements which "propose on a macro scale support for claims and responses to counterarguments" (Bacha, 2010, p. 231). The teaching approach resulted in the production of more counterarguments and rebuttals in students' essays, and these "organizational improvements" demonstrate the efficacy of explicitness in the teaching of argumentation schemata (Bacha, 2010). Bacha's conclusion echoes the findings of Gleason (1999), Yeh (1998), Nussbaum and Schraw (2007) and Wolfe et al. (2009) on the effectiveness of explicit instruction in promoting students' argumentative writing. However, since that interventional research has been conducted in the L2 context, it bears more significance in the exploration of enhancing critical thinking in L2 argumentative writing.

Another interventional study on critical-thinking using enhanced L2 instruction was conducted by Yang and Gamble (2013). The study involved experimental and control groups of first-year non-English majors at a Taiwan university. The experimental group instruction featured guided activities such as peer critiques and debates, while the control group activities comprised group presentations and process writing. At the

end of the intervention, students were required to write an essay on a controversial topic. Students' overall English proficiency and academic achievement were also gauged with respective tests. The writing of the two groups was measured in terms of critical thinking using the Holistic Critical Thinking Scoring Rubric (Facione & Facione, 1994). The results showed that the experimental group significantly outperformed the control group in critical thinking as displayed in their argumentative writing. The notable gain of the experimental group was also demonstrated in their English proficiency and academic achievement.

The review of the above literature indicates that in the L2 context, explicit instruction on good argumentative reasoning is needed in order to foster critical thinking in the composition classrooms.

2.9 Argumentative Writing and Critical Thinking of English Majors at Chinese Universities

2.9.1 Critical Thinking Deficiency in the Argumentative Writing of English Majors at Chinese Universities

A large-scale survey conducted by a Chinese national research project team - WTO Entry and Educational Reform and Development of Foreign Language Disciplines in Chinese Universities (hereafter the Project)[4]- revealed that 31% of the respondents

---

[4] This research project was commissioned by the Higher Education sector of the Chinese Education Ministry, aimed at preparing China for a better position in the World Trade Organization.

(both teachers and students) consider the graduates of the English language programs poorer in critical thinking performance than graduates from other university disciplines; 52% of the respondents from the English discipline hold a similar view. These survey results mirror earlier observations by Chinese scholars in the EFL field (e.g., He et al., 1999; Huang, 1998, 2010). As to the roots of this problem, researchers have offered various perspectives. Atkinson (1997) and others (e.g., Fox, 1994) contended that critical thinking was mainly a Western notion which is unteachable to L2 students due to varied cultural backgrounds. Paton (2005), however, disagreed by arguing that critical analysis has existed in the history of science in China since ancient times; therefore, culture should not be the reason for the lack of critical thinking in Chinese students. He concluded that insufficient language proficiency and subject knowledge may have led to Chinese students' lack of critical thinking.

According to the research of He et al. (1999), Huang (1998) and the Project (2001), the main cause of the critical thinking deficiency among English majors in China is the teaching emphasis on language learning rather than thinking competence. In a recent study, Wen et al. (2010) used a critical thinking measurement they designed for assessing 2318 undergraduates from 14 disciplines at 11 Chinese universities. Their findings revealed that the critical thinking abilities of first-year English majors were actually superior to those of the students in other social science disciplines. However, after three years of university education, English majors' critical thinking development was insignificant compared with other students. The above empirical studies suggest that critical thinking education for the English majors at the university level may not be successful. As such, more explicit, goal-oriented instruction on

critical thinking may be desirable. In summary, the instructional approach adopted in the current study using the argumentative reasoning skills as means of fostering critical thinking may go some distance towards addressing perceived deficiencies in critical thinking among English majors.

2.9.2 Empirical Studies Addressing the Critical Thinking Issue of English Majors in China

What are the actual performance and features of English majors' argumentative reasoning among Chinese undergraduates? Having realized the lack of research in this area, Wen and Liu (2006) performed an exploratory study on the traits of English majors' "abstract thinking" in English argumentative essays. The researchers designed four parameters, characterizing the thinking processes when composing, to analyze the student writers' abstract thinking: "relevance, explicitness, coherence and sufficiency" (Wen & Liu, 2006). They found that students are comparatively weak in sufficiency of argumentative reasoning. Although the four parameters are not scientific measurement of critical thinking in argumentative writing, this study remarkably started a key area in the research on English argumentative writing by Chinese students. As the authors pointed out, in the past decade, there has been a growing interest in the L2 writing research in China, including pedagogical research on L2 writing. However, few studies have entailed thinking skills as one of the factors influencing writing performance (Wen & Liu, 2006). Wen and Liu (2006) stressed the necessity of instruction on logical thinking abilities in developing the writing performance of university students.

Another study focusing on English majors' argumentative writing in China was conducted by Qin and Karabacak (2010) which was described in section 2.6.3. They found that the quantity of key argumentative elements such as counterarguments and rebuttals were significantly correlated with the overall quality of an argumentative essay. The results of their empirical study could provide useful clues for instructional studies on argumentative writing.

One essential dimension of argumentative writing instruction is the pedagogical materials. Liu (2005) compared instructional materials on argumentative compositions available on American and Chinese websites. The comparative analysis identified two salient differences between the two groups. Firstly, while the American materials regard "anticipating the opposition" as an essential element in written argumentation, the Chinese materials either did not mention it or treated it as an optional technique. The second difference is the "epistemological emphasis in Chinese material and its absence in American materials" (Liu, 2005, p. 13). Liu noted that Chinese materials were based on Marxist philosophy as the overwhelming world view and urged students to argue in this way, i.e., students are told to understand the world with "basic principles of Marxism and Maoism" learned in their politics courses and to argue in this way (Liu, 2005, p. 9). Despite the utility of this observation, more research is needed to analyze the impact of composition textbooks on argumentative writing pedagogy because the textbooks used in composition classes are the mainstream instructional materials. This need, again, highlights the significance of the present research, which will investigate via classroom observations, questionnaires and interviews how the textbook is used in

argumentative writing classes and how the textbook, along with other factors will influence students' performance in and perceptions of argumentative writing.

Su (2011) reported a comparative study on explicit and implicit teaching of critical thinking in written argumentation in a Chinese university. The 18-week intervention on 205 fourth-year English majors yielded certain positive effects. While both explicit and implicit instructions were effective in improving students' argumentative writing as well as critical thinking skills, "there [was] a dramatic discrepancy in the effectiveness of these two methods with the explicit one having an advantage" (Su, 2011, p. 150). However, some problems stand out with this study. First, the author assessed critical thinking in writing mainly from three aspects: whether there was relevant, sufficient thesis support; whether the evidence (data) was sound and powerful; whether there was good logic in arguing. These criteria, however, cannot fully measure the fundamental elements of critical thinking. For instance, acknowledging and refuting counterargument, an eminent feature of critical thinking, was not considered in the assessment of critical thinking in students' argumentative essays. Second, there was not a detailed data analysis of students' performances on each of the three parameters of critical thinking.

Despite the studies described above in the Chinese context, there is a relative dearth of empirical studies on the critical thinking cultivation of English majors in their argumentative writing.

2.10 The Role of Culture in Critical Thinking and Argumentative Writing

Culture can be defined as the characteristic profile of a society with regards to its values, norms and institutions (Lytle, Brett, Barsness, Tinsley, & Janssens, 1995). Since a considerable amount of literature has made connections between culture and a lack of CT in East Asian cultures, some coverage of the literature in this area appears appropriate. In the following text, I will review some literature concerning what cultural factors shape critical thinking and how culture influences argumentation. In fact, critical thinking and argumentation are so closely connected that they are discussed as one issue in some articles (e.g., Chan, Ho, & Ku, 2011; Kuhn & Crowell, 2011). However, in this section, in order to distinguish between what factors shape this conceptual process of critical thinking and how these factors influence the product of critical thinking and informal reasoning, i.e., argumentation, the section below is divided into two subsections: culture and critical thinking, and culture and argumentation.

2.10.1 Culture and Critical Thinking

Culture and culture-related factors constitute one of the most debated elements that can influence students' critical thinking abilities (Manalo, Kusumi, Koyasu, Michita, & Tanaka, 2013a). In their attempts to sort out the impact of culture on critical thinking, researchers tend to distinguish between Western culture and non-Western cultures, e.g., Asian cultures. In section 2.8.1, I delineated two conflicting views held by researchers pertaining to whether critical thinking is teachable and whether it is

exclusive to native speakers. At the center of the debate is whether Asian students show a lower level of critical thinking compared to their Western counterparts. While there is no indication that this debate has run its course, a number of empirical studies demonstrate that culture and culture-related factors do influence students' critical thinking performance (e.g., Chan et al., 2011; Lun, Fischer, & Ward, 2010; Manalo et al, 2013a). In this section, the focus is on what and how culture and culture-related factors exert an impact on the critical thinking performance of Western and non-Western students. Empirical evidence reveals that cross-cultural differences in factors such as language proficiency (Floyd, 2011; Lun et al, 2010; Manalo, Watanabe & Sheppard, 2013b), epistemic beliefs (Chan et al, 2011; Weinstock, Neuman, & Glassner, 2006), self-contrual (Manalo et al, 2013a), self-regulatory mode (Manalo et al, 2013a), and self-efficacy (Manalo et al, 2013a) contribute to variance in students' performance in critical thinking. These factors fall into two broad categories: psychology and cognition. Related studies on these factors and their impact on critical thinking are outlined below.

A few studies (e.g., Floyd, 2011; Lun et al., 2010; Manalo et al 2013b) have explored reasons for the observed differences in critical thinking between cultural groups and found that language proficiency accounted at least partially for the differences. One recent study was conducted by Lun, Fischer and Ward (2010), which consisted of two stages. Participants in the study were 70 university students comprising 24 Chinese, 35 New Zealand European and 11 New Zealand aborigines; however, only data solicited from the first two groups were analyzed. In the first stage, the researchers tested whether differences in critical thinking could be measured between Asian and

Western students. In the second stage, they explored the roles of English language ability and dialectical thinking style in explaining the observed differences. Findings revealed that it was English proficiency, not thinking style, that could partially explain the differences in critical thinking between the two cultural groups. Manalo et al. (2013b) designed an intervention study with 111 Japanese students to examine whether language structure or language proficiency might influence students' critical thinking. They found that particular language structure (Japanese in that case) did not affect students' critical evaluation. However, it was found that students' English proficiency test scores were positively correlated with production of evaluative statements in English, indicating that inadequate second language proficiency could restrain critical evaluation use (Manalo et al., 2013b). Floyd's research (2011) also echoed these findings.

Personal epistemology is defined as one's beliefs about the nature of knowledge and knowing (Hofer & Pintrich, 2002). Epistemic beliefs which influence how people handle and use knowledge have been proposed by researchers to have an impact on one's critical thinking performance (Brabeck, 1983; Gallagher, 1998; Jones, Merritt, & Palmer, 1999; Kuhn, 1999; Kurfiss, 1988; Mines, King, Hood, & Wood, 1990; Schraw, 2001). For instance, in both Brabeck's (1983) and Mines et al's (1990) studies, students' critical thinking level was found to be significantly positively related to their reflective judgment (epistemological understanding). Meanwhile, empirical studies in both Western (e.g., Feucht & Bendixen, 2010) and non-Western contexts (e.g., Chan & Elliot, 2004) have noted cross-cultural differences in epistemic beliefs, and these cultural differences in epistemic beliefs influence a certain people's

critical thinking, according to empirical data (e.g., Chan et al, 2011).

Chan et al. (2011) designed two studies to inquire into the interaction between culture, epistemic beliefs and critical thinking. In the first study, the epistemic beliefs, cognitive ability, thinking dispositions and critical thinking of 138 Chinese undergraduate students were measured with established scales, the results showing epistemic beliefs were the most related to critical thinking performance of participants. In the second study, 111 undergraduates read an argumentative text on a controversial issue, and were asked to provide their stance and supporting evidence. They were then presented with counterarguments and required to refute them. In this study, participants' epistemic beliefs were also measured using the same scales as in the first study. Data analysis was conducted to investigate to what extent epistemic beliefs predicted key argumentation skills such as two-sided argument skills and tackling of counterarguments. The results revealed that the belief in "knowledge being certain" significantly predicted both weaker two-sided argumentation and a greater tendency to devalue reasonable counterarguments. The authors suggested that the Chinese-Confucian culture might account for the beliefs of knowledge being certain which subsequently undermined the critical thinking performance of participants. Similarly, Weinstock et al. (2006) found that epistemological levels, among two other factors, predicted students' ability to identify informal reasoning fallacies, which is also an aspect of critical thinking.

Manalo et al. (2013a) investigated the role of two cultural factors, namely, independent-interdependent self-construals and regulatory mode, in university

students' critical thinking capability. In addition, two other factors were examined for their possible links to critical thinking ability: self-efficacy and location. The participants were 276 undergraduates studying in two cities in Japan and 87 undergraduates studying in Auckland, New Zealand. The two cohorts were considered to be operating in non-Western (Asian) and Western cultures. The data collection instrument used in the study was a questionnaire comprising various sets of items to assess conditions of critical thinking, independent and interdependent self-construals, regulatory mode and self-efficacy. Correlational analysis was then conducted with the data. Findings from their study suggest that cultural factors (self-construal, regulatory mode and self-efficacy) do influence students' critical thinking competence. In view of findings from this study, a brief overview of the scholarship on self-construal, regulatory mode and self-efficacy is included below.

A well-established theory that focuses on the psychological differences between Western and non-Western cultures is Markus and Kitayama's (1991) theory of independent and interdependent self-construals. According to Markus and Kitayama, culture plays a crucial role in people's view of their own self, i.e., self-construals. Markus and Kitayama (1991) proposed that people from Western societies have more independent self-construals, with individuals seeking to maintain their independence from others by attending to the self. In contrast, they asserted, people from Asian cultures have more interdependent self-construals, with individuals emphasizing harmonious interdependence with others by attending to others. Markus and Kitayama (1991) pointed out: "a consideration of the social context and the reaction of others may also shape some basic, non-social cognitive activities such as

categorizing and counterfactual thinking" (p 231). It is believed that an individual's self-construals could have notable consequences on their cognition and critical thinking (Markus & Kitayama, 1991; Manalo et al., 2013a).

Another cultural factor that could be correlated with an individual's critical thinking ability is "regulatory mode" (Higgins, Pierro, & Kruglanski, 2008; Kruglanski et al., 2000). According to the regulatory mode theory, people are pulled by two basic self-regulatory functions: assessment and locomotion. Assessment refers to "the comparative aspect of self-regulation concerned with critically evaluating entities or states, such as goals and means, in relation to alternatives in order to judge relative quality,' and locomotion refers to "the aspect of self-regulation concerned with movement from state to state and with committing the psychological resources that will initiate and maintain goal-related movement in a straightforward and direct manner" (Kruglanski et al., 2000, p. 794). It was proposed that culture shapes the distribution of an individual's tendency towards the assessment mode or locomotion mode (Higgins et al., 2008). If these researchers' proposal is correct, then culture could have a bearing on one's critical thinking as well (Manalo et al., 2013a).

Empirical studies outlined in this section indicate that cultural or culture-related factors such as English language proficiency, epistemic belief, self-construal, self-regulatory mode and self-efficacy may exert influence on students' critical thinking competence. This review of literature is helpful in understanding and explaining the critical thinking performance of the participants in this project. Unlike most of the empirical research described above that used certain scales to gauge

students' critical thinking ability, this project is concerned with the solid product of students' critical thinking skills and disposition—their argumentative essays. As such, the next section particularly explores how culture impacts on students in argumentative writing by reviewing relevant scholarship.

2.10.2 Culture and Argumentation

The impact of culture on argumentation has been discussed in the fields of intercultural communication, contrastive rhetoric and argumentation research. Intercultural communication literature reports that argumentation and reasoning vary across cultures (Fisher, 1980; Walker, 1986); contrastive rhetoric studies reveal that intercultural differences exist in what is considered as effective writing. Philosophers and argumentation theorists who uphold multiculturalism emphasize the importance of cultural difference in argument appraisal (Feldman, 1994; Goldberg, 1994). They contend that the quality of argumentation depends upon culturally specific beliefs, values and presuppositions. What can be inferred is that the criteria for soundness of argumentation are culturally driven (Usyal, 2012) and what is behind these cultural differences in argumentation is an epistemic matter (Siegel, 1999). Some empirical studies are reviewed below outlining how culture influences argumentation.

Empirical studies have been conducted to inquire into the specific cultural differences in argumentation. In a series of intervention studies by Chan et al. (2011), students of Chinese-Confucian background were found to hold the "knowledge being certain" epistemic belief. In other words, people with such beliefs tend to draw a conclusion

quickly on a controversial topic without speculating on alternative views. Chan et al. (2011) perceived this epistemic belief as absolutist and "naive" (p.74). This belief is said to have a negative effect on the specific process of critical thinking and argumentation of Chinese students, withholding them from counterarguing and refuting (Chan et al., 2011). Chan et al.'s (2011) research may shed light on the absence of counterargumentation elements in Chinese students' argumentative essays by providing an epistemological perspective. In another study, Mason and Scirica (2006) scrutinized the interaction between epistemological understanding and eighth-grader's counterargumentation skills. They found that students' level of epistemological understanding is a significant predictor of the production of arguments, counterarguments and rebuttals. Hornikx and Hoeken (2007) examined cultural differences in the evidence type and evidence quality to persuade readers. This study, instead of focusing on the cultural differences between the East and West, investigated the variation between two Western European countries (France and The Netherlands) in argument quality. Results showed that the two countries differ in regard to the persuasiveness of different evidence types and evidence quality.

Giebels and Taylor (2009) examined the issue of persuasive arguments and cultural differences from the angle of (hostage) crisis negotiation. The researchers analyzed 25 cases of crisis negotiation from low-context and high-context cultures. According to Hall (1976), low-context communication featuring explicit and direct messages is predominant in individual cultures, and high-context communication featuring hidden and context-dependent messages is dominant in collectivistic cultures. It was found that compared with low-context perpetrators, high-context ones use fewer persuasive

arguments, and tend to reciprocate persuasive arguments more immediately in the first half of negotiations (Giebels & Taylor, 2009). Giebels and Taylor (2009) assumed that this was because persuasive arguments are less central to the high-context argumentation style so that high-context perpetrators might have taken the negotiators' persuasive arguments more superficially. The researchers further assumed that high-context cultures might respond to arguments with fewer counterarguments. The results of this study may have some bearing for this research project because in Chinese culture, which is generally agreed to be high-context, students' argumentation skills may be influenced.

Contrastive rhetoric research pertains to the cultural influence on writing, a large fraction of which examines the differences in the structure of written argumentation across cultures. Initiated by Kaplan's (1966) seminal study, scholars compared cross-cultural features in expository and argumentative writing by ESL writers in terms of rhetorical patterns. For instance, Kaplan (1966), by distinguishing between the Western linear thinking style and Eastern indirect thinking style, found that Asian students tended to present their stance towards the end of an essay. The rhetorical patterns of students' essays have since been compared and contrasted across cultures or across language and within cultures in many other articles (e.g., Atkinson & Ramanathan, 1995; Clyne, 1987; Petric, 2005). These studies are not reviewed in detail in this dissertation because this project focuses more on essential argumentation skills, i.e., counterarguing and refuting; however, they are mentioned here in order to provide the background to the connection between culture and argumentation.

The aforementioned literature centres around cross-cultural differences in argumentation. However, in other literature, the tendency to merge both Eastern and Western culture in argumentative style is noted. Durkin (2008) employed grounded theory and case study to explore East Asian master students' perception of Western norms of argumentation in both classroom debate and writing assignment. Over a four year period students and lecturers were interviewed at three case sites: 42 East Asian students in the UK, 18 Chinese students at a Chinese university, and six British students at a UK university, for comparison purposes. Sixteen lecturers were also interviewed at the three sites. Durkin found that, instead of fully acculturating into the Western norms, the majority of East Asian participants adopted a "Middle Way" that combines their traditional academic values and the Western style of critical argumentation. Durkin (2008) referred to the East Asian and Western approaches to argumentation as "conciliatory dialogue" and "wrestling debate" respectively. Evidence from that research suggests that East Asian students opted for a "Middle Way" that incorporates the Western adversarial approach to debate in their approach to argumentation, so allowing ample space for diversity of opinions (Durkin, 2008).

You (2010a) argued that the inclusion of other people's views has been a practice especially in traditional Chinese political debates. By analyzing a famous debate discourse in ancient China, You (2010a) revealed that Confucian debaters used varied rhetorical devices to argue. You claimed that this was "a proactive response to the multiethnic society that China had increasingly become" (p.370). He suggested that an inclusive approach to alternative views as demonstrated in that political debate were required to meet the needs in a multi-culture and multi-ethnic society in China.

By revealing the complexities of the Chinese argumentative tradition, You (2010a) criticized some researchers as oversimplifying the issue. For instance, researchers (e.g., Cai, 1999; Connor, 1996) found a lack of argumentative tradition in China because people tend to avoid confrontation and are indirect in expression especially among modern Chinese.

Wu and Rubin (2000) compared the impact of collectivism and individualism on argumentative writing of Chinese and North American college students. Participants in the study were 40 Taiwanese undergraduates and an equal number of U.S university students. Students were asked to write on two familiar topics. Taiwanese students wrote one essay in Chinese and the other in English; American students wrote both essays in English. Student essays were then coded and analyzed for writing features linked to collectivist or individualist orientations: directness, personal disclosure, use of proverb and other canonical expressions, collective self and assertiveness. Comparisons were made across nationalities and languages, across nationalities and within languages, and across language alone. One of the results showed that American students writing in English, compared to Taiwanese students writing in Chinese, were more direct and disclosed more personal anecdotes. The results in Wu and Rubin's (2000) study were found to be associated with certain cultural writing traditions (such as the expressive norms of humaneness and collective virtues) and a Chinese language writing preference (using proverbs). Wu and Rubin (2000) also concluded that Taiwanese are becoming Westernized in their writing.

In conclusion, there is considerable evidence that the appraisal of argument quality is

culture-specific (Siegel, 1999; Connor, 2002), but there may be a tendency towards a decrease in cultural differences in argumentative writing (Wu & Robin, 2000). A shift of the focus has been made from intercultural differences to transcultural writing. The trend is that people in the East are rhetorically Westernized, and mixed patterns in argumentative writing (Wu & Rubin, 2000), collaborative approach to argumentation (Durbin, 2008) and the transfer between L1 and L2 argumentative writing of ESL learners (Uysal, 2012) perhaps represent the trend of transcultural academic writing. Although culture clearly has some impact on argumentative style, other forces also play a role. One of these is the washback from language tests which is discussed in the next section.

2.11 Washback of High-Stakes Writing Tests on Argumentative Writing Pedagogy

There is a great deal of literature on the washback effect of language tests, so reviewed here is a selected relevant bibliography. This research project concerns how the writing section of high-stakes English tests both worldwide and domestically in China influence Chinese students' argumentative writing and critical thinking, and what implications the findings of this project have on writing pedagogy. For this purpose, this literature review section comprises four aspects: 1) a brief overview of the research on washback, 2) empirical studies on the use of tests in curriculum innovation, 3) the washback of high-stakes tests on the learning and teaching of English writing, and 4) whether critical thinking competence can be enhanced via writing assessment.

2.11.1 A Brief Overview of the Research on Washback

In language education, the impact of testing on language teaching and learning has been termed "washback" or "backwash" and has been the target of a great number of studies since the seminal work of Alderson and Wall (1993). Weigle (2002), in her book, *Assessing Writing*, claims that washback from tests on curricula and pedagogy can be positive or negative. Empirical studies have been conducted to find out both the beneficial and detrimental consequences of tests (e.g., Cheng, 2005; Hamp-Lyons, 1997; Xie & Andrews, 2013). While in the earlier years of research on washback, most discussions focused on the harmful effects of examinations (e.g., Frederiksen, 1984), in recent years, increasing attention has been paid to turning the powerful impact of tests to an advantage in the classroom (e.g., Andrews, 2004; Shohamy, 2007).

A substantial body of washback literature concerns national or international high-stakes language tests (e.g., IELTS, TOEFL) either in the L1 context (e.g., Frederiksen, 1984; Herman & Golan, 1993; Scott, 2007) or in the L2 context (e.g., Cheng, 1997; Andrews, Fullilove, & Wong, 2002). While these studies vary from one another, there is a consensus that high-stakes language tests exert a great impact not only on schools and universities, but also on the whole society. Research on washback of Chinese high-stakes tests centres around the National Matriculation English Test (NMET) (e.g., Gu, 2013; Qi, 2005, 2007), the CET 4 & 6 (Cheng, 2008), and the TEM 4 & TEM8 (Zou, 2003, 2010). Much of this research is concerned with the impact of the overall test, rather than focusing on the writing component of the

tests. In an article with conceptual links to the present study, Zou (2003) reviewed the TEM8 test development and principles for the test design and analyzed the national teaching syllabus for English majors to reveal the interaction between the two. She concluded that the TEM8 test was developed under the guidance of the Syllabus; in turn, the feedback from the test resulted in revision of the Syllabus, which was meant to exert a substantial impact on classroom instruction. In a more recent article, Zou (2010) claimed that the TEM test centre, i.e., the committee overseeing the development of the test, had realized the significance of investigating the washback effects of the TEM4 & TEM8, and declared that researching the test washback "is now a top priority on the working agenda for enhancing the overall validation of the TEM battery" (p.16).

2.11.2 The Use of Tests in Curriculum Innovations in China

The use of large-scale, high-stakes tests to achieve intended positive feedback has been a general practice in many societies, especially examination-oriented societies (Andrews, 2004; Orafi & Borg, 2009; Qi, 2005; Shohamy, 2001). As such, language tests can potentially serve as mechanisms or agents for "creating more valid and real language policies that mediate and negotiate between ideology and practice" (Shohamy, 2007, p. 117). However, empirical data reveal varied, even conflicting results as to the extent to which test design and content can produce direct and desirable curricular changes in the Chinese contexts (e.g., Cheng, 1997; Qi, 2007).

To take just one of many examples, in one empirical study, Cheng (1997) examined

the washback effect of a high-stakes English test on the English language pedagogy in Hong Kong secondary schools. Cheng employed three data-collection techniques in the study: questionnaires, interviews and classroom observations. The results indicated that the test impacted the teaching materials in a straightforward and fast way, but the teaching methodology in a slower and somewhat difficult way. Cheng assumed that the limited effectiveness of the test washback on teaching methods was due to the lack of pedagogical freedom in local schools. In another large-scale study on washback conducted in Hong Kong, Cheng (2005) produced similar findings. Nevertheless, this study, with its rich quantitative and qualitative data, is of much value to researchers who query whether testing is an effective agent for pedagogical change.

Another similar example is an empirical study in mainland China, in which Qi (2005) reported the unsuccessful fulfillment of the intended washback of a high-stakes test on the secondary school English education in China. According to Qi (2005), the NMET test writers designed the test to produce changes in classroom instruction, but fell short of their goals. Qi (2005) analyzed the reasons why the NMET failed to bring about the intended washback using interview and questionnaire data collected from the NMET test writers, teachers of English, students and English subject inspectors. She found that the conflicting functions of the test itself contributed to the failure of its intended washback. The test served two functions: selecting students for post-secondary educational institutions, and promoting pedagogical innovation. However, according to Qi (2005), in practice the first function was amplified, and thus impeded the second function.

Whilst Cheng's (1997) study supplies partial evidence for the impact of tests in producing pedagogical changes, i.e., the test washback immediately changed the teaching content, but only slowly changed the teaching method, Gu's (2013) empirical research demonstrated the washback effect of the NMET in China. Gu interviewed an experienced secondary teacher of English and video-taped her teaching activities over a period of three weeks. The findings revealed that in the real teaching context, "textbooks provide the backbone of the what; teaching experience guides the how; high-stakes examinations define the what and the how" (p.1). These results coincide with Cheng and Qi's (2006) observation that the NMET in China actually plays the role of the syllabus and is a decisive factor in the curricular issues such as the teaching material and time allotment. Parallel findings can be found in the literature.

The literature reviewed above suggests that high-stakes tests can have a powerful influence on teaching and learning. However, to bring about intended curricular innovation via assessment is a complicated and long process that involves many other factors (Wall, 2000). Despite this, if tests are "properly conceived and implemented", then aligning teaching with what is assessed can facilitate curricular innovation (Popham, 1987, p. 679).

2.11.3 The Washback of High-Stakes Writing Tests on English Writing Pedagogy

While the aforementioned studies are concerned with the washback of an overall test on pedagogical issues, a small number of studies have inquired about the impact of

the writing assessment on the teaching and learning of English writing. Green (2006), for example, in order to explore the washback of the academic writing section of IELTS, investigated the relationship between writing test design features and pedagogical practices via the observation of the behavior of 197 learners and 20 teachers over 51 classroom hours. Green (2006) found that the IELTS tasks featured: 1) impersonal topics, 2) a limited range in terms of rhetorical function, and 3) composition based on personal opinions, and he made predictions according to these features. While his chief purpose was to examine how the washback of the IELTS writing task predicted classroom activities, Green also compared data collected from the writing-focused IELTS preparation classes with that of English for Academic Purposes (EAP) classes at U.K. universities. Classroom observation data showed that 58% of the time in IELTS preparation classes and 48% of the time in EAP classes was spent practicing impersonal topics conforming to the IELTS standards. In addition, topics of study were comparable in both IELTS and EAP classes. Most notably, little time was spent on personal topics or academic topics in either IELTS preparation classes or EAP classes. However, topics were discussed in more depth in some EAP classes. Evidence in Green's study (2006) suggested considerable washback of IELTS writing tasks on the writing pedagogy conducted in the classes he viewed.

In another washback study, Qi (2007) explored the effectiveness of a high-stakes test in its role as an agent for pedagogical change in the Chinese context. Qi (2007) compared the actual and anticipated influences of the writing task of the NMET test on secondary school teaching, with data collected via interviews, classroom observations, and questionnaires. The findings showed that the writing task design in

the NMET had not been an efficient agent for generating changes in the secondary school classroom. The communicative features of the writing design in the test did not exert the anticipated impact on the teachers when they prepared students for the NMET test. In the actual classrooms, both teachers and students put emphasis on the grammatical correctness and the "assumed preferences of markers", but neglected the communicative approach, which highlighted simulated authentic writing as advocated by the test constructors (Qi, 2007, p.51).

2.11.4 Can Critical Thinking Competence Be Enhanced via Writing Assessment?

In view of the function of tests as a language policy tool, and given the close interaction between critical thinking and written or oral language, one question to be raised at the end of this section is: can critical thinking education be encouraged via writing assessment? Some indication in the literature comes from Eckstein and Noah (1993) who contend that one crucial function of tests is "encouraging higher levels of competence and knowledge" (p.11). Similarly, Alderson and Wall (1993) argue that "tests are held to be powerful determiners of what happens in the classroom" (p. 115). If the active engagement of critical thinking is underscored in the language tests, teachers and learners are more likely to include practice to this end. Following this line of thought, Shirkhani and Fahim (2011) propose that EFL learners' critical thinking abilities should be enhanced through assessment practice by outlining some specific suggestions for test designers and English language teachers. However, to date, no empirical study has been conducted to explore the idea of enhancing critical thinking competence through writing assessment. In this sense, Study 3 of this project

is based on a sound theoretical framework, but may be the first empirical study in this particular area.

In summary, high-stakes language tests are capable of influencing the national educational policies in general, and what the students learn and how teachers instruct their students in a language classroom in particular (Shohamy, 2007), although there are other factors affecting the test consequences (Andrews, 2004; Wall, 2000). With regards to promoting critical thinking in the argumentative writing of mainland Chinese students, the use of high-stakes English tests could be salient. A review of literature here not only highlights the implications that Study 2 of this project may have on pedagogy in argumentation, but also justifies the importance of investigating the writing prompts of high-stakes tests in Study 3.

2.12 The Writing Prompt and its Effect on Writing Performance

Writing prompts constitute a significant factor in the study of L2 argumentative writing because the components of a writing prompt can have considerable impact on students' writing performance in terms of rhetorical and linguistic features as well as critical thinking ability. With the overarching purpose of the present project being the investigation of critical thinking in students' argumentative writing, the third study of this project was designed to explore the relationship between two components of the writing prompt, i.e., the object of enquiry and rhetorical function, and the critical thinking of student writers. This section reviews the relevant literature.

2.12.1 The Subject Matter (Object of Enquiry) of a Writing Prompt in High-Stakes Tests

The subject matter of writing prompts concerns the topic or content area that test-takers are required to write about (Weigle, 2002). This "subject matter" is also referred to as the "object of enquiry" of a piece of writing (e.g., Moore and Morton, 2005). The subject matter of writing prompts can have considerable bearing on the linguistic and rhetorical features in a student writer's responses (Hinkel, 2002; Yu, 2010). Hinkel (2002) claimed that students should have certain exposure to the subject matter of the prompts, but not necessarily be personally involved. She observed that students' response to a prompt that is close to their personal experience tends to be simpler and less sophisticated than their responses to impersonal prompts. In other words, writers do not always demonstrate the full range of their writing abilities in essays about personal topics. Conversely, when forced to write on impersonal topics, ESL student writers tend to exhibit more native-like language features (Hinkel, 2002). Similarly, Yu (2010) found that impersonal topics resulted in higher lexical diversity. This personal/impersonal dichotomy of the subject matter in a writing prompt was also explored by Hoetker and Brossell (1989) and Spaan (1993), the results of both of which revealed that most participants performed without significant differences on the two writing prompt types as measured by holistic scores, but Spaan discovered that advanced-proficiency-level participants performed better responding to impersonal argumentative prompts. Other researchers (e.g., He & Shi, 2012; Lee & Anderson, 2007; Tedick, 1990) inquired about the specific/general issue of the subject matter in a writing prompt. Tedick (1990) found that students generally

performed better on topics related to their fields of study. Lee and Anderson (2007) also found that topic generality (whether topics are answerable regardless of students' academic major) affected participants' writing performance. As to what effect the prompts with different subject matter cause in a written response in empirical studies, researchers investigated the effect that differing prompts have on L2 texts in terms of modal verb uses (Hinkel, 2009), lexical diversity (Yu, 2010), and general performance and overall score (Cho, Rijmen & Novak, 2013; He & Shi, 2012; Lim, 2010; Lee & Anderson, 2007). The significant findings of these investigations are outlined below.

Yu's (2010) study was designed for two purposes: 1) to verify the relationship between lexical diversity, holistic quality of written or spoken discourse, and test-takers' language proficiency; 2) to explore the effect of the topic in the writing prompts on the lexical diversity in test-takers' essays. For the purpose of the present study, reviewed here are the second research aim and its relevant results. Yu (2010) defined lexical diversity as the range of vocabulary displayed in a text following Duran, Malvern, Richards and Chipere (2004). The data comprised 200 essays from the archived data of an international language test. The essays were written on five different topics: two personal and three impersonal. The lexical diversity of the essays was measured with the parameter D which is employed by an increasing number of researchers and also suitable for short texts. Findings revealed that the lexical diversity was significantly positively correlated with the subject matter in the prompts, even after controlling for the two variables of essay score and overall language test score. Specifically, lexical diversity was reported to be higher in test-takers' response

to impersonal topics.

Hinkel (2009) explored the effects of essay topics on modal verb uses in L1 and L2 academic writing. Hinkel (2009) analyzed the uses of two types of modal verbs, i.e., possibility/ability and obligation/necessity, in 718 essays by students of various degree programs at four American universities, 523 of which were written by nonnative English speakers. These essays were written on five prompts, four of whose subject matter was culturally bound and of a personal nature, e.g., parents, choice of academic majors, while one was less personal. Hinkel found that the mean frequency of modal verbs in L2 essays was significantly influenced by the subject matter in the prompt. In particular, in the L2 writing of Asian students, the use of obligation/necessity modals in essays about cultural and personal matters was higher. The results echoed and complemented earlier findings by Hinkel herself (1995) that a dominant use of certain types of modal verbs could be culture-and-topic dependent. The methodology in Hinkel's (2009, 1995) studies is of certain interest to Study 3 of this project, which explore the effect of prompts with distinct subject matter and object of enquiry on the critical thinking of students, which are partially reflected in their use of verbs.

Lee and Anderson (2007) measured the effect of three factors (topical content, writer's major, and writer's language proficiency) on the writing performance of test-takers on a large-scale American writing test. The data were 2,888 essays written by graduate students on three topics: brain, ethics, and trade. The findings revealed that different topics affected test-takers' writing scores while controlling for the

English proficiency. However, students' majors had no effect on their writing performance. On the other hand, Lim's (2010) study yielded reverse findings, revealing that prompt dimensions, including topic domain, generally did not affect writing performance on a high-stakes test.

In a recent study, He and Shi (2012) examined the effects of topical knowledge on the writing performance of 50 ESL students of varied levels of English proficiency at a Canadian university. The students were asked to write timed-impromptu essays on two prompts modeled after writing prompts of a high-stakes English test in Canada, with one prompt requiring general knowledge and the other pertaining to specific knowledge. Findings indicated that the writing performance of students across three proficiency levels on the general topic were significantly better than that of the specific topic. In another study regarding whether clear information is conveyed in the subject matter of a writing prompt, Cho et al (2013) discovered that the distinctness of ideas within the prompt had an effect on test-takers' writing performance on TOEFL.

To summarize, to date most scholarship on the subject matter of writing prompts is confined to the effect of two dichotomies, i.e., personal/impersonal prompts, or general/specific prompts on the rhetorical or linguistic features of students' writing. Whilst these empirical studies provide a useful foundation and information for the present study, I find that new endeavors are lacking to further explore the variation of subject matter in a writing prompt. For instance, some domains of topic content related to decision-making may promote the logical thinking of students. As such, the

subject matter of writing prompts could impact students' critical thinking as well. In this sense, the third study of this project is exploratory building on the literature, but advancing it by comparing the effects of two prompts with subject matters belonging to distinct domains.

2.12.2 Rhetorical Function of Writing Prompts and Its Possible Influence on L2 Writing

Rhetorical function is the purpose of written or spoken discourse, i.e., whether the discourse is to describe, or explain, or compare and contrast, etc. (Selinker, Trimble & Trimble, 1976; Selinker, Trimble & Trimble, 1978). While the writing of various genres has different rhetorical functions, the genre does not equal the rhetorical function because genre is concerned with the style of a piece of writing. However, the distinction between the two has not been always clear. In some studies on university writing (e.g., Gardner & Nesi, 2013), the two concepts are not particularly distinguished.

Starting with Selinker et al.'s (1976, 1978) work in the 1970s, the research on rhetorical function has flourished over the past four decades. In recent years, studies on rhetorical function have been undertaken in various specified areas. For instance, Sowards (2012) explored the rhetorical functions of letter writing; Charles (2007) investigated the use of corpora to help students "make the connection between general rhetorical purposes and specific lexico-grammatical choices" (p.289); Petric (2007) identified nine rhetorical functions of citations in ESL master theses:

attribution, exemplification, further reference, statement of use, application, evaluation, establishing links between sources, comparison of one's own findings or interpretation with other sources, and other.

In a study that perhaps comes closest to the present one, Moore and Morton (2005) compared two sets of writing prompts: 155 writing prompts used as university assignment tasks across disciplines in Australia and 20 IELTS essay prompts. The two sets were compared over four dimensions: genre, source of information, rhetorical function and object of enquiry. For the purposes of the present study, reviewed here are the between-group differences and similarities in rhetorical function and object of enquiry. Moore and Morton (2005) identified 9 rhetorical functions in the writing prompts, namely, evaluation, description, summarization, hortation, comparison, explanation, recommendation, prediction and instruction, with evaluation and description being the two most frequent in university assignments, and evaluation and hortation the two most frequent in IELTS essay writing. With regard to object of enquiry, Moore and Morton (2005) categorized all prompts in the data into two broad types: phenomenal and metaphenomenal, following Halliday (1994). The phenomenal type required students to write about the concrete entities in the real world, such as events, situations and practices; the metaphenomenal required students to write about abstract entities of ideas, theories and methods. It was found that the objects of enquiry in all IELTS prompts fell into the phenomenal types, while university tasks involve both phenomenal and metaphenomenal. Moore and Morton's (2005) study is of particular interest to the present study because their classification scheme for writing prompts appears to best fit the research context. Therefore, the two terms -

"rhetorical function" and "object of enquiry" (subject matter) - were adopted in the third study of this research project.

Substantial empirical evidence indicates that different rhetorical tasks might elicit different kinds of linguistic production (Ginther & Grant, 1997; Hinkel, 2002; Hoetker, 1982; Reid, 1990; Spaan, 1993). Reid (1990), for example, compared writers' responses to two types of topic: 1) comparison/contrast and take a position (C/C) and 2) describe and interpret a chart or graph (G). Student writers were native speakers of Arabic, Chinese, Spanish and English. Findings revealed that different topic tasks elicited responses that were measurably different in linguistic features. Specifically, writers used significantly more content words in the C/C topics than in the G topics, while they used significantly longer words and significantly more pronouns in the G topics. However, when responding to the two types of topics, student writers' syntax did not change. Reid (1990) also compared differences among the four language backgrounds.

In a similar study, Spaan (1993) investigated the impact of both the rhetorical function and the content area on student's writing performance. Participants were 88 students consisting of 61 undergraduates and 27 graduates and, in terms of writing competence, 34 beginning, 27 intermediate and 27 advanced writers. All participants were randomly assigned a set of two prompts representing two topic types: narrative/personal or argumentative/impersonal. In addition to noting the overall score of an essay, Spaan (1993) analyzed both the linguistic and rhetorical features of the data, with the linguistic features examining over three dimensions: fluency,

syntactic sophistication and accuracy. The results revealed that there were no significant differences in general writing performance for the two groups who wrote on the two different prompt sets. However, the evidence meriting attention in this study is that the advanced writers scored slightly higher on the argumentative/impersonal prompt.

Rhetorical function in the writing prompt may also have an effect on students' critical thinking. In an empirical study, Wang (2010) compared Chinese students' choice of prompts and high-stakes test prompts, and found that variation existed between the two. The preferences in students' topic selection to some extent are linked to the rhetorical function of the writing prompt. Notably, Wang's research (2010) indicated that there is a relationship between students' topic selection and their critical thinking development. Wang (2010) commented that students' prompt selection reflected problems in the classroom instruction and assessment practices in mainland China. Wang (2010) concluded the writing prompts have the potential to engage students in higher order thinking skills such as critical thinking, analysis and problem solving.

In conclusion, the literature reveals that both the subject matter (object of enquiry) and rhetorical function of a writing prompt can have a considerable impact on students' writing performance in terms of rhetorical and linguistic features. Further, the design and selection of writing prompts can have a bearing on students' critical thinking development.

This literature review chapter provides a theoretical basis which guides and supports

the present research. In the next chapter, I will report on the research methods and

procedures used to answer the research questions in the project.

CHAPTER 3

METHODOLOGY

3.1 Purpose, Design and Methodology of the Project

The purpose of this research project is to address the critical thinking development of Chinese undergraduate English majors in their argumentative writing. The research was motivated by my initial observation as a teacher of English composition that one essential element of critical thinking, i.e., counterargumentation, was absent in students' argumentative essays at a mainland Chinese university. In addition, the stereotypical writing prompts used in high-stakes tests seemed to generate formulaic responses among students, which might have undermined students' critical thinking.

3.1.1 Links Between the Three Studies

Three studies were conducted among undergraduate English majors at a mainland Chinese university. Study 1 examined students' performance in and perceptions of written argumentation. It further inquired about what the contributing factors might be and how these factors might have shaped students' performance and perceptions. Study 1 aimed to answer three research questions:

RQ1: How do a group of mainland Chinese undergraduates respond to a typical argumentative writing prompt? In particular, do they produce any counterarguments and rebuttals?

RQ2: What are these students' perceptions of argumentative writing?

RQ3: What factors contribute to students' performance in and perceptions of argumentative writing? How are counterargumentation and critical thinking represented in these factors?

Study 2 investigated the effect of an instructional intervention on students' performance in and perceptions of written argumentation. Informed by the modified Toulmin model of argumentation, the intervention aimed to improve students' argumentative strategies, especially counterargumentation skills. Four argumentative elements (claim, data, counterargument and rebuttal) adapted from Toulmin's model (1958) and developed in empirical studies (e.g., Nussbaum & Schraw, 2007; Qin & Karabacak, 2010; Stapleton, 2001) were used in the classroom instruction and were identified when analyzing students' argumentative essays. Study 2 adopted a quasi-experimental design with pretest and post-test on experimental and control groups. The impact of the intervention on students' critical thinking ability was also measured. Four research questions were answered in Study 2:

RQ4: How does an intervention in counterargumentation affect students' written argumentation performance in terms of quantity of counterarguments, rebuttals, and overall quality of an argumentative essay?

RQ5: Is there a correlation between the extent of counterargumentation and the overall score of an essay?

RQ6: What are the students' perceptions of argumentative writing after the intervention?

RQ7: How does the intervention impact students' critical thinking?

While Study 2 focused on the pedagogical intervention, Study 3 was designed to investigate the influence of assessment on the development of students' argumentative writing and critical thinking. The writing prompts of three high-stakes tests were surveyed in terms of the object of enquiry and rhetorical function. Based on the survey, two types of writing prompts, namely, conventional and exploratory prompts, were generated. Then, two groups of students were asked to respond to the two prompts. The responses of the two groups were compared and analyzed in order to see how differentiated prompts might influence students' written argumentation and critical thinking. Two research questions were asked in Study 3:

RQ8: What patterns do the writing prompts in three high-stakes tests, IELTS, TOEFL and TEM4 display in terms of their rhetorical function and object of enquiry?

RQ9: In what ways do the essays of students responding to a "conventional" writing prompt differ from those of an exploratory, non-conventional prompt?

3.1.2 Overall Methodology

To answer the nine research questions in the three studies, a mixed-methods approach was adopted with both quantitative and qualitative data collected. In essence, the mixed method approach has particular value when we want to achieve a fuller understanding of a multi-faceted phenomenon (Mertens, 2010). As Fraenkel, Wallen, and Hyun (2012) point out, the mixing of quantitative and qualitative research

methods can not only help to clarify relationships between variables, but also allow in-depth exploration between variables. In addition, there is substantial evidence in the literature that mixed methods can "open up fruitful new avenues for research in the social sciences" (Dornyei, 2007, p163). In the case of this research project, the adoption of mixed methods is necessary. Since the issue of critical thinking abilities in the written argumentation of mainland Chinese undergraduates is embedded in complex educational and social contexts, the inquiry into this issue was inevitably challenging, which was why three studies were designed to address the issue. For the same reasons, multiple methods and a variety of sources of data were employed to expand the understanding of the issue. The first study used a pretest, a questionnaire, interviews, and classroom observation to collect data; content analysis was conducted to analyze data. The main part of the second study employed a quasi-experimental design with a pretest and a post-test including treatment and control groups (intact classes) because in real learning environments with natural class groups, quasi-experimental inquiries can have solid external validity (Dornyei, 2007). The third study was composed of surveys and another quasi-experiment. The data collection instruments in the three studies are outlined in Tables 1, 2 and 3 in the following section.

3.1.3 Overview of the First Study

The first study, addressing RQs 1-3 (see Table 3.1), began with a pretest (Appendix 1) to examine the performance of students' argumentative writing in both experimental and control groups. In order to assess participants' argumentative writing proficiency,

a content analysis was conducted on the pretest scripts to investigate students' skills in producing counterarguments and rebuttals; a questionnaire (Appendix 2) and follow-up interviews were administered to collect further data on: 1) how students perceived argumentative writing; 2) what factors the students and teachers considered may have influenced students' performance in the perceptions of argumentative writing. After that, the researcher carried out document analysis and observation to explore how the Syllabus (2000), the current high-stakes test writing prompts, the writing textbooks and the classroom instruction influence students' performance in and perceptions of argumentative writing.

Table 3.1

Study 1 research questions matching data collection and analysis methods

RQ1: How does a group of mainland Chinese undergraduates respond to typical argumentative writing prompts? In particular, do they produce any counterarguments and rebuttals when presenting prompts requiring them to take a stand?		
Data collection method: Pretest	Source of data: Pretest scripts	Data analysis method: Content analysis
RQ2: What are these students' perceptions of argumentative writing?		
Data collection method: Pre-questionnaire & pre-interview	Source of data: Questionnaire & interview responses	Data analysis method: Content analysis
RQ3: What factors contribute to students' performance in and perceptions of argumentative writing? How are counterargumentation and critical thinking represented in these factors?		
Data collection method: Pre-questionnaire; Interview with teachers; Classroom observation	Source of data: Questionnaire & interview responses; documents; Observation protocol and notes	Data analysis method: Content analysis

3.1.4. Overview of the Second Study

A 12-week intervention was implemented during which the experimental groups received instruction including counterargumentation skills while the control groups received normal instruction in written argumentation. The intervention was followed by a posttest (Appendix 1) and a questionnaire (Appendix 2; this questionnaire was almost the same as the one used before the intervention). The results from the pretest and posttest in both experimental groups and control groups were compared to find out whether there was any improvement in argumentative writing skills and whether there was any difference in this respect between experimental and control groups. Statistical data collected through text analysis on students essays in the pretest and posttest of the experimental groups and control groups was compared with regard to two aspects: (1) the frequency count of counterarguments and rebuttals in essays between the two groups; (2) the overall scores of essays between the experimental and control groups. The results from the first and second questionnaires were compared to find out the changes in students' perception of the argumentative writing. Small group interviews (semi-structured) were conducted following the first and the second questionnaires to reveal students' perceptions of argumentation in more depth and opinions about the instructional intervention.

In the last stage of the study, students' essays from the pretest and posttest in both experimental and control groups were evaluated again using a critical thinking rubric (Appendix 3) adapted from Stapleton (2001) and Washington State University Critical Thinking Project (hereinafter WSU CT Project, 2009) to see whether there was any

variation in the development of critical thinking between the two groups. The rubrics generated by Stapleton (2001) and WSU CT Project (2009) have been the most adopted in assessing critical thinking in students' written argumentation. The rubric used in this study is an attempt to combine the strength of both rubrics. Table 3.2 provides a general depiction of Study 2. The procedures of the study and data-generating instruments will be described in detail in section 3.3.

Table 3.2

Study 2 research questions matching data collection and analysis methods

RQ4: How does an intervention in counterargumentation affect students' written argumentation performance in terms of quantity of counterarguments and rebuttals, and the overall quality of an argumentative essay?		
Data collection methods: Intervention posttest	Source of data: Posttest scripts	Data analysis methods: Content analysis
RQ5: Is there a correlation between the extent of counterargumentation and the overall quality of an argumentative essay?		
Data collection methods: Posttest	Source of data: Posttest scripts	Data analysis methods: Statistical analysis (Pearson correlation)
RQ6: What are students' perceptions of argumentative writing after the intervention?		
Data collection methods: Post-questionnaire Post-interview	Source of data: Questionnaire and interview responses	Data analysis methods: Content analysis
RQ7: How does the intervention impact students' critical thinking?		
Data collection methods: Pretest Posttest	Source of data: Pretest and posttest scripts	Data analysis methods: Statistical analysis (ANCOVA)

3.1.5 Overview of the Third Study

The third study was carried out in two phases. In the first phase, the prompts of three high-stakes English tests, namely, TOEFL, IELTS and TEM4 (a Chinese test) were analyzed. Six sets of 20 (n= 120) recent writing prompts were gathered randomly from the past tests of each of the three tests[5]. The prompts were categorized according to their rhetorical function and object of enquiry, following Moore and Morton (2005), in order to see whether these prompts displayed certain patterns. "Rhetorical function" here refers to what the discourse unit, or essay, "is trying to do" (Trimble, 1985), e.g., evaluating components of an argument, or explaining the cause of an entity (Moore & Morton, 2005). "Object of enquiry" refers to the nature of the variable introduced in the writing prompt, broadly speaking, the topic under discussion (Moore & Morton, 2005).

Moore and Morton (2005) found that writing prompts of high-stakes English tests tended to be restricted to narrow fields and thus might impact negatively on university students. Based on the results from Moore and Morton (2005) and the findings from phase 1 of the third study which are aligned with Moore and Morton (2005), I hypothesized that Chinese undergraduates' lack of critical thinking in their written argumentation was in part due to the convergent writing prompts which I termed conventional writing prompts. It was further hypothesized that writing prompts with different rhetorical function and object of enquiry would stimulate more

---

[5] Each of the three tests comprises two writing tasks: essay writing and non-essay writing; hence six sets of writing prompts were gathered for investigation.

critical thinking in students' writing. With these hypotheses, an exploration was designed in the second phase of this study.

The second phase was an exploratory study in which two groups of undergraduate students responded to two different argumentative prompts. The control group responded to a conventional writing prompt; in other words, the prompt was similar to those found in major English tests. The experimental group responded to a prompt whose object of enquiry and rhetorical function were quite unlike those found in generic prompts. First, the object of enquiry was in an area other than education, technology or social issues, which were found to be overwhelmingly represented in the high-stakes tests (as explained in the Results). Second, the rhetorical function required more analyzing and problem-solving instead of evaluation and hortation, which were found to be the two most common rhetorical functions in the high-stakes tests surveyed in the first phase. Scripts were collected and transformed into electronic data. Both corpus and discourse analyses were performed on the data.

Table 3.3

Study 3 research questions matching data collection and analysis methods

RQ8: What patterns do the writing prompts in three high-stakes tests, IELTS, TOEFL and TEM4, display in terms of their rhetorical function and object of enquiry?		
Data collection methods: Survey	Source of data: Prompts from IELTS, TOEFL and TEM4	Data analysis methods: Content analysis
RQ9: In what ways do the essays of students responding to a "conventional" writing prompt differ from those of an experimental, non-conventional prompt?		
Data collection methods:	Source of data:	Data analysis methods:

| Quasi-experiment | Scripts of experimental and control groups | Corpus analysis & Discourse analysis |

3.2 Settings of the Studies

The three studies were conducted in Jiangxi Normal University, a provincial educational university (public) in central China. Studying and living at the university are approximately 30,000 undergraduate students of whom around 1,300 are English majors. Students in the university are admitted from all parts of China; in recent years some programs have been opened to international students. A total number of 254 first-year and second-year English majors and four teachers of English Writing participated in the three studies, with Studies 1 and 2 involving 125 second-year students and four teachers and Study 3 129 first-year students. While this section provided a general description of the background of participants focusing on the English Writing course, more detailed participant information will be provided accordingly in the sections below.

The English Writing courses, comprising English Writing I, II, III and IV, are among the basic courses for English majors from the first, second and third years. English Writing I is provided in the second semester of the first year. English Writing II and English Writing III are provided in the first and second semesters of the second year respectively, and English IV is provided in the first semester of the third year (see Table 3.4 for more details of English writing courses at the research site). The course textbook used for English Writing I, English Writing II and English Writing III is *A Handbook of English Writing* (Ding et al., 1994) which is also used in the Beijing

Foreign Studies University and many other universities. A table of contents of the textbook can be found in Appendix 4. This textbook is used as the fundamental teaching material for English Writing I, English Writing II and English Writing III. Supplementary materials are generally also used by teachers in the university which, mainly adapted from media and overseas publications, are decided by individual English writing teachers since collective lesson preparation sessions are not available.

Table 3.4

English Writing courses for undergraduate English majors at the research site

Year	Semester	Course	textbook	curriculum
one	one			
	two	English Writing I	*A Handbook of English Writing*	Diction, sentences and paragraphs
two	one	English Writing II	*A Handbook of English Writing*	Different types of writing: description, narration, exposition and argumentation, with a focus on argumentation
	two	English Writing III	*A Handbook of English Writing*	Practical writing: letters, note, notice, resume, reading report, etc.
three	one	English Writing IV	Instructor's Self-prepared materials	Thesis writing
	two			
four	one			
	two			

Briefly, English Writing I deals with punctuation rules, diction, sentence and

paragraph writing in English. In the English Writing II class, students practise essay writing in light of the four types of discourses, namely, description, narration, exposition and argumentation; however, the focus is largely on expository and argumentative writing. Since the distinction between the two types of writing is marginal (Ding et al, 1994), in practice, "argumentative writing" is used as equivalent to expository writing and persuasive writing in class discussions. In essence, argumentative writing is the curricular focus of English Writing II. English Writing III is offered in the second semester for the second year English majors. Because by the end of that semester students are to take the TEM4 test, teachers help students prepare for the TEM4 writing test which consists of the writing of (1) a short note; and (2) an argumentative essay. The focus of English Writing IV is thesis writing. The English writing curricula outlined above appear to be not unlike those in many other Chinese universities (You, 2010b). English writing pedagogy in Chinese universities has traditionally followed a uniform curriculum to some extent for two main reasons: first, the teaching of all courses of English programs is under the guidance of the Syllabus (2000) across the country because most Chinese universities are public universities which are run by the Ministry of Education, and the Syllabus (NACFLT, 2000) was written by a commission entrusted with the task by the Ministry of Education. Second, there are not many available instructional materials for English writing (You, 2010b).

3.3 The First Study

3.3.1 Questionnaires

While the first procedure in Study 1 was a pretest designed to collect data for answering Research Question 1, this pretest is regarded as one part of the quasi-experiment, which consisted of pretest, intervention and posttest; it will therefore be presented in section 4.

A questionnaire (Appendix 2) with eight close-ended questions and one open-ended question was distributed to the 125 participants in the quasi-experiment to obtain some essential information about how the participants perceived written argumentation. The primary purpose of the questionnaire is to elicit data to answer Research Question 2. However, among the nine questions, Question 9 was designed specifically to help answer the first half of Research Question 3, i.e., What factors contribute to students' performance in and perceptions of argumentative writing? The students were asked to answer this question only before the intervention. I designed the questions which met the needs of this particular study. Another function of the questionnaires was to elicit general impressions in order to facilitate the follow-up interviews. In addition, the responses to the questions were to provide useful information for the instructional intervention. In fact, the questionnaire data was used to develop a teaching schedule for the intervention.

In light of Dornyei's (2010) guidelines, the questions were written concisely, with

each question written in both English and Chinese. A pilot test on the instrument was given to 63 students at the research site before it was used in the actual study. The internal consistency of the questions was measured with Cronbach's alpha. An acceptable reliability (0.76) was resulted for the questionnaire items.

The same questionnaire was administered twice (in a pre-post mode) in order to compare the changes in participants' perceptions of counterargumentation and argumentative writing before and after the intervention, but Question 9 was only answered in the pre-questionnaire. The pre-questionnaire was administered in class to all the 125 participants at the beginning of the 12-week intervention, and 119 responses were collected. The post-questionnaire was administered in class to all the participants, and 117 responses were collected. Since the questionnaire respondents were also participants in the quasi-experiment, a description of the respondents' demographics will be given in section 3.3.5.1.

3.3.1.1 Coding and Analyzing of the Questionnaire Data

I collated and coded the questionnaire data together with a helper who had appropriate qualifications. With a master's degree in English language education, the helper had been a teacher of English for nearly nine years at the university where the research took place. In addition to her professional qualifications, she showed considerable interest in this research project and communicated well with me. She served as the helper for all the questionnaire and interview data analysis. Data coding and analysis of the pre-questionnaire were conducted immediately after receiving the

completed questionnaire in order to provide important information for succeeding research procedures. The post-questionnaire data analysis was carried out a few months later. However, data analysis for both pre-questionnaire and post-questionnaire is reported together to avoid redundancy. A similar procedure is followed with the presentation of the data analysis for the pretest and posttest, and the pre-interview and post-interview.

For data coding of the close-ended responses, the "rows-and-columns approach" (Brown, 2001, p.95) was adopted. Each question number was listed as a column heading, and each row was for one person's responses. Then each possible answer was given a code, for instance, "A Description" equaled "1" and "D Argumentation" equaled "4" on question 1. Recording the data involved the following steps. First, the responses were manually transferred from the paper questionnaires onto a SPSS spreadsheet. For the pre-questionnaire, two spreadsheets were made, one for the control group and the other for the experimental group. The post-questionnaire was done likewise. When I completed the data entry, I invited the helper to conduct a check. The checking was done when both of us agreed that all the data were present and correct. A statistical data analysis was then conducted by the helper and myself. The coding for the ninth question, which asked the respondents to rank the factors influencing their argumentative writing from 1 to 6 (with 1 indicating the most significant and 6 the least) was performed as follows. The frequencies of ranking for every factor were counted using the COUNTIF function in Excel. Then the mean rank was calculated and compared. For instance, for option A (classroom instruction at college), 62 respondents ranked it "1", 28 ranked it "2", 17 ranked it "3", 6 ranked it

"4", 6 ranked "5", and nobody ranked it "6". Rank "1" was given a value of 6 points and "6" a value of 1 point because of their respective significance. The mean rank for option A is hence $(62 \times 6 + 28 \times 5 + 17 \times 4 + 6 \times 3 + 6 \times 2 + 0)/119 = 5.18$. In this way, a list of the first four most influential factors was tentatively drawn up. The results were later reconsidered when combined with data from the interview data with writing teachers concerning the same question. In this way, I was able to finally identify the four factors considered by the student respondents and teacher informants as influencing the argumentative writing the most.

As to the coding of open-ended responses, the norms of qualitative content analysis were adopted. For the pre-questionnaire, approximately 14 percent (n = 17) of respondents answered the last question. For the post-questionnaire about 16 percent (n=19) did so. All the responses were converted into an electronic file for data analysis, which entailed a process of systematically sorting and classifying the answers (Bogdan & Biklen, 1997). For this part of the data, the analysis was done by myself and the helper independently. The analysis was data-driven in nature, that is, we had no presumptions before the analysis. All the categories to which the responses were assigned emerged in the course of data analysis. As stated by Dornyei (2010), content analysis of open-ended questions is an iterative process. Thus, each response was read with key contents highlighted, and initial categories were formed. The categories were then cross-checked by looking again at the responses. After that, these categories were examined to form broader categories. Eventually, with the joint work of both coders, all the responses were classified into four general categories, namely, "expression", "provision of supporting evidence", "structure" and "horizon".

For instance, under the category of "horizon", students expressed their wish to broaden their horizons and enrich their background knowledge in order to write more impressive main arguments. Definitions and examples of each category are reported in the Findings chapter. Upon the completion of data analysis, the inter-coder agreement coefficient was calculated using Cohen's Kappa. Inter-coder agreement of assigning responses to the four categories was: "expression" K=.92, "provision of supporting evidence" K=.89, "structure" K=.91, "horizon" K=.90.

3.3.2 The Interviews

Focus group semi-structured interviews with students and with teachers of English writing were conducted after the questionnaire. As stressed by researchers (e.g., Bryman, 2012; Dornyei, 2007), the strengths of focus group interviews lie in the inter-group interaction, which facilitates the yield of rich data. Focus group interviews usually center on a specific theme; in this study, the theme is written argumentation and counterargumentation. The construction of the interview questions was informed by the questionnaire data, and in turn the interview data provided more in-depth insight into the questionnaire responses. The actual interview was done in Chinese as requested by the informants. Four focus groups (one from each of the four classes in this study) of six student informants participated in a 35-minute interview before and after the instructional intervention. A stratified sampling was adopted in choosing the student informants. Procedures of the sampling were as follows: in each class, students were categorized into three groups according to their writing proficiency. They were then labelled with a number. For the six informants from each

class, two were from the top proficiency, two from the intermediate proficiency and two from the low proficiency groups. In this way, I assumed that voices of students at each proficiency level were represented. In addition to the student interviews, one focus group comprising three teachers took part in a 40-minute interview. The three teachers who agreed to participate in the interview were engaged in the teaching of English argumentative writing. Their demographics are set forth in Table 3.5.

Table 3.5

Demographics of teacher informants in the interview

	Gender	Age	Academic degree	Teaching experience	Years of teaching writing
Informant A	female	35	MA	9 years	7
Informant B	male	30	PhD	3 years	3
Informant C	female	41	MA	15	10

The interviews with students were administered in a pre-post mode in order to compare the change in informants' conceptions of written argumentation and counterargumentation before and after the intervention. Both the pre-interview and post-interview centered around two general questions:

● What approaches do you adopt to make your argumentative writing more persuasive?

● What are the difficulties you encountered in making your writing persuasive?

These two questions set boundaries for the interviews which were intentionally administered in a relaxed atmosphere in order to elicit more data.

The interview with teachers concerned one topic: what factors exert an impact on the

teaching and learning of argumentative writing. At the beginning of the interview, teachers were presented with the following question:

In your opinion, which of the following factors influence undergraduate English majors' performance in and perceptions of English argumentative writing?

A. Teachers' classroom instruction at college

B. What students have learned in the middle school about English argumentative writing

C. Relevant stipulations in The National Syllabus on the Teaching of English Majors

D. The English writing textbook used for classroom instructions on argumentative writing

E. The requirements and criteria in the writing prompts and rubrics of high-stakes tests such as TEM4 & TEM8

F. The argumentative essays students read in newspapers and other media

Teachers were asked to rank the six influencing factors according to their significance. For instance, "1" indicates the most influential factor. After the ranking, they were asked to explain their choices and provide other points of view about the issue.

I was the moderator for all nine interviews, which were recorded. Upon completion of the focus group interviews, all the sound data were transcribed into written data in electronic files. Three electronic files were set up during the data transcription for the student pre-interview and post-interview respectively: one for interview question 1; one for interview question 2, and all other information was put into the third file. In each file, data from the experimental group and the control group were entered into

separate sections. An electronic file was also set up for the teacher interview.

3.3. 2.1 Coding and Analyzing of the Interview Data

The interview data were synthesized by content preceding the actual coding, which was carried out by myself and the helper. As stated in section 3.3.1.1, the pre-interview and the post-interview data were coded and analyzed separately, right after the event. As the methods of coding and analyzing the two sets of data were the same, they are presented as a whole in this section. The approach we adopted was still a data-driven one. The process of coding entailed nine steps: coding sample data independently, raw categories being suggested, discussion, raw categories being modified and broad categories being formed, coding independently, assessment of reliability, managing discrepancy, final independent coding, and assessment of reliability. Actually, the steps of independent coding, establishing categories and assessing intercoder reliability were conducted several times in iteration. First, samples were randomly taken from data; each coder read through the transcripts intensively. Once a piece of useful information was spotted, it was highlighted and a possible category was proposed appearing in the margin of the electronic file. The suggested categories were termed raw categories because they needed further modification and generalization, and they were large in number. In the following stage of discussion, the two coders searched for patterns within the raw categories to form broader categories. Once broad categories were established, the two coders coded independently. When the independent coding was completed, the intercoder agreement was measured using Cohen's Kappa. Discrepancies were managed through

negotiation between the two coders. The reliability was measured again in Kappa after the final coding. The coefficients of agreement are reported in the Findings chapter. For the pre-interview, the preliminary coding resulted in 7 categories for question 1, 6 categories for question 2 and 2 others from coder 2, and 6 categories for question 1, 6 categories for question 2 and 3 others from myself. After negotiation, we eventually established 3 categories for question 1, 4 categories for question 2 and 2 categories of other information. For the post-interview, the final version of coding involved 4 categories for question 1, 3 categories for question 2 and 2 categories of other information. The inter-coder reliability was measured using Cohen's Kappa after the data coding which is reported in the results chapter.

Notably, a conflict was identified between the pre-questionnaire data and pre-interview data. Participants' responses to the question "Have the two terms "counterarguments" (opposing views) and "rebuttals" (responses to opposing views) been discussed in your English writing class?" were fairly positive (16%) in the pre-questionnaire. However, during the interviews, almost all student informants appeared to know little about the techniques of recognizing counterarguments and refuting them. In view of this, I assumed that most participants in the intervention had no previous training or practice in counterargumentation. Thus, in designing the intervention materials, I paid special attention to cater for the needs of student writers learning to counterargue.

3.3.3 Qualitative Document Analysis on the Syllabus, Prompts and Rubrics, and Textbook

As revealed in the Findings chapter, questionnaire and interview data have shown that the nationally approved Syllabus, nationwide high-stakes test writing prompts and rubrics, composition textbooks and classroom teaching instruction were deemed by students and teachers to have exerted considerable impact on students' argumentative writing. In this section, qualitative document analysis was used as instrument to answer the second half of Research Question 3, i.e., How are counterargumentation and critical thinking represented by these factors? What this research question concerns is how these factors may have influenced undergraduate English majors' performance in and perceptions of argumentative writing. To this end, analysis of relevant documents (the Syllabus), writing prompts and textbooks was conducted to collect data. Moreover, analysis of the Syllabus, writing prompts and textbooks, together with classroom observation, provided essential triangulation to check the results from the quasi-experiment and the exploratory study.

Qualitative document analysis, also referred to as ethnographic content analysis, is an emergent research method being increasingly used to explore the depth of text (Altheide, Coyle, DeVriese & Schneider, 2008; Bryman, 2012). A document may be defined as any symbolic representation that can be retrieved for analysis (Altheide et al., 2008). Documents used as sources can be personal or official (Bryman, 2012); in this research, official documents were analyzed. A qualitative approach focuses on describing and tracking discourse, including words, meaning and themes (Altheide et

al., 2008). In this study, qualitative document analysis was used to inquire into (1) the Syllabus, (2) high-stakes test writing prompts and rubrics; (3) composition textbooks. Since qualitative document analysis is qualitative content analysis by nature [6] (Bryman, 2012), in the rest of the thesis, content analysis is used to refer to this part of data analysis.

3.3.3.1 The Teaching Syllabus for English Majors

The Syllabus (NACFLT, 2000) was written by a commission authorized by Chinese Ministry of Education. It provides the most authoritative guidelines and policies for English teaching and testing for undergraduate English majors at mainland Chinese universities. As such, it was supposed to exert considerable influence on students' English learning, either through its direct impact on pedagogy or through the washback of assessment. Because this research project concerned the critical thinking in argumentative writing, relevant entries or paragraphs in the Syllabus became the focus of content analysis. It is believed that content analysis on documents could be revealing and valuable (Rapley, 2007). Qualitative document analysis was employed here to explore how the Syllabus stipulates critical-thinking-related instruction and English writing instruction.

The original version of the Syllabus, which is in Chinese, was used for analysis since there is no official English version. The text was read through, texts about "critical

---

[6] Qualitative document analysis is qualitative content analysis by nature. However, qualitative content analysis is a broader term than qualitative document analysis. Therefore, I chose to term the method as qualitative document analysis for particularity.

thinking" and "English writing" being highlighted for closer examination to develop systematic analysis of the data. The 12-page Syllabus comprises six sections, namely, General Objectives, Curricular Arrangement, Instructional requirements, Pedagogical Principles, Pedagogical Methods, and Assessment. As Schreier (2012) pointed out, segmentation in qualitative content analysis is important. As a procedure for segmenting the document, each of the six sections in the Syllabus was employed as a unit of analysis. In this section, extracts pertaining to the two themes are outlined and analysis of these contents will be reported in the Findings section.

The General Objectives section outlines the educational purpose for undergraduate English majors, with the cultivation of independent thinking and creativity being listed as equally important as knowledge acquisition. The Curricular Arrangement section classifies all the courses into three categories: courses on specialized skills, courses on specialized knowledge, and courses on specialty-related knowledge. Under this classification scheme, English writing courses are compulsory courses on specialized skills, which are advised to be taken in the second and third academic years by students (it is advised that an English writing course focusing on thesis writing be provided in the fourth academic year). The instructional requirements for English writing, as for other skills of the English language, are made at five levels which are detailed in section 4.2.1 of the Findings chapter.

The content directly related to critical thinking development of English majors appears in the section of Pedagogical Principles. The last of the five paragraphs in that section reads:

Students' thinking ability and creativity shall be cultivated. Through the instruction, we should train our students' abilities of analyzing, summarizing, synthesizing and abstract thinking. They should learn to approach an issue from multiple perspectives, and they should develop creativity through discovering and solving problems. We should maintain the balance between teaching language skills and cultivating thinking abilities and creativity since both are equally important pedagogical goals. (the Syllabus, NACFLT, 2000, p. 12)

It should be noted that "to approach an issue from multiple perspectives" is deemed a fundamental skill that should be trained in the English writing class. Some content in the Assessment section was also noted and interpreted. It reads:

To evaluate the implementation of this syllabus, the National Advisory Committee for Foreign Language Teaching, authorized by the Ministry of Education, organizes Test for English Majors Band 4 and Test for English Majors Band 8 at the fourth semester and Eighth semester for English majors during their undergraduate program. (the Syllabus, NACFLT, 2000, p. 14)

This paragraph may not appear directly relevant to critical thinking or English writing; however, it is highlighted and analyzed to facilitate other research stages in this study because it makes two points clear. Firstly, TEM4 & TEM8 are the only two official, nation-wide tests for undergraduate English majors in mainland China. Secondly,

TEM4 & TEM8 are expected to reflect the stipulation of the Syllabus. In other words, the two tests are criterion-referenced.

3.3.3.2 Analyzing Writing Prompts and Rubrics of Domestic High-Stakes English Tests for English Majors

As was stated in Section 3.3.3.1, TEM4 & TEM8 are the two national tests for English majors at Chinese universities. TEM4 is mandatory for all second year students in the English program; although not mandatory, TEM8 is taken by mostseniors. Since the two tests were officially launched 20 years ago, they have gained considerable social recognition in respect of both academic qualifications and the job market (Jin & Fan, 2011). The TEM4 & TEM8 results have been used as a benchmark for an increasing number of companies and institutions to recruit employees in Mainland China. For the above reasons, TEM4 & TEM8 were selected in order to evaluate the influence of writing prompts and rubrics of high-stakes English tests on English majors' perceptions on and performance in argumentative writing. To be specific, the goal instructions, rather than the object of enquiry, or the rhetorical function in the content of the prompts were analyzed at this stage of the study, while the rubric of TEM4 and the rubric of TEM8 underwent content analysis respectively as a complete document.

Content Analysis of Goal Instructions of TEM4 & TEM8 Writing Prompts

Goal instructions are the instructions in writing prompts which tell students what they

should accomplish in their writing tasks (Nussbaum & Kardash, 2005), and the task variable is one of the six vital variables that encompass a writing prompt (Kroll & Reid, 1994). Goal instructions in writing prompts have been the focus of some recent studies (e.g., Ferretti et al., 2009; Nussbaum & Kardash, 2005). In this study, the investigation into the goal instructions was prompted by the empirical studies of Ferretti et al. (2009) and Nussbaum and Kardash (2005).

Prompts of TEM4 & TEM8 from past years were collected[7]. Both TEM4 and TEM8 are given once a year (in April and March respectively). TEM8 writing prompts from 1997 to 2012 were gathered and analyzed. A preliminary survey on the 16 prompts revealed that they all served one sole rhetorical function: argumentation, and the goal instructions remained the same (see an example of TEM8 prompts in Appendix 5). Unlike TEM8, the writing section of TEM4 consists of two parts: essay writing and note (or notice) writing. Only the essay writing part, which involves genres of argumentation, exposition, and occasionally description, was within the scope of this investigation. Since this research concerned solely argumentative writing, only prompts of essay writing in argumentation were gathered. After reading through 17 essay writing prompts of TEM4 from 1996 to 2012, 13 prompts were found to have involved mainly argumentation. These prompts were then read again for their goal instructions. Similar to the TEM8 writing prompts, the goal instructions of the TEM4 writing prompts remained virtually unchanged (see an example of TEM4 prompts in Appendix 6).

---

[7] Both TEM4 & TEM8 underwent a major change in 2005. However, the change was in the increase of essay length. For instance, word requirements on TEM4 essays rose from 150 words to 200 words, and TEM8, 300 to 400. TEM4 has two writing tasks. In Study 2, only the task concerning essay writing was examined.

The goal instructions in TEM8 writing prompts read:

In the first part of your essay you should state clearly your main argument, and in the second part you should support your argument with appropriate details. In the last part you should bring what you have written to a natural conclusion or make a summary.

The goal instructions in TEM4 writing prompts read:

You are to write in three parts.

In the first part, state specifically what your opinion is.

In the second part, provide one or two reasons to support your opinion.

In the last part, bring what you have written to a natural conclusion or make a summary.

These goal instructions were specifically noted and analysed in respect of their capabilities of accommodating alternative views. The role they might play in developing students' counterargumentation skills and critical thinking abilities is illustrated in the Findings and Discussion chapters respectively.

Content Analysis of TEM4 & TEM8 Rubrics

TEM8 and TEM4 have their respective rubrics (Appendices 7 & 8) which were developed by the test commission and used as the criteria for marking schemes by

teachers. The total score for a TEM4 essay is 15 points while it is 20 for TEM8. However, both rubrics specify criteria at five different levels, with Band 1 the lowest and Band 5 the highest. The criteria are for evaluating essays from two aspects: task completion and essay structure (in particular, whether there are sufficient supporting details for the proposition); vocabulary, syntax, punctuation and expression. Every band of either rubric was carefully read for any trace of elements related to critical thinking and counterargumentation.

3.3.3.3 Typical Composition Textbooks

Textbooks provide essential materials for classroom teaching, thus playing a vital role in the teaching and learning process. Investigating the use of English writing textbooks helped shed light on understanding students' performance in argumentative writing in this study.

As You (2010b) has claimed, the writing pedagogy for English majors was uniform in terms of the use of textbook and the content of teaching in China before 1986. The chief reason for that may have been the scarcity of English composition textbooks (You, 2010b). For the past three decades, *A Handbook of English Writing* has been one of the most popular and widely used composition textbooks for English majors at Chinese universities (You, 2010b). However, it is unclear to what extent this text is still being used in Mainland China undergraduate writing courses. Accordingly, because there appears to be little discussion in the literature about this uniformity, I performed a preliminary survey on the use of English writing textbooks for

undergraduate English majors at different Chinese universities. Such a survey is challenging because the use of textbooks in university courses in China is not made public to those outside a given university. With the help of the vice dean of the faculty where I have worked, we sent messages to 42 available contacts[8] via QQ[9] and eventually from the 41 responses I received, I obtained information about more than a dozen universities. The use of English writing textbooks in these universities falls into the following six categories:

- No textbooks are designated by the school authority. Teachers prepare course materials by themselves. The English writing courses are taught mostly, but not necessarily, by native speakers of English. Xiamen University, Tianjing Foreign Studies University and Shandong University are in this category.

- The English writing course is combined with the reading course on a pedagogical reform basis. Hence the textbook is for both reading and writing and in many cases self-prepared by teachers. An example for this category is Ningbo University, Hangzhou Normal University.

- The textbook used is written by the staff and published by the university publishing house. Two cases in point are Nanjing University and Guangdong University of Foreign Studies.

- Using the traditional and the most popular textbook written by Prof. Ding and his colleagues. For example, Jiangxi Normal University, Central China Normal University, South China Normal University and Beijing Foreign Studies

---

[8] These contacts are the people in charge of the undergraduate programs of the English department at each university.

[9] QQ is an instant messaging software service developed by Tencent Holdings Limited and used extensively in mainland China. Statistics show that as of March 2013, there are 798 million active QQ accounts.

University.

- Taking advantage of the flourishing textbook market, and using the most recent textbooks. For example, Shanghai International Studies University, Suzhou University, and Beijing Foreign Studies University.

- Using English textbooks published and used in the US, UK, or other countries. Two examples for this category are Nanjing Normal University and Nanjing University of Finance and Economics.

Due to the preliminary nature of the survey, and the relatively small proportion of this survey in Study 1, I did not conduct further investigation into the particular textbook use of the universities listed above. Instead, I focused on and conducted a content analysis of *A handbook of English Writing* (Ding et al., 1994) because it is the textbook employed at the research site and because this textbook may have exerted the greatest influence on students' writing with its wide use during the past 30 years.

This textbook has both English and Chinese versions. Although it is beyond my capacity to obtain the number of universities or people who use the textbook, we can get some idea of how widely it is used by its printings ever since it was first published in 1984. The second edition, which underwent major changes and was issued in 1994, is the edition used at the research site[10]. In the preface of its third edition published in 2009, the printing of the first and second editions of the book was reported to be more than a million copies.

---

[10] This edition is more widely used than the third edition to the thesis writer's knowledge. As of April 2013, it had been printed 69 times.

As can be seen from the Table of Contents of Ding et al.'s (1994) textbook (See the Table of Contents of the textbook in Appendix 4), the first four parts display a hierarchy of linguistic layers, starting from words and ending with passages. In a book depicting the history of English composition in China, You (2010b) commented that such structure reflected traces of current-traditional rhetoric which refers to the textbook-based methods of composition instruction. It was under the influence of current-traditional rhetoric that discourse is classified into four modes: description, narration, exposition and argumentation. In addition, it is Ding et al.'s (1994) point of view that exposition is to a great extent similar to argumentation. For the purposes of this research, the argumentation section was singled out for content analysis.

The section of argumentation comprises three parts: main discussion on written argumentation; models essays; exercises. An interesting first-sight observation was how brief this section was. A quick count told me that the main discussion was around 1050 words. The authors begin by defining the function of an argumentative essay as "to make the reader agree with its point of view and support it, to persuade him to change his mind or behavior, and to approve a policy or a course of action that it proposes" (Ding et al., 1994, p. 234). In the following, the authors delineate six requirements when writing argumentative essays: (1) a debatable point; (2) sufficient evidence; (3) good logic; (4) clear logic; (5) good use of the other three types of writing; (6) an honest and friendly attitude. Under *a debatable point*, it was advised that the topic for an argumentative essay should be "something which can be viewed from more than one angle and is therefore open to dispute" (Ding et al., 1994, p. 234). Under *good logic*, two ways of reasoning - inductive and deductive reasoning - were

discussed briefly. Under *clear logic*, it reads:

> *A typical argumentation essay consists of three parts: an introduction*
> *which identifies the issue to be discussed and explains the importance of*
> *such a discussion; a body which presents the evidence; and a conclusion*
> *in which the proposition, if it is stated at the beginning, is reaffirmed.*
> *In the body, it is advisable that you devote the first one or two paragraphs*
> *to a consideration of the other side of the case before stating and*
> *amplifying your own view. Then in each paragraph that follows, you add*
> *something new and important to your argument. (Ding et al., 1994, p.237)*

Noticeably, the other side of the case is mentioned. However, as the main concern here is the structure of an essay, it gives the impression that considering the other side is merely a strategy of setting a target for a writer to aim at, rather than to encourage critical thinking. The key terms such as counterarguments and rebuttals do not appear in the text. Moreover, the text fails to provide explicit instruction in counterargumentation skills despite the fact that counterargumentation is essential to enhance persuasiveness of argumentative essays and a hallmark of critical thinking. This will be discussed in greater detail in the Findings chapter.

3.3.4 Classroom observation

Classroom observation can help illuminate even the most familiar events (Wragg, 2011). In order to find out how written argumentation is taught in the composition

classroom and to answer Research Question 3, I conducted classroom observation at Jiangxi Normal University. I obtained permission to sit in two classes of second year English majors (excluding the two experimental classes and the two control classes in the intervention) which were taught by two different teachers. The observation took place from Oct. 12 to Nov. 23 in 2013; it was done at this time because I was told by the two instructors that they were to teach argumentation as a type of writing during this period. During the 6 weeks, I observed 11 teaching periods[11] (each period contains two lessons of 50 minutes each) using a classroom observation form (Appendix 9) focusing on the instruction in argumentation and counterargumentation. The two instructors are labelled as Instructor A and B (see details of the instructors in Table 3.6).

Table 3.6 Instructors involved in the classroom observation

	Gender	Age	Academic degree	Teaching experience	Years of teaching writing
Instructor A	female	35	MA	9 years	7
Instructor B	male	32	MA	6 years	6

Structured, quantitative observation schemes are the standardized data collection method associated with classroom research (Dornyei, 2007). However, as Allright and Bailey (1991) argued, closed techniques may easily miss insights provided by the language classroom. In this research, an open-ended classroom observation protocol was constructed highlighting four points: teaching objectives; pedagogical activities; argumentation skills discussed in class; use of materials. The advantage of an

---

[11] There should have been 12 teaching periods by the two instructors. However, in one case, the teaching period was rescheduled by one of the instructors and I failed to sit in that period.

open-ended protocol is that it allowed more detailed information to be written down in larger quantity. For most of the time, I took notes sitting at the back of the classroom, except for occasional chats with the instructors or students during the breaks. Upon the completion of each observation, an electronic file was established and named "Classroom Observation No.1", etc. After all the observation sessions, the content in the 11 files was copied and pasted into four new files, each for a different point of observation, for later analysis. The results of the analysis are reported in the next chapter.

3.4 The Second Study

3.4.1 The Instructional Intervention

3.4.1.1 The Participants and Instructor

Participants were 125 second year English majors from four intact classes, of which two were experimental groups, and two were control groups. Each class had around 31 students aged from 18 to 20, with five-sixths being female students. These students admitted to the English language discipline generally scored slightly higher in English in the national matriculation test, with an average score of around 120 (the maximum being 150). The 12-week instructional intervention took place in the first semester of the second academic year for the English majors in their English Writing II class. The timing was determined by the following three factors. 1) After the students had completed the English Writing I course, they had a foundation in English

writing; hence, they were ready to take more intensive training at tertiary-level argumentative writing. 2) Normally, the curricular focus of English Writing II was argumentative writing. As such, the pedagogical experiment was aligned with the regular teaching arrangement. In other words, the experimental classes and the control classes were synchronized in the curriculum. 3) Since these students had to prepare for the TEM-4 test that immediately followed the semester, the writing task of this test, which is generally an argumentative essay, provided the students with intrinsic motivation for learning and practicing argumentative writing skills. The writing class met for two hours once a week.

One teacher, Angel (a pseudonym), served as the instructor for both the experimental and control classes. Angel, who had nine years of experience in teaching English writing, was also the regular teacher of the English writing course in these four classes. This research was a quasi-experimental one, so the sampling was basically nonrandom. The selection of instructor symbolized the quasi-experimental nature of the study. Angel was the first teacher to respond to a recruitment notice to English Writing II teachers and displayed enthusiasm for the project; thus, she was chosen out of a total of four teachers who responded to the notice. But, in order to maximize the representativeness of the sample, Angel and I assigned her four classes to either the experimental groups or control groups. Before the application of the intervention, there were training sessions during which Angel was familiarized with the content of the research until she was fully qualified for the instruction. This training was necessary because during her eight-year experience of teaching English writing, Angel had not discussed counterargumentation skills in her class. The training lasted

over two months during which I had regular meetings with Angel. She was first provided with fundamental readings in the Toulmin (1958) model, informal logic and counterargumentation, her questions being answered in detail by myself. She was then trained in how to use the materials developed by myself for use in the experimental classes. In addition, she was advised to teach the control classes normally, with the same materials used in other English writing classes. One of the primary concerns as well as ethical procedures in this project was that the ordinary curricular arrangements would not be disturbed. The knowledge and skills that students were supposed to acquire would not be reduced.

In this study as well as in many of the argumentative writing studies where instructional interventions have been applied (e.g., Nussbaum & Schraw, 2007; Wolfe et al., 2009), the researchers avoid being the instructors themselves in order to eliminate researcher bias. Researcher bias should be more exactly termed "experimenter expectancy" which refers to the researcher's subtle manipulation of conditions favoring expected findings (Suter, 2012). That is, unconsciously the researcher may pay more attention to the experimental groups, which will eventually affect the outcome of the experiment and damage the internal validity of the research. In this intervention study, Angel had no direct interest in the research; in other words, what she cared about was providing high quality teaching. Accordingly the experimenter expectancy problem was less a concern. In addition, the design of using one instructor for both experimental and control group neutralized any concerns about manipulating independent variables, e.g., teaching methods or teaching experience. To be specific, if two instructors were used, individual instructor's teaching methods,

teaching experience, even gender or age can affect the outcome of the experiment. Methodological theorists (e.g., Fraenkel, Wallen, & Hyun, 2012; Creswell, 2009) suggest "holding certain variables constant" as a means of controlling extraneous variables in order to minimize internal validity threats. Using one instructor for both control and experimental groups could effectively control for the variables of teacher's teaching experience, gender, age and personality. In other words, these variables were held constant.

3.4.1.2 The Writing Prompts and Rubric for the Pretest and Posttest

The pretest and posttest prompts (Appendix 1) were modeled after the TEM8 with regard to the timing (50 minutes for both the pretest and posttest) and argumentative nature of the topic. However, two changes were included to fit the research design. First, unlike the TEM8, there were no goal instructions regarding the structure of the essay. Rather, students were free to construct the most persuasive essays within their capacity. If the goal instructions were left in, this might imply that counterargumentation was unnecessary. Second, there was no word limit on the essay to provide students with maximum opportunity to expand their argumentation within the time allowed.

Two scorers (excluding myself) blindly scored the essays using the TEM8 writing rubric (see Appendix 7). The rationale for using this rubric was to investigate whether the inclusion of counterargumentation in the responses alone would result in a higher score, even though this rubric also considered elements such as organization,

vocabulary and grammar. To this end, the two raters were not informed of the nature of the intervention before the rating. As teachers of English writing, they had at least six years of essay marking experience using a similar rubric. When the formal scoring began, scripts from both experimental and control groups were given an opaque code and then shuffled into one stack to be distributed to each scorer. After the scoring was completed, the interrater reliability was calculated.

3.4.1.3 Procedures of the Intervention

The control group received typical classroom instructions in argumentative writing which encompassed three main stages. In the first stage, students were introduced to the structure of an argumentative essay: the introduction, the body, and the conclusion. It was stressed that the writer's claim should appear at the beginning, and in the body part this claim should be supported with sufficient evidence. The teacher provided sample essays in addition to using the models in the textbook. In the second stage, students were taught how to write each part of an argumentative essay. For instance, they learned nine ways of beginning an argumentative essay. They were also told to memorize a glossary of useful words and sentence patterns for argumentative writing. They were required to practice these skills both in class and after class. In the third stage, the teacher taught the students how to apply these skills in response to the writing prompts in TEM4 and TEM8. Notably, there was no mention of counterargumentation throughout the instructions.

The experimental group received additional instruction in counterargumentation apart

from the typical instructions described above. A teaching schedule (Appendix 10) for the 12-week intervention was prudently made in advance on which the teaching for the experimental group was based. At the beginning of the intervention, students in the experimental classes received instruction on the elements of persuasive argumentation. The instruction was carried out in a step-by-step manner: 1) model essays and visual aids; 2) brainstorming and planning (8 topics); 3) joint writing; and 4) independent writing. During the first phase, the instructor demonstrated why and how argumentation comprising claim, data, counterarguments and rebuttals meets the standards of good argumentation. The instructor then provided model essays (adapted from English newspapers and composition textbooks or written by the researchers) for students to analyze. Using a diagram as a visual aid (Appendix 11), as in Nussbaum & Schraw (2007), the instructor discussed and worked with students to identify the four argumentative elements. One focus of the instruction was on why each of these elements is essential for effective persuasion, and how these argumentative elements help the author achieve the desired purpose of the essay.

Analyzing model essays helped students understand and familiarize themselves with counterargumentation and refutation. In the second phase, the instructor collected a number of written topics on current and controversial issues and used them in class. The students were encouraged to brainstorm and come up with ideas on these topics. Specifically, the instructor listed 12 topics (see Appendix 12) and asked students to choose eight that interested them. These eight topics were used to train students in arguing and counterarguing skills. For each of the topics, students were required to work in groups of four or five, state their positions to group members and produce as

many arguments, counterarguments and rebuttals as possible. Since individual students held varied, or opposing views towards a controversial topic, group members worked together to generate their own data, counterarguments and rebuttals. The instructor also participated in the group work. After group discussions, a presenter from each group reported their claims, supporting data, counterarguments and rebuttals. The instructor wrote down the main argumentative elements of each group in an electronic file that was sent to each student for later use.

For the first three topics, joint writing followed the brainstorming and planning. In this way, the instructor and the students experienced the construction of a whole argumentative essay together. During this process, negotiation of meaning, construction of arguments and decisions on structures of an essay took place between the instructor and students. Joint writing scaffolded the students' argumentative writing in terms of counterarguments and rebuttals. After students grasped the essential elements of argumentation, they were required to write independently on the other topics, while referring to the electronic file resulting from the brainstorming and planning phase.

3.4.1.4 Scoring and Coding of the Pretest and Posttest Scripts

Another coder and I then conducted a text analysis of the pretest and posttest essays. Four argumentative elements (claim, data, counterargument and rebuttal) adapted from Toulmin's (1958) model and developed in empirical studies (e.g., Bacha, 2010; Nussbaum & Schraw, 2007; Qin & Karabacak, 2010; Stapleton, 2001) were identified

and coded using different colors. Specifically, claims were highlighted in red, data in green, counterarguments in blue and rebuttals in purple (see an example of a coded essay in Appendix 13). Definitions and examples of the four argumentative elements are given in Appendix 14. A frequency count was generated after coding each script. When discrepancies of greater than one were noted in the frequency counts in any one category of the four elements in a script, the two coders revisited the script and resolved the differences through discussion. When there were differences of one, the two scores were averaged. The inter-rater reliability was calculated in terms of Cohen's (1960) Kappa, which is suitable for nominal scale correlations in reducing the risk of chance agreement (Hayes & Hatch, 1999). The correlation between the uses of Toulmin-like elements and overall quality of argumentative essays was analyzed using Pearson's correlation test.

In the process of coding, indicator words and semantic structures and prepositional phrases were noted. For example, "I think…," and "as far as I am concerned," usually indicated a claim. "First of all," "then," "furthermore" and "finally" suggested that a sequence of arguments was forthcoming. Indicators of counterarguments included "Some people may argue that…," and "Other people may hold different views," while "however" and "I totally disagree with them" indicated a rebuttal. Naturally, other less transparent indicators were also present and had to be scored carefully, often requiring discussion between scorers when disagreement arose.

During the process of coding the four argumentative elements in the pretest and posttest essays of both groups, some challenges emerged. Although the "claim" is the

foremost argumentative element outlining the writer's standpoint on an issue, we found that in approximately 10% of pretest essays of both groups, the writer held a non-commital view to the topic question. This was coded as "zero claim." For instance, one student wrote "Every coin has two sides." "On one hand, urbanization does harm to our life." "On the other, urbanization promotes economy and benefit[s] us." And his/her conclusion was "We can believe that urbanization will make our life better in the future." This type of response became clearer in the posttest essays of both groups, suggesting that either the argumentative writing instruction with or without mentioning counterargumentation was effective in helping students produce stronger claims.

Some previous studies (e.g., Qin & Karabacak, 2010; Stapleton, 2001) coded counterargument claims and counteragument data as separate argumentative elements. In this study, however, we did not distinguish between them. It was noted that in some cases, there were both counterargument claims and data in one sentence (e.g., "some people hold the idea that traditional Chinese culture is out of date and cannot keep the [sic] pace with modern society for there were so many flaws of [C]onfucianism.") In this sentence, "data" was coded as the section starting with the word "cannot" because what follows until the end of the sentence is support for the counterclaim in the first part of the sentence. In other cases, counterargument claims and data were in separate sentences. In a couple of cases, counterargument claims stood alone without any counterargument data (e.g., "some people don't think it beneficial for young people to make more efforts to preserve traditional Chinese culture"). In this study, counterarguments were coded regardless of whether

counterargument claims and data were in one sentence, in separate sentences, or the data were missing entirely. The decision was made for the following reasons: first, the focus of this study is whether participants were aware of alternative views other than their own views, and whether they can effectively refute those alternative views; second, when analyzing emergent writers' counterargumentation we deemed it unnecessary to distinguish between counterargument claims and counterargument data.

One of the difficulties in coding counterarguments was to differentiate counterarguments and opinions from reservations. For instance, one student's claim was "Traditional Chinese culture needs our preservation." However, in one of the following paragraphs, the author wrote, "But it not [sic] saying that we should preserve all things of traditional Chinese culture." In such cases, this was not coded as a counterargument, but regarded as the author's own opinion with reservation or compromise.

Rebuttals were coded in a similar way; however, it should be noted that some rebuttals did not effectively refute the counterargument, or were not aligned in terms of content with the counterargument. Such rebuttals were not counted. In an essay that argued for preserving traditional Chinese culture, a student wrote as a counterargument, "Traditional culture is regarded as useless by some young people." The following sentence, "Do you agree with them?" was not coded as an effective rebuttal because it did not refute the counterargument. In another case, in response to a counterargument, "but some people was [sic] afraid that traditional culture would

harm the economic development," the writer merely added, "I think they must read more history to know the charming tradition." Both coders determined that the second sentence (presumably the rebuttal) was not logically in line with the first sentence. Therefore, no rebuttal was coded in this case.

3.4.2 Assessing Critical Thinking in Argumentative Writing

This section describes the method used to answer Research Question 7, i.e., how does the intervention impact students' critical thinking? For this purpose, participants' essays from the pretest and posttest were measured again using a critical thinking rubric (Appendix 3) to see the variation in critical thinking development of both the experimental group and the control group before and after the intervention.

3.4.2.1 The Rubric for Assessing Critical Thinking in Essays

The literature reveals many critical thinking rubrics, but few rubrics associate critical thinking with argumentation. The critical thinking rubric (Appendix 3) in this project was generated by myself following Stapleton (2001) and WSU CT Project (2009). Built on previous empirical studies (e.g., Crammond, 1998), Stapleton (2001) outlined four key critical thinking elements: argument, evidence, recognition of opposition, and fallacies for raters to identify in the texts of second language learners' writing. Stapleton's (2001) study was significant in having established a feasible critical thinking measurement for written argumentation, which has been referred to considerably. However, as Stapleton (2001) pointed out, it proved to be murky for the

two raters in that study to reach agreement on the number of each element, and even the classification of an element.

The Guide to Rating Critical & Integrative Thinking (hereinafter WSU CT rubric) developed by the WSU CT Project (2009) addressed this problem by incorporating the key critical thinking elements in seven items. The seven items are: problem at issue, writer's own perspective or position, other salient perspectives or positions, key assumptions, quality of supporting data, consideration of context and audience, and conclusion. Each item is claimed to identify a prominent area of critical thinking based on scholarly works of Toulmin (1958), Paul (1993) and Facione (1990). In each item, students' performance in critical thinking is depicted at seven different levels: absent, minimal, emerging, developing, competent, effective and mastering. Furthermore, the performance at various levels is scored ranging from 0 to 6 points. With the WSU CT rubric, the evaluation of students' critical thinking in writing might become more operational because students' critical thinking ability was quantified with a relatively reasonable scoring guide. The 2009 WSU CT rubric has been refined through the 1997, 2001, 2004 and 2006 versions, and has been used to measure the critical thinking development of students from multiple disciplines at the university and beyond.

The rubric developed for this study was to a large extent based on Stapleton (2001) and the WSU CT rubric (2009). However, modifications were made in this rubric in two respects. First, the "fallacies" element in Stapleton (2001) was not included. "Fallacies", in essence, depicts various kinds of errors in reasoning (Stapleton, 2001).

In the critical thinking measurement of this study, "fallacies" were not identified as a separate item in the rubric but dealt with in another way, i.e., if a reason did not support a conclusion, it was deemed invalid. As such, the concept of "fallacies" actually permeated every item of this rubric. The concept was not only used in Item Three of the rubric which identifies appropriate supporting data, but could also be used in other items. For instance, Item 4 identifies "other salient perspectives and writer's refutation"; if any fallacies existed, then the essay would score low on this item. Second, the item of "key assumptions" in the WSU CT rubric was not considered in this rubric because of the obscure nature of that element. The concept of assumptions has evolved from "warrants" in the Toulmin model that are often assumed or unstated positions. This "layer of complexity" is often "unnecessary for analytical purposes" (Nussbaum & Kardash, 2005, p.70). Therefore, in a number of studies on students' argumentative writing, "warrants" or "assumptions" were excluded to enable more reliable classification of argumentative elements (e.g., Nussbaum & Kardash, 2005; Nussbaum & Schraw, 2007; Qin & Karabacek, 2010; Stapleton, 2001; Varghese & Abraham, 1998). Likewise, it would pose difficulty and cause confusion to analyse "assumptions" as a critical thinking element in this study. For the afore-mentioned reasons, "assumptions" were not listed in this critical thinking rubric.

3.4.2.2 The Scoring and Analyzing of Pretest and Posttest Scripts for Critical Thinking

The scoring and analyzing was done by myself and a second rater with appropriate

qualifications (a master's degree in English language education and six years of teaching experience). Each item of the rubric evaluates one essential element of students' critical thinking ability at four levels, with 1 point indicating the lowest level and 4 points the highest, the maximum being 24 points. Before the scoring started, we familiarized ourselves with the rubric and had detailed discussions about each of the six items. These discussions were helpful in clarifying definitions and meanings of the terms used in the rubric and we exchanged ideas as to how we interpreted the scoring criteria. In fact, these discussions led to modifications of the original rubric resulting in its final version. Then we "trial ran" 10 scripts. For each script, we noted down the points for a particular item and the overall CT score. At the end of the trial assessment, each rater explained how the scoring decision was made based on the criteria given in the rubric. We then discussed further the correct use of the rubric till satisfactory agreement was reached. After that, the two raters did the scoring independently. When the scoring was completed, the inter-rater reliability was calculated using Pearson correlation coefficient, resulting in an alpha of .71. It was found that discrepancy of over 3 points in the overall score existed in 23 scripts. These scripts were reviewed until disagreement was reduced via discussion. After negotiation, the CT score of the pretest, posttest scripts of both experimental and control groups was then entered into SPSS for an ANCOVA test to see whether there was any difference in critical thinking development between the two groups. The results are reported in the next chapter.

3.5 The Third Study – Exploratory Research

The third study was a logical extension of the second study for two reasons. On the one hand, the second study investigated the impact of high-stakes test writing prompts on undergraduates' critical thinking performance. However, writing prompts were not used as an independent variable in the quasi-experiment of the first study. I realized that a third study was necessary because writing prompts was a field worth exploring, especially in regard to critical thinking cultivation. On the other hand, the second study was designed to reveal the effect of critical-thinking-embedded classroom instruction on students' critical thinking ability in writing, which emphasized the pedagogical side of the critical thinking problem. I thought it would be interesting to further the research by looking at the assessment side of the problem. Therefore, the third study focused on the influence of writing prompts on students' critical thinking ability.

3.5.1 Phase One of the Third Study

Phase one of the third study aimed to answer RQ 8, i.e., What patterns do the writing prompts in three high-stakes tests, IELTS, TOEFL and TEM4, display in terms of their rhetorical function and object of enquiry?

In this study, three tests were chosen because of their high-stakes nature and the large numbers of test takers: IELTS, TOEFL and TEM4. All of these tests have sections testing the four basic skills of reading, listening, speaking and writing, although TEM4 assesses speaking separately in a delayed and optional manner. In 2011, 1.7

million people took the IELTS in 130 countries (IELTS Annual Review, 2011). ETS, the parent company of TOEFL, claims to have tested 27m students in total and has locations in 165 countries (ETS, 2013). In contrast, the TEM4 is confined to Chinese test takers and is written by about 270,000 students a year.[12] It should be noted that this Chinese test was chosen, despite its lack of international exposure, because the participants in the script-generating study come from China. In addition to the high-stakes nature and large number of participants that these three tests have, all three contain two writing tasks: a section with prompts that require an essay-like response, and a section with prompts that require a shorter and non-essay-like response. In the present study, more focus is put on the essay prompt. However, in order to make our investigation more comprehensive, we also surveyed the non-essay-like prompts.

In the IELTS writing section, Task 1 requires test-takers to describe a diagram or table within about 150 words. The essay prompt is Task 2 of the Academic test aimed at those about to undertake university studies. The prompt requires a 250-word

> *response to a point of view, argument or problem… [and] candidates are*
> *assessed on their ability to present a solution to a problem; to present and*
> *justify an opinion; to compare and contrast evidence, opinions and*
> *implications; to evaluate and challenge ideas, evidence or an argument.*
> *(IELTS, n.d., p.4)*

---

[12] An English test called CET (College English Test) in China includes a much greater number of participants (9.5 million annually (Cheng, 2008)); however, since the writing prompts are in Chinese and the expected length of responses are short, this test was not selected.

In Task 1 of the TOEFL, test-takers need to summarize opposing views on a topic provided either in written or audio form in 150-225 words. The essay prompt, like the IELTS, is also the second writing task. Test takers

> *write an essay in response to a question that asks you to express and support your opinion about a topic or issue. Your essay will be scored on the quality of your writing. This includes the development of your ideas, the organization of your essay, and the quality and accuracy of the language you use to express your ideas.*

As for TEM4, the writing section similarly consists of two tasks. The section concerning essay writing is "Section A Composition," in which test-takers are required to write an argumentative essay addressing the issue in the prompt. "Section B Note-writing" of the TEM4 asks test-takers to write a note of 60-80 words.

In order to answer RQ8, six sets of 20 (n= 120) recent writing prompts were gathered randomly from the past tests of each of the three tests. IELTS prompts were extracted from the official websites. However, since there were no past papers housed at official TOEFL and TEM4 sites, prompts were gathered from different secondary sources, such as TOEFL textbooks and unofficial sites on the Internet. I then categorized the prompts according to their rhetorical function and object of enquiry following Moore and Morton (2005). A rigorous process of double blind coding was deemed unnecessary because indicator words and terms in the prompts (noted below) pointed to clear distinctions between categories. For example, for evaluation, I looked

for keywords in the goal instructions, such as "To what extent do you agree with these views" or "How far do you agree or disagree with the above views," which require candidates to evaluate and choose between different ideas. For hortation, I looked for keywords related to the urging of an action, i.e., "should," "need" or "it is better to…" Examples of hortation included "Governments should make more effort to promote alternative sources of energy," or "Working from home should be encouraged as it is good for workers and employers." Similarly, I coded the prompts for object of enquiry whose content categories were equally transparent as will be apparent in the next chapter.

3.5.2 Phase Two of the Third Study

Phase two of the third study addressed RQ 9, i.e., In what ways do the essays of students responding to a "conventional" writing prompt differ from those of an experimental, non-conventional prompt? An exploratory design with control group was adopted to collect data. Scripts from both experimental and control groups were converted into electronic data, and then coded and analysed for linguistic features.

3.5.2.1 Participants

One hundred and twenty-nine first year undergraduate students ranging in age from 18 to 21 participated in the study. They were all English majors studying at the same provincial public university in Mainland China (but different from the students in the first and second studies). At the time of the study, they were taking English writing I

course that met once a week for two hours, the curricular focus of the writing course being sentence and paragraph writing. The rationale of using first-year students in Study 3 was that they had not yet received tertiary-level instruction in written argumentation, hence their responses to different prompts were supposed to be more related to the prompts and less associated with the impact from class instruction. While they were from different parts of the country, they all met the admission standards of the program, which required a minimum grade of 120 on a matriculation English exam. Given their similarity in educational background and English proficiency, they could be regarded as a homogeneous group. As these students were randomly assigned to two classes, one class was designated the control group and the other the experimental.

3.5.2.2 Procedures

For RQ9, two groups of students were required, one control and one experimental, along with distinct writing prompts for each. The control group (n=64) was assigned the following prompt specially developed for this study to be in line with the prompts from the writing sections of the three tests described above. It is argumentative, evaluative and with a goal instruction similar to those found in most of the 60 samples of essay writing prompts.

Prompt for the control group

China has recently experienced a great expansion of the population and

size of its cities. This social process is called urbanization. A recent study

by the Asian Development Bank and the National Development and

Reform Commission estimates that cities in China will grow by about 15

million people each year and by a total of 230 million over the next 15

years. Do you think urbanization is making people's lives better?

You will be allowed 35 minutes. Please write on the answer sheet a

composition of about 200 words on the following topic:

Is urbanization making our life better?

The prompt for the experimental group (n=65) was informed by the field of behavioral economics. Briefly, behavioral economics focuses on areas of human behavior that help explain the irrationality of economic decisions that standard economics fails to explain. In other words, it explores other domains, such as emotion and shortsightedness that tend to play a significant role in decision-making (Ariely, 2008). I chose this field because of its rich focus on both intuition and logic, as well as its potential for creative and critical thinking. Here, one could argue that economics is a specialized field unsuitable for a writing prompt; however, the same could be said for socio-cultural prompts categorized above that stray into sociology, technology and education, equally specialized fields. In fact, behavioral economics explores questions about everyday human behavior and decision-making that cross disciplines.

The experimental prompt was devised so that it would meet the conditions recommended in the literature for prompt content, namely that it be contextualized, authentic, and accessible (Kroll & Reid, 1994). However, more importantly, I also

endeavored to create a context that would stimulate original and critical thought, unencumbered by concerns over the participants' uneven schema. Additionally, the prompt had to avoid the evaluative and hortative rhetorical functions customarily found in writing prompts. In essence, the chosen prompt attempted to move beyond the "take a stance on a socio-cultural issue" type of task to new ground that required more than simply giving reasons to support a point. In effect, in devising a new kind of prompt, the goal was to explore the type of language and organization the participants used to see if it differed from the control group. An equally important goal was to investigate the type of problem solving and connection-making that participants would generate that may be absent in "take a stand" essays because of concerns about prescriptive responses (see below) or hackneyed arguments in the public domain.

Prompt for the experimental group

A day care center in Shanghai had a problem. Although parents were supposed to pick up their children each day by 4 PM, often they would arrive late. This meant that some children and teachers had to wait, causing stress for both. In order to try to solve this problem, two economists offered their help. They created a plan that charged parents a 20 RMB penalty each time they arrived later than ten minutes after 4 o'clock. This penalty was added to the parents' monthly fee of 1500 RMB. Soon after the new system started, however, a strange thing happened. The number of late pick-ups actually increased. In the end, the number of parents arriving late more than doubled. Instead of improving the

situation, the economists' plan made things worse.

Explain why you think the plan had the reverse effect and how you would change it. You will be allowed 35 minutes. Please write on the answer sheet a composition of about 200 words.[13]

3.5.2.3 Data analysis

The scripts of the writing test were collected from both the experimental and control groups and each was assigned a code. All the data were then converted to electronic text to facilitate data analysis. In view of the exploratory and open-ended nature of the RQ9, I adopted a data-driven approach with an aim to trace meaningful differences in both sets of responses to the two prompts. Data analysis comprised two stages.

In the first stage, preliminary discourse analysis was performed via an automated text analyzer for general linguistic features of the data. The purpose of the preliminary discourse analysis was to obtain fundamental indices of the length and overall quality of essays written by the experimental and control groups. To this end, a free online text analyzing program accessed via the website www.usingenglish.com was employed. All the scripts in the experimental group were copied and pasted into one text file. Since the total number of words exceeded 10,000 in this file, a program

---

[13] This prompt was based on, and modified from a study in Haifa, Israel, appearing in the following paper: Gneezy, U., & Rustichini, A. (2000). A fine is a price. *Journal of Legal Studies, 29*, 1-17.

entitled "Advanced Text Analyzer" was selected to perform the analysis. The text was then entered and analyzed. The control group essays were dealt with in the same way.

In a brief version of the output of the automated discourse analysis, three indices were listed, namely, "Length", "Lexical Density" and "Fog Index". The average length of essays was calculated by dividing the words by the number of essays. Both Lexical Density and (Gunning) Fog Index are readability tests. Lexical Density measures the proportion of the content (lexical) words to the grammatical (functional) words. The Fog Index estimates the number of years of education that readers need to understand a given text. It is measured by calculating the average number of words per sentence and the number of difficult words and generating an index from these measures.

After the preliminary discourse analysis, further discourse analysis was conducted for syntactic complexity, lexical diversity and word frequency which are considered as the most predictive linguistic features of argumentative writing quality. Following McNamara, Crossley and McCarthy (2010), and using the computer program Coh-Metrix 3.0, the syntactic complexity was measured by the number of words before the main verb; lexical diversity was measured by the Measure of Textual Lexical Diversity (MTLD); word frequency was measure by Celex, log frequency of for all words. In this study, length, lexical density and Fog Index, together with syntactic complexity, lexical diversity and word frequency were utilized as six indicators of students' performance in argumentative writing. The outputs of discourse analyses for the experimental and control groups were compared for resemblance and variation in students' responses to the two sets of writing prompts.

In the second stage, qualitative and corpus analyses were performed focusing on: 1) metadiscourse, 2) organization, and 3) lexical items. Following Crismore, Markkanen, and Steffensen (1993) and Hyland (1998), two categories of metadiscourse were coded: textual metadiscourse and interpersonal metadiscourse. Under the textual metadiscourse category, I focused on three subcategories: transition markers, frame markers and code glosses. Under the interpersonal categories, two subcategories were coded: hedges and attitude markers. These five subcategories were chosen because I was informed by the initial coding that these were the most frequent in L2 academic writing. Table 3 illustrates the coding scheme for metadiscourse use in both the experimental and control group essays. I coded the metadiscourse in all the scripts (N=129). A second coder, with a master's degree in applied linguistics and having an appropriate research background, coded 15% of the scripts. The intercoder reliability, calculated with Cohen's Kappa, was .93, .92, .96, .89, and .91 for transition markers, frame markers, code glosses, hedges and attitude markers respectively. Disagreements were then resolved through negotiation.

Table 3.7

Coding scheme for metadiscourse use in both experimental and control group essays

Category	Function	Examples
Textual metadiscourse		
Transition markers	express semantic relation between main clauses	in addition / but / therefore / thus / and
Frame markers	explicitly refer to discourse acts or text stages	firstly/secondly/ finally / all in all/ to conclude
Code glosses	help readers grasp meanings of ideational	namely / e.g. / in other words / such as

		material
Interpersonal metadiscourse		
Hedges	withhold writer's full commitment to statements	might / perhaps / it is possible / about
Attitude markers	express writer's attitude to propositional content	in my opinion/ as far as I am concerned/I agree

The organizational structure of the essays was analyzed by noting the layout of paragraphs in each script, and reading the first three to five sentences of the first paragraph and comparing the two groups. The reason for reading the first three to five sentences of the first paragraph was based on the fact that students rarely used more than five sentences to introduce a topic. The reading enabled me to find out if there was any difference in how participants began their essays, in particular, whether they used a lead-in. Customarily, a writer begins an argumentative essay with a preamble of a few sentences to introduce the topic, which is termed a "lead-in." In this study, I looked at the first three to five sentences of each essay to code this lead-in and tried to find patterns within the essays of the two groups. In view of the plain nature of the coding, I considered it unnecessary for a second coder to participate in the whole coding process. Instead, the second coder checked 15% of the coding upon completion. The agreement between the two coders was 93%.

The purpose of noting the layout of paragraphs was to examine how participants arrange their argumentative essays. In particular, one effort was to search for evidence as to whether participants of the two groups followed a three-part argumentation model in their essays, namely, an introduction-body-conclusion model, since this model is the most commonplace structure adopted in argumentative essays. Another

effort was to code where and how the writer's claim occurs in an essay. For example, conventionally the writer's claim was presented in the introduction, and this claim was supported by evidence in the body; in the conclusion, the claim was restated to a greater or lesser extent. For this part of the coding, the second coder also coded 15% of the scripts, and the intercoder agreement by percentage was .87.

After coding and analyzing the organizational structure of the scripts, a corpus analysis using the AntConc software (Anthony, 2012) was performed. The "Top 100 lexical words" and "Top 100 functional words" were generated for both groups. Leaving aside those words that are directly related to the subject of the two writing prompts, e.g., "daycare" and "urbanization", differences and similarities of the words in the lists were noted for the two groups. These differences or similarities were decided by the frequency with which a word was used rather than the rank of the word in the corpus analysis list. For instance, the word "think" ranked 10th and 12th in the experimental and control groups respectively. However, the number of times they were used in students' essays was very different. With the concordance function of the software, the context in which the words were used was also examined to note variation or resemblance between the two groups.

In this chapter, the methods of data collection and analysis in order to answer the nine research questions in the three studies are delineated. The findings are reported in the next chapter.

CHAPTER 4

FINDINGS

4.1 Overview of the Chapter

This chapter presents the findings of the research project in answer to the nine research questions. The first study addresses three research questions:

RQ1: How does a group of mainland Chinese undergraduates respond to a typical argumentative writing prompt? In particular, do they produce any counterarguments and rebuttals?

RQ2: What are these students' perceptions of argumentative writing?

RQ3: What factors contribute to students' performance in and perceptions of argumentative writing? How are counterargumentation and critical thinking represented in these factors?

Four research questions are asked in the second study:

RQ4: How does an intervention in counterargumentation affect students' written argumentation performance in terms of quantity of counterarguments and rebuttals, and overall quality of an argumentative essay?

RQ5: Is there a correlation between the extent of counterargumentation and the overall score of an essay?

RQ6: What are the students' perceptions of argumentative writing after the intervention?

RQ7: How does the intervention impact students' critical thinking?

The two research questions in the third study are:

RQ8: What patterns do the writing prompts in three high-stakes tests, IELTS, TOEFL and TEM4, display in terms of their rhetorical function and object of enquiry?

RQ9: In what ways do the essays of students responding to a "conventional" writing prompt differ from those of an experimental, non-conventional prompt?

These questions, however, will not be answered in order because the answers will be organized in a more natural way based on their connection to other data. The following is an overview of the structure of this chapter.

Section 4.2 encompasses part of the findings of Study 1 used to answer RQ 3, which focuses on how counterargumentation and critical thinking are represented in the four factors influencing students' argumentative writing and critical thinking. The results that answer RQs 1 & 2 are reported in Section 4.3 because the pretest data for RQ 1 and the pre-questionnaire & pre-interview data for RQ 2 are to be compared with the posttest data and the post-questionnaire & post-interview data. In order to avoid repetition and to achieve data unity, the data are reported as a whole in Section 4.3, answering four of the research questions: RQs 1 & 4, RQs 2 & 6. Section 4.2.1 presents part of the questionnaire data and the teacher interview data which identified the four factors contributing to students' performance in and perceptions of argumentative writing. In Section 4.2.2, the representation of critical thinking and counterargumentation in the Syllabus and its possible impact on the argumentative writing of undergraduate English majors is presented. Section 4.2.3 reports the representation of counterargumentation and critical thinking in the writing prompts

and rubrics of TEM4 & TEM8 and their possible impact. Section 4.2.4 presents the representation of counterargumentation and critical thinking in composition textbooks and their probable influence. Section 4.2.5 presents the findings of the classroom observation.

Section 4.3 reports on the results of Study 2 in answer to questions from RQ1 to RQ7 except RQ3 which has been answered in Section 4.2. Section 4.3.1 demonstrates the effect of the intervention on students' written argumentation by comparing their argumentative essays before and after the intervention; hence, this section answers both RQ 1 and RQ 4. The relationship between the quantity of counterarguments and rebuttals and the overall score of an essay is presented in Section 4.3.2, and this serves to answer RQ 5. In Section 4.3.3, differences of students' perceptions on argumentative writing before and after the intervention are revealed in answer to RQ 2 & RQ 6. Section 4.3.4 reports the effect of the intervention on students' critical thinking in response to RQ 7.

Section 4.4 covers the findings of Study 3. Findings of the survey on writing prompts of TOEFL, IELTS and TEM4 addressing RQ 8, and findings of the quasi-experiment addressing RQ 9 are presented in Sections 4.4.1 and 4.4.2 respectively.

4.2 Findings of Study 1

In this section, I will use part of the pre-questionnaire and pre-interview data, and the document analysis data to answer RQ3. The responses to RQ 1 & 2 will be included

in section 4.3 for the reasons stated above.

4.2.1 The Factors Influencing Argumentative Writing and Critical Thinking

As described in Sections 3.3.1 & 3.3.2 of Chapter 3, a questionnaire responded to by students and interviews with teachers were administered to identify which factors exert the greatest impact on students' argumentative writing. The particular item in the questionnaire pertaining to this was Question 9, which asked students to rank the six factors possibly contributing to their performance in, and perceptions of, argumentative writing. The frequency and proportion of the ranking given to each factor and the mean rank for each factor are shown in Table 4.1.

Table 4.1

Ranks of factors influencing argumentative writing by pre-questionnaire respondents (n=119)

Rank / Factor	1	2	3	4	5	6	Mean Rank
A classroom instruction	62(52%)	28(24%)	17(14%)	6(5%)	6(5%)	0(0%)	5.18
B middle school instruction	11(9%)	12(10%)	6(5%)	17(14%)	39(33%)	34(29%)	2.63
C the Syllabus	6(5%)	12(10%)	22(19%)	51(43%)	11(9%)	17(14%)	3.15
D composition textbook	17(14%)	34(29%)	62(52%)	6(5%)	0(0%)	0(0%)	4.52
E writing prompt and rubric	12(10%)	34(29%)	13(11%)	22(19%)	17(14%)	21(18%)	3.49
F argumentative writing in media	12(10%)	0(0%)	0(0%)	17(14%)	45(38%)	45(38%)	2.19

According to the data shown, factors A, D, E, and C, i.e., classroom instructions at college, English composition textbooks used in class, writing prompts and rubrics of high-stakes English tests, and the Syllabus were considered as the first four factors exerting an impact on students' argumentative writing. This list was regarded as tentative and had to be reconsidered in conjunction with the data from the teacher interview.

In the semi-structured focus group, three teachers of English writing focused on the same task of identifying factors influencing students' performance in and perceptions of argumentative writing. Presented below are teachers' choices among the six factors, namely, A. classroom instruction at college, B. middle school instruction, C. the Syllabus, D. the composition textbook used for argumentative writing instruction, E. high-stakes writing prompts and rubrics, and F. argumentative essays in the media. Informant A's ranking of the six factors was (from the most to the least significant): D, A, E, C, F, and B. To her, the composition textbook is the most influential factor because teachers' instructions are mainly based on it, and classroom instruction is the major means by which students learn how to write argumentative essays. She also considered the writing prompts and rubrics used in TEM4 & TEM8 and the Syllabus as exerting a considerable impact on the teaching and learning of argumentative writing.

Informant B's ranking was: E, D, A, F, C, and B. He thought that we had to admit that in the present society our teaching is exam-oriented, thus the requirements in the writing prompts and rubrics actually set the rules for the argumentative writing

pedagogy. He said that while textbooks and classroom instruction were important to train students in argumentative skills, argumentative essays in newspapers and other media also provide useful supplements. As to the Syllabus, he said it was supposed to shape the pedagogy and assessment of argumentative writing, but that stipulations in the Syllabus were not implemented in practice for various reasons.

Informant C's ranking was: C, A, D, E, B, and F. She considered the Syllabus as the most influential factor because she believed that the making of textbooks and writing prompts was based on it. In summary, the three members of the focus group considered that the composition textbook, classroom instruction, writing prompts and the Syllabus as the first four factors contributing to students' argumentative writing performance and perception.

Taking both the questionnaire and focus group data into consideration, I identified four factors collectively considered by students and writing teachers as exerting a substantial impact on the argumentative writing of undergraduate English majors. Again, they are: the composition textbook, classroom instruction, writing prompts and the Syllabus. Upon identifying these factors, the representation of counterargumentation and critical thinking in these factors was investigated. The sections 4.2.2, 4.2.3, and 4.2.4 below report the findings.

4.2.2 The Representation of Critical Thinking and Argumentative Writing in the Syllabus and the Possible Impact

As delineated in Section 3.3.3 of Chapter 3, a qualitative content analysis was conducted on the text of the Syllabus (NACFLT, 2000) to uncover the representation of critical thinking and English writing in the document, and its possible impact on undergraduate English majors. The Syllabus comprises six sections, namely, General Objectives, Curricular Arrangement, Instructional requirements, Pedagogical Principles, Pedagogical Methods, and Assessment. Each of these sections was used as a unit of analysis. Therefore, findings resulting from the content analysis are presented by the order of section. Unlike common qualitative content analysis, no categories were developed during the data analysis process. Instead, the findings of qualitative content analysis in this section, as well as in the following two sections, are presented basically in the form of description and interpretation which is based on systematic analysis of the document (Bryman, 2012). My readers may develop the idea that some of the content in this section and the following two sections seems more at home in the Discussion chapter. However, since qualitative content analysis requires some interpretation that is part and parcel of those contents, certain parts of these three sections are analytical. In addition, the implications of the findings will be covered in more depth in the Discussion chapter.

In the section of General Objectives, it is stipulated that the education of undergraduate English majors is aimed at "qualifying students in fundamental English language skills and knowledge, and cultivating their ability in independent thinking

and creativity" (p. 1). This statement demonstrates that nurturing thinking competence has been deemed a goal as equally important as transmitting specialized knowledge in the education of undergraduate English majors at Chinese universities. Given the overarching nature and brief style of this section, the generalized term "independent thinking and creativity" is used, which entails fundamental properties of critical thinking, because thinking independently and creatively is essential for a critical thinker (Facione, 1990). This is in fact an important feature of the revised edition[14] of the Syllabus in order to meet the requirements of the globalization and knowledge economy of the twenty-first century (The NACFLT, 2000).

The main part of the Curricular Arrangement section is a classification scheme that categorizes specialized courses for undergraduate English majors into three types: Type 1 - courses on specialized skills, Type 2 - courses on specialized knowledge, and Type 3 - courses on specialty-related knowledge. Under this scheme, English writing courses are courses on specialized skills, which are advised to be taken in the second and third academic years by students. It is further stipulated that English Writing courses are compulsory for which students meet at least two hours a week for four, or even five semesters. Given the significant position and substantial time allocation of the English Writing courses, and because the cultivation of independent thinking and creativity is a primary curricular objective, it could be deduced that the cultivation of independent thinking and creativity should be a theme permeating the writing classroom of undergraduate English majors.

---

[14] This edition was based on The Teaching Syllabus for Junior-Level Undergraduate English Majors and The Teaching Syllabus for Senior-Level Undergraduate English Majors which came into effect from the end of the 1980s and the beginning of the 1990s.

Instructional requirements for various linguistic skills (e.g., reading, writing, and speaking) are specified in the third section of the Syllabus at five levels: entry, Band 2, Band 4, Band 6, and Band 8. Table 4.2 shows the instructional requirements for writing.

Table 4.2

Instructional requirements of English writing courses for English majors

Level	Requirements
Entry	Students shall be able to write 100-word essays with clear expression and without severe grammatical mistakes within 20 minutes; they can rewrite based on texts; they can write simple letters, notes and notices appropriately, without obvious mistakes in form and diction.
Band 2	Students shall be able to write essays of 120-150 words within 30 minutes, and the essays should be pertinent in content, with clear organization and appropriate expression; they can rewrite or summarize texts; they can write notes and notices.
Band 4	Students shall be able to write essays of 150-200 words within 30 minutes on the provided topics, outlines, graphs or statistics, and the essays should be pertinent in content, with logical organization, correct grammar and good expression; they can write notes or notices of 60 words within 10 minutes.
Band 6	Students shall be able to write summary, book report, course paper and formal letters; the writing should be idiomatic in language and reflect certain depth of thinking; they can write essays of 250-300 words within 30 minutes.
Band 8	Students shall be able to write essays of various genres with sufficient content, standard language and smooth expression; they can write 300-400 words within 30 minutes; they can write a thesis of 3,000-5,000 words with clear logic, adequate content and standard language.

According to the Curricular Arrangement (that English writing courses shall be

provided for four or five semesters starting from the second semester of the first academic year), Band 4 should be the requirement for the English writing course for sophomores in their second semester, while Band 8 outlines requirements for third-year students in their second semester. These requirements concern three aspects: length of writing within a certain time limit, writing tasks (or genres), and evaluating criteria. Of importance to this research are the evaluating criteria. In the statements about evaluating criteria, four phrases, namely, "clear organization", "logical organization", "depth of thinking", and "clear logic", related to critical thinking appear in Band 2, Band 4, Band 6 and Band 8 respectively. These are the only phrases related to critical thinking, and they are restricted to only the structure of a piece of writing. In addition, "reflect certain depth of thinking" is a vague expression that would be hard for teachers to act on in class. These vague terms suggest that critical thinking, as it has been defined in this study and in the literature, might be underrepresented in the instructional requirements. One possible consequence is that composition teachers would ignore the cultivation of critical thinking when they teach writing. Therefore, it is at this point that instructional requirements for teachers, as displayed through these rubrics, contradict the educational goals for English majors.

The content directly related to critical thinking development of English majors appears in the section of Pedagogical Principles. The last of the five paragraphs in that section reads:

> *Students' thinking ability and creativity shall be cultivated. Through*
>
> *the instruction, we should train our students in abilities of analyzing,*

summarizing, synthesizing and abstract thinking. They should learn to
approach an issue from multiple perspectives, and they should
develop creativity through discovering and solving problems. We
should maintain the balance between teaching language skills and
cultivating thinking abilities and creativity since both are equally
important pedagogical goals. (the Syllabus, NACFLT, 2000, p. 12)

It can be argued that the message conveyed in this paragraph is strong and specific. The statements highlight the significance of critical thinking development and what teachers should do in class to achieve it. In this paragraph, the cultivation of critical thinking is illustrated at two levels. Firstly, "thinking ability" is specified as "abilities of analyzing, summarizing, synthesizing and abstract thinking". Secondly, and notably, "to approach an issue from multiple perspectives" is deemed a fundamental skill that should be trained in the English class. "To approach an issue from multiple perspectives" is aligned with the concept of "considering alternative views" which has been widely acknowledged as a hallmark of critical thinking (e.g., Palmer, 2012; Perkins et al., 1991).

In the Pedagogical Methods section, however, very little content is dedicated to the means of developing critical thinking in the writing classroom. In contrast, students are encouraged to undertake various extracurricular activities to develop their abilities in thinking and creativity as well as other aspects. The statement concerned reads:

Aligned to the classroom teaching, extracurricular and practical

activities shall be able to stimulate students to study and to cultivate their abilities in the comprehensive use of the English language, organizing, socializing, thinking and creativity. These activities can take the form of reading, speech, debating, book reporting, interviewing, etc. (the Syllabus, NACFLT, 2000, p. 13)

Since the cultivation of critical thinking has been established both in the General Objectives and in the Pedagogical Principles sections, there should be more guidance for teachers as regards how to foster critical thinking in class. However, this element is absent in the Pedagogical Methods section. This is another point where guidance for action contradicts the stipulation of educational goals.

No relevant content was found in the Assessment section. Considering the significant washback of testing on both teaching and learning, critical thinking development, as a primary educational goal, should have been represented here. This is the third point of contradiction in the Syllabus concerning critical thinking education.

However one statement was noted. It reads:

To evaluate the implementation of this syllabus, the National Advisory Committee for Foreign Language Teaching, authorized by the Ministry of Education, organizes Test for English Majors Band 4 and Test for English Majors Band 8 in the fourth semester and Eighth semester for English majors during their undergraduate program. (the

Syllabus, NACFLT, 2000, p. 14)

This paragraph may not appear directly relevant to critical thinking or English writing; however, it is highlighted to facilitate other research stages in this study because it makes two points clear. Firstly, TEM4 & TEM8 are the only two official, nation-wide tests for undergraduate English majors in mainland China. Secondly, TEM4 and TEM8 are expected to reflect the stipulation of the Syllabus. In other words, the two tests are criterion-referenced.

To summarize, the Syllabus recognizes the importance of critical thinking development in the education of undergraduate English majors. However, critical thinking development is more conceptual than operational. It is represented, to various extents, in sections of General Objectives, Curricular Arrangement and Pedagogical Principles, while absent in the sections of Instructional Requirements, Pedagogical Methods and Assessment. The result of this content analysis suggests that critical thinking development exists as abstract guidelines in the Syllabus, but when it comes to practice, inadequate direction has been given to teachers. Therefore, with regard to developing critical thinking of undergraduate English majors, the Syllabus serves as too general a guideline; further, definite guidance is lacking as to the appropriate classroom instruction and teaching methods for teachers to follow.

4.2.3 Representation of Counterargumentation and Critical Thinking in Writing Prompts and Rubrics of TEM4 and TEM8 and the Possible Impact

The results of the content analysis of writing prompts were based on 16 TEM 8 writing prompts and 13 TEM4 prompts (see the selecting criteria in 3.3.3.2), and the respective rubric of TEM8 (see Appendix 7) and TEM4 (Appendix 8). The data analysis was qualitative in nature. In terms of writing prompts, only the goal instructions were investigated (the content of writing prompts was investigated in Study 2 of this research project), while the two complete rubrics were analyzed.

The goal instructions in TEM8 writing prompts read:

> *In the first part of your essay you should state clearly your main argument, and in the second part you should support your argument with appropriate details. In the last part you should bring what you have written to a natural conclusion or make a summary.*
>
> *Marks will be awarded for content, organization, language and appropriateness. Failure to follow the above instructions may result in a loss of marks.*

The goal instructions in TEM4 writing prompts read:

> *You are to write in three parts.*
> *In the first part, state specifically what your opinion is.*

In the second part, provide one or two reasons to support your opinion.

In the last part, bring what you have written to a natural conclusion or make a summary.

Marks will be awarded for content, organization, language and appropriateness. Failure to follow the above instructions may result in a loss of marks.

High resemblance was found between the goal instructions in TEM4 and TEM8 writing prompts. They were therefore analyzed concurrently. The goal instructions contain two dimensions. The first dimension specifies how test-takers should structure their written response to the prompt. The second dimension encompasses two general and brief statements about the marking criteria of the written argumentation with a focus on language proficiency. It is the first dimension that is of particular interest to the study.

According to the first dimension of the goal instructions, test-takers are to write the essay in three parts, each part presenting claim ("main argument", or "opinion"), support for that claim, and summarization respectively. Such instructions quite definitely condition test-takers when they respond to the writing prompts. To some extent, they restrict students' writing especially in terms of logical reasoning and essay construction. Such goal instructions leave little room for considering alternative or opposite opinions to one's own. Thus, it can be argued that counterargumentation - the key concept of critical thinking – is not encouraged because there is no

expectation that alternative viewpoints are required.

Both the rubrics of TEM4 and TEM8 are composed of five bands referring to the different quality of the written response of test-takers, with Band 5 the highest and Band 1 the lowest. Like the goal instructions in writing prompts, the two rubrics resemble each other. In order to avoid redundancy, and given the great consistency in the content of each band, it is deemed unnecessary to go through each band in reporting the results of content analysis of the rubrics. Band 5 of both TEM4 and TEM8 is taken as an example below.

Band 5 of TEM 4 reads:

EFFECTIVE COMMUNICATION WITH ACCURACY

The writing effectively addresses the writing task. It demonstrates a well developed logical organizational structure with clearly stated main ideas and sufficient supporting details. It has almost no errors of vocabulary, spelling, punctuation or syntax, and it displays an adequate ability to use the language with appropriacy. No difficulty is experienced by the reader.

Band 5 of TEM8 reads:

DESCRIPTION: EFFECTIVE   COMMUNICATION   WITH ACCURACY

The writing effectively addresses the writing task. It demonstrates a well developed logical organizational structure with clearly stated main ideas and sufficient supporting details. It has almost no errors of vocabulary, spelling, punctuation or syntax, and it displays an adequate ability to use the language with appropriacy. No difficulty is experienced by the reader.

Content analysis resulted in the following findings. Firstly, the illustration of marking standards in the rubrics echoes the goal instructions in the writing prompts to a considerable extent. Specifically, the rubrics consist of two dimensions as in the writing prompts, with the first and second dimensions related to the organization and expression of an essay respectively. The goal instructions require students to include claim, supporting details and summary in three parts. The rubric restates this in the sentence, "[effective writing] demonstrates a well developed logical organizational structure with clearly stated main ideas and sufficient supporting details". This statement implies that "a well developed logical organizational structure" entails only the writer's own claim and evidence to support this claim. Accordingly, the acknowledgement of alternative viewpoints seems to be ignored in such a rubric.

Second, the criteria on "vocabulary, spelling, punctuation or syntax" and "use the language with appropriacy" suggest that much importance should be attached to the language proficiency of test-takers when teachers evaluate the essays. An argument can be made here that this focus on accuracy undermines the Syllabus, which stipulates that cultivating thinking ability is as important as improving language

proficiency. Some may argue that criteria on the essay organization in the first dimension do reflect the requirements on thinking. However, requirements on thinking as reflected in the rubrics of TEM4 and TEM8 are not only hard to operationalize, but they also ignore key elements of critical thinking.

4.2.4 Representation of Counterargumentation and Critical Thinking in Composition Textbooks and the Possible Impact

A qualitative content analysis was conducted on one English composition textbook - *A handbook of English writing* (Ding et al., 1994) - to find out the possible impact of composition textbooks on the critical thinking ability in the argumentative writing of undergraduate English majors. The reason for focusing on one textbook, instead of more, was made on the following four grounds: first, this was the textbook used on the research site in this project; second, it has been the predominant textbook used by undergraduate English majors for nearly 30 years and most teachers were trained to teach English writing by using this book (You, 2010b); third, in this project, the results of a preliminary survey on composition textbook use at over 30 universities reveal that *A handbook of English writing* was the selected textbook at approximately one third of the universities. Fourth, due to the limitation of this research and given that investigation into composition textbooks accounts for a relatively small proportion of the whole project, I decided to focus on *A handbook of English writing* as the source of data analysis in this section.

Reflecting traces of current-traditional rhetoric, the textbook is organized by different

levels of language structure (see a table of contents of the textbook in Appendix 4) of word, sentence, paragraph and passage. In terms of writing genres, description, narration, exposition and argumentation are discussed in separate sections. Since this project aimed at the critical thinking ability in argumentative writing of English majors, the section on argumentation was the target of the content analysis, the results of which are reported below.

The argumentation section is comprised of three parts: main discussion on argumentation, sample essays and exercises. The main discussion is approximately 1,050 words in length, followed by four models (two articles by skilled writers and two by students) and three exercises.

In the beginning, the purpose of argumentation is defined as a passage "to convince" others (Ding et al., 1994, p. 234). More specifically, an argumentative essay is "to make the reader agree with its point of view and support it, to persuade him to change his mind or behavior, and to approve a policy or a course of action that it proposes" (Ding et al., 1994, p. 234). That is to say, a successful argumentative essay depends on how persuasive it could be. As such, what counts the most is the means of supporting a view and convincing others, which gives the subsequent text the primary task of guiding students to achieve persuasiveness in written argumentation.

In order to write a good argumentative essay, the student writers are advised to meet six requirements: (1) discuss a debatable point; (2) provide sufficient evidence; (3) include good logic; (4) write with clear logic; (5) include good use of the other three

types of writing; (6) keep an honest and friendly attitude. *A debatable point* is defined as "something which can be viewed from more than one angle and is therefore open to dispute" (Ding et al., 1994, p. 234). This definition highlights a fundamental element of argumentation, i.e., it clarifies the fact that argumentation, oral or written, starts from a controversial issue and aims at convincing others. This is a good starting point to elaborate the dialogic nature of argumentation and the interaction between argumentation and critical thinking in succeeding content. Written argumentation is like a verbal exchange between the writer and the reader, each stating their view and supporting it in order to persuade the other. The reader, although mute, should be given the chance to voice alternative views; the writer can refute the alternative views. Nevertheless, acknowledging alternative views has not been properly addressed in the succeeding text. Under *sufficient evidence*, examples, statistics, and other evidence are listed as means of support for the viewpoint in an essay. Under *good logic*, two ways of reasoning - inductive and deductive reasoning - are discussed briefly.

It is under *clear logic* that acknowledging alternative views is mentioned. It reads:

> *A typical argumentation essay consists of three parts: an introduction*
> *which identifies the issue to be discussed and explains the importance*
> *of such a discussion; a body which presents the evidence; and a*
> *conclusion in which the proposition, if it is stated at the beginning, is*
> *reaffirmed.*
> *In the body, it is advisable that you devote the first one or two*
> *paragraphs to a consideration of the other side of the case before*

stating and amplifying your own view. Then in each paragraph that

follows, you add something new and important to your argument.

(Ding et al., 1994, p.237)

Noticeably, considering the other side of the case is suggested as a method of introducing one's own view here. As this part of the text concerns the structure of an essay, it gives the impression that considering the other side is merely a strategy of setting a target for a writer to shoot at, rather than encouraging fairness (a dimension of critical thinking). Notably, the key terms "counterarguments" and "rebuttals" do not appear in the text. These two terms are essential if the concept of counterargumentation is to be taught. Therefore, the text fails to provide explicit instruction in counterargumentation skills despite the fact that counterargumentation is essential for enhancing persuasiveness of argumentative essays and a hallmark of critical thinking. Explicit instruction in counterargumentation entails scaffolding in various forms before students can develop counterargumentation skills. For instance, argumentation schemata adapted from the Toulmin model could be utilized to introduce terms of counterargument and rebuttal. Figures listing arguments on one side and counterarguments on the side could be used as further material to explain the interaction between the two.

A review of the four model essays (two written by students and two by established writers) revealed that counterargumentation appeared in only two of the essays (by students). The provision of models is an important function of composition textbooks since analyzing and imitating models is one of the primary classroom activities

employed by composition teachers. Consequently, model essays, especially the ones by established writers, are supposed to demonstrate how counterargumentation can be skillfully woven into an argumentative essay to make it more convincing. From this perspective, this part of the textbook does not fulfill its function competently, and is not fully aligned with relevant statements in the main discussion.

In terms of exercises, only one of the five questions in Exercise I asks the students to "read the model essays and discuss: does the writer consider the other side of the argument before stating and amplifying his or her views" (Ding et al., 1994, p 260). This simple yes-or-no question obviously lacks the strength of stimulating critical thought.

In addition, Exercise III reads:

> *Choose your own topic and write an argumentative paper. See to it that your theme is debatable, or else the essay will fail completely. Provide sufficient, well-arranged evidence and present it in a composed and friendly tone. (Ding et al., 1994, p 261)*

There is a contradiction in the instructions of this exercise. If the theme is a debatable one, the readers, although mute when reading, might have different views from the writer; then it would be reasonable for the writer to voice their views and refute them instead of merely providing evidence to support their own view. Furthermore, this exercise is misleading in itself. The condition in which a writer ignores alternative

views is termed myside bias by critical thinking theorists (e.g., Perkins, 1985; Wolfe & Britt, 2008). In order to avoid myside bias and enhance critical thinking in argumentative writing, more practice is needed for students to incorporate counterarguments and rebuttals in writing.

The findings of this content analysis of a textbook suggest that more specific instructions and scaffolding exercises are needed on counterargumentation skills in the Argumentation section of *A handbook of English Writing* because, as revealed in recent literature, written argumentation in general involves substantial practice; counterargumentation skills in particular entail intensive training (Ferritti et al., 2000; Gleason, 1999; Nussbaum & Kardash, 2005).

4.2.5 Findings of the Classroom Observation

Eleven sessions of classroom observation were conducted with an open-ended observation scheme (Appendix 9) focusing on teaching objectives, pedagogical activities, argumentation skills discussed in class, and use of materials. The findings of the classroom observation are presented point by point.

4.2.5.1 Teaching objectives

As illustrated in section 3.3.4 of the thesis, argumentation became the curricular focus during the period of time when I was allowed to observe the two composition classes taught by Instructor A and Instructor B respectively. Table 4.3 lists the teaching

objectives of every observed session. What should be noted here is that classroom instruction in both classes centered on TEM 4 writing tasks, which entailed argumentation for around three months until the TEM4 test.

Table 4.3

Teaching objectives of the two observed composition classes

	Instructor A	Instructor B
Session 1	Overview of argumentation and argumentative essays	An introduction to argumentative writing
Session 2	Skills of writing argumentative essays	The structure of argumentative essays
Session 3	Skills of writing argumentative essays	How to write the introductory part of argumentative essays
Session 4	TEM4 and argumentative writing	How to write the body part of argumentative essays
Session 5	TEM4 and argumentative writing	How to write the concluding part of argumentative essays
Session 6	*Rescheduled and missed by the observer*	How to write argumentative essays for TEM4 test

Observed data show that both instructors attempted, in various ways, to elaborate basic skills of written argumentation. Both instructors, after briefly discussing fundamentals of written argumentation, moved on to the close connection between TEM4 and argumentative writing. Such a move indicates an exam-oriented pedagogy in composition instruction at the research site.

4.2.5.2 Pedagogical activities and use of materials

Pedagogical activities taking place in the classroom of the eleven observed sessions and use of material are described in Table 4.4.

Table 4.4

Pedagogical activities and use of materials of the observed composition classes

Instructor A
Session 1: The instructor read together with students the main discussion part of the Argumentation section in the textbook to introduce argumentation as a type of writing. Use of material: textbook Students identified good topics for argumentative writing within the topics provided by the instructor; use of material: self-prepared Students conceptualized good topics for argumentative essays
Session 2: The instructor discussed with students the basics of writing an argumentative essay. Use of material: textbook The instructor analyzed model essays to emphasize distinct features of and important skills employed in argumentative writing. Notably, the self-prepared model essays do not include counterargumentation. Use of material: textbook/ self – prepared (articles from publications)
Session 3: The instructor explained various ways of supporting one's proposition. Instructor provided words and expressions used frequently in argumentative essays. Instructor provided argumentative writing templates for students. Use of material: self-prepared
Session 4: The instructor discussed the TEM 4 standard structure of an argumentative essay. Students practiced constructing an argumentative essay by making outlines for two or three topics. Then group discussion about the feasibility of the outlines. Teacher commented on several outlines. Use of material: self-prepared
Session 5: Analyzing writing prompts in past TEM4 tests to find patterns of the content by the instructor and students. The instructor presented a model answer to one of the writing prompts.

Use of material: self-prepared and TEM4 reference books
Instructor 2
Session 1: the instructor introduced argumentation as a type of writing and illustrated the interrelation between argumentation and exposition, using examples. Instructor discussed with students the requirements of good argumentative essays. Source of material: textbook and self-prepared material
Session 2: the instructor analyzed the structure of argumentative essays with model essays among which the self-prepared ones do not include counterargumentation. Students practiced by making outlines for topics of argumentative essays provided by the instructor. Use of material: textbook and self-prepared material
Session 3: the instructor presented writing the introductory part of an argumentative essay. Joint writing by instructor and students. Independent writing by students. Use of material: textbook and self-prepared material
Session 4: the instructor presented writing the body part of an argumentative essay. Joint writing by instructor and students. Independent writing by students. Use of material: textbook and self-prepared material
Session 5: the instructor presented writing the concluding part of an argumentative essay. Joint writing by instructor and students. Independent writing by students. Use of material: textbook and self-prepared material
Session 6: the instructor listed the writing prompts of TEM4 tests over the past 10 years and summarized the features. The instructor analyzed the model answer to one of the TEM4 writing prompts. Use of material: TEM4 reference books.

4.2.5.3 Argumentation skills discussed

Whilst Table 4.4 provides a detailed description of the pedagogical activities in the observed composition classes, summarized here are particular instructions in argumentation skills observed in the two classes.

The structure of an argumentative essay was the focus of discussion in respect of

written argumentation skills. Both instructors explained that an introduction with thesis statement, a body with two or three supporting details and a conclusion with restated claim constituted the standardized argumentative essay. To facilitate learning, one instructor provided templates for students to adopt; the other instructor adopted a process approach by teaching each part separately by means of gradual improvement (scaffolded writing to independent writing). In addition, lists of words and phrases frequently used in written argumentation were provided to assist students.

It was noted that both instructors closely associated argumentative writing with the TEM4 test and other high-stakes tests. For instance, both instructors articulated that written argumentation was very important because it was the genre most frequently employed in response to a writing prompt of TEM4 or TEM8. Furthermore, when discussing how to write argumentative essays, both instructors explicitly referred to the TEM4 writing prompts and rubrics.

One remarkable aspect of the findings is that neither instructor used the term counterargumentation in class, nor was the notion of acknowledging alternative views linked to critical thinking, despite the fact that the textbook, as mentioned above in section 4.2.4, covered it, albeit inadequately. As a matter of fact, a critical discussion of alternative viewpoints was never heard during the whole period of observation. "The other side of the case" was briefly mentioned because it was in the textbook by one instructor who treated this as a strategic way of beginning one's essay. The other instructor completely ignored it.

Findings from the classroom observation suggest that the instructions in argumentation at the research site were in line with the guidelines in the Pedagogical requirements in the Teaching Syllabus for English Majors. The textbook – *A handbook of English writing* (Ding et al., 1994) - was used with complementary materials in class. Thus, it appeared that washback from TEM4, which excludes notions of counterargumentation in its goal instruction was evident. In other words, the instructions in argumentation were exam-oriented. Acknowledging alternative viewpoints was not included in the classroom instruction; there was no discussion on the interaction between written argumentation and critical thinking in the classes.

At the end of the observation sessions, I talked to each of the two instructors separately to probe into the reasons why they glossed over counterargumentation in their teaching. Instructor A admitted that she herself had no idea of counterargumentation, and that it was just "natural" for her to curtail the mentioning of "the other side of the case". She thought that the reasons might be her educational background and cultural influence. Instructor B said that benefiting from his experience as an instructor at an English language training company, he knew that counterargumentation was a fundamental concept in argumentative writing. However, he felt obliged to align his teaching to TEM 4 test requirements. In addition, he did not know how to teach students to counterargue because there was limited content in the textbook.

4.3 Findings of Study 2 (and Part of the Findings of Study 1)

4.3.1 Differences of Students' Perceptions of Argumentative Writing Before and After Intervention

The findings in this section answer RQ 2 and RQ 6, namely, what the students' perceptions of argumentative writing are before and after the intervention.

4.3.1.1 Findings from questionnaires

Questionnaires with seven close-ended multiple-choice questions [15] and one open-ended question were conducted in a pre-post manner before and after the instructional intervention to compare students' perceptions on written argumentation and counterargumentation. Tables 4.5 and 4.6 show the frequency and percentage of responses to the close-ended questions in the pre-questionnaire and post-questionnaire from both groups. According to responses to Q1, in the experimental group, before the intervention, 39.9% of participants considered argumentation (option D) as the most challenging genre of English writing to them; the percentage dropped to 20.3% after the intervention, whilst for the control group, the percentage decreased from 27.4% to 22.4%. A two-sample t-test between

---

[15] One closed-ended question which asked respondents to rank the factors influencing their argumentative writing was only answered before the intervention. The results of that part of data are not reported here but in section 4.2.1.

percentages[16] was performed for the experimental and control groups respectively to determine whether there was significant variation in students' perception of argumentation. For the experimental group, the t-statistic was significant at the .05 alpha level, t (234)=2.349, p=.0197. For the control group, the t-statistic was insignificant at the .05 alpha level, t (234)=0.888, p=.3755. Therefore, after the intervention, in the experimental group the number of students who perceive argumentative writing as difficult and challenging dropped significantly (p<.05). The data indicated that students' confidence in the argumentative writing of the experimental group was enhanced more than that of the control group.

Concerning Q4 and Q5, in the experimental group, 96.6% of students claimed that they had discussed counterarguments and rebuttals in their English writing class and 100% of them deemed counterargumentation important in written argumentation after the intervention, compared to 15.8% and 44.7% before the intervention respectively. In the control group, 16.1% of students claimed they had discussed counterarguing and refuting in the composition class; the percentage of students who practiced the skills increased slightly to 17.5% after the intervention. However, in this group the percentage of students who perceive counterargumentation as an essential skill in writing argumentative essays fell from 47.9% before the intervention to 44.5% after the intervention. T-statistics indicated that for the experimental group, the numbers of students who considered counterargumentation as an essential argumentative skill and had practiced it in class had significantly increased (p<.05).

---

[16] A software program named Statistics Calculator was used to compare two percentages to determine whether there is a significant difference between them. In the cases of this study, a two-sample t-test was proper to compare the data from two groups.

As to the way in which they usually convinced their readers (Q6), before the intervention, 33.3% of students in the experimental group claimed that they would include counterarguments and rebuttals in addition to main arguments and supporting evidence; however, after the intervention, 84.7% claimed they would use that approach. In the control group, 32.3% of students before the intervention thought that including counterarguments and rebuttals made their argumentative essays more convincing; after the intervention, 31.6% of students believed in this approach. For the experimental group, a significantly greater proportion of students (*p<.05)* would use counterarguments and rebuttal to persuade in their written argumentation after the intervention.

Table 4.5

Response frequency (with percentage) in the experimental group

	Pre-questionnaire (n=57)		Post-questionnaire (n=59)	
Q1	A 27(47.4%)	B 4 (7.0%)	A 28 (47.5%)	B 3 (5.1%)
	C 4(7.0%)	D 22 (39.9%)	C 16 (27.1%)	D 12 (20.3%)
Q2	A 1(1.8%)	B 1(1.8%)	A 3 (5.1%)	B 1 (1.7%)
	C 4 (7.0%)	D 51 (89.5%)	C 6 (10.1%)	D 49 (83.1%)
Q3	A 2 (3.5%)	B 3 (5.3%)	A 4 (6.8%)	B 7 (11.9%)
	C 23 (40.3%)	D 29 (50.9%)	C 23 (38.9%)	D 25 (42.4%)
Q4	A 9 (15.8%)	B 48 (84.2%)	A 57(96.6%)	D 2 (3.4%)
Q5	A 54 (44.7%)	B 3 (55.3%)	A 59 (100%)	D 0 (0%)
Q6	A 0 (0%)	B 17 (29.8%)	A 0 (0%)	B 2 (3.4%)
	C 21 (36.8%)	D 19 (33.3%)	C 7(11.9%)	D 50 (84.7%)
Q7	A 16 (28.1%)	B 30 (52.6%)	A 20 (33.9%)	B 28 (47.5%)
	C 3 (5.3%)	D 8 (14.0%)	C 2 (3.4%)	D 9 (15.3%)

Table 4.6

Response frequency (with percentage) in the control group

	Pre-questionnaire (n=62)	Post-questionnaire (n=58)
Q1	A 19(30.6%) B 2(3.2%) C 24 (38.7%) D 17(27.4%)	A 21 (36.2%) B 5 (8.6%) C 19 (32.8%) D 13 (22.4%)
Q2	A 7(11.3%) B 5(8.1%) C 4 (6.5%) D 46 (74.2%)	A 4 (6.9%) B 1 (1.7%) C 1 (1.7%) D 52 (89.7%)
Q3	A 4 (6.5%) B 10 (16.1%) C 19 (30.6%) D 29 (46.8%)	A 2 (3.4%) B 6 (10.3%) C 23 (39.7%) D 27 (46.6%)
Q4	A 10 (16.1%) B 52 (83.9%)	A 10(17.2%) B 48 (82.8.8%)
Q5	A 57 (91.9%) B 5 (8.1%)	A 49 (84.5%) B 9 (15.5%)
Q6	A 0 (0%) B 15 (24.2%) C 27 (43.5%) D 20 (32.3%)	A 0 (0%) B 18 (31.6%) C 21(36.8%) D 18 (31.6%) a missing value appeared in this question
Q7	A 18 (29.0%) B 31 (50.0%) C 4 (6.5%) D 9 (14.5%)	A 15 (25.7%) B 30 (51.7%) C 3 (5.2%) D 10 (17.2%)

The interview data from the closed questions indicate that the intervention was effective in shaping students' perception of argumentative writing. For the experimental group, the intervention significantly enhanced students' confidence in argumentative writing; significantly greater proportions of students claimed to have practiced counterargumentation in class, perceive counterargumentation as an essential skill in written argumentation, and would use counterargumentation in their argumentative essays ($p<.05$).

The open-ended question, Question Eight, was an extension of Question Seven, i.e., if your answer is "Others" to the last question, please specify, so apart from the three situations spelled out in Question Seven, students could note down other conditions

that they considered the most difficult aspect of writing an argumentative essay. 17 responses (8 from the experimental group and 9 from the control group) and 19 responses (9 from the experimental group and 10 from the control group) were collected from the pre-questionnaire and post-questionnaire respectively.

Qualitative content analysis resulted in five categories for the 36 responses to the open-ended question in both the pre-questionnaire and post-questionnaire. These categories were, namely, "English expression," "lack of background knowledge," "topic interpretation," "main argument" and "logic and thinking". Table 4.7 lists the definitions and examples for every category. For the experimental group, "lack of background knowledge" was claimed to be the most difficult aspect of argumentative writing (3 cases) in the pre-questionnaire, while in the post-questionnaire, most students (4 cases) thought that "logic and thinking" was the hardest for written argumentation. For the control group, "English expression" (3 cases) and "topic interpretation" (3 cases) were noted down as the most difficult thing in writing argumentative essays in the pre-questionnaire and post-questionnaire respectively.

As indicated by the data, no obvious patterns were noticed in the responses from the two groups before and after the intervention. However, as this part of the data is small in quantity, any patterns observed are not meant to be generalizable but indicative only.

Table 4.7

Definition and examples for categories of responses to the open question

Category	Definition	Example
English expression	The process of transforming ideas into English	It's hard to express myself. I think of something, but it becomes another when I write down my thoughts.
lack of background knowledge	Insufficient knowledge relating to the topic of essays	Having no idea about the topic, so don't know what to write.
topic interpretation	The understanding of the topic given in the writing prompt	To me it's hard to grasp the focus of a topic.
main argument	The statement of a writer's main point of view on an issue.	I find it difficult to write the thesis statement.
logic and thinking	The logical connection between sentences and paragraphs.	There's no good logic in my essay.

4.3.1.2 Findings from interviews

Focus group (with 6 students in each group) interviews were administered before and after the intervention centering on two questions:

- What approaches do you adopt to make your argumentative writing more persuasive?

- What are the difficulties you encountered in achieving the persuasiveness in your writing?

Two coders coded the interview data. For the first question, one category was

generated in the pre-interview and three in the post-interview from the experimental group responses; the interrater reliability for coding the four categories in terms of Cohen's Kappa was .92, .85, .89, and .83 respectively. The control group responses to the first questions were classified in two categories for the pre-interview and still two for the post-interview, the interrater reliability being $K=.81$, $K=.90$, $K=.86$, and $K=.89$ respectively. For the second question, the experimental group responses fell into three categories in the pre-interview and two in the post-interview, with Kappa values being .90, .84, .95, .93 and .84 correspondingly. The control group responses to the second question were coded into two categories for the pre-interview and three for the post-interview, Kappa values being .94, .96, .84, .88, and .91 respectively.

For the first question, responses of the experimental group in the pre-interview focused on one aspect: "structure of an argumentative essay." Several informants claimed that the opinion-support approach is the best way to write a written argumentation. One informant said, "I will describe the problem, analyze the problem and solve the problem. In the end, I'll conclude. This is what I learned from high school. This is also how we write argumentation in Chinese." Another articulated, "When I write an argumentative essay, I adopt this approach: what the issue is - comment on both sides of it - my opinion." Still another informant said that she had not learned much about English argumentative writing, which, together with the first quotation, revealed that participants had a fairly weak foundation in written argumentation in English. Before they undertook the course English Writing II, what they could resort to was the knowledge of Chinese argumentative writing and the instructions from their high school lessons. The second quotation suggested the wide

spread washback of College English Test (CET) batteries, because that approach conformed to the goal instructions in the writing prompts of CET tests.

The responses of the experimental group in the post-interview to this question increased in number and expanded in content; they fell into three categories, namely, "counterargumenation skills," "structure," and "supporting evidence." Concerning "counterargumentation skills", one informant shared her view, "We should consider both our side and the opposite side, and choose the strongest point in the opposite side to refute. If we can do so, we can be very persuasive". As to "structure", one informant said, "Now when I get a topic, I know the procedures: my claim-evidence-counterargument-rebuttal-conclusion". About "supporting evidence", one informant said, "Experience in life will help us produce more supporting evidence so we can write better".

Responses of the control group to the first question in the pre-interview were categorized in two types: "supporting evidence" and "counterargumentation". Much was said about the first category. One informant said, "I think that a piece of evidence should be presented from general to specific, because specific evidence is more persuasive". As to the second category, one informant said, "If I need to write a long essay, I will consider the opposite side". In the post-interview, the control group responses fell into two categories: "main argument" and "supporting evidence". When it came to "main argument", one informant said, "We should put forward the main argument straightforwardly in the beginning, and the main argument should be extraordinary, thus making our essay more attractive". As is indicated by the data,

some students in the control group knew about counterargumentation from other sources, but since counterargumentation skills had not been discussed in the English writing class, students tended not to associate counterargumentation with the persuasiveness of written argumentation.

For the second question, responses of the experimental group in the pre-interview were coded into three categories: "expression", "structure", and "supporting evidence". In the post-interview, two categories were identified from responses of this group, namely, "expression" and "logical thinking". When informants discussed difficulties in writing argumentative essays, many of them complained that they could not express themselves well due to limited vocabulary. This problem continued after the intervention, suggesting that instruction focusing on argumentative skills was not effective in improving students' language proficiency. As a matter of fact, the improvement of language proficiency is a slow process, and normally cannot be achieved within a few months. However, the data indicate that the intervention facilitated students' structure of written argumentation in English. In general, the students' responses lacked depth, indicating that they had not seriously considered what challenges faced them before.

Among responses of the control group to the second question in the pre-interview, two categories were identified: "expression" and "supporting evidence". In the post-interview, responses of this group to the same question fell into three categories, namely, "structure", "topic interpretation" and "supporting evidence". About "topic interpretation", one informant said, "Once I get a topic, it's hard to find the real focus

of it". It is noted that after 12 weeks of ordinary instruction in argumentation, there were fewer perceptions on expression as a difficult aspect of argumentative writing so they do not form a separate category.

Although between-group differences were observed in responses to the two questions in the pre-interview, it is more interesting to investigate the between-group variation after the intervention. For the first question concerning approaches adopted to enhance persuasiveness in argumentative essays, the post-interview responses of the experimental group involved "counterargumentation skills", "structure" and "supporting evidence", while the control group only listed "main argument" and "supporting evidence". It became evident that after receiving instructions in counterargumentation, students in the experimental group acquired such skills. This is clear indication of the efficacy of the intervention. In contrast, the methods the control group could resort to were fewer and restricted without appropriate instruction.

For the second question vis-à-vis difficulties in achieving persuasiveness in argumentative writing, the experimental group responses fell into categories of "expression" and "logical thinking", while those of the control group enumerated "structure", "topic interpretation" and "supporting evidence". It was noted that students in the control group perceived supporting evidence as both their major way and the big challenge to write persuasive essays. This, in a sense, showed that normal instruction in argumentation was inefficient to assist students to attain persuasiveness. Markedly, the experimental group came to think about logical thinking of an

argumentative essay, presumably triggered by the concept of counterargumentation.

4.3. 2 The Effect of the Intervention on Students' Written Argumentation

After the 12-week intervention was completed, the pretest and posttest scripts from both the experimental and control groups were collected. Two raters blindly rated the scripts, and then two coders conducted text analysis of the scripts to identify and code four argumentative elements in the essays: claims, data, counterarguments and rebuttals. The subsequent paragraphs present the results of statistical analysis of the data which answer RQ1 and RQ 4.

Preliminarily, a descriptive analysis of the data from the rating and coding was conducted to provide basic information about the outcome variables. Table 4.8 lists five variables, i.e., the overall score of an essay and the four argumentative elements in the essay, with the means and standard deviations of each variable.

Table 4.8

Descriptive statistics of outcome variables

	Group							
	Experimental (N=63)				Control (N=62)			
	Pretest		Posttest		Pretest		Posttest	
Variable	M	SD	M	SD	M	SD	M	SD
Claims	0.94	0.20	1.00	0.00	0.95	0.20	1.00	0.00
Data	2.21	1.03	2.28	0.76	2.01	1.02	2.43	0.84
Counterarguments	0.06	0.30	0.98	0.44	0.09	0.28	0.03	0.18
Rebuttals	0.12	0.59	1.39	0.78	0.13	0.45	0.05	0.28
Score of essay	12.65	1.49	14.92	1.34	12.45	1.21	13.11	1.04

The mean scores for pretest essays were 12.65 and 12.45 (out of a total score of 20) for the experimental and control groups respectively. A T-test (two-tailed) showed that there was no significant difference between experimental and control classes in argumentative writing proficiency before the intervention.

As shown in Table 4.8, for the experimental group, the mean frequency of counterarguments rose from 0.06 in the pretest to 0.98 in the posttest while that of rebuttals rose from 0.12 to 1.39. The mean score rose significantly ($p<.01$) from 12.65 to 14.92, while the mean frequency of arguments, i.e., the number of supporting reasons, went up slightly from 2.21 to 2.28. For the control group, however, no increase was noted in the production of counterarguments and rebuttals in posttest essays. The mean overall score increased significantly ($p<.05$) from 12.45 in the pretest to 13.11 in the posttest; and the mean frequency of arguments grew from 2.01 to 2.43. Thus, the post-test scores for the experimental group were significantly higher ($p<.05$) than those of the control group.

The initial interrater reliability for coding the four argumentative elements between the two coders was determined by calculating the Kappa coefficient. For the pretest scripts, the results were claim, K=.79; data, K=.80; counterargument, K=.94; and rebuttal, K=.80; for the posttest scripts, no statistics was computed for claim because the two coders assigned identical coding to each of the scripts in this category; the other coefficients were data, K=.78; counterargument, K=.91; and rebuttal, K=.83. Under the Kappa interpretation guidelines provided by Landis and Koch (1977), the two coders reached substantial agreement for coding the pretest and posttest data.

It was found that before the intervention, and in students' usual argumentative essays, there was virtually no inclusion of counterarguments and rebuttals; however, after the intervention, the numbers of counterarguments and rebuttals, and the overall quality of the argumentative essays significantly improved for the experimental group. About 10% of the pretest scripts of both experimental and control groups had zero claims, i.e., the writer's response to the topic question was non-committal. However, after the intervention, every script in both the groups contains a definite claim, which shows that the argumentative writing instructions for both groups were effective in helping students make stronger claims. As to the place of occurrence of the claims, in the pretest scripts, 83% of the claims for both groups appear at the beginning of the essay; the proportion rose to 100% in the posttest scripts for the experimental group and 98% for the control group.

4.3.3 The Relationship Between the Quantity of Counterarguments and Rebuttals and the Overall Score of an Essay

In order to uncover the relationship between the use of counterarguments and rebuttals and the overall score of an essay in answer to RQ 5, I performed Pearson correlation tests using the posttest data of the experimental and control groups. Table 4.9 shows that the overall scores of the experimental group essays were significantly positively correlated with frequency of arguments ($p<.05$), counterarguments ($p<.05$), and rebuttals ($p<.01$). Table 4.10 shows that the mean scores of the control group essays were significantly positively correlated with frequency of arguments ($p<.01$). There was also a positive correlation between the mean scores and the frequency of

counterarguments (p<.05) and rebuttals (p<.05). Thus, there are clear indications that the inclusion of counterarguments and rebuttals enhanced the overall quality of an essay as well as its persuasiveness. A correlation coefficient was not available between the claim and the overall score in both groups because every essay contained one claim.

Table 4.9

Pearson correlations between the use of four argumentative elements and the overall score of essays (experimental group)

	claim	data	counterargument	rebuttal
Overall score of the essay	N/A	0.35*	0.38*	0.65**

*p<.05; **p<.01.

Table 4.10

Pearson correlations between the use of four argumentative elements and the overall score of essays (control group)

	claim	data	counterargument	rebuttal
Overall score of the essay	N/A	0.73**	0.25*	0.26*

*p<.05; **p<.01.

As a reliability check, I examined the posttest cases in the experimental group that did not include any counterarguments or rebuttals, and the posttest cases in the control group that did include counterarguments or rebuttals. The results further confirmed the findings mentioned above. In the experimental group, the average score of the six scripts without counterarguments or rebuttals was 12.51, much lower than the mean (14.92) for the experimental group. On the other hand, the two scripts in the control group that did include counterarguments or rebuttals scored 13.94, much higher than the mean (12.45) for that group.

It is worth mentioning that students in both the control and experimental groups generated more arguments after a 12-week period of instruction in argumentative writing. Nevertheless, I discovered that in essays with counterargumentation elements, the amount of data supporting initial claims tended to be less than in those without counterargumentation elements. Descriptive statistics in Table 4.8 further support this. For the experimental group, the mean of arguments increased from 2.21 in the pretest to 2.28 in the posttest, while for the control group, the mean of arguments rose from 2.01 to 2.43.

4.3.4 The Effect of the Intervention on Students' Critical Thinking

Two raters independently rated the pretest and posttest scripts for critical thinking according to the rubric (Appendix 3) developed for this study. The inter-rater reliability coefficient was. 92 after negotiation. After the rating was completed, descriptive analysis was performed with the data. Table 4.11 lists the mean and

standard deviation (SD) of critical thinking scores for both groups in the pretest and posttest. Before the intervention, the mean scores for the experimental and control groups were 13.81 and 13.85 (out of a total of 24) respectively. After the intervention, the mean score for the experimental group and the control group rose to 16.31 and 14.71 respectively.

Table 4.11

Mean and standard deviation (SD) of CT score for pretest and posttest scripts

Maximum	Mean		SD				Minimum	
	Pre	Post	Pre	Post	Pre	Post	Pre	Pos
Experimental group 19.5 (n=63)	13.81	16.31	1.47	1.77		11	11	17
Control group 15.5 (n=62)	13.85	14.71	1.92	1.12		10	10.5	17

I observed that there was a difference of 0.04 points in the pretest between the two groups, but I didn't know whether the difference was statistically significant. In other words, I was not sure whether the two groups were significantly different in critical thinking ability before the intervention. I also observed that the experimental group increased more in the posttest mean score than the control group did. However, it was yet to be ascertained whether the variation in the critical thinking development of the two groups was statistically significant.

Under the situations, an ANCOVA test on the SPSS 21 was performed because an ANCOVA test enables us to compare post-test scores of two groups while controlling for the pre-test scores as the covariate (Dornyei, 2007). Therefore, instead of performing two T-tests as I did in section 4.2.7 (one to determine whether there was significant difference between the two groups in the pretest, the other to compare means of the two groups in the posttest), one statistical test served the purpose.

Table 4.12 presents the ANCOVA test output. Three participants in the experimental group took the pretest but were absent from the posttest, and one participant in the control group missed the pretest but took the posttest. The four cases were treated as missing values during the SPSS procedure. Therefore, in the ANCOVA test, the numbers of experimental and control groups were 60 and 61 respectively. As can been seen, the between-group difference in the pretest was insignificant ($p<.05$). The row starting with "group" reports a significance value of .000, which means there was a statistically significant difference ($p<.0005$) between the mean scores of the experimental and control groups. In addition, the ANCOVA output shows an effect size of .358. As indicated by the inferential statistics, the intervention in counterargumentation was effective in enhancing experimental group students' critical thinking ability in their argumentative writing.

Table 4.12

SPSS ANCOVA test outputs

Between-Subjects

Factors

		N
group	1.00	60
	2.00	61

Tests of Between-Subjects Effects

Dependent Variable: post

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	146.670[a]	2	73.335	34.159	.000	.367
Intercept	308.641	1	308.641	143.764	.000	.549
pre	7.577	1	7.577	3.529	.063	.029
group	141.337	1	141.337	65.834	.000	.358
Error	253.330	118	2.147			
Total	28456.250	121				
Corrected Total	400.000	120				

a. R Squared = .367 (Adjusted R Squared = .356)

4.4 Findings of Study 3

4.4.1 Findings of the Survey on Writing Prompts of TOEFL, IELTS and TEM4

In order to answer RQ 8 in this study, i.e., what patterns do the writing prompts of

three high-stakes tests, IELTS, TOEFL, TEM4, display in terms of rhetorical function

and object of enquiry, six sets of 20 (n=120) recent writing prompts from the three

tests were gathered for content analysis. The rhetorical function and object of enquiry of the prompts were coded and categorized following a scheme developed by Moore and Morton (2005). The analyses of the essay writing prompts and non-essay writing prompts were reported separately in the following.

4.4.1.1 Analysis of Essay Writing Prompts

For the 60 essay writing prompts, seven categories of rhetorical function were generated from the coding, with evaluation (44%) and hortation (23%) being the two largest categories. In regard to the object of enquiry, the 60 essay writing prompts covered ten fields of topics, with education and technology accounting for approximately 50 percent of the total. Table 4.13 shows the frequencies of objects of enquiry mapped against rhetorical functions as percentages.

Table 4.13

Frequencies of objects of enquiry mapped onto rhetorical functions in essay writing prompts of the three tests

Rhetorical Function	Objects of enquiry (no. of occurrences in parentheses)		
	IELTS (n=20)	TOEFL (n=20)	TEM4 (n=20)
Evaluation	11 (55%)	11 (55%)	4 (20%)
	Education (5)	Education (3)	Environment (1)
	Technology (3)	Technology(3)	Inter-personal
	Globalization (1)	Social issue (2)	relationship (1)
	Quality of life (1)	Inter-personal	Living style (1)
	Social issue(1)	relationship (1)	Volunteering (1)
		Quality of life (1)	
		Travel (1)	
Hortation	4 (20%)	8 (40%)	2 (10%)

	Career (1) Environment (1) Social issue (1) Technology (1)	Education (4) Inter-personal relationship (2) Environment (1) Technology (1)	Education (1) Environment (1)
Prediction			(2/18 – 10%) Inter-personal relationship (1) Other: Career (1)
Comparison	1 (5%) Technology (1)	1 (5%) Education (1)	3 (15%) Education (2) Other: Media (1)
Explanation	3 (15%) Education (1) Technology (1) Social issue (1)		3 (15%) Technology (1) Travel (1) Volunteering (1)
Recommendation	1 (5%) Social issue (1)		3 (15%) Education (1) Environment (1) Living style (1)
Description			3 (15%) Education (1) Living style (1) Other: Culture (1)

Figure 4.1 illustrates the breakdown of rhetorical functions in essay prompts of the three tests. Table 4.14 includes examples of actual prompts for each rhetorical function. Similar to Moore and Morton (2005), evaluation (44%) and, to a lesser extent, hortation (23.3%) were the most common rhetorical functions among the 60

prompts. Others were explanation (11.7%), comparison (8.3%), description (6.7%) recommendation, (6.7%) and prediction (3.3%). Some notable differences between the three tests included the TOEFL's almost complete focus on evaluation and hortation when compared to the other two tests, which had a greater variety of rhetorical functions. Also, TEM4 was the only test with prompts demanding a descriptive response.
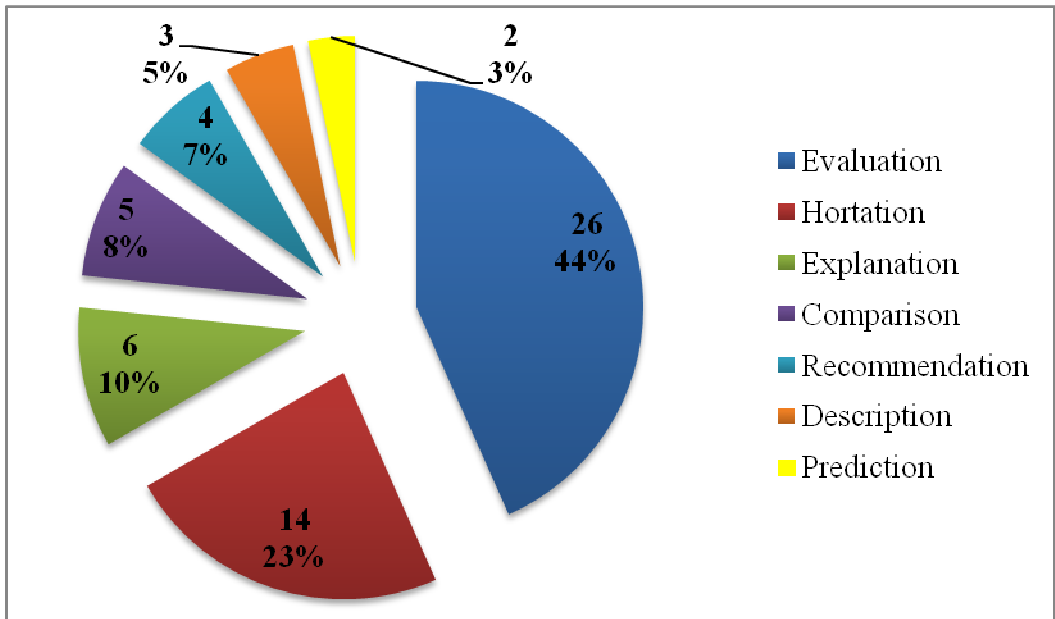
Figure 4.1 Rhetorical function of essay writing prompts

Table 4.14

Example essay writing prompts for rhetorical functions

Rhetorical function	Example Prompt
Evaluation	Using a computer every day can have more negative than positive effects on your children. Do you agree or disagree?
Hortation	Governments should make more effort to promote alternative sources of energy.

	To what extent do you agree or disagree with this opinion? Give reasons for your answer and include any relevant examples from your own knowledge or experience.
Explanation	Illiteracy has traditionally been viewed as largely a third world problem. However it is becoming apparent that in countries such as the USA and Australia, illiteracy is on the increase. Discuss possible causes for this and its effect on society.
Comparison	Some people think that children should begin their formal education at a very early age and should spend most of their time on school studies. Others believe that young children should spend most of their time playing. Compare these two views. Which view do you agree with? Why?
Recommendation	Countries such as China, India and Japan have unsustainable population growths. In fact many experts are of the opinion that the population 'explosion' which is now a very worrying concern, is the most serious threat to life on this planet. Give some suggestions to address this problem
Description	The Dragon Boat Festival is one of the important national festivals in China. Write … a composition of about 200 words on the following topic: The Dragon Boat Festival
Prediction	Nowadays young people tend to phone more often

than write to each other. So, some say that phones will kill letter writing. What is your opinion? Write … a composition of about 150 words on the following topic:

Will Phones Kill Letter Writing?

Figure 4.2 shows the breakdown of the objects of enquiry in essay writing prompts. Notable similarities between the three tests include the repeated instances in all three tests of the following themes: education, technology, social issues, the environment, interpersonal relations, living style, quality of life, travel, volunteering and globalization. Remarkable among the 60 prompts were the goal instructions, of which over half (n=33) contained variations on the question, "Do you agree or disagree?" invoking a claim-taking position. Figure 4.3 lists the common goal instructions. While most of these goal instructions are meant to produce a response that is evaluative in nature, a few prompts lay outside of this function.
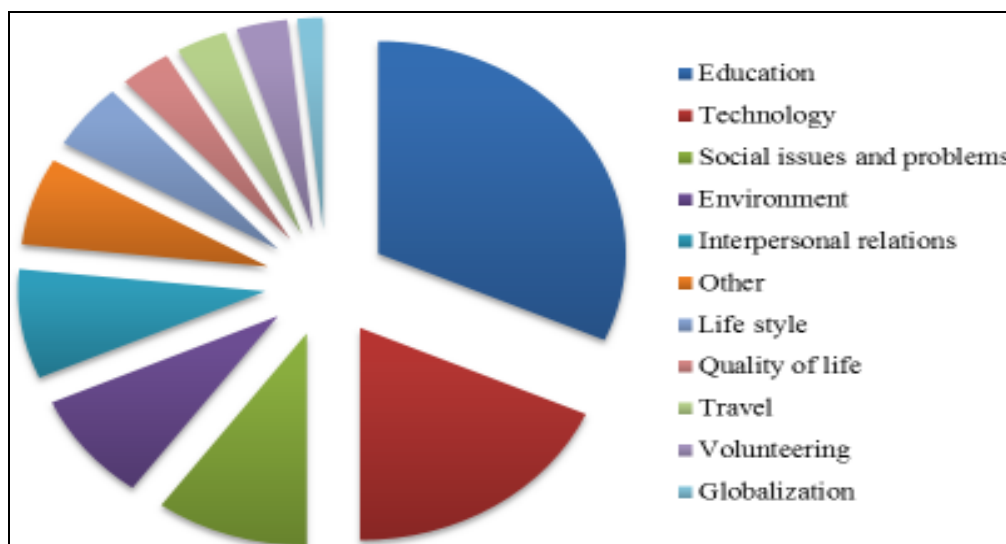
Figure 4.2 Objects of enquiry in essay writing prompts
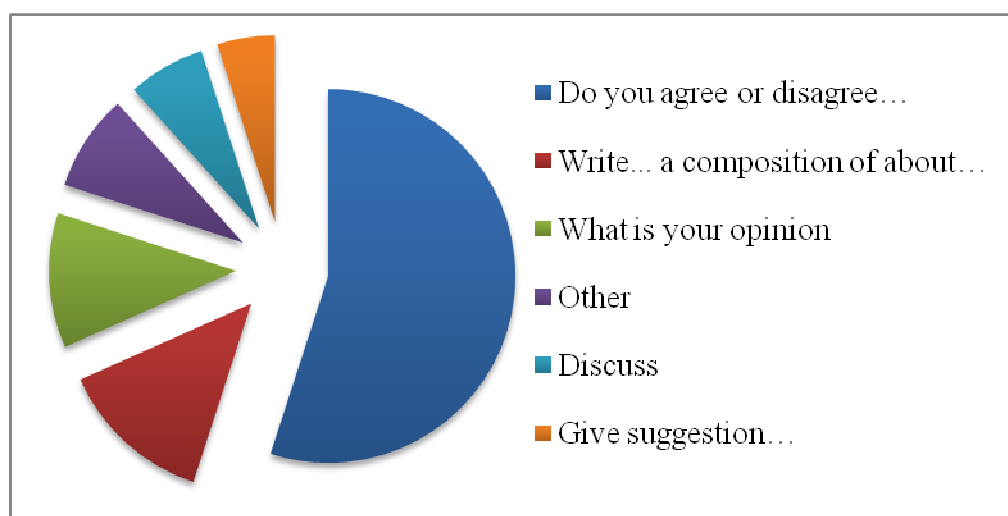
Figure 4.3 Keywords and directions in essay writing prompt goal instructions

Tangential to both rhetorical function and object of enquiry is the notion of agency, in other words, who are the responsible parties for taking action (if any) assumed in the prompts? Figure 4.4 shows that over half of the prompts involved an assumed responsible agent. For example, a prompt from IELTS concluded with "all cigarette advertising should be banned. Discuss." Clearly, the government is the agent here. Those where no agent was involved are typified by the following prompt: "…there is no immediate or predictable likelihood of machines taking over [translation] from humans. Do you agree or disagree?"
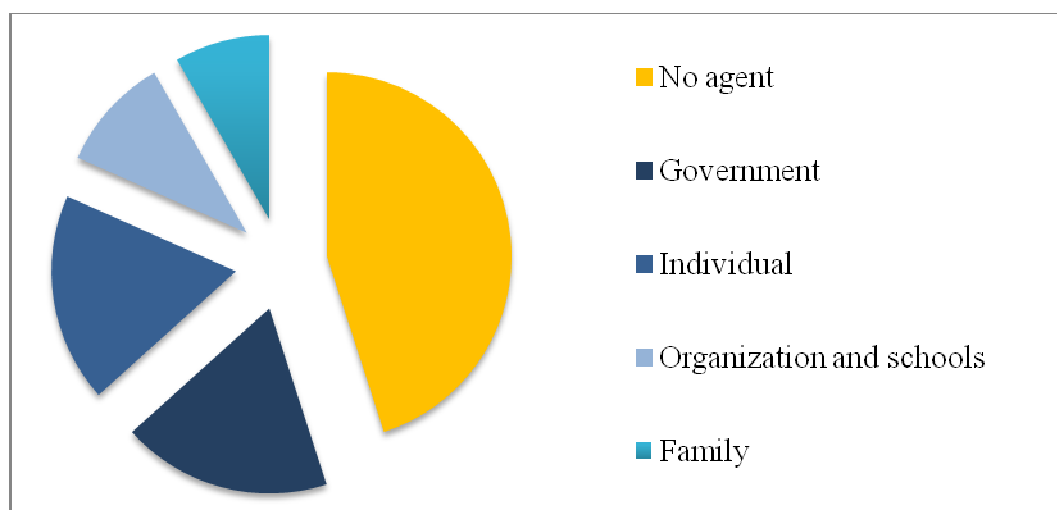
Figure 4.4 Assumed agents in writing essay prompts

4.4.1.2 Analysis of Non-essay Writing Prompts

As to the non-essay prompts, in terms of rhetorical function, the IELTS, TOEFL and TEM4 non-essay prompts required description, summarization and note-writing responses respectively and exclusively. As for the object of enquiry, IELTS non-essay prompts were classified into only one category – socio-cultural. The classification of TOEFL Task 1 prompts had one more category in addition to the socio-cultural one – nature. TEM4 note-writing prompts were categorized into one broad category - socialization, including making or declining an invitation, making an inquiry, recommending, thank-you notes, etc.

4.4.2 Findings of the Second Phase of Study 3

In the second phase of Study 3, the experimental and control groups wrote argumentative essays on two writing prompts of a distinct nature respectively. Scripts

were transferred to electronic files. Data analysis of the scripts of both groups comprised two stages. In the first stage, preliminary text analysis was performed via an automated text analyzer for general linguistic features of the data; after the preliminary text analysis, further discourse analysis was conducted for syntactic complexity, lexical diversity and word frequency which are considered as the most predictive linguistic features of argumentative writing quality.

Table 4.15

Results of automated text analysis

Group	Average length (no. of words)	Lexical density	Fog index
Experimental (n=65)	228.35	10.34%	8.6
Control (n=64)	242.11	11.54%	9.8

Table 4.16

Further indices of essay quality for experimental and control groups

Group	Syntactic complexity M (SD)	Lexical diversity M (SD)	Word frequency M (SD)
Experimental (n=65)	4.10(0.34)	75.69(3.98)	3.12(0.02)
Control (n=64)	3.64(0.31)	68.85(5.60)	3.07(0.03)

As shown in Table 4.15, the control group performed at a higher level than the experimental group in average essay length, lexical density and Fog index. However, Table 4.16 indicates that the experimental group outperformed the control group in syntactic complexity and lexical diversity; but in terms of word frequency, the control group writers demonstrated more proficiency because they use words that occur less frequently. The six indices in Tables 4.15 & 4.16 are all meant to indicate the writing quality of the essays, but showing contrasts in scores. Actually, it is possible that an essay has high lexical diversity, but low lexical density (Johansson, 2008). I do not regard this as a disparity caused by a different text analyzing tool. In addition, there are profound reasons caused by the choice of writing prompts in the high-stakes tests and writing classrooms. The ostensible contradiction between text analyzer outputs and Coh-Metrix outputs will be discussed further in the next section.

In the second stage, qualitative and corpus analyses were performed focusing on: 1) metadiscourse use, 2) essay organization and 3) lexical items. Table 4.17 shows the use of two categories of metadiscourse, namely, textual metadiscourse and interpersonal metadiscourse by both the experimental group and control group.

Table 4.17

Descriptive statistics of metadiscourse use (per essay) for both experimental and control groups

	Group				
	Experimental (N=65)			Control (N=64)	
Variable	M	SD		M	SD
Textual metadiscourse					

Transitive markers	2.80	1.51	3.05	1.36
Frame markers	2.30	1.84	1.71	1.45
Code glosses	0.30	0.47	0.30	0.46
Interpersonal metadiscourse				
Hedges	1.66	1.43	0.85	0.93
Attitude markers	1.60	0.75	1.25	0.79

As Table 4.17 indicates, the experimental group used significantly more hedges in their essays ($p<.05$). Hedges that were used in the experimental group essays include: may, might, maybe, perhaps, seem to, to some degree, personally. Hedges used by the control group include: maybe, to some extent, in a way. The experimental group also used more frame markers and attitude markers. However, the control group used more transitive markers. The means of code glosses used by the two groups were the same, and these were the least frequently used among the five subcategories of metadiscourse by both groups.

The organizational structure of the essays was also compared by noting the number of paragraphs in each script and reading the first three sentences of the first paragraph and comparing the two groups. For example, we noticed a marked contrast in the way participants began their essays. Customarily, a writer begins an argumentative essay with a preamble of a few sentences to introduce the topic. In this study, we looked at the first three sentences of each essay to code this lead-in. It was found that in 75% (49 out of 65) of the experimental group essays, the topic was introduced without any preamble, e.g., "In my opinion, I think there are three reasons about this phenomenon;" in contrast, in the control group, 52 out of 64 essays (81%) began in a less straightforward, and arguably, more sophisticated way by hooking the reader with

the provision of some background information, e.g.,

Recently, there is a heated discussion about whether urbanization makes

our life better. As we all know, China is developing very fast. As a

consequence, the rate of urbanization is increasing. This changes our lives

greatly. In my opinion, urbanization makes our life better in many aspects.

As for the organizational structure of the scripts, in the control group, 98% of the essays (n=64) followed a three-part argumentation model, namely, an introduction-body-conclusion model. In almost all cases, the writer's claim was presented in the introduction and this claim was supported by evidence in the body. In the conclusion, the claim was restated to a greater or lesser extent. For the experimental group, however, the three-part structure was adopted in only 15% (n=10) of the essays. Sixty percent (n=39) included two or three paragraphs with no introduction or conclusion, and the remaining 25% (n=16) either lacked an introduction or a conclusion.

Beyond the analysis of the organizational structure of the scripts, a corpus analysis was conducted. Using the AntConc software (Anthony, 2012), the "Top 100 lexical words" and "Top 100 functional words" were generated for both groups. Tables 4.18 and 4.19 respectively show some of the more notable differences and similarities in the two groups leaving aside those words that were directly related to the content of the two prompts, e.g., "daycare" and "urbanization." These differences or similarities were decided by the frequency with which a word was used rather than the rank of the word in the corpus analysis list. For instance, the word "think" ranked 10th and 12th

in the experimental and control groups respectively. However, the number of times they were used in students' essays was very different. "Think" was used 115 times in the experimental group compared to 67 in the control group. As to the context in which the word was used, the concordance revealed, there was also significant variation. Seventy-seven percent of the control group's usage of "think" was to express personal opinions or ideas, e.g., "I don't think urbanization is making our life better." For the experimental group, however, notably only 45% of cases were used in this way, while 55% of the instances of "think" were used in an attributive sense, i.e., ascribing other people's thoughts or beliefs, e.g., "this would lead to a misunderstanding of parents who think that money will solve everything." Thus, the word "think" was used by a significantly higher percentage of students in the experimental group (p<.05) to project thoughts onto others.

Table 4.18

Selected lexical differences between experimental and control groups

Word type	Experimental Group			Control Group	
	Word	Rank	Frequency	Rank	Frequency
Lexical	think	10	115	12	67
Functional	if	25	88	49	33
	should	27	82	36	47
	first	40	41	58	23

Table 4.19

Selected lexical similarities between experimental and control groups

Word type	Experimental Group			Control Group	
	Word	Rank	Frequency	Rank	Frequency
Lexical	feel	67	16	84	14
Functional	can	8	211	11	201
	because	39	44	42	42
	however	74	20	56	24

Perhaps the most noteworthy frequency difference was regarding the usage of the word "if." For the experimental group, "if" was the 25th most common functional word and was used 88 times, while for the control group, "if" was used only 33 times, ranking 49th. A rereading of the contexts of sentences where the word "if" appeared in the scripts of the experimental group revealed that in almost all cases (n=87), the writer was presenting hypothetical arguments; in the control group, the frequency of hypothetical arguments with the word "if" (31 cases) or in other forms (2 cases of hypothesis with modals) was much lower.

In this chapter, findings are reported in answer to the nine research questions in the three studies of this project. In the next chapter, these results and findings are interpreted and discussed for prominent issues.

CHAPTER 5

DISCUSSION

In this chapter, key issues emerging from the findings of the three studies in the present project are discussed. In order to organize this chapter into a coherent whole, I determined that the results of the nine research questions in the three studies can be best analyzed and interpreted by forming natural groupings based on their content. This means that, organizationally, it is best not to simply discuss each of the nine questions in sequential order. Therefore, I have organized the sections below according to certain common themes. For example, RQ2 and RQ6 form a natural pair because they concern qualitative findings about the students' perceptions of argumentative writing before and after the intervention. Another example is RQ1 and RQ4; they are juxtaposed in section 5.4 for a discussion about the effect of explicit instruction on students' argumentative writing as well as critical thinking development. Table 5.1 illustrates how each section in this chapter maps onto the research questions and the themes of discussion.

Table 5.1

Sections mapped onto research questions and themes of discussion

Section	Theme	Pertaining Research Question
5.1	Possible impact of the Syllabus, textbooks, high-stakes test writing prompts and classroom instruction on students' argumentative writing and critical thinking	RQ3: What factors contribute to students' performance in and perceptions of argumentative writing? How are counterargumentation and critical thinking represented in these factors?
5.2	Impact of the intervention on	RQ2: What are these students'

		students' perceptions of argumentative writing	perceptions of argumentative writing? RQ6: What are the students' perceptions of argumentative writing after the intervention?
5.3		The impact of culture on the Chinese students' performance in argumentative writing and critical thinking	RQ1: How does a group of mainland Chinese undergraduates respond to a typical argumentative writing prompt? In particular, do they produce any counterarguments and rebuttals? RQ4: How does an intervention in counterargumentation affect students' written argumentation performance in terms of quantity of counterarguments and rebuttals, and overall quality of an argumentative essay? RQ7: How does the intervention impact students' critical thinking?
5.4		Explicit instruction in argumentative writing pedagogy and critical thinking pedagogy	RQ5: Is there a correlation between the extent of counterargumentation and refutation and the overall score of an essay? RQ4 & RQ7
5.5		On the Quasi-experimental Design of Study 2	
5.6		Limitations of Study 2	
5.7		The convergence of writing prompts in high-stakes English tests in terms of rhetorical function and object of enquiry	RQ8: What patterns do the writing prompts in three high-stakes tests, IELTS, TOEFL and TEM4, display in terms of their rhetorical function and object of enquiry?
5.8		The cookie cutter effect of typical writing prompts	RQ9: In what ways do the essays of students responding to a

		"conventional" writing prompt differ from those of an exploratory, non-conventional prompt?
5.9	The role of writing prompts in cultivating students' argumentative writing and critical thinking abilities	RQ9
5.10	Limitations of Study 3	

5.1 Representation of Counterargumentation and Critical Thinking in Four Factors and the Possible Impact on Undergraduate English Majors' Argumentative Writing

The first and second research questions of this project inquired into how two groups of students respond to an argumentative writing prompt and how they perceive argumentative writing. Informed by the answers to RQ1 and RQ2, the third research question investigated what factors contributed to students' performance and perceptions, and how counterargumentation and critical thinking are represented in these factors. Questionnaire and interview data revealed that students and teachers considered four factors, namely, the Syllabus, the classroom instruction at college, the composition textbook, the high-stakes writing prompts and rubrics, as having substantial impact on their argumentative writing. Qualitative content analysis (of the Syllabus, textbook, prompts and rubrics) and classroom observation (of instruction in argumentative writing) were then conducted. In this section, I interpret the findings to see how these factors might influence students' written argumentation pertaining to RQ3.

5.1.1 The Syllabus

In general, the salient status of the Syllabus lies in two dimensions: first, it lays out the curriculum and pedagogy for undergraduate English programs in mainland China; second, it lays the foundation for high-stakes tests to refer to. In this sense, it is presumed to be the conceptual source for the other three factors influencing students' argumentative writing. Therefore, the discussion of the influence of the four factors starts with the Syllabus. Students in all likelihood have not read the Syllabus; this may be why most of them listed it as ranking fourth or fifth among the influential factors in the questionnaire. On the other hand, during interviews, the teachers were more aware of how the Syllabus shapes the curricular activities (and ranked it among the top four influential factors). Although they did not directly refer to the Syllabus in their teaching, they fully realized the significant role it plays in the education of undergraduate English majors.

Some researchers contend that the function of a syllabus may be limited because what is written in the document is one thing, but what is actually done is another (Cheng, 1997; Qi, 2005). Other researchers, however, have found that English writing pedagogy for both English majors and non-English majors is under the unified guidance of the syllabi in China (Cheng, 2008; You, 2004; Zou, 2003). Indeed, in the Chinese context, the impact of the Syllabus cannot be underestimated because it is the only state-sanctioned document which regulates the teaching and testing of all courses for undergraduate English majors. In an empirical study, remarkable alignment was found between the Syllabus and the TEM8 test (Zou, 2003). Given the

exam-oriented settings in China and the washback of tests on pedagogy, the impact of the Syllabus on the students' argumentative writing merits attention. The implication is that even if the Syllabus fails to mandate directly what happens in the classroom, it exerts an impact on the pedagogy via test washback.

A noticeable contradiction in the Syllabus is that it stipulates critical thinking development as one of the primary educational goals for undergraduate English majors. However, it fails to provide specified guidance as to the implementation of critical thinking in the curriculum. Critical thinking is mentioned in a generalized fashion in three sections, General Objectives, Curricular Arrangement and Pedagogical Principles, but it is missing in the sections of Instructional Requirements, Pedagogical Methods and Assessment in the Syllabus. The lack of clear direction with regard to implementing critical thinking skills in the curriculum is understandable given the widespread interpretations of the term (Stapleton, 2011). In some senses, "critical thinking" is a term whose meaning may often be assumed as self-evident, when in fact, as explained in Chapter 2, there are quite specific definitions, some of which may not be self-evident at all. This may explain in part the absence of counterargumentation and critical thinking elements in the textbook, writing prompts and rubrics and classroom instruction, which will be discussed in more detail below.

Another area that warrants attention is that the Pedagogical Method section in the Syllabus, instead of directing teachers to incorporate critical thinking in their pedagogical methods, encourages students to participate in extracurricular activities

that will contribute to their "thinking ability and creativity" (the Syllabus, NACFLT, 2000, p. 13). This may further discourage teachers from including critical thinking in their classroom instruction. In the case of English writing courses, the consequences are that the writing classroom has not been utilized as an arena of higher order thinking. As noted by researchers, however, English writing courses at Chinese universities are characterized by exam-oriented training of short compositions (Li, 2007), which focuses on form rather than content, and memorizing model essays (You, 2004). These indicate a reality distant from the prescription in the General Objectives of the Syllabus.

To sum up, there is some representation of critical thinking in the Syllabus but specific direction of carrying it out in the classroom is lacking. This might account for the disparity found in the writing classroom practice and the Syllabus speculation. Because of the absence of more specified discourse in the Syllabus, the Syllabus to some extent has failed to fulfill its role of guidance in pedagogy for courses including the English writing courses. Nonetheless, the impact of the Syllabus is also manifested through the writing assessment which will be discussed in the following section.

5.1.2 The Writing Prompts and Rubrics of TEM4 and TEM8 and Their Washback on Argumentative Writing Pedagogy

The TEM4 & TEM8 are criterion-referenced tests designed in line with the stipulation in the Syllabus, but the writing prompts and rubrics of both tests were

found to have insufficient representation of critical thinking in that they do not encourage counterargumentation. This section discusses how this insufficient representation may have influenced the writing pedagogy which eventually shaped students' performance on argumentative writing. In the Syllabus, it is stressed that the development of logical thinking and creative thinking shall be strengthened (NACFLT, 2000) and, according to the Pedagogical Principles in the Syllabus, students should be trained to evaluate, analyze and synthesize, and to "approach an issue from multiple perspectives" (the Syllabus, NACFLT, 2000, p 12).

Taken literally, "critical thinking" is not mentioned by name here. However, critical thinking is embedded in logical thinking (Norris & Ennis, 1989). More importantly, "to evaluate, analyze and synthesize" are key elements in various definitions of critical thinking (e.g., Fisher & Scriven, 1997; Norris & Ennis, 1989). "To approach a matter from multiple perspectives" implies that one should consider not only one's own view, but also alternative views. In this sense, "to approach a matter from multiple perspectives" encompasses counterargumentation. As such, this statement, underscoring the importance of approaching matters from multiple perspectives, is at the heart of Study 2 the main purpose of which was inquiring how an intervention in counterargumentation impacts students' argumentative writing and critical thinking. Although TEM4 & 8 are designed in accordance with the educational goals specified in the Syllabus, and the Syllabus highlights the significance of cultivating critical thinking, as revealed in the findings, the writing prompts and rubrics evidently do not encourage counterargumentation which is considered a key element of critical thinking in an argumentative essay. In essence, the intervention in Study 2 was

designed to highlight this contradiction.

This neglect of counterargumentation (and subsequent critical thinking) is particularly manifest in the goal instructions of the TEM4 & TEM8 writing prompts that specifically request support for only viewpoints taken, as opposed to engaging in a discussion of alternative opinions. In effect, the complete absence of counterargumentation noted here makes it unsurprising that the TEM4 & TEM8 rubrics also include no descriptors for rewarding the mention and refutation of alternative viewpoints.

The impact of the TEM4 & TEM8 prompts on students' argumentative writing is mainly exerted through the washback of the tests on teaching and learning. Washback can be "counterproductive in terms of student learning" (Weigle, 2002, p. 54), and this may well be the case with the writing prompts in TEM4 & 8. However, the flipside of this negative washback is Ferretti et al.'s (2009) assertion that "[g]oals may affect performance when they provide clear direction about what needs to be included in the essay" (p.577). Unfortunately, in the case of the present study, it is only the negative washback effect of the goal instructions in TEM4 & TEM8 writing prompts that may be leaving an impression on students, i.e., that an argumentative essay is solely evaluated on the merits of how well the writer's position is backed up with data.

The goal instructions in TEM4 & TEM8 prompts may be having an influence on the way in which teachers instruct students in argumentative writing. When an exam's

goal instruction is confined to claims and data, without any mention of counterargumentation, the impact on pedagogy can be considerable, especially in an exam-oriented society such as the Chinese one. Therefore, one conclusion that can be drawn from the findings is that there is a need to introduce counterargumentation into the goal instructions and grading rubrics of the TEM4 & TEM8. Here, I would contend that if goal instructions that encouraged counterargumentation were introduced in the TEM4 & TEM8, they would most likely result in enhanced persuasiveness in students' written responses. Moreover, given the high-stakes nature of TEM4 & 8 for English-major graduates in China, such a change could have a powerful washback effect on the teaching of argumentative writing in China's undergraduate classrooms. Such a change could also enhance critical thinking, one of the stated objectives of the national curriculum (the Syllabus, NACFLT, 2000), by way of encouraging the consideration of alternative viewpoints.

It should be noted at this point that counterargumentation and critical thinking should not be conflated with each other. Counterargumentation by way of its consideration of alternative viewpoints is only one aspect of critical thinking; other aspects of critical thinking include analytical skills for problem-solving as well as dispositional propensities such as open-mindedness. However, these skills and propensities are best developed within a mindset of viewing issues from multiple perspectives. Furthermore, it was found that counterargumentation was lacking in Chinese students' argumentative essays. This therefore suggests that counterargumentation is the aspect of critical thinking that should be underscored among Chinese students.

The rhetorical function and object of enquiry of TEM4 & 8 writing prompts also have remarkable bearings on students' critical thinking development, which is the inquiry focus of Study 3 of this project. Therefore, the issue is discussed in sections 5.8 and 5.9.

5.1.3 The Composition Textbook and the Classroom Instruction

Since the composition textbook and classroom instruction are closely interrelated, they are discussed jointly in this section. While the results of the classroom observation suggest that standard instruction in argumentative writing neglects the teaching of counterargumentation and refutation, this finding may extend to undergraduates in mainland China in general. Corroborating this is the fact that Ding et al.'s (1994) book, which devotes minimal attention to counterargumentation, has been used as the standard course book for English majors at scores of universities in China. Further, in her study comparing online instructional materials on argumentative writing for Chinese and American school writers, Liu (2005) notes that, in contrast to American materials, Chinese materials do not treat "anticipating the opposition" as a crucial element in argumentative essays. I do not intend to exclude the possibility that certain teachers go out of their way to introduce the concept of counterargumentation in their classes; however, given the exam-driven context in China, these cases are probably few.

You (2010b) noted that the English composition pedagogy at mainland Chinese universities was uniform due to the implementation of the nationwide Syllabus and

the concurrent use of composition textbooks. He claimed that Ding et al.'s (1994) textbook was "most widely adopted" in colleges (p, 119). In another article, You (2004) found that composition teachers at Chinese universities were left with few choices but to cater for students' need to pass exams. The pedagogy was, therefore, examination-oriented and old-fashioned (Gu, 2013; You, 2004). The findings in this project align in part with You's (2010b) observation. Based on my content analysis of the textbook and findings from the classroom observation, I found that both the composition textbook and the classroom instruction were not training students' counterargumentation skills in their argumentative writing, and as a result, critical thinking skills training was also insufficient. As mentioned above, the ability to counterargue is only one attribute of a good critical thinker, but it is the one that should be highlighted in Chinese students' argumentative writing. Mapping the findings of this study onto what has been found by You (2004, 2010b) and Gu (2013), one conclusion that can be drawn is that the deficiency of critical thinking elements in argumentative writing instruction is likely to be nationwide, and the reason is probably the washback from high-stakes writing tests.

This project uncovered some further evidence for the deficiencies noted here. The project involved six composition teachers who enabled me to obtain both formal and informal data as well as direct and indirect data pertaining to the role of the instructor in this issue. These data suggest that teachers' lack of proper training in counterargumentation might be another factor contributing to students' argumentative writing performance. During the classroom observation, my conversations with the two teachers indicate that the instructors themselves were ignorant about

counterargumentation. This was also the reason why intensive training was applied to Angel (the instructor of both the experimental and control groups) before the instructional intervention in counterargumentation. Based on these data, it is suggested that formal training is needed for teachers in order to gain both knowledge about and skills in teaching counterargumentation. If the teachers themselves do not realize the significant role of counterargumentation in argumentative writing and critical thinking cultivation, and if they do not grasp counterargumentation skills, they will certainly have difficulty teaching it.

5.1.4 Limitations of Study 1

In examining the influence of textbooks on students' perceptions of and performance in argumentative writing, a survey was first conducted to uncover the use of composition textbooks in a dozen of universities in mainland China. However, the content of only one composition textbook was intensively analyzed. Although justifications have been given concerning this decision, the analysis of one single textbook is evidently a limitation for this study, along with other limitations, e.g., the translation of the Syllabus done by myself instead of an accredited professional. In future studies, both qualitative and quantitative content analysis can be conducted on several composition textbooks most adopted in Chinese universities to obtain a panorama of the textbook impact in this issue.

5.2 Impact of the Intervention on Students' Perceptions of Argumentative Writing - Implications from the Questionnaire and Interview Data

This section discusses the findings of the two research questions (RQs 2 & 6) that dealt with the students' perceptions of written argumentation. As suggested by the findings, the intervention in counterargumentation enhanced students' confidence in written argumentation. It also motivated them in terms of critical thinking.

5.2.1 Counterargumentation Skills Enhance Students' Confidence in Written Argumentation

By comparing the results of the questionnaire administered before and after the intervention, it was found that after the intervention the proportion of students who perceived argumentation as the most challenging type of writing in the experimental group decreased more significantly than that of the control group ($p<.05$). My interpretation of this finding is that the intervention in counterargumentation resulted in more confidence of students in argumentative writing. An observed common phenomenon is that when students have to translate their ideas and thoughts into written arguments they tend to feel frustrated, even if they can argue competently in a real conversation (Graff, 2003). Many educators contend that argumentation is the most difficult type of writing at school (Gleason, 1999). The complexity of written argumentation may have deprived students of their confidence in this type of writing because it requires students to skillfully shift from the dialogue situation to a monologue one (Gleason, 1999). In addition, it is difficult to organize an

argumentative essay "conceptually and structurally" (Gleason, 1999, p. 81). Indeed, researchers have reported poor performance of both L1 and L2 students at varied levels on argumentative writing (e.g., Applebee, Langer, Jenkins, Mullis, & Foertsch, 1990; Applebee, Langer, Mullis, Latham, & Gentile, 1994; Gleason, 1999; Qin & Karabacak, 2010; Varghese & Abraham, 1998). Students' lack of confidence in written argumentation may be both the consequence and the cause of the reported poor performance. An intervention in counterargumentation skills seem to have facilitated the writing of students of the experimental group, and given them some confidence in coping with written argumentation. In particular, counterarguing and refuting somehow may have helped students break the constraint of seeking evidence for only one side. As writers, they learned a fundamental skill that can strengthen their own position by acknowledging and refuting opposing views. In a sense, the process of counterargumentation may have expanded their mind.

This assumption was affirmed by the data of respondents' answers to the question about the most difficult aspect of English argumentative writing. The proportion of students claiming "finding enough evidence to support the thesis" as their biggest challenge dropped after the intervention for the experimental group, compared to a slight increase in that proportion of the control group. Understandably, if a writer only focuses on their own proposition, their thoughts are probably confined to a relatively narrow scope. So when the writer tries to support their proposition they may not be able to generate sufficient arguments either quantitatively or qualitatively. In reality, people hold varied, even conflicting viewpoints about any controversial topic. If the writer's mind is open to accommodating alternative views other than their own, they

actually demonstrate and would eventually benefit from this essential element of argumentative reasoning. In essence, considering other people's views in one's written argumentation and making efforts to refute those views make one's arguments more powerful, and the writer's confidence can be built upon this. Interpretations of the interview data presented below to some extent confirm this.

The follow-up interview data further illustrate the role that counterargumentation skills could play in enhancing students' confidence in argumentative writing. In response to the question "what difficulties have you encountered in achieving persuasiveness in your argumentative writing", informants from the experimental group no longer listed "finding supporting evidence" as barriers to writing good argumentative essays (achieving persuasiveness) in the post-interview. This implies that teaching counterargumentation concepts and skills to students can be compared to opening up a window to the outside world for them, which inspired them to furnish (at least in their own opinions) better responses. In fact, as the post-interview data revealed, when discussing how to make their argumentation more persuasive, the informants from the experimental group highlighted the mastering of skills in counterargumentation, while the control group still stressed the importance of adequate supporting details.

This analysis also dovetails with my own personal experience. When I taught the English writing course to undergraduates at a mainland Chinese university, I often heard students complain that in responding to an unfamiliar topic of impromptu argumentative writing, the biggest headache is that their mind turned blank. This

personal experience was underpinned by the interview responses from student participants in this project.

5.2.2 Counterargumentation Practice Stimulating Critical Thinking among Students

In response to the second interview question, i.e., what difficulties have you encountered in achieving persuasiveness in argumentative writing, before the intervention the informants from the experimental group were concerned about "expression", "supporting details" and "structure of argumentative essays" as the encountered difficulties in making an argumentative essay more persuasive. However, after the intervention, their responses to the same question were categorized into "expression" and "logical thinking". This change suggests that the intervention in counterargumentation might have been effective in helping students strengthen their thesis because they may have gained more strategies to support their position by either finding sufficient supporting details or refuting others' arguments, and that the intervention materials, especially the diagram of argumentation schema, might have helped their understanding of the proper structure of an argumentative essay. More importantly, the intervention helped students shift their attention to a higher level of thinking. Notably, students in the experimental group began to think about their own thinking processes after the intervention. The mere occurrence of such a process suggests that the intervention to some extent motivated critical thought among students. However, for the control group, after 12 weeks of instruction in argumentative writing, students spared no thought for critical-thinking-related

problems. For example, their concerns were confined to sufficiency of evidence and correct expression.

It should be stated here that, the effect of a 12-week intervention is limited; such a short amount of time can only introduce counterargumentation as a set of skills and stimulate students to think more about informal reasoning. In this sense, the completion of this project was just a start towards critical thinking education via the argumentative writing classroom. As such, the counterargumentation intervention in this project can be regarded as an endeavor to promote "education for thinking" (Kuhn, 2005, p12). Education for undergraduate English majors at mainland Chinese universities has been criticized as overemphasizing the instruction on language proficiency while neglecting the education of thinking (Huang, 1998; WTO Entry and Educational Reform and Development of Foreign Language Disciplines in Chinese Universities Project team, 2001). As a consequence, undergraduate English majors generally demonstrate a deficiency in critical thinking. Furthermore, little research, let alone curricular endeavors, have been made to address this issue (Wen et al., 2010).

The significance of this research project lies in highlighting this deficiency while concurrently advocating "education for thinking" through writing pedagogy. Evidently, writing is a process of engaging in ideas both for and against one's position. As such, the writing classroom should be an arena for integrating writing, reasoning and meaningful learning (Bean, 2011). Among different genres of writing, argumentative writing is most closely related to reasoning. In fact, argumentation is

perceived as equivalent to informal reasoning (Voss and Means, 1991) and a critical thinker must be a skillful reasoner. Therefore, teaching students sound and effective argumentation schemata which incorporates the basic schema and an elaborated one, as has been done in this project (see a diagram of argumentation schemata in Appendix 11), may eventually facilitate students' critical thinking development.

Despite its efficacy in enhancing students' critical thinking ability, the intervention seemed less effective than normal instruction in improving students' language proficiency. Before and after the intervention, interview informants in both groups mentioned their difficulty in expressing their thoughts in English. This suggests that the intervention in counterargumentation seemed to have brought no gain vis-a-vis the English language proficiency of students. For the experimental group, the proportion of students who considered their language proficiency as the most difficult aspect of argumentative writing slightly increased after the intervention, while that of the control group slightly dropped. This might be because typical instruction in argumentative writing pays more attention to the language, but less to the writing process and thinking skills. In effect, it is known that the improvement of language proficiency is a slow process, and normally cannot be achieved within a few months.

This section discusses how the intervention in counterargumentation influenced students' perceptions of argumentative writing. It is suggested that counterargumentation skills improved students' confidence in written argumentation and further stimulated them in critical thinking. However, students' performance and perceptions of argumentative writing and instruction are also shaped by a more

complex factor – culture. The next section provides a discussion of this issue based on both literature and findings from this project.

5.3 Culture and the Argumentative Writing and Critical Thinking of Chinese Students

In this section, the influence of culture on Chinese students' argumentative writing and critical thinking will be discussed separately as it was in the Literature Review chapter. While argumentative writing and critical thinking are closely interrelated, and argumentative writing is a product that embodies the process of critical thinking, they are distinct by nature. The fact that argumentative writing may be viewed as a concrete entity while critical thinking is abstract determines that researchers must adopt different ways to deal with them. Therefore, section 5.3.1 covers the influence of culture on students' argumentative writing and section 5.3.2 discusses the impact of culture on students' critical thinking ability. In addition, the two sections focus on different dimensions, with section 5.3.1 stressing how students' practice in counterargumentation is affected by the Chinese argumentative tradition, and with the focal point of section 5.3.2 being whether critical thinking is teachable to Chinese students. Additionally, section 5.3.3 discusses the trend of transcultural argumentative writing.

5.3.1 How Chinese Culture Impacts on Students' Argumentative Writing

This section is concerned with results in answer to RQ1, which inquired into how two

groups of students responded to an argumentative writing prompt before an intervention. The pretest scripts revealed that students normally did not include counterarguments and rebuttals in their argumentative essays. The findings are not surprising, and dovetail with another study on Chinese undergraduates' argumentative writing (Qin & Karabacak, 2010). The subsequent research question (RQ3) further inquired about what and how factors contribute to students' performance on argumentative writing, the findings of which are discussed in 5.1. This section discusses a more complex factor – culture – which plays a role in written argumentation (Mercier, 2013). This factor was not explicit in any of the research questions, yet I contend it is an implicit contributory factor in argumentative writing and critical thinking of students. In addition to the effect of textbooks, high-stakes writing prompts and classroom instruction, the fact that Chinese students do not counterargue in their argumentative essays might also be due to cultural influence. In discussing the impact of culture on the argumentative writing of undergraduate English majors in mainland China, I mainly focus on what role Chinese culture plays in shaping their practice and perception of counterargumentation. Below I will first discuss some previous research in fields that became relevant as the findings of this study were generated: Chinese argumentative tradition, epistemic beliefs, and contrastive rhetoric, then try to compare and contrast the results of those studies with the findings of this project, in order to shed light on this issue.

5.3.1.1 Chinese Argumentative Tradition and Counterargumentation

One issue warranting consideration is how students' performance on

counterargumentation is related to the Chinese tradition of argumentation. To figure out this, I will focus on two questions: first, do Chinese people have an argumentative tradition; second, does counterargumentation exist in traditional Chinese argumentative writing? A number of researchers have explored the field of Chinese argumentative tradition. Researchers (e.g., Becker, 1986; Cai, 1999; Connor, 1996), especially in the early research on English argumentative essays by Chinese people, were inclined to the assumption that Chinese do not have a tradition of argumentation and debate because Chinese people prefer reconciliation to confrontation. Holding similar assumptions, Peng and Nisbett (1999) observed that in Chinese culture, people preferred a dialectical way of arguing, i.e., seeking a middle way between two opposing views instead of counterarguing and refuting. In another article, Nisbett, Peng, Choi and Norenzayan (2001) went deeper into this issue. They argued that people of East Asian culture tend to hold a holistic view of the world – regarding the world as an interrelated whole, in contrast with the Western analytical view – attending to the individual entity. This holistic cognition lays the foundation for the dialectical reasoning of East Asians, and the tradition of avoiding conflicts shapes the way of argumentation and rhetoric of East Asians (Nisbett et al., 2001).

Nonetheless, more recent research advances evidence that tells another story. Via the scrutiny of traditional Chinese rhetoric, especially policy essays, You and Liu (2009) and You (2010a) claim that Chinese do have a tradition of argumentation. You and Liu (2009) asserted that there were three genres of the expository and persuasive writing in ancient China: policy essays, eight-legged essays and discourse essays. Policy essays were less familiar to foreign researchers. Through rhetorical analysis of

classical policy essays, You and Liu (2009) identified essential features of argumentation and concluded that people under Confucian culture "love to argue" (p. 41). Similarly, Mercier (2013) contends that there exists universality of argumentative reasoning in both Eastern and Western cultures. In particular, Mercier (2013) ascertained elaborated forms of argumentation in the traditional Chinese rhetoric.

More pertinent to the focus of this project is whether counterargumentation elements exist in traditional Chinese argumentative writing. At least one researcher affirmed such existence, referring to two well-known Chinese writings. You (2010a), after scrutinizing *Discourses on Salt and Iron*[17], contended that the consideration of other people's thoughts and opinions was a commonplace practice in traditional Chinese political debates. He assumed that this inclusive attitude was developed to meet the needs of the Chinese society - a multi-ethnic, multi-cultured one. In addition, You cited a book by Chinese philosopher Mencius, who used a large number of counterarguments and rebuttals when debating political issues, as an example of Chinese argumentation tradition. However, it should be noted that what You (2010a) analyzed was a transcription of a famous oral debate over a political issue in ancient China. As to the book, *The Mencius*[18], which was a dictation of Mencius' talk by his students, it was written by means of question-and-answer, recording Mencius' conversations with kings of the time. Therefore, *The Mencius* mainly comprises long dialogues of argumentation which include counterarguments and rebuttals. The two

---

[17] A famous book mainly comprising records of a court debate in 81 BCE, which reflects the political and economic conditions of the Western Han period in Chinese history.

[18] Some Chinese scholars think that Mencius himself wrote the book or at least participated in the writing of the book.

classical Chinese writings might constitute evidence of counterargumentation elements in traditional Chinese argumentative writing, but both of the writings are records of oral debates, which distinguishes them from typical written argumentation.

5.3.1.2 Epistemic Beliefs and Counterargumentation

The presence of counterargumentation may not be rare in traditional Chinese argumentative writing. However, when it comes to students' written argumentation in the present time, the absence of counterargumentation is evident. The students' epistemic belief may be contributing to this absence. A person's epistemic belief is influenced by the cultural context, which in turn impacts on their argumentation (Siegel, 1999). Researchers (e.g., Mason & Scirica, 2006) found that people's epistemological understanding was a significant predictor of their production of arguments, counterarguments and rebuttals when presented with controversial topics. An empirical study by Chan et al. (2011) revealed that in Chinese Confucian culture, people tend to accept established views and knowledge; hence, they usually do not have the mindset to perceive an issue from alternative angles. Chan et al. (2011) maintained that this might illustrate the lack of counterargumentation in Chinese students' argumentative writing.

5.3.1.3 Argumentative Reasoning as Reflected in Contrastive Rhetoric

Researchers in contrastive rhetoric have investigated the characteristics of expository and persuasive writing in Chinese culture. These findings may not be related to

students' counterargumentation ability, but are concerned with students' reasoning in written argumentation. The findings of early contrastive rhetoric research could be generalized into three categories. First, Kaplan (1966) claimed that argumentative essays written by Chinese students followed a rhetorical pattern that is circular and indirect. Second, Chinese writers were claimed to be indirect in presenting their individual opinions, thus lacking a personal voice in their argumentative essays (Shen, 1989). Third, the eight-legged essays were regarded as having greatly impacted the way Chinese people write argumentative essays (Cai, 1999; Kaplan, 1972; Matalene, 1985).

However, the present research yielded contrary results to those assertions. The findings of this project suggest that the aforementioned three traits were not found in participants' scripts. Findings of Study 2 reveal that 83% of pretest scripts have thesis statements that contain a definite personal opinion in the first paragraph, which suggests that even before an intervention in counterargumentation, a majority of Chinese students adopted deductive reasoning and a linear way of thinking in their argumentative essays.

In summary, China has an argumentative tradition that possesses a fairly elaborate argumentative schema. But this tradition does not characterize typical counterargumentation unless in oral debates over political issues. It is likely that the Chinese argumentative tradition is compatible with argumentative traditions in other cultures. Given proper conditions (such as classroom instruction or hints in prompts), the Chinese argumentative tradition could accommodate counterargumentation. In the

text below I will interpret the data to a greater extent based on the research revisited above.

5.3.1.4 Interpretations of Counterargumentation Performance of Participants in This Project

Before applying the aforementioned research findings to the interpretation of the participants' performance on counterargumentation in this project, I consider it essential to make four distinctions when studying argumentation by Chinese students. First, there is a need to distinguish between argumentation for academic purposes and argumentation for other purposes. For instance, there exists a remarkable difference between the argumentation in exam essays and argumentation in political debates. Second, we need to distinguish between written argumentation and oral debates. More sophisticated skills are entailed in written argumentation compared to the oral one. Third, a distinction must be made between the argumentative writing in one's native language and that in English as a foreign language (EFL). Fourth, we also should distinguish between written argumentation in the current social cultural context and that in the past. The Chinese argumentative rhetoric is not static (Stapleton, 2002b), so findings made two or three decades ago may not apply to the present situation. That is also why the results of this project may have important implications for the research on written argumentation in academic contexts of Chinese higher education.

Accordingly, I bear in mind the following three points in the data interpretation

procedure. Firstly, the participants in this project are undergraduate English majors; little of the previous research into Chinese rhetoric has explored the influence of culture on university students' argumentative writing. Even more unclear is the argumentative writing performance of English majors, the education for whom stresses language proficiency, rather than thinking competence. Moreover, in this project, I focused on the impromptu essay, which is perhaps the most adopted form in the academic context, especially in high-stakes exams. Another focus of this project is written argumentation. It is probable that students have mastered counterargumentation skills in debate contests and in everyday oral argumentation. However, as the data in the pretest reveals, their performance on counterargumentation in argumentative essays is a different story.

Before an intervention in counterargumentation, students in both experimental and control groups basically did not counterargue. I assume that, apart from the direct impact of curricular factors, there should be the influence of culture, which makes it a habitual action not to address the alternative views and refute them. This influence is both direct and indirect. Directly, as has been discussed above, the argumentative tradition and epistemic beliefs peculiar to Chinese may contribute to students' unwillingness to counterargue. Indirectly, culture may exert influence via the curriculum-related documents or activities; that is, the syllabus, textbooks, writing prompts, and classroom instructions are all subject to the particular Chinese cultural context. One contradiction stands out here, namely, the literature reveals that the Chinese argumentation tradition could accommodate counterargumentation. If we assume there is cultural influence on students' argumentative writing, why then was

counterargumentation almost completely absent in students' initial essays? Here I argue that the impact of curricular factors is more proximate than the broad cultural influence. This very fact once again highlights the peculiarity of written argumentation and the necessity of training in counterargumentation skills in argumentative writing instruction.

After receiving the intervention, most students in the experimental group included counterargumentation in their essays. As this part of the data suggests, participants' performance on written argumentation is more influenced by classroom instruction. But this again shows that in traditional culture, students regard the teacher as the authoritative knowledge transmitter, and therefore follow the classroom instruction. In addition, the post-treatment findings imply that counterargumentation is not in conflict with the Chinese argumentative condition. Otherwise, students could not have grasped the counterargumentation skills in a fairly smooth manner.

The contrast between the results of the data in this project and the findings of the literature on contrastive rhetoric suggest that the previous findings no longer apply to the current argumentative writing of students. This perhaps is due to the gradual change that is brought to Chinese society by the open and reform policy starting from the early 1980s. With the Chinese society becoming more open to the outside world, English learning has become increasingly important to students and they now have more access to varied sources of English materials, which may eventually have some bearing on their thinking pattern and their argumentative writing structure. As the findings reveal, before the intervention, the structure of the majority of students'

essays demonstrated the traits of linear and deductive thinking. This finding is in line with those of Wu and Robin (2000) and Durkin (2008), which will be discussed in more detail in section 5.3.3.

To conclude, culture exerts both a direct and an indirect impact on students' argumentative writing. In the direct path, Chinese argumentative tradition, which does not emphasize counterargumentation, and Chinese students' epistemic beliefs may have contributed to students' lack of counterargumentation in their essays. By the indirect path, such an impact could be seen through curricular documents and activities which are all subject to the cultural context. However, I should point out that the Chinese argumentative tradition could accommodate counterargumentation given appropriate conditions such as classroom instruction or goal instructions in writing prompts.

5.3.2 How Chinese Culture Might Impact Students' Critical Thinking Ability

The role of culture in critical thinking has been regarded as a very debatable and complicated issue (Manalo et al., 2013a). The impact of culture and culture-related factors on multiple aspects of students' critical thinking has received considerable attention and study in the literature, with varied findings. The data of the three studies, especially Study 2 and Study 3, sometimes echoed and at other times contrasted with the findings of other empirical studies.

This section focuses on the influence of culture on students' critical thinking

competence via interpreting the findings of RQ 7 in an effort to disentangle the myth that critical thinking is unteachable to Chinese students (e.g., Atkinson, 1997). Findings of Study 2 suggest that students of Chinese culture can be trained to write and think critically. Specifically, the findings of Study 2 indicate that an instructional intervention in counterargumentation resulted in significantly more production of counterarguments and rebuttals in students' argumentative essays. In addition, the experimental group had a significant gain in their critical thinking ability after the intervention measured by an established critical thinking rubric. Findings of Study 3 reveal that in students' responses to an experimental prompt, elements of high-level analytical thinking appear more frequently. These findings suggest that critical thinking competence of undergraduate English majors at Chinese universities can be encouraged through both writing pedagogy and prompts. To some extent, these findings constitute evidence against Atkinson's assumption that critical thinking is a non-Asian trait.

Fox (1994) contends that critical thinking which is essential for effective writing is an outcome of American culture. She further claims that few cultures share this way of thinking. Fox's viewpoint is upheld by Atkinson (1997) who maintained that critical thought is more of a social practice than a teachable set of behaviors. These beliefs triggered vigorous debate throughout the 1990s that still resonate today. The present study, using a critical thinking instrument created by an American university and adopted by a number of American universities, revealed that an intervention in counterargumentation in the writing classroom resulted in enhanced critical thinking among undergraduate English majors in a Chinese university. This appears to call into

question the assertion that critical thinking is a social practice unique to Western people. Instead, critical thinking could be improved among non-Western students via pedagogical activities.

In general, the outcomes of Study 2 and Study 3 in this project indicate that critical thinking is a teachable set of skills to Chinese undergraduate students. These findings concur with Benesch (1999) whose research has great significance in critical thinking education for L2 students. According to her, contemporary conjecture of critical thinking should be dialogic critical thinking. Atkinson (1997), Atkinson and Ramanathan (1995), and Ramanathan and Kaplan (1996a, 1996b) share the view that critical thinking acquisition is an unconscious socializing process which takes place during childhood; therefore, critical thinking is an unteachable concept especially after adolescence. However, Benesch (1999) thinks their position is wrong in two respects. First, they blurred the disparity between monologic critical thinking and dialogic critical thinking. Monologic critical thinking is often the basis for U.S. skills-based curricula (Benesch, 1999). The proponents of monologic critical thinking either overlook the consideration of alternative viewpoints, or deal with alternative views at a superficial level. For instance, Ramanathan and Kaplan (1996b) define the third dimension of critical thinking as "the ability to look for hidden assumptions and fallacies in every argument" (p. 226), whereas dialogic critical thinking proposes examining and debating assumptions (Gieve, 1998). Second, Benesch (1999) argues that critical thinking is teachable and should be included in school curricula in both L1 and L2 settings.

It is of considerable importance to differentiate between monologic and dialogic critical thinking in terms of the critical thinking development for L2 learners. Benesch's approach to critical thinking education for L2 learners might help avoid simplistic cultural essentialism regarding this issue. Atkinson and others perceive critical thinking as unteachable based on a narrow perception of critical thinking. It seems more likely for Chinese students to develop critical thinking in a dialogic way. In effect, the outcomes of this project ascertain the possibility and significance of the critical thinking education for university students in China.

5.3.3 The Trend of Transcultural Argumentative Writing (Culture and the Appraisal of Argument Quality)

Argumentation is culturally influenced (Connor, 2002; Kaplan, 1966) and the appraisal of argument quality is said to be culturally specific (Siegel, 1999). However, people in the East are becoming academically westernized and simultaneously tending to take on rhetorical patterns used in the West (Wu & Robin, 2000) including strategies in argumentation (Durkin, 2008). Literature on argumentation research suggests that a shift of the focus has been made from intercultural differences to transcultural writing in the field (e.g., Durkin, 2008; Wu & Robin, 2000). In other words, the disparity between the East and the West in terms of rhetorical patterns and strategies of argumentation may be becoming less significant.

Data yielded in this study support this tendency. After a few weeks' intervention of explicit instruction in counterargumentation, almost all of the experimental group

essays contained proper counterarguments and rebuttals. The word "proper" is used here because in the process of coding and analyzing, those counterarguments and rebuttals that did not conform to the standard were not counted. Generally speaking, the production of sound counterarguments was slightly higher than the production of suitable and logically correct rebuttals. The reasons for this might be that finding exact evidence to refute a counterargument requires a higher level of logical thinking and argumentation skills. The post-interview data provide some triangulation for this. A few students mentioned the difficulty they encountered in refuting an opposite viewpoint.

The organization of the essays also reflected the trend. In the pre-test essays, which reflected students' normal argumentation performance, approximately 83% of students adopted a structure of deductive reasoning, i.e., the thesis statement was made in the introduction followed by supporting evidence (and counterargumentation for the experimental group). In the post-test essays, a higher percentage of essays in both the experimental and control groups exhibited deductive reasoning. The outcomes suggest that there exist some basic elements in argumentation by people in both the East and the West. In addition, with increased globalization, transcultural argumentative writing is likely to become the new focus for researchers.

5.4 Explicit Instruction in Argumentative Writing Pedagogy and Critical Thinking Pedagogy

Study 2 of this project employed an intervention of explicit instruction in

counterargumentation for the experimental group. This design was developed with theoretical and empirical evidence and yielded positive effects in enhancing the counterargumentation skills and critical thinking ability of participants in the experimental group. However, it is worth noting that Study 2 of this project is the first one that I know of that employed an intervention in counterargumentation to investigate into university students' argumentative writing ability in the Chinese context. In the text below, I discuss some literature which justifies the use of explicit instruction in developing students' argumentation and critical thinking ability. These studies either generated findings without an instructional intervention or used an intervention in the L2 context other than the Chinese one. In this sense, the present study was based on a solid theoretical and empirical foundation but has special significance in this area.

As Graff (2003) contended, arguing is a skill that children acquire at an early age, but arguments in the classroom seem so remote from oral arguments with parents or friends that there seems little carryover in these practices. Highlighting the importance of pedagogical practice is an empirical study by Moss (2000) who remarked that "informal literacies do not act as a powerful resource within schooled settings" (p.62). Similarly, the study by Chandrasegaran (2008) suggested that explicit instruction may be desirable for students to exploit their "informal literacies" and achieve a higher level of writing competence. In another empirical study, Goh (2008) maintained that factors contributing to argumentation skills were classified into individual, procedural and social factors. The existence of these factors indicates that students' performance on written argumentation varies widely (Goh, 2008). The

divergence of students' argumentation ability justified further the adoption of explicit instruction in Study 2.

How is the explicit instruction in counterargumentation related to students' critical thinking development? Voss and Means (1991) argued that argumentation is at the core of critical reasoning, emphasizing that instruction in argumentation was fundamental to the development of students' reasoning and thinking skills. But since instructions can be both explicit and implicit, which is the more appropriate way to provide training in good reasoning? In the teaching of argumentative reasoning, there is evidence that explicit instruction is more effective than implicit instruction in both the L1 and L2 context. Many studies (e.g., Osborne, 2010; Su, 2011) have shown that teaching explicitly specific strategies helps make critical reasoning an overt feature of education, hence improve students' reasoning ability.

The results of Study 2 indicate the efficacy of explicit instruction on written argumentation. In particular, the data show that explicit instruction in counterargumentation is essential for Chinese students even at the tertiary level. This finding is aligned with the conclusion of both theoretical and empirical research on argumentative writing. Empirical studies by Zohar and Nemet (2002) and Ferritti et al. (2000) showed explicit instructions were effective measures to train students to reason and to produce arguments and counterarguments in the L1 context. Findings of this study support the assertion that argumentation can be improved through practice in the L2 context. Findings of Study 2 also dovetail with Reznitskaya et al.'s (2007) conclusion that group discussion of controversial topics and explicit instruction in

argumentation schema benefited students' argumentative writing.

The finding that explicit instruction in counterargumentation is effective in the Chinese context may have important implications for the argumentative writing pedagogy at Chinese universities.

5.5 On the Quasi-experimental Design of Study 2

One area that is subject to possible challenge is the quasi-experimental design of Study 2 in this project. It may be argued that since the experimental group received explicit instruction in counterargumentation while the control group did not, the results should be obvious. When a skill is taught in the classroom, the natural expectation is that students will perform better in that skill after the instruction. However, claims in the literature, noted above, express doubt that Chinese students can display critical thinking ability; thus, the findings that a simple intervention enabled students to produce counterarguments, a widely recognized signal of effective critical thinking, are justified. Moreover, these results serve to highlight the present deficiency in the current classroom instruction, which is also reflected in the Syllabus, the composition textbooks, and the writing prompts and rubrics of high-stakes exams.

5.6 Limitations of Study 2

In the spirit of counterargumentation, which is the central focus of this thesis, I would

like to seriously consider a major drawback of the data analysis in Study 2 of this project. When analyzing students' argumentative essays collected from the pretest and posttest, the stress was on the soundness of the argumentation schemata, which was measured in this study by the presence of essential argumentative elements. The downside is that the actual quality of arguments was not considered. In other words, an emphasis on the structural framework of written argumentation has resulted in the neglect of the substance of written argumentation. While the focus of analysis in Study 2 was on frequency counts of arguments, counterarguments and rebuttals, it did not take the actual quality of those elements into consideration. Although the findings of this study suggest that counterarguments and rebuttals enhance the overall quality of an essay, clearly, the persuasiveness of any argument also depends on the intrinsic message being conveyed (Sampson & Clark, 2008; Simon, 2008). In other words, if an essay contains arguments that are inaccurate, lacking in logic or include fallacies, it will not be persuasive, regardless of whether it follows good argumentative form with a full complement of counterarguments and rebuttals. In this sense, the present study succeeds in identifying key features of persuasiveness (counterarguments and rebuttals), but does not cover all its dimensions.

Another note of caution in interpreting the findings of this study concerns the suggestion to include goal instructions that encourage counterargumentation in TEM8. Here, some caution or even more research appears prudent because any additional goal instructions have the potential to result in unintended repercussions. Nussbaum and Kardash (2005), for example, note that in their study, goal instructions demanding counterarguments tended to decrease the number of arguments ("primary

claims" was the term used in that study). While this did not occur in the present study – arguments actually increased – the possibility that arguments would reduce in number or quality remains, and such a concern requires more monitoring and research.

While Study 2 has focused on the goal instructions in the prompts of TEM4 & TEM8, it has not investigated how the actual issues raised in the prompts used in this study or the TEM4 & TEM8, i.e., controversial socio-cultural topics, have a bearing on the argumentation skills of students. It is possible that differing subject matter arouses varying levels of argumentation ability (Stapleton, 2001) and may be having effects on the persuasiveness of students' essays.

These limitations of the study point towards potential future areas of investigation, not only with regard to the TEM battery, but also writing prompts in high-stakes exams in general. While good argumentative form demands the inclusion of supported claims, counterarguments and rebuttals, it is only via a careful review of those features that the actual quality of the support, as opposed to its mere existence, can be assessed. Some studies (e.g. Simon, 2008; Zohar & Nemet, 2002) have made efforts to assess the quality of arguments, but these studies have mostly been investigating the writing of students in the natural sciences. Since the writing prompts in mass-market language tests tend to include only socio-cultural topics, some means to differentiate the quality of arguments in responses is needed.

5.7 The Convergence of Writing Prompts in High-stakes English Tests in Terms of Rhetorical Function and Object of Enquiry

RQ8 concerned whether the writing prompts in IELTS, TOEFL and TEM4 display any patterns in terms of their rhetorical function and object of enquiry and indeed patterns were found. Similar to Moore and Morton (2005), evaluation and hortation appeared most frequently, while explanation, comparison, prediction and recommendation were represented only sporadically. Although it is understood that evaluation and hortation are common functions in academic contexts, the persistent focus on these two functions, to the exclusion of other functions, directs washback into rather narrow confines, and even neglects a stated goal "to facilitate a variety of writing processes" (Cumming, Kantor, Powers, Santos, & Taylor, 2000, p. 9) for one of the tests. Particularly notable from Cumming et al. (2000) is its expressed goal of including tasks of a problem-solution nature: "identify a problem and analyze it and/or /propose a solution to it" (p. 12). Note that the problem-solving skill is considered as an essential element of critical thinking (e.g., Halpern, 1998). And such problem-solution in the writing tasks has important educational implications for critical thinking development. However, this rhetorical function was entirely absent in the TOEFL prompts we examined and seldom apparent in the IELTS prompts, although the TEM4 did display a wider variety of rhetorical functions.

As for the object of enquiry, again a rather narrow set of content areas was generated with over half focusing on just two broad topics (education and technological change), and most having elements of social responsibility of one sort or another often with

agents stated or implicit. This restrictive nature is often mentioned in the literature (Moore & Morton, 2005; Reid & Kroll, 1995; Weigle, 2002), and even in documents attached to the tests. For example, Cumming et al. (2000) list eleven considerations (p. 13) when choosing topics, and IELTS, in their Information for Candidates, states that "[t]he issues raised are of general interest to, suitable for and easily understood by candidates entering undergraduate or postgraduate studies" (p. 5).

Thus, the choice of rhetorical function and object of enquiry are clear challenges for test makers who not only need to consider the schema of those who respond to the prompt, but also the scoring scheme that has to be applied consistently from test to test. Given these limitations, the convergence to an apparent narrow range noted among the 120 prompts is to some extent understandable. Nevertheless, the diverse requirements of academic environments where there is a broad spectrum of disciplines with equally wide rhetorical and content characteristics begs the question whether there is an alternative, and this brings us to our next research question (RQ9).

5.8 The Cookie Cutter Effect of Typical Writing Prompts

The findings in Study 3, especially those concerning RQ9, can be interpreted in two ways. Using standardized indexes, the control group outperformed the experimental group in aspects including essay length, lexical density, Fog index and word frequency. In other words, the control group collectively wrote longer essays that were more academic in nature based on their lexical density (ratio of content to functional words), Fog index (sentence length and complex words) and word

frequency (the use of common words). This may have been due to the students' familiarity with the prompt, having frequent practice with evaluative and hortative rhetorical functions and the object of enquiry (socio-cultural issue). Studies have shown that topic familiarity and prompt type exert significant positive effects on the writing performance of students across proficiency levels (Stapleton, 2001; He & Shi, 2012). Similarly, the more sophisticated lead-ins in the scripts of the control group again may have been products of repeated preparation.

However, the experimental group essays demonstrated a higher level in terms of lexical diversity (use of different word types) and syntactic complexity (number of words before the main verb), which is suggestive of a higher cognitive demand on the part of the writer. Qualitative analysis of the scripts revealed that the control group used more similar sentences in their essays. Some writers in the control group copied sentences from the writing prompt. Many students in the control group, for example, began their essays with a similar, even identical sentence. This may have caused the low scores in syntactic complexity and lexical diversity for the control group. The same may be said of the collective uniformity of organizational patterns noted in the paragraphs of the control group scripts. In contrast, (the second interpretation) receiving an unfamiliar task and topic, which the exploratory prompt was meant to be, may have disadvantaged the writing performance of the students in the experimental group. In light of the shorter length, lower scores on the four standard measures and inconsistent organization, the experimental group was clearly outperformed.

However, this does not mean that the students in the experimental group were

hindered in their thinking and arguing abilities. The higher frequency of use of hedges by the experimental group could suggest that the prompt-solving is more stimulating in terms of critical thinking development. Although hedges serve to reduce an author's commitment to a statement, they reflect the author's thinking process which involves speculation, imaging other people's minds and interpreting other people's behavior, which typifies high-order thinking. Another illustration of this was the greater use of the words "if" and "think" by the experimental group. "If" is often used to make a prediction or generate a hypothetical situation. Take-a-stand-oriented prompts may be less likely to stimulate this type of thinking or their associated linguistic structures. Likewise, the marked difference in the way the two groups used the word "think" is revealing. Rather than simply using this word to express personal beliefs as the control group tended to do, a significantly higher percentage of writers in the experimental group projected thoughts onto others - certainly a more challenging mental exercise both logically and linguistically. Undeniably, the isolation of two words where differences appeared between the two groups is not sufficient to make any generalizations beyond the present study; however, in the exploratory spirit, these examples illustrate how areas of omission, both logical and linguistic, can emerge when prompts gravitate towards the one-size-fits-all approach apparently taken by the tests, especially TOEFL and IELTS, surveyed in RQ8 of this study.

In other words, because of the narrowness of rhetorical function and object of enquiry in high-stakes tests, it appears that student writers are not being exposed in composition classes to as full a range of genres, topics or linguistic structures as they

could be. In the composition class at the university where the experiment took place, the subject matter of a typical prompt is usually sociocultural, and the task type always requires students to take a stand on an issue and support it, i.e., evaluative and/or hortative. Much exposure to this type of prompt and sufficient practice in responding to it certainly enhances the performance of student writers. However, it is contended that the high frequency of using writing prompts with narrow subject matter and formulaic tasks brings about a cookie cutter effect. As a consequence, students end up writing formulated structures on a limited set of ideas that do not necessarily reflect their chosen discipline, nor do they expand cognitive functions much beyond simply supplying reasons to support claims. Underscoring this point is the widespread use of writing templates among Mainland Chinese EFL learners at different levels. These templates, which essentially are generic model answers, are often memorized by students in preparation for high-stakes tests in the knowledge that at least some chunks of the template will provide a respectable answer to the prompt they will soon respond to. He (2010), for example, explains how test-takers are coached in Mainland China:

> *To prepare the trainees for the writing tasks, the coaching programs/schools provide several examples of essays for them to use no matter what social problem or phenomenon is being discussed… The vagueness of the issue under discussion [in these essays] renders the samples suitable for whatever social problem or phenomenon is involved. (p. 153-4)*

In this way, it can be argued that the writing templates, to some extent, reflect the narrowness in rhetorical function and subject matter of prompts prevailing in writing assessment and instruction. This narrowness in turn may be leading to an entirely unintended washback, i.e., the memorized template.

This concern about high-stakes test washback developing only a narrow set of writing skills for study in university has been frequently noted (e.g., Read & Hirsh, 2005; Turner, 2006). One notable observation study by Green (2006) of EAP (English for Academic Purposes) and IELTS preparation classes at universities and language schools, found there was considerable overlap of writing content in the two class types, and sometimes even direct mention of "IELTS," in the EAP classes. However, determining what is washback from high-stakes tests, and what is simply determined to be necessary for coping in an academic environment is difficult to uncover. Nevertheless, the findings from both studies performed here indicate that test prompts do converge on certain topics and rhetorical functions that in turn may be ignoring a whole range of genres and ways of communicating ideas. When considering the academic purposes of students coming from a wide range of disciplines, confining writing instruction and practice to such a narrow range may not be in their best interests.

5.9 The Role of Writing Prompts in Cultivating Students' Argumentative Writing and Critical Thinking Abilities

All the three studies in this project indicate the salient role that writing prompts play

in cultivating students' argumentative writing and critical thinking abilities. Study 1 revealed that both students and teachers recognized writing prompts, especially those in the high-stakes English tests, as a contributory factor of students' performance and perceptions on argumentative writing. Study 2 indicated that the lack of counterargumentation representation in the writing prompts and rubrics of TEM 4 & TEM8 (the two most important English proficiency tests for undergraduate English majors) might have exerted negative washback on students' performance of written argumentation. The findings of Study 3 suggest that the object of enquiry and rhetorical functions in writing prompt not only impact on rhetorical and linguistic features of students' argumentative writing, but also have an effect on their critical thinking. Compared with the students who responded to a traditional prompt, those who responded to an experimental prompt demonstrated more critical thinking through their use of the language. However, students' responses to a traditional prompt exhibit higher performance in essay length, lexical density, Fog Index, and word frequency, implying that students handle such prompts more competently. One possible explanation is that traditional prompts have a very limited range, and this allows students to respond in ways that they are familiar with, i.e., the cookie cutter effect discussed in Section 5. 8.

In one study by Wang (2010), Chinese university students were invited to select prompts that they would like to respond to from a collection of topics. Wang found that although students' selection differentiates to some degree from actual prompts used in high-stakes test prompts, both students' prompt choice and test prompts were confined to a narrow scope of society and life experience. This narrow scope may

reflect test washback on classroom instruction as well as assessment practice in mainland China. Despite the negative effect outlined here, the flipside is the positive potential that the writing prompt has if new objects of enquiry and rhetorical functions are considered for the purposes of enhancing students' analysis, problem-solving and other critical thinking abilities.

In summary, the findings of Study 3 imply that writing prompts can be used as an effective means of enhancing students' critical thinking performance in their argumentative writing.

5.10 Limitations of Study 3

Any conclusions arising from Study 3 are constrained by certain assumptions and limitations related to the exploratory nature of the research. One potential criticism of this study concerns the writing prompts generated for the two groups in RQ9, which compared responses to typical and exploratory prompts. Clearly, the two sets of scripts resulting from these prompts are distinct, and any patterns noted between the two groups are idiosyncratic to a certain extent and, as is customary in exploratory research, can be taken as indicative only. Clearly, the isolation of a few words in the corpuses of the responses to the two prompts cannot serve to make definitive statements about critical thinking ability, or lack thereof, in one group or the other. Nevertheless, the differences noted may serve as indicators for possible further study in the spirit of the exploratory nature of this study.

The convergence to a cookie-cutter type of prompt, i.e., take a stand on a social issue and defend it, similar to convergences seen in many aspects of life, e.g., the shapes of commercial aircraft or mobile phones, reflects constraints caused by several factors. For one, high-stakes tests require large numbers of scorers who need to be trained to score scripts arising from multiple prompt responses. Thus, it is understandable that prompts and rubrics are aligned within certain boundaries related to rhetorical function and object of enquiry. The experimental prompt in the present study explores responses from students, largely ignoring the practical issue of scoring reliability. Indeed, it could also be argued that straying into an alternative object of enquiry, i.e., behavioral economics, approaches territory beyond the realm of language, although similar arguments could be made about those presently used, i.e., education and technology. In essence, Horowitz (1991) some time ago summed up the difficult task facing test designers:

> *the concern of those who create writing assessments – to mitigate examinees differences in background knowledge – contrasts sharply with the concern of designers of academic writing tasks, whose precise purpose in creating questions is to place on the examinee the burden of proving mastery of a specific body of knowledge and a specific disciplinary approach to that knowledge.* (p. 74)

This chapter discusses important implications of the three studies in this project. For instance, it provides some deliberation pertaining to the effect of an intervention on students' argumentative writing and critical thinking

development. It also discusses the role of writing prompts in cultivating students' argumentative writing and critical thinking abilities. A succinct conclusion of the three studies and their respective significance is provided in the next chapter.

CHAPTER 6

CONCLUSION

Three studies have been conducted in this research project to investigate critical thinking in the argumentative writing of undergraduate English majors in mainland China. Presented in this chapter is a summary of the main findings and implications of the research outcome.

6.1 Findings of Study 1 and Study 2

Study 1 investigated undergraduate English majors' typical performance in and perceptions of argumentative writing. It further examined what factors contributed to students' performance and perceptions, and how counterargumentation and critical thinking were represented in these factors. Research methods used to collect data in Study 1 included a writing test, a student questionnaire, four focus group interviews with students and one focus group interview with university teachers of English writing. The writing test data showed that students produced virtually no counterarguments and rebuttals in their argumentative essays. The questionnaire and interview data revealed that over half of the participant students were unfamiliar with counterargumentation and most of them had received no proper training in it. Results from the questionnaire with students and interviews with teachers of English writing showed that four factors influence students' argumentative writing, namely, 1) the Syllabus; 2) writing prompts and rubrics of TEM4 and TEM8; 3) composition textbooks; and 4) classroom instruction. For the first three factors, qualitative content

analysis was conducted, and classroom observation was used for the final factor to explore how counterargumentation and critical thinking were represented.

The investigation into the four factors indicated that the Syllabus does encourage critical thinking development for undergraduate English majors, but provides insufficient guidance for its implementation in the curriculum, especially in terms of the English Writing course. Therefore, the Syllabus was found to be unhelpful in developing students' counterargumentation and critical thinking skills. In addition, the writing prompts and rubrics of TEM4 and TEM8, the composition textbooks and the classroom instructions were found to lack any contribution towards counterargumentation in undergraduate English majors' argumentative writing.

Informed by the findings of Study 1, a second study, Study 2, was devised. Study 2 explored the effect of explicit instruction in counterargumentation skills on the overall quality of students' argumentative writing and their critical thinking ability. A 12-week instructional intervention using an adapted Toulmin model of argumentation (1958) was implemented. After the intervention, participants were given a post-test, the results of which were compared with that of the pretest (the writing test aforementioned) to evaluate the impact of the intervention on students' argumentative writing; questionnaires and interviews with students were also used in the manner of before-and-after mode to compare participants' perceptions of argumentative writing. In addition, students' critical thinking ability demonstrated in their argumentative essays was measured using a rubric developed in the study and variations were noted before and after the intervention.

Questionnaire and interview data indicated a positive effect of the intervention on participants' knowledge and confidence in argumentative writing for the experimental group. More importantly, the intervention was found to have significantly enhanced: 1) the numbers of counterarguments and rebuttals produced in students' argumentative essays ($p<.01$); 2) the overall quality of the argumentative essays ($p<.01$). There was also a significant increase in the mean score of posttest essays for the control group ($p<.05$). However, no increase was noted in the production of counterarguments and rebuttals in posttest essays of the control group. Notably, the posttest scores for the experimental group were significantly higher ($p<.01$) than those of the control group. Furthermore, the instructional intervention in counterargumentation improved students' critical thinking ability significantly for the experimental group ($p<.01$).

The findings demonstrated that the classroom instruction was effective in helping Chinese students incorporate alternative views in their argumentative essays. Using an experimental design, it was also confirmed that presenting and refuting alternative views was significantly positively correlated with the overall quality of an argumentative essay.

6.2 Implications of Study 1 & Study 2

The findings of Studies 1& 2 may have important implications for writing prompts and rubrics, composition textbooks as well as argumentative writing pedagogy in China and beyond. It is proposed that counterargumentation be considered in

high-stakes English examination writing prompts and rubrics, and that critical thinking elements be better embedded in textbooks and classroom instruction on argumentative writing. In view of the research findings, it is proposed that the goal instructions and descriptors in writing prompts and rubrics respectively of high-stakes tests be reconsidered so as to encourage counterargumentation. It is also proposed that the notion of counterargumentation be included in the writing curriculum for mainland Chinese undergraduates because "approaching a matter from multiple perspectives" (the Syllabus, NACFLT, 2000, p.12) and thinking critically is stipulated as one of the chief pedagogical goals.

Therefore more attention and research is needed with regard to the content and goal instructions in argumentative writing prompts as they have the potential for considerable impact on pedagogy via washback and the subsequent thinking processes of students. In a broader sense, it is anticipated that this study will stimulate trends in society towards inclusiveness and respecting the views of others, a key part of which is to engender a mindset of open-mindedness and fairness. All of these lead back to critical thinking.

6.3 Findings of Study 3

Study 3 was an extended study based on the results of Study 1 and Study 2. It explored the effect of two types of writing prompts with distinct natures on students' critical thinking ability. Carried out in two phases, the focus of Study 3 was the object of enquiry and rhetorical function of writing prompts. In the first phase, 120 prompts

from three high-stakes English tests, namely, TOEFL, IELTS and TEM4, were analyzed in order to see whether these prompts displayed certain patterns. It was found that evaluation and hortation were the two most common rhetorical functions. As for the object of enquiry, notable similarities were found among the three tests with three themes of a sociocultural nature – education, technology and social issues – dominating. Based on these findings, a prompt was deemed conventional in this study if its rhetorical function was evaluation or hortation and its object of enquiry was sociocultural in nature.

In the second phase of Study 3, two groups of undergraduate students responded to two argumentative prompts different in object of enquiry and rhetorical function. The control group wrote responses to a conventional prompt. For the experimental prompt, however, the object of enquiry was in an area other than education, technology or society, and the rhetorical function focused on analyzing and problem-solving (as opposed to evaluation and hortation). Text analysis was first performed on the data to produce standardized indexes of writing quality. The results indicated that the control outperformed the experimental group in essay length, lexical density, Fog index and word frequency, but the experimental group did better in syntactic complexity and lexical diversity. Then content and corpus analyses were performed focusing on the use of metadiscourse, essay organization, and certain lexical items. It was found that the experimental group used more hedges, which suggested more engagement of higher-order thinking. Further corpus analysis of the data found that the experimental group referred to hypothetical situations more frequently, suggesting greater analytical thinking. In terms of organizational structure, a much higher percentage of

the control group essays used a preamble to introduce the topic, and the control group demonstrated a more elaborated way of presenting an argument.

6.4 Implications of Study 3

The findings of Study 3, despite being a small-scale, exploratory study, suggest that the writing prompts in three high-stakes language tests result in test-takers responding in a fashion that is confined to a narrow range of rhetorical functions and objects of enquiry. Further, an analysis of the scripts from a control group who responded to a prompt similar to those found in this narrow range revealed that, although this group performed better in terms of certain linguistic indexes, the test-takers' organization and language use was restricted to certain patterns and usages when compared to an experimental group that responded to an exploratory prompt. These findings raise concerns about the impact that high-stakes tests may be having on classroom instruction; specifically, there is a concern that, via washback from the tests, students may be learning to write English essays using a narrow range of organizational patterns and language, and even a restricted form of reasoning, all while writing on an exceeding limited range of topics. It is contended that a wider variety of prompts that takes better consideration of the many specific purposes encountered in academic contexts may broaden the scope of written language and forms of critical reasoning to the benefit of students.

Whilst the three studies focused on different dimensions of critical thinking, with Study 1 and Study 2 investigating counterargumentation and Study 3 emphasizing

problem-solving, they all targeted the critical thinking development demonstrated in students' argumentative writing. The findings are anticipated to draw attention to the cultivation of students' critical thinking ability and disposition in both the pedagogy and assessment of argumentative writing.

The findings of the three studies may have important implications for writing assessment as well as argumentative writing pedagogy in China and beyond. It is proposed that counterargumentation be considered in the writing prompts and rubrics of high-stakes English tests, and included in classroom instruction on argumentative writing. It is also contended that a wider range of prompts may broaden the scope of written language and forms of critical reasoning to the benefit of students.

REFERENCES

Alderson, J. C., & Hamp-Lyons, L. (1996) TOFEL preparation courses: a study of washback, *Language Testing, 13*, 280–297.

Alderson, J. C., & Wall, D. (1993) Does washback exist? *Applied Linguistics, 14(2)*, 115–129.

Allright, D., & Bailey, K. (1991). *Focus on the language classroom: An introduction to language classroom research for language teachers*. New York: Cambridge University Press.

Altheide, D., Coyle, K., DeVriese, K., & Schneider, C. (2008). Emergent Qualitative document analysis. In Hesse-Biber, S. N., & Leavy, P. (eds), *Handbook of emergent methods*. New York: The Guilford Press.

Andrews, S. (2004). Washback and Curriculum Innovation. In L., Cheng & Y., Watanabe (Eds), *Washback in language testing: Research contexts and methods* (pp. 37-50). Mahwah, NJ: Laurence Erlbaum.

Andrews, S., Fullilove, J., & Wong, Y. (2002). Targeting washback: a case study. *System, 30(2),* 207-223.

Anthony, L. (2012). AntConc (ver. 3.3.5) [software]. Retrieved from http://www.antlab.sci.waseda.ac.jp/

Applebee, A.N., Langer, J.A., Jenkins, L.B., Mullis, I.V.S., & Foertsch, M.A. (1990). *Learning to write in our nation'sschools:Instruction and achievement in 1988 at Grades 4, 8.and 12*.Princeton:Educational Testing Service.

Applebee, A.N., Langer, J.A., Mullis, I.V.S., Latham, A.S., Gentile, C.A. (1994). *NAEP 1992: Writing Report Card*. Washington, DC: U.S. Government Printing Office.

Ariely, D. (2008). *Predictably irrational: The hidden forces that shape our decisions*. New York: HarperCollins.

Atkinson, D. (1997). A critical approach to critical thinking in TESOL. *TESOL Quarterly, 31(1),* 71–94.

Atkinson, D., & Ramanathan, V. (1995). Cultures of writing: An ethnographic comparison of L1 and L2 university writing/language programs. *TESOL Quarterly,* 29, 539-568.

Bacha, N. N. (2010). Teaching the academic argument in a university EFL environment. *Journal of English for Academic Purposes, 9,* 229-241.

Barnet, S., & Bedau, H. (2011). *Critical Thinking: Reading and Writing* (7th ed). Boston: St. Martin's.

Baron, J. (1985). *Rationality and intelligence.* New York, NY: Cambridge University Press.

Baron, J. (1991). Beliefs about thinking. In J. F. Voss, D. N. Perkins, & J. W. Segal (Eds.), *Informal reasoning and education* (pp.169-186). Hillsdale, NJ: Lawrence Erlbaum.

Baron, J. (1995). Myside bias in thinking about abortion. *Thinking and Reasoning, 1,* 221–235.

Bean, J. C. (2011). *Engaging ideas: The professor's guide to integrating writing, critical thinking, and active learning in the classroom.* San Francisco: Jossey-Bass Publishers.

Becker, C. B. (1986). Reasons for the lack of argumentation and debate in the Far East. *International Journal of Intercultural Relations, 10*, 75-92.

Benesch. S. (1999). Thinking Critically, Thinking Dialogically. *TESOL Quarterly,* 33(3), 573-580.

Berelson, B. (1952). *Content analysis in communication research.* New York: Free Press.

Billig, M. (1987). *Arguing and thinking: A rhetorical approach to social psychology.* Cambridge: Cambridge University Press.

Blair, J. A., & Johnson, R. H. (1987). The current state of informal logic. *Informal Logic, 9 (2&3)*, 147-151.

Bogdan, R.C., & Biklen, S.K. (1997). *Qualitative research for education: an introduction to theory and methods* (3rd ed.). Boston: Allyn & Bacon.

Brabeck, M. (1983). Critical thinking skills and reflective judgment development: Redefining the aims of higher education. *Journal of Applied Developmental Psychology, 4*, 23-34.

Brockriede, W. (1975). Where is argument? *Journal of the American Forensic Association, 11*, 179-182.

Brockriede, W. (1980). Argument as epistemological method. In D.A. Thomas (Ed.), *Argumentation as a way of knowing.* Falls Church, VA: Speech Communication Association.

Brookfield, S. D. (1987). *Developing critical thinkers: Challenging adults to explore alternative ways of thinking and acting.* San Francisco: Jossey-Bass.

Brown, J. D. (2001). *Using Surveys in Language Programs.* Cambridge: Cambridge University Press.

Brown, M. N., & Keeley, S. M. (2012). *Asking the right questions: A guide to critical thinking* (10th ed.). Boston: Pearson.

Bryman, A. (2012). *Social research methods* (4th ed.). Oxford: Oxford University Press.

Burkhalter, N. (1995). A Vygotsky-based curriculum for teaching persuasive writing in the elementary grades. *Language Arts*, *72*, 192–199.

Cai, G. (1999). Texts in contexts: Understanding Chinese students' English compositions. In C. Cooper & L. Odell (Eds.), *Evaluating Writing: The Role of Teachers' Knowledge about Text, Learning, and Culture*. Urbana, IL: National Council of Teachers of English, 279-297.

Chan, N. M., Ho, I. T., & Ku, K. Y. L. (2011). Epistemic beliefs and critical thinking of Chinese students. *Learning and Individual Differences, 21*, 67-77.

Chan, K. W., & Elliot, R. G. (2004). Epistemological beliefs across cultures: Critique and analysis of beliefs structure studies. *Educational Psychology, 24*, 123−142.

Chandrasegaran, A. (2008). NNS students' arguments in English: Observations in formal and informal contexts. *Journal of Second Language Writing, 17,* 237-254.

Charles, M. (2007). Reconciling top-down and bottom-up approaches to graduate writing: Using a corpus to teach rhetorical functions. *Journal of English for Academic Purposes, 6,* 289-302.

Cheng, L. (1997). How does washback influence teaching? Implications for Hong Kong. *Language and Education, 11(1)*, 38-54.

Cheng, L. (2005). *Changing language teaching through language testing: A washback study* (Vol. 21). Cambridge: Cambridge University Press.

Cheng, L. (2008). The key to success: English language testing in China. *Language Testing*, *25*, 15–37.

Cheng, L., & Qi, L. (2006). Description and examination of the national matriculation English test. *Language Assessment Quarterly, 3(1)*, 53-70.

Cho, Y., Rijmen, F., & Novak, J. (2013). Investigating the effects of prompt characteristics on the comparability of TOEFL iBT^{TM} integrated writing tasks. *Language Testing, 30(4)*, 513-534.

Clyne, M. G., (1987). Cultural differences in the organization of academic texts: English and German. *Journal of Pragmatics, 11*, 211-247.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37-46.

Connor, U. (1990). Linguistic/rhetorical measures for international persuasive student writing. *Research in the Teaching of English, 24(1)*, 67-87.

Connor, U. (1996). *Contrastive Rhetoric: Cross-Cultural Aspects of Second-Language Writing*. Cambridge: Cambridge University Press.

Connor, U. (2002). New directions in contrastive rhetoric. *TESOL Quarterly, 36,*493-510.

Crammond, J. G. (1998). The uses and complexity of argument structures in expert and student persuasive writing. *Written Communication, 15*(2), 230-268.

Creswell, J.W. (2009). *Research design: qualitative, quantitative, and mixed methods approaches* (3rd ed). Los Angeles: SAGE.

Crismore, A., Markkanen, R., & Steffensen, M. S. (1993). Metadiscourse in persuasive writing: A study of texts written by American and Finnish university students. *Written Communication, 10(1)*, 39-71.

Crowhurst, M. (1991). Interrelationships between reading and writing persuasive discourse. *Research in the Teaching of English, 25(3)*, 314-338.

Cumming, A., Kantor, R., Powers, D., Santos, T., & and Taylor, C. (2000) TOEFL 2000 writing framework: A working paper. *TOEFL Monograph Series.* Princeton, NJ: Educational Testing Service.

Dewey, J. (1910). *How we think*. Lexington: D. C Heath.

Ding, W., Wu, B., Zhong, M. & Guo, Q. (1994). *A Handbook of English Writing* (2nd ed). Beijing: Foreign Language Teaching and Research Press.

Dornyei, Z. (2010). *Questionnaires in second language Research: construction, administration, and processing* (2nd ed.). New York, London: Routledge.

Dornyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.

Dowdy, N. S. (1998). *The effects of teaching rebuttal skills on the written arguments of college students*. Michigan: A Bell & Howell Company.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education, 84(3)*, 287-312.

Duran, P., Malvern, D., Richards, B., & Chipere., N.(2004). Developmental trends in lexical diversity, *Applied Linguistics 25(2)*, 220–242.

Durkin, K. (2008). The middle way: East Asian master's students' perceptions of critical argumentation in UK universities. *Journal of Studies in International Education*, 12(1), 38-55.

Duschl, R. A., Ellengoben, K., & Erduran, S. (1999). Understanding dialogic argumentation among middle school science students. Invited paper at the annual meeting of the American Educational Research Association (AERA), Montreal, April 1999.

Eckstein, M.A., & Noah, H.J. (1993). *Secondary School Examinations: International Perspectives on Policies and Practice*. New Haven: Yale University Press.

Ennis, R.H. (1962). A concept of critical thinking. *Harvard Educational Review, 32,* 81-111.

Ennis, R.H. (1987). A taxonomy of critical thinking dispositions & abilities. In J. Baron & R. Sternberg (Eds.), *Teaching thinking skills: Theory and practice*, (pp.9-26). New York: W. H. Freeman.

Ennis, R.H. (2003). Critical Thinking Assessment. In Fasko, D. (Ed). *Critical thinking and reasoning: Current research, theory, and practice*. Cresskill, NJ.: Hampton Press.

Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education, 88(6)*, 915–933.

ETS. (2013). About the TOEFL iTB test. Retrieved from http://www.ets.org/toefl/ibt/about?WT.ac=toeflhome_ibtabout2_121127

Facione, P. A. (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction.

Facione, P. A., & Facione, N.C. (1994). *The Holistic Critical Thinking Scoring Rubric.* Millbrae, CA: The California Academic Press.

Feldman, R. (1994). Good argument. In F. F. Schmitt (ed.), Socializing Epistemology: The Social Dimensions of Knowledge, Lanham, MD: Roman and Littefield, pp. 159-188.

Felton, M. (2004). The development of discourse strategies in adolescent argumentation. *Cognitive Development, 19,* 35 - 52.

Ferretti, R. P., Andrews-Weckerly, S. & Lewis, W.E. (2007). Improving the argumentative writing of students with learning disabilities: description and normative consideration. *Reading and Writing Quarterly, 23*, 267-285.

Ferretti, R. P., Lewis, W.E. & Andrews-Weckerly, S. (2009). Do goals affect the structure of students' argumentative writing strategies? *Journal of Educational Psychology, 101(3)*. 577-589.

Ferretti, R. P., MacArthur, C. A., & Dowdy, N. S. (2000). The effects of an elaborated goal on the persuasive writing of students with learning disabilities and their normally achieving peers. *Journal of Educational Psychology, 92*(4), 694-702.

Ferris, D. R. (1994). Rhetorical strategies in student persuasive writing: Differences between native and non-native English speakers. *Research in the Teaching of English*, 28, 45-65.

Ferris, D. R. (2005). Tricks of the trade: The nuts and bolts of writing research. In P. Matsuda & T. Silva (Eds.), Second language writing research (pp. 219–230).

Mahwah, NJ: Lawrence Erlbaum.

Feucht, F. C., & Bendixen, L, D. (2010). Exploring similarities and differences in personal epistemologies of U.S. and German elementary school teachers. *Cognition and Instruction, 28(1),* 39-69.

Fisher, G. (1980). *International Negotiations: A cross-cultural perspective.* Boston: Intercultural Press.

Fisher, A., & Scriven, M. (1997). *Critical thinking: Its definition and assessment.* Norwich, UK: Centre for Research in Critical Thinking.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication, 32*(4), 365-387.

Floyd, C. B. (2011). Critical thinking in a second language. *Higher Education Research and Development, 30*, 289-302.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education* (8th ed). Bosten: McGraw-hill Higher Education.

Frederiksen, N. (1984). The real test bias: influences of testing on teaching and learning. *American Psychologist, 39*, 193-202.

Fox, H. (1994). *Listening to the world: Cultural issues in academic writing*. Urbana, IL.: National Council of Teachers of English.

Fulkerson, R. (1996a). *Teaching the argument in writing*. Urbana, IL: National Council of Teachers of English.

Fulkerson, R. (1996b). The Toulmin model of argument and the teaching of composition. In Emmel, B., Resch, P., & Tenney, D. (Eds). *Argument revisited, argument redefined: Negotiating meaning in the composition classroom* (pp.45-72). Thousand Oaks, Calif.: Sage Publications.

Gallagher, S. A. (1998). The road to critical thinking: The Perry scheme and meaningful differentiation. *NASSP Bulletin, 82(595)*, 12−20.

Gardner, S., & Nesi, H. (2013). A classification of genre families in university student writing. *Applied Linguistics, 34(1),*25-52.

Genishi, C. (1992). Developing the foundation: Oral language and communicative competence. In C. Seefeldt (Ed.), *The early childhood curriculum: A review of current research,* (pp. 85-117). New York: Teachers College Press.

Giebels, E., & Taylor, P, J. (2009). Interaction patterns in crisis negotiations: persuasive arguments and cultural differences. *Journal of Applied Psychology, 94(1),* 5-19.

Gieve, S. (1998). Comments on Dwight Atkinson's "A Critical Approach to Critical Thinking in TESOL": Another reader reacts…. *TESOL Quarterly, 32,* 123-129.

Ginther, A., & Grant, L. (1997). The influence of proficiency, language background, and topic on the production of grammatical form and error on the Test of Written English. In A. Huhta, V. Kohonen, L. Kurki-Suonio, & S. Luoma (Eds.), *Current Developments and Alternatives in Language Assessment: Proceedings of LTRC 96* (pp. 385-397). JyvaÈskylaÈ: University of JyvaÈskylaÈ.

Gleason, M. M. (1999). The role of evidence in argumentative writing. *Reading and Writing Quarterly, 15*(1), 81-106.

Graff, G. (2003). Clueless in academe: How schooling obscures the life of the mind. New Haven, CT: Yale University Press.

Green, A. (2006). Watching for washback: Observing the influence of the international English language testing system academic writing test in the classroom. *Language Assessment Quarterly, 3(4)*, 333-368.

Grice, H. P. (1975). Logic and Conversation. In Donald Davidson and Gilbert Harman (Eds.). *The Logic of Grammar* (pp. 64-75). Encino, California: Dickenson.

Goh, W. W.L., (2008). *Dialogic argumentation as a path to enhancing individual argumentive reasoning in academically disadvantaged 8th –graders* (Unpublished PhD thesis). Columbia University, Columbia.

Goldberg, D. T. (1994). Introduction: Multicultural conditions. In D. T. Goldberg (ed.), *Multiculturalism: A critical reader* (pp. 1-41). Oxford: Blackwell.

Gu, Y. (2013). The unbearable lightness of the curriculum: what drives the assessment practices of a teacher of English as a Foreign Language in a Chinese secondary school? *Assessment in Education: Principles, Policy & Practice*, ifirst article. DOI: 10.1080/0969594X.2013.836076

Hairston, M. C. (1981). *Successful Writing: A Rhetoric for Advanced Composition*. New York: Norton.

Hall, E. T. (1976). *Beyond culture.* Garden City, NY: Anchor Press.

Halliday, M. A. K. (1994). An introduction to functional grammar (2nd ed.). London: Edward Arnold.

Halpern, D. F. (1998). Teaching for critical thinking: Helping college students develop the skills and dispositions of a teaching critical thinking for transfer across domains: Dispositions, skills, structure training, and metacognitive monitoring. *American Psychologist, 53*, 449-455.

Halpern, D. F. (2003). *Thought and knowledge: An introduction to critical thinking* (4th ed). Mahwah, NJ: Lawrence Erlbaum.

Hamp-Lyons, L. (1997). Washback, impact and validity: ethical concerns. *Language*

Testing, 14(3), 295-303.

Hayes, J. N., Hatch, J. (1999). Issues in measuring reliability: correlation versus percentage of agreement. *Written Communication, 16*, 354-367.

He, L. (2010). The graduate school entrance English examination. In L. Cheng & A. Curtis (Eds). *English language assessment and the Chinese learner* (pp 145-157). London: Routledge.

He, L., & Shi, L. (2012). Topical knowledge in ESL writing. *Language Testing*, *29*, 443-464.

He, Q., Yin, T., Huang, Y. & Liu. H. (1999). On the reform of Undergraduate Foreign Language Disciplines. *Foreign Language Teaching and Research, 31 (1).* 24-28.

Herman, J. L., & Golan, S. (1993). The effects of standardized testing on teaching and schools, *Educational Measurement: Issues and Practice, 12(4),* 20-25, 41-42.

Higgins, E. T., Pierro, A., & Kruglanski, A. W. (2008). Re-thinking culture and personality: How self-regulatory universals create cross-cultural differences. In R. M. Sorrentino, & S. Yamaguchi (Eds.), *Handbook of motivation and cognition across cultures* (pp. 161–190). New York: Elsevier.

Hinkel, E. (1995). The Use of Modal Verbs as a Reflection of Cultural Values. *TESOL Quarterly 29*, 325–343.

Hinkel, E. (2002). Second language writers' text: Linguistic and rhetorical features. Mahwah, NJ: Lawrence Erlbaum.

Hinkel, E. (2009). The effects of essay topics on modal verb uses in L1 and L2 academic writing. *Journal of Pragmatics, 41*, 667-683.

Hitchcock, D & Verheij, B. (2005). The Toulmin model today: Introduction to the special issue on contemporary work using Stephen Edelston Toulmin's layout of arguments. *Argumentation, 19*. 255-258.

Hitchcock, D & Verheij, B. (2006). Introduction. In D. Hitchcock and B. Verheij (Eds.). *Arguing on the Toulmin Model: New Essays in Argument Analysis and Evaluation* (pp1-23). Dordrecht: Springer.

Hoetker, J. (1982). Essay examination topics and students' writing. *College Composition and Communication, 33(4)*, 377-392.

Hoetker, J., & Brossell, G. (1989). The effects of systematic variations in essay topics on the writing performance of college freshmen. *College Composition and Communication, 40(4)*, 414-421.

Hofer, B. K., & Pintrich, P. R. (Eds). (2002). *Personal epistemology: The psychology of beliefs about knowledge and knowing*. Mahwah, NJ: Erlbaum.

Hornikx, J., & Hoeken, H. (2007). Cultural differences in the persuasiveness of

evidence types and evidence quality. *Communication Monographs, 74*, 443-463.

Horowitz, D. (1991). ESL writing assessments: Contradictions and resolutions. In L. Hamp-Lyons (Ed.), *Assessing second language writing in academic contexts* (pp. 71-85). Norwood, NJ: Ablex.

Huang, Y. (1998). On the critical thinking deficiency. *Foreign Languages and Their Teaching, 7*. 16-19.

Huang, Y. (2010). English programs need a through reform: Rethinking the critical thinking deficiency. *Foreign Language World, 136*. 11-16.

Hughes, W., Lavery, J. A., & Doran, K. (2010). *Critical thinking: An introduction to the basic skills* (6th ed.). Peterborough, Ont.: Broadview Press.

Hyland, K. (1998). Persuasion and context: The pragmatics of academic metadiscourse. *Journal of Pragmatics, 30*, 437-455.

IELTS Annual Review (2011). Retrieved from www.ielts.org.

IELTS: Information for candidates (n.d.). Retrieved from www.ielts.org.

Inch, E. S., & Warnick, B. (2010). *Critical thinking and communication: The use of reason in argument* (6th ed.). Boston, Mass.; Hong Kong: Allyn & Bacon.

Jin, Y. & Fan. J. (2011). Test for English Majors (TEM) in China. *Language Testing 28(4)*. 589-596.

Johansson, V. (2008). Lexical diversity and lexical density in speech and writing: a developmental perspective. *Lund Working Papers for Linguistics, 53*, 61-79.

Johnson, R. H. (2000). *Manifest rationality: A pragmatic theory of argument*. Mahwah, NJ: Lawrence Erlbaum.

Jones, P. C., Merritt, J. Q., & Palmer, C. (1999). Critical thinking and interdisciplinarity in environmental higher education: The case for epistemological and values awareness. *Journal of Geography in Higher Education, 23(3)*, 349−357.

Kaplan, R.B. (1966). Cultural thought patterns in intercultural education. *Language Learning, 16(1)*, 1-20.

Kaplan, R.B. (1972). *The Anatomy of Rhetoric: Prolegomena to a Functional Theory of Rhetoric*. Philadelphia: Center for Curriculum Development.

Kaya, E. (2013). Argumentation practices in classroom: Pre-service teachers' conceptual understanding of chemical equilibrium. *International Journal of Science Education, 35(7)*, 1139-1158.

Kneupper, C. W. (1978). Teaching argument: an introduction to the Toulmin Model. *College Composition and Communication 29*. 237-241.

Knipper, K. J., & Duggan, T. J. (2006). Writing to learn across the curriculum: Tools

for comprehension in content area classes. *Reading Teacher,* 59, 462–470.

Knudson, R. (1991). Effects of instructional strategies, grade, and sex of students' persuasive writing. *Journal of Experimental Education*, *59*, 141–152.

Knudson, R. (1992). Analysis of argumentative writing at two grade levels. *Journal of Educational Science 85(3).* 169-179.

Knudson, R. E. (1994). An analysis of persuasive discourse: Learning how to take a stand. *Discourse Processes, 18*(2), 211-230.

Kroll, B., & Reid, J. (1994). Guidelines for designing writing prompts: Clarifications, caveats and cautions. *Journal of Second Language Writing, 3,* 231-255.

Kruglanski, A. W., Thompson, E. P., Higgins, E. T., Atash, M. N., Pierro, A., Shah, J. Y., et al. (2000). To 'do the right thing' or to 'just do it': Locomotion andassessment as distinct self-regulatory imperatives. *Journal of Personality & Social Psychology, 79*, 793–815.

Kuhn, D. (1991). *The skill of argument.* Cambridge: Cambridge University Press.

Kuhn, D. (1999). Metacognitive development. In L. Balter, & C. S. Tamis-LeMonda (Eds.), *Child psychology: A handbook of contemporary issues* (pp. 259−286). Philadelphia, PA: Psychology Press.

Kuhn, D. (2005). *Education for thinking*. Cambridge, Mass.: Harvard University Press.

Kuhn, D. & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young adolescents' thinking. *Psychological Science, 22,* 545-552.

Kuhn, D., & Udell, W. (2003). The development of argument skills. *Child Development, 74*, 1245-1260.

Kuhn, D. & Udell, W. (2007). Coordinating own and other perspectives in argument. *Thinking & Reasoning, 13(2).* 90-104.

Kurfiss, J. G., Association for the Study of Higher Education, The George Washington University, & The George Washington University. (1988). *Critical thinking: Theory, research, practice, and possibilities.* Washington, D.C.: The George Washington University.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.

Lee, H., & Anderson, C. (2007). Validity and topic generality of a writing performance test. *Language Testing, 24*, 307-330.

Leitão, S. (2003). Evaluating and selecting counterarguments. *Written Communication, 20*(3), 269-306.

Leki, I. (1992, March). Changing perspectives on L2 reading and writing. *Plenary*

speech presented at TESOL conference.

Li, Y. (2007). Apprentice scholarly writing in a community of practice: An intraview of an NNES graduate student writing a research article. *TESOLQuarterly, 41*(1), 55-79.

Lim, G. S. (2010). Investigating Prompt Effects in Writing Performance Assessment. *Spaan Fellow Working Papers in Second Language Assessment. 8*, University of Michigan English Language Institute.

Liu, L. (2005). Rhetorical education through instruction across cultures: a comparative analysis of select online instructional materials on argumentative writing. *Journal of Second Language Writing, 14(1),* 1-18.

Lu, L. (2011). Metadiscourse and genre learning: English argumentative writing by Chinese undergraduates. *HKU Theses Online (HKUTO).*

Lun, V. M. C., Fischer, R., & Ward, C. (2010). Exploring cultural differences in critical thinking: Is it about my thinking style or the language I speak? *Learning and Individual Differences, 20,* 604-616.

Lunsford, K. J. (2002). Contextualizing Toulmin's model in the writing classroom. *Written Communication, 19(1)*. 109-174.

Lytle, A. L., Brett, J. M., Barsness, Z., Tinsley, C. H., & Janssens, M. (1995). A paradigm for quantitative cross-cultural research in organizational behavior. In B. M. Staw & L. L. Cummings (Eds.), *Research in organizational behavior* (Vol. 17, pp. 167–214). Greenwich, CT: JAI Press.

Manalo, E., Kusumi, T., Koyasu, M., Michita, Y., & Tanaka, Y. (2013a). To what extent do culture-related factors influence university students' critical thinking use? *Thinking Skills and Creativity*, *10*, 121-132.

Manalo, E., Watanabe, K., & Sheppard, C. (2013b). Do language structure or language proficiency affect critical evaluation? In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), Proceedings of the 35th Annual Conference of the Cognitive Science Society (pp. 2967–2972). Austin, TX: Cognitive Science Society.

Markus, H.R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98, 224-253.

Mason, L., & Scirica, F. (2006). Prediction of students' argumentation skills about controversial topics by epistemological understanding. *Learning and Instruction, 16*, 492-509.

Matalene, C. (1985). Contrastive rhetoric: An American writing teacher in China. *College English, 47(8)*, 789-809.

McCann, T. M. (1989). Student argumentative writing knowledge and ability at three grade levels. *Research in the Teaching of English, 23,* 62-76.

McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. *Written Communication, 27(1)*, 57-86.

McPeck, J. E. (1981). *Critical thinking and education.* NY, New York: St. Martin's Press.

Mercier, H. (2013). Introduction: Recording and explaining cultural differences in argumentation. *Journal of Cognition and Culture, 13*, 409-417.

Mertens, D.M. (2010). Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods (3rd ed.). Thousand Oaks, Calif.: Sage.

Mines, R. A., King, P. M., Hood, A. B., & Wood, P. K. (1990). Stages of intellectual development and associated critical thinking skills in college students. *Journal of*

College Student Development, 31(6), 538−547.

Moore, T., & Morton, J. (2005). Dimensions of difference: A Comparison of university writing and IELTS writing. Journal of English for Academic Purposes, 4, 43-66.

Moss, G. (2000). Informal lietracies and pedagogic discourse. *Linguistics and Education, 11(1),* 47-64.

National Advisory Commission on Foreign Language Teaching in Higher Education. (2000). *National Teaching Syllabus for College English Majors in China*. Shanghai: Shanghai Foreign Language Education Press.

Nielsen, J. A. (2013). Dialectical features of students' argumentation: A critical review of argumentation studies in science education. *Research in Science Education, 43(1)*, 371-393.

Nisbett, R. E., Peng, K., Choi, I. and Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review, 108*, 291-310.

Norris, S. P. (2003). The meaning of critical thinking test performance: The effects of abilities and dispositions on scores. In D. Fasko Jr. (Ed.), *Critical Thinking: Current Research, Theory, and Practice*. Dordect, The Netherlands: Kluwer.

Norris, S. P., & Ennis, R. H. (1989). *Evaluating critical thinking.* Pacific Grove, CA: Critical Thinking Books and Software.

Nussbaum, E. M. (2008). Using argumentation vee diagrams (AVDs) for promoting argument-counterargument integration in reflective writing. *Journal of Educational Psychology, 100*(3), 549.

Nussbaum, E.M. (2011). Argumentation, dialogue theory, and probability modeling: alternative frameworks for argumentation research in education. *Educational Psychologist, 46(2)*, 84-106.

Nussbaum, E. M., & Kardash, C. A. M. (2005). The effects of goal instructions and text on the generation of counterarguments during writing. *Journal of Educational Psychology, 97*(2), 157.

Nussbaum, E. M., & Schraw, G. (2007). Promoting argument-counterargument integration in students' writing. *The Journal of Experimental Education, 76*(1), 59-92.

O'Keefe, D. J. (1999). How to handle opposing arguments in persuasive messages: A meta-analytic review of the effects of one-sided and two-sided messages. *Communication Yearbook*, 22, 209-249.

Orafi, S. M. S., & Borg, S. (2009). Intentions and realities in implementing communicative curriculum reform. *System, 37 (2)*, 243-253.

Osborne, J. (2010). Arguing to learn in science: The role of collaborative, critical discourse. *Science, 328*, 463-466.

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching, 41(10)*, 994-1020.

Palmer, W. (2012). *Discovering arguments:An introduction to critical thinking, writing, and style* (4th ed.). Boston, Mass.; Hong Kong: Prentice Hall.

Pandis, M., Ward, A., & Matthews, S. R. (Eds.). (2005). *Reading, writing, thinking*: *Proceedings of the 13th European conference on reading*. Newark, DE: International Reading Association.

Paton, M. (2005). Is critical analysis foreign to Chinese students? In E. Manalo & G.Wong-Toi (Eds.), *Communication skills in university education: The international dimension (pp. 1–11)*. Auckland: Pearson Education.

Patraglia, J. (1995). Writing as an unnatural act. In J. Petraglia (Ed.), *Reconceiving writing, rethinking writing instruction* (pp. 79-100). Mahwah, NJ: Erlbaum.

Paul, R. W. (1993). *Critical thinking: What every person needs to survive in a rapidly changing world.* Santa Rosa, CA: Foundation for Critical Thinking.

Paul, R. W. (1995). *Critical thinking: How to prepare students for a rapidly changing world.* Santa Rosa, CA: Foundation for Critical Thinking.

Paul, R.W, Elder, L., & Bartell, T. (2004). *Study of 38 public universities and 28 private universi ties to determine faculty emphasis on critical thinking in instruction. Executive summary*. Retrieved August 10, 2005, from www.criticalthinking.org/about/ centerforCT.shtml.

Peng, K., & Nisbett, R. E. (1999). Culture, dialectics and reasoning about contradiction. *American Psychologist, 54*, 741-754.

Perkins, D. N. (1985). Postprimary education has little impact on informal reasoning. *Journal of Educational Psychology,* 77, 562–571.

Perkins, D. N., Farady, M., & Bushey, B. (1991). Everyday reasoning and the roots of intelligence. In J. F. Voss, D. N. Perkins, & J. W. Segal (Eds.), *Informal reasoning and education* (pp. 83–106). Hillsdale, NJ: Erlbaum.

Perkins, D. N., Jay, E., & Tishman, S. (1993). Beyond abilities: a dispositional theory of thinking. *Merrill-Palmer Quarterly, 39,* 1-21.

Perkins, D. N., & Tishman, S. (2001). Dispositional aspects of intelligence. In J.M. Collis & S. Messick (Eds). *Intelligence and Personality: Bridging the Gap in Theory and Measurement* (pp. 233 - 257). Mahwah, NJ: Lawrence Erlbaum Associates.

Petraglia, J. (1995). Writing as an unnatural act. In J. Petraglia (Ed.). *Reconceiving writing, rethinking writing instruction* (pp.79-100). Mahwah, NJ: Erlbaum.

Petric, B. (2005). Contrastive rhetoric in the writing classroom: a case study. *Journal of English for Specific Purposes*, *24*, 213-228.

Petric, B. (2007). Rhetorical functions of citations in high-and low-rated master's theses. *Journal of English for Academic Purposes, 6,* 238-253.

Piolat, A., Roussey, J.Y., & Gombert, A. (1999). The development of argumentative schema in writing. In G. Rijlaarsdam & E. Espéret (Series Eds.) & J. Andriessen & P. Coirier (Vol. Eds.), *Foundations of argumentative text processing (pp. 117–135).* Amsterdam: Amsterdam University Press.

Pithers, R. T., & Soden, R. (2001). Critical thinking in education: A review. *Educational Research, 42(3)*, 237-249.

Popham, J. (1987). The merits of measurement-driven instruction. *Phi Delta Kappa May*, 679-682.

Porte, G., & Richards, K. (2012). Focus article: Replication in second language writing research. *Journal of Second Language Writing, 21*, 284-293.

Qi, L. (2005). Stakeholders' conflicting aims undermine the washback function of a high-stakes test. *Language Testing, 22(2)*, 142-173.

Qi, L. (2007). Is testing an efficient agent for pedagogical change? Examining the intended washback of the writing task in a high-stakes English test in China. *Assessment in Education: Principles, Policy & Practice, 14(1)*, 51-74.

Qin, J., & Karabacak, E. (2010). The analysis of toulmin elements in chinese EFL university argumentative writing. *System, 38*(3), 444-456.

Qu, W. (2005). *Introducing argumentation*. Beijing: Higher Education Press.

Qu, W. (2013). *Introducing critical thinking and argumentation*. Beijing: Higher Education Press.

Raimes, A., & Zamel, V. (1997). Response to Ramanathan and Kaplan. *Journal of Second Language Writing, 6,* 79-81.

Ramage, J. D., Bean, J. C., & Johnson, J. (2010). *Writing arguments: A rhetoric with readings* (8th ed.). New York; Hong Kong: Pearson Longman.

Ramanathan, V., & Kaplan, R. (1996a). Audience and voice in current L1 composition tests: Some implications for ESL student writers. *Journal of Second Language Writing, 5,* 21-33.

Ramanathan, V., & Kaplan, R. (1996b). Some problematic "channels" in the teaching of critical thinking in current L1 composition textbooks: Implications for L2 student-writers. *Issues in Applied Linguistics, 7,* 225-249.

Rapley, T. (2007). *Doing conversation, discourse and document analysis*. London: Sage.

Read, J., & Hirsh, D. (2005). *English language levels in tertiary institutions (Export Levy Programme Project E4)*. Wellington: Education New Zealand.

Reid, J. (1990). Responding to different topic types: A quantitative analysis from a contrastive rhetoric perspective. In B. Kroll (Ed.), Second language writing: Research insights for the classroom (pp. 191-209). Cambridge: Cambridge University Press.

Reid, J., & Kroll, B. (1995). Designing and assessing effective classroom writing assignments for NES and ESL students. *Journal of Second Language Writing, 4,* 17-41.

Reznitskaya. A., Anderson, R. C., & Kuo, L.J. (2007). Teaching and Learning Argumentation. *The Elementary School Journal, 107(5)*, 449-472.

Rothery, J. (1996). Making changes: Developing an educational linguistics. In R. Hasan & G.Williams (Eds.), *Literacy in society* (pp. 86–123). New York: Addison Wesley.

Rottenberg, A. T., & Winchell, D. H. (2009). *Elements of argument: A text and reader* (9th ed.). Boston, Mass.: Bedford/St. Martins.

Sadler, T. D., & Fowler, S. R. (2006). A threshold model of content knowledge transfer for socioscientific argumentation. *Science Education, 90(6)*, 986–1004.

Sampson, V., & Clark, D. B. (2008). Assessment of the ways students generate arguments in science education: Current perspectives and recommendations for future directions. *Science Education, 92(4)*, 447-472.

Santos, C. M. M., & Santos, S. L. (1999). Good argument, content and contextual dimensions. In G. Rijlaarsdam & E. Espéret (Series Eds.) & J. Andriessen & P. Coirier (Vol. Eds.), *Foundations of argumentative text processing* (pp. 75–95). Amsterdam: Amsterdam University Press.

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. Mahwah, NJ: Lawrence Erlbaum.

Schraw, G. (2001). Current themes and future directions in epistemological research: A commentary. *Educational Psychology Review, 13(4)*, 451−464.

Schreier, M. (2012). *Qualitative content analysis in practice*. Los Angeles: SAGE.

Scott, C. (2007). Stakeholder perceptions of test impact. *Assessment in Education: Principles, Policy & Practice, 14(1)*, 27-49.

Sears, A., & Parsons, J. (1991). Toward critical thinking as an ethic. *Theory and Research in Social Education, 19,* 45–46.

Selinker, R. M., Trimble, T. & Trimble, L. (1976). Presuppositional rhetorical information in EST discourse. *TESOL Quarterly, 10(3)*, 281-290.

Selinker, R. M., Trimble, T. & Trimble, L. (1978). Rhetorical function – shifts in EST discourse. *TESOL Quarterly, 12(3)*,311-320.

Shirkhani, S. & Fahim, M. (2011). Enhancing critical thinking in foreign language learners. *Procedia – Social and Behavioral Sciences, 29,* 111-115.

Shen, F. (1989). The classroom and the wider culture: Identity as a key to learning English composition. *College Composition and Communication, 40 (4)*, 459-466.

Shohamy, E. (2001). *The power of test: A critical perspective of the use of language tests.* London: Pearson.

Shohamy, E. (2007). Language tests as language policy tools. *Assessment in Education: Principles, Policy & Practice, 14(1)*, 117-130.

Siegel, H. (1988). *Educating reason: Rationality, Critical Thinking and Education.* New York: Routledge.

Siegel, H. (1999). Argument quality and cultural differences. *Argumentation, 13*, 183-201.

Simon, S. (2008). Using Toulmin's argument pattern in the evaluation of argumentation in school science. *International Journal of Research & Method in Education 31(3).* 277-289.

Sowards, S. K. (2012). Rhetorical functions of letter writing: Dialogic collaboration, affirmation, and catharsis in Dolores Huerta's letters. *Communication Quarterly, 60(2),* 295-315.

Spaan, M. (1993). The Effect of Prompt in Essay Examinations. In D. Douglas & C. Chapelle (Eds.), *A new decade of language testing research* (pp. 98-122). Alexandria, VA: TESOL.

Spurgin, Sally Dewitt. (1985). *The Power to Persuade*. Englewood Cliffs, NJ: Prentice Hall.

Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking: the argument evaluation task. *Journal of Educational Psychology, 89,* 342-357.

Stanovich, K.E. & West, R.F. (2008). On the failure of cognitive ability to predict myside and one-sided thinking biases. *Thinking and Reasoning, 14(2).* 129-167.

Stapleton, P. (2001). Assessing critical thinking in the writing of Japanese university students. *Written Communication*, 18(4), 506–548.

Stapleton, P. (2002a). Critical thinking in Japanese L2 writing: rethinking tired constructs. *ELT Journal 56(3).* 25—257.

Stapleton, P. (2002b). Critiquing voice as a viable pedagogical tool in L2 writing: Returning the spotlight to ideas. *Journal of Second Language Writing, 2002(11)*, 177-190.

Stapleton, P. (2011). A survey of attitudes towards critical thinking among Hong Kong

secondary school teachers: Implications for policy change. *Thinking Skills and Creativity, 6*, 14-23.

Su, Y. (2011). A comparative study on explicit and implicit teaching of critical thinking in exposition and argumentation. *Journal of Beijing Institute of Technology (Social Sciences edition), 13(2).*150-156

Suter, W. N. (2012). *Introduction to educational research: a critical thinking approach.* Thousand Oaks, Calif.: Sage.

Taylor, L. (2005). Washback and impact. *The ELT Journal, 59(2)*, 154-155.

Tedick, D. (1990). ESL writing assessment: subject-matter knowledge and its impact on performance. *English for Specific Purposes, 9*, 123-143.

The National Advisory Committee for Foreign Language Teaching. (2000). *The teaching syllabus for English majors*. Shanghai: Shanghai Foreign Language Teaching and Research Press.

Tian, J. (2008). *The influence of undergraduate learning contexts on Chinese graduate students' argumentation and critical thinking in writing*. (Unpublished PhD thesis). University of York, York.

Tian, J., & Low, G. D. (2011). Critical thinking and Chinese university students: A

review of the evidence. *Language, Culture and Curriculum, 24*(1), 61-76.

Toplak, M. E., & Stanovich, K. E. (2003). Associations between myside bias on an informal reasoning task and amount of post-secondary education. *Applied Cognitive Psychology, 17*, 851-860.

Toulmin, S., (1958). *The uses of argument.* Cambridge: Cambridge University Press.

Toulmin, S. E., Rieke, R. D., & Janik, A. (1984). *An introduction to reasoning* (2nd ed.). New York: Macmillan.

Trimble, L. (1985). *English for science and technology: A discourse approach.* Cambridge: Cambridge University Press.

Turner, Y. (2006). Students from mainland China and critical thinking in postgraduate business and management degrees: Teasing out tensions of culture, style and substance. *International Journal of Management Education, 5(1)*, 3–11.

Uysal, H. H. (2012). Argumenation across L1 and L2 writing: Exploring cultural influences and transfer issues. *Vigo International Journal of Applied Linguistics, 9,* 133-159.

Van Eemeren, F. H. (2009). The study of argumentation. In R. A. Eberly, A. A. Lunsford & K. H. Wilson (Eds.), *The sage handbook of rhetorical studies* (pp. 109-124). Thousand Oaks, Calif.: Sage Publications.

Van Eemeren, F.H., & Grootendorst, R. (1988). Rationale for a pragma-dialetical perspective. *Argumentation, 2*, 271-291.

Van Eemeren, F. H., Grootendorst, R., & Henkemans, F.S. (1996). *Fundamentals of argumentation theory: A handbook of historical backgrounds and contemporary developments*. Mahwah, NJ: L. Erlbaum.

Van Eemeren, F. H., Grootendorst, R., Jackson, S., & Jacobs, S. (1993). *Reconstructing argumentative discourse*. Tuscaloosa, AL: University of Alabama Press.

Varghese, S. A., & Abraham, S. A. (1998). Undergraduates arguing a case. *Journal of Second Language Writing, 7*(3), 287-306.

Venville, G. J., & Dawson, V. M. (2010). The impact of a classroom intervention on grade 10 students' argumentation skills, informal reasoning, and conceptual understanding of science. *Journal of Research in Science Teaching, 47(8)*, 952–977.

Vesterman, W. (2006). *Reading and Writing Short Arguments* (5th ed). Mountain View, CA: Mayfield.

Voss, J. F. (2005). Toulmin's model and the solving of ill-structured problems. *Argumentation, 19*. 321-329.

Voss, J.F., & Means, M.L. (1991). Learning to reason via instruction in argumentation. *Learning and Instruction*, 1, 337-350.

Walker, G. B. (1986). "Communicating across cultures: Argument and International Negotiation." In F. Eemeren, R. Grootendorst, A. Blair, & C. Willard (eds). *Argumentation: Analysis and Practices* (pp. 238-250). Dordrecht-Holland: Foris Publications.

Wall, D. (2000). The impact of high-stakes testing on teaching and learning: can this be predicted or controlled? *System, 28(4),* 499-509.

Walton, D. N. (1989). Dialogue theory for critical thinking. *Argumentation,* 3, 169-184.

Walton, D. N. (1998). *The new dialectic: Conversational contexts of argument.* Toronto: University of Toronto Press.

Walton, D. N. (2007). *Dialog theory for critical argumentation.* Amsterdam: John Benjamins Publishing Company.

Wang, D. (2010). Chinese students' choice of writing topics: A comparison between their self-selected topics and the writing prompts in large-scale tests. *The Journal of Asia TEFL, 7(3)*, 165-187.

Washington State University Critical Thinking Project (2009). *Guide to rating Critical & integrative thinking.* Retrieved from *https://my.wsu.edu/pls/portal/docs/PAGE/CTLT/CITRUBRIC/CIT%20RUBRIC%202009%20WITH%20RATIO%20SCALE%202009%20FINAL.PDF* on Nov 6, 2013.

Weigle, S.C. (2002). Assessing writing. Cambridge: Cambridge University Press.

Weinstock, M. P., Neuman, Y., & Glassner, A. (2006). Identification of informal reasoning fallacies as a function of epistemological level, grade level, and cognitive ability. *Journal of Educational Psychology, 89(2),* 327-341.

Wen, Q. & Liu, R. (2006). An exploratory study on features in English majors' abstract thinking in English argumentative compositions. *Journal of Foreign Languages*, 162. 49-58.

Wen, Q., Wang, H., Wang, J., Zhao, C. & Liu, Y. (2010). A comparative study of critical thinking skills between English and other liberal arts majors. *Foreign Language Teaching and Research, 42(5),* 350-355.

Wenzel, J. W. (2006). Three perspectives on argument: Rhetoric, dialectic, Logic. In R. Trapp & J. E. Schuetz (eds). *Perspective on argumentation: Essays in honor of Wayne Brockriede*. New York: International Debate Education Association.

White, E. M. (1993). Assessing higher-order thinking and communication skills in

college graduates through writing. *The Journal of General Education, 42(2),* 105-122.

Wolfe, C. R., & Britt, M. A. (2008). The locus of the myside bias in written argumentation. *Thinking & Reasoning, 14*(1), 1-27.

Wolfe, C. R., Britt, M. A., & Butler, J. A. (2009). Argumentation schema and the myside bias in written argumentation. *Written Communication, 26*(2), 183.

Wood, N. V. (2009). *Perspectives on argument* (6th ed.). Upper Saddle River, N.J.: Prentice Hall.

Wragg, E.C. (2011). *An introduction to classroom observation* (3rd ed.). London: Routledge.

WTO Entry and Educational Reform and Development of Foreign Language Disciplines in Chinese Universities Project team. (2001). Reflections on the curricular and instructional reform of foreign language disciplines in Chinese universities. *Foreign Language World,* 85. 9-15.

Wu, S. Y., & Rubin, D. L. (2000). Evaluating the impact of collectivism and individualism on argumentative writing by Chinese and North American college students. *Research in the Teaching of English,* 35(2), 148-178.

Xie, Q., & Andrews, S. (2013). Do test design and uses influence test preparation? Testing a model of washback with structural equation modeling. *Language Testing, 30(1)*, 49-70.

Xu, Y. (2006). *Identifying and controlling writing task difficulty factors in English tests.* Unpublished PhD thesis. Shanghai: Shanghai International Studies University.

Yan, W. & Sun, Y. (2012). The use of cohesive devices in argumentative writing by Chinese EFL learners at different proficiency levels. *Linguistics and Education,* 23, 31-48.

Yang, Y.T.C., & Gamble, J. (2013). Effective and practical critical thinking-enhanced EFL instruction. ELT Journal, 67(4). 398-412.

Yeh, S. S. (1998). Empowering education: Teaching argumentative writing to cultural minority middle school students. *Research in the Teaching of English, 33,* 49–83.

You, X. (2004). "The choice made from no choice": English writing instruction in a Chinese University. *Journal of Second Language Writing, 13*(2), 97-110.

You, X. (2010a). Building Empire through Argumentation: Debating Salt and Iron in Western Han China. *College English, 72*, 367-384.

You, X. (2010b). *Writing in the devil's tongue: A history of English composition in China*. Carbondale: Southern Illinois University Press.

You, X. & Liu, Y. (2009). Confucians love to argue: Policy essays in ancient China. *College Composition and Communication, 60*, 41-50.

Yu, G. (2010). Lexical diversity in writing and speaking task performances. *Applied Linguistics, 31(2)*, pp. 236-259.

Zhang, M. (2000). Cohesive features in the expository writing of undergraduates in two Chinese universities. *RELC Journal, 31,* 61–95.

Zhang, Z. (2011). *English Composition: From creative thinking to critical thinking*. Beijing: Foreign Language Teaching and Research Press.

Zhou, Y. (2004). *Exploring into comparability of two national EFL tests (CET-6 and TEM-4) in China.* Shanghai: Shanghai Foreign Language Education Press.

Zhu, W. (2001). Performing argumentative writing in English: Difficulties, processes, and strategies. *TESL Canada Journal/Revue TESL Du Canada,19(1),* 34-50.

Zohar, A. & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching, 39(1)*. 35-62.

Zou, S. (2003). The alignment of language teaching syllabus and language testing: TEM8 test development and administration. *Foreign Language World, 98(6)*, 71–78.

Zou, S. (2010). Looking back and ahead – In commemoration of the 20th anniversary of TEM4 and TEM8. *Foreign Language World, 141 (6)*, 9-25.

Appendices

Appendix 1: Prompts for pretest and posttest in study 1

Writing prompt for pretest
China has recently experienced a great expansion of the population and size of its cities. This social process is called urbanization. A recent study by the Asian Development Bank and the National Development and Reform Commission estimates that cities in China will grow by about 15 million people each year and by a total of 230 million over the next 15 years. Do you think urbanization is making people's life better?
Write on the answer sheet a composition on the following topic:
Is urbanization making our life better?
There is no word limit on the essay.

Writing prompt for posttest
The revival of traditional culture has become a hot topic in China today. In the face of globalization and fast economic growth, there is a tendency that traditional culture is marginalized, especially among youngsters. Do you think young people should make greater efforts to preserve traditional Chinese culture?
Write on the answer sheet a composition on the following topic:
Should young people make greater efforts to preserve traditional Chinese culture?
There is no word limit on the essay.

Appendix 2: Questionnaire
1. Which one of the following genres of English writing is the most challenging to you?
以下哪一种英文写作文体对你来说最难？
    A. Description   B. Narration   C. Exposition   D. Argumentation
    A. 描述   B. 叙述   C. 讲解   D. 议论

2. Which one do you think is the most important genre of writing at the tertiary education level?
你认为大学阶段哪一种英文写作文体最重要？

A. Description   B. Narration   C. Exposition   D. Argumentation

A．描述   B．叙述   C．讲解   D．议论

3. Which one do you think is the most useful in your future academic study or career?

你认为对你未来的学业和职业最有用的文体是？

A. Description   B. Narration   C. Exposition   D. Argumentation

A．   描述   B．叙述   C．讲解   D．议论

4. Have the two terms: counterarguments (opposing views) and rebuttals (response to opposing views), or terms of the like, been discussed in your English writing class?

以下这两个词语在你的英语写作课上讨论过吗？抗辩（不同观点）和反驳（回应不同观点）

A. Yes   B. No

A．有   B．没有

5.   When you take a stand on a controversial issue for an argumentative essay, and provide evidence to support your view, do you think it is necessary to consider the opposition to your view and respond to it?

写议论文时，当你对一个有争议性的话题给出了自己的立场并提供了论据来支持这个立场，你是否认为有必要考虑与你的立场不同的观点并做出回应？

A. Yes   B. No

A．有必要   B．没必要

6. When you write an argumentative essay, which one of the following ways do you adopt to convince your readers?

你写议论文时，会采用下列哪一种方法来说服你的读者？

A. proposition (thesis statement, or main argument)

A. 给出立场（即文章的中心论点）

B. proposition + arguments

B. 立场+论据

C. proposition + arguments + counterarguments

C. 立场+论据+抗辩

D. proposition + arguments + counterarguments + rebuttals

D. 立场+论据+抗辩+反驳

7. What's the most difficult aspect of writing an English argumentative essay to you?

你在写英语议论文时遇到的最大困难是什么？

   A. English language proficiency

   A. 英语总体水平差（词汇，语法不够）

   B. Finding enough evidence to support the thesis

   B. 不能想出足够的论据来支持论点

   C. Considering the opposite side and responding to it

   C. 没能考虑反方观点，或觉得很难回应

   D. Others.

   D. 其他。

8. If you have chosen others in Question 7, please specify the difficulty

_____.

如果你在第七个问题中选了 D，请具体指出你遇到的困难

_____.

9. In your opinion, which of the following factors influence your performance in and perceptions of English argumentative writing? Please use Arabic numbers to indicate order of significance in the box. For instance, "1" indicates the most influential factor.

你认为下列哪些因素对你的英语议论文写作能力与认知有较大的影响？请在方框内用阿拉伯数字表明其重要性。例如，1 表示最重要的影响因素。

A. Teachers' classroom instructions at college □

A. 大学课堂上老师对英语议论文的教学和指导。

B. What I have learned in the middle school about English argumentative writing □

B. 在中学所学到的英语议论文写作方法。

C. Relevant stipulations in The National Syllabus on the Teaching of English Majors □

C. 全国英语专业本科生教学大纲的有关规定和要求

D. The English writing textbook used for classroom instruction in argumentative writing □

D. 在英语议论文课堂上使用的英语写作教材

E. The requirements and criteria in the writing prompts and rubrics of high-stakes such as TEM4 & TEM8 □

E. 在托福，雅思，专四和专八这样的重要考试中的议论文题目中的要求和标准。

F. The argumentative essays I read in newspapers and other media □

F. 在报纸或其他媒体上看到的议论文的写法。

Appendix 3: Rubric for critical thinking in writing

(1) Identifies and explains *issue/topic at hand*

• Does not ID or explain main issue/topic at hand; is confused (1 point)

• IDs main issue, does not explain clearly (2 points)

• IDs main issue/topic clearly, explains in limited fashion (3 points)

• IDs main issue/topic clearly, explains fully by discussing subsidiary and/or other relevant issues (4 point)

(2) Identifies and presents the *student's own perspective/analysis* regarding the issue at hand

• Fails to ID and state his/her own perspective/analysis on the issue/topic at hand (1 point)

• IDs and states own perspective/analysis, but fails to clarify own perspective versus other salient perspectives (2 points)

• IDs and states own perspective/analysis, but does so in a limited fashion (3 points)

• IDs and states own perspective/analysis, and considers it in light of other salient perspectives (4 point)

(3) Presents, assesses appropriate supporting data/evidence.

• Fails to support own perspective with appropriate data (1 point)

• Supports own perspective with limited, insufficient data (2 points)

• Supports own perspective with data, but data are weak (3 points)

• Supports own perspective with sufficient, convincing data (4 points)

(4) Identifies and considers *other salient perspectives/analyses* regarding issue/topic at hand

• Does not cite or utilize other perspectives/analyses regarding the topic/issue (1 point)

• Cites and utilizes other perspectives/analyses that are of limited value (2 points)

• Cites and utilizes other salient perspectives/analyses, but does so in a limited fashion (3 points)

• Cites and utilizes other salient perspectives/analyses, and brings them to bear on the issue/topic at hand (4 points)

(5) Identifies *conclusions and implications* of the issue/topic at hand

• Fails to ID conclusions/implications of the issue/topic (1 point)

• IDs conclusions/implications, but within a single context (2 points)

• IDs conclusions/implications as having connections to other contexts, but in a limited fashion (3 points)

• IDs conclusions/implications relative to the contexts important to the issue/topic at

hand (4 points)

(6) Communicates effectively, considers the influence of the contexts on the issue with a sense of the audience.

• Fails to communicate effectively and/or discuss the problem only in egocentric terms (1 point)

• Communicate fluently, but discuss the problem only in egocentric terms (2 points)

• Communicate effectively, and considers social/cultural or scientific/technological contexts of the issue, but without an assessment of the audience (3 points)

• Communicate effectively, considers social/cultural or scientific/technological contexts of the issue with an assessment of the audience (4 points)

Appendix 4: Table of contents of Ding el al's (1994) *A Handbook of English Writing* (Excerpt)

Part One Manuscript Form

Part Two Diction

Part Three The Sentence

Part Four The Paragraph

Part Five The Whole Composition

Appendix 5: TEM8 Writing Prompt 2012

A recent survey of 2,000 college students asked about their attitudes towards phone calls and text-messaging (also known as Short Message Service) and found the students' main goal was to pass along information in as little time, with as little small talk, as possible. "What they like most about their mobile devices is that they can reach other people," says Naomi Baron, a professor of linguistics at American University in Washington, D.C., who conducted the survey. "What they like least is that other people can reach them." How far do you agree with Professor Baron?

In the first part of your essay you should state clearly your main argument, and in the second part you should support your argument with appropriate details. In the last part you should bring what you have written to a natural conclusion or make a summary.

You should supply an appropriate title for your essay.

Marks will be awarded for content, organization, language and appropriateness. Failure to follow the above instructions may result in a loss of marks.

Appendix 6: TEM4 writing prompt 2011

Recently government agencies in some big cities have been studying the possibility of putting a "pollution tax" on private cars. The amount of tax private car owners would have to pay would depend on the emission levels, i.e. engine or vehicle size. This has caused quite a stir among the public. Some regard it as an effective way to control the number of cars and reduce pollution in the city. But others don't think so. What is your opinion?

Write on ANSWER SHEET THREE a composition of about 200 words on the following topic:

Should Private Car Owners Be Taxed for Pollution?

You are to write in three parts.

In the first part, state specifically what your opinion is.

In the second part, provide one or two reasons to support your opinion.

In the last part, bring what you have written to a natural conclusion or make a summary.

Marks will be awarded for content, organization, language, and appropriateness. Failure to follow the instructions may result in a loss of marks.

Appendix 7: TEM8 Composition Marking Scheme

BAND 5 SCORE 20-18

DESCRIPTION: EFFECTIVE COMMUNICATION WITH ACCURACY

The writing effectively addresses the writing task. It demonstrates a well developed logical organizational structure with clearly stated main ideas and sufficient supporting details. It has almost no errors of vocabulary, spelling, punctuation or syntax, and it displays an adequate ability to use the language with appropriacy. No difficulty is experienced by the reader.

BAND 4 SCORE 17-15

DESCRIPTION: GOOD COMMUNICATION WITH FEW INACCURACIES

The writing adequately addresses almost all of the writing task, though it deals with some parts more effectively than others. It demonstrates a generally well developed logical organizational structure with main ideas and sufficient supporting details. It has relatively few significant errors of vocabulary, spelling, punctuation or syntax, and it displays an ability to use the language with appropriacy. Very little difficulty is

experienced by the reader.

BAND 3 SCORE 14-12

DESCRIPTION: PASSABLE COMMUNICATION WITH SOME INACCURACIES

The writing adequately addresses most of the writing task. On the whole, it demonstrates an adequately developed logical organizational structure, though there may occasionally be a lack of relevance, clarity, consistency or support. It has occasional errors of vocabulary, spelling, punctuation or syntax, which may, from time to time, obscure meaning, and for the most parts it displays some ability to use the language with appropriacy. Occasional difficulty is experienced by the reader.

BAND 2 SCORE 11-9

DESCRIPTION: PROBLEMATIC COMMUNICATION WITH FREQUENT INACCURACIES

The writing only addresses some of the writing task. It demonstrates an inadequate organizational structure, and there may quite often be a lack of relevance, clarity, consistency or support. It has frequent errors of vocabulary, spelling, punctuation or syntax, and it displays a limited ability to use the language with appropriacy. Some difficulty is experienced by the reader.

BAND 1 SCORE 8-6

DESCRIPTION: ALMOST NO COMMUNICATION

The writing almost completely fails to address the writing task. It has neither an organizational structure nor coherence. Almost all sentences contain errors of vocabulary, spelling, punctuation or syntax, and it displays no ability to use the language with appropriacy. Even after considerable effort on the part of the reader, the text is largely incomprehensible.

Appendix 8: Marking Scheme for TEM4 Composition

band	score	description
5	15	EFFECTIVE COMMUNICATION WITH ACCURACY The writing effectively addresses the writing task. It demonstrates a well developed logical organizational structure with clearly stated main ideas and sufficient supporting details. It has almost no errors of vocabulary, spelling, punctuation or syntax, and it displays an adequate ability to use the language with appropriacy. No difficulty is experienced by the reader
4	14-12	GOOD COMMUNICATION WITH FEW

		INACCURACY
		The writing adequately addresses almost all of the writing task, though it deals with some parts more effectively than others. It demonstrates a generally well developed logical organizational structure with main ideas and supporting details. It has relatively few significant errors of vocabulary, spelling, punctuation or syntax, and it displays an ability to use the language with appropriacy. Very little difficulty is experienced by the reader.
3	11-9	PASSABLE COMMUNICATION WITH SOME INACCURACIES The writing adequately addresses most of the writing task. On the whole, it demonstrates an adequately developed organizational structure, though there may occasionally be a lack of relevance, clarity, consistency or support. It has occasional errors of vocabulary, spelling, punctuation orr syntax, which may, from time to time, obscure meaning, and for the most part it displays some ability to use the language with appropriacy. Occasional difficulty is experienced by the reader.
2	8-6	PROBLEMATIC COMMUNICATION WITH FREQUENT INACCURACIES The writing only addresses some of the writing task. It demonstrates an inadequate organizational structure, and there may quite often be a lack of relevance, clarity, consistency or support. It has frequent errors of vocabulary, spelling, punctuation or syntax, and it displays a limited ability to use the language with appropriacy. Some difficulty is experienced by the reader.
1	5-3	ALMOST NO COMMUNICATION The writing almost completely fails to address the writing task. It has neither an organizational structure nor coherence. Almost all sentences contain errors of vocabulary, spelling, punctuation or syntax, and it displays no ability to use the language with appropriacy. Even after considerable effort on the part of the reader, the

		text is largely incomprehensible.

Appendix 9: Classroom Observation Scheme

Instructor_____ **Date of Observation**_____

1. Teaching objectives:
2. Pedagogical activities:
3. Argumentation skills discussed in class:
4. Use of materials:

Appendix 10: Teaching schedule for the 12-week intervention

Week	Teaching objectives	Materials, activities
First week	Introducing good argumentation criteria and critical thinking	Toulmin model of argumentation; How to do critical writing
Second week	Introducing good argumentation criteria	The diagram of argumentation schema; key language for argumentative writing
Third week	Reading as a non-believer	Challenging reasons and evidence; Whales CT exercise
Fourth week	Analyzing model essays and identifying argumentative elements	The clothes that teachers wear
Fifth week	Generating counter opinions	JXNU is a wonderful place
Sixth week	Generating rebuttals	Refuting explanation exercise
Seventh week	Inclusion of counterargument and rebuttal in essays	Female teachers exercise; Pocket money exercise
Eighth week	Brainstorming and outlining	Metro construction exercise; social responsibility exercise
Ninth week	Joint writing	Group work, teacher-student joint writing

Tenth week	Independent writing and teacher comments	Teacher feedback in the form of conferencing
Eleventh Week	Independent writing and teacher comments	Teacher feedback in the form of conferencing
Twelfth week	review	Summarizing critical writing

Appendix 11: Diagram to illustrate complete argumentation
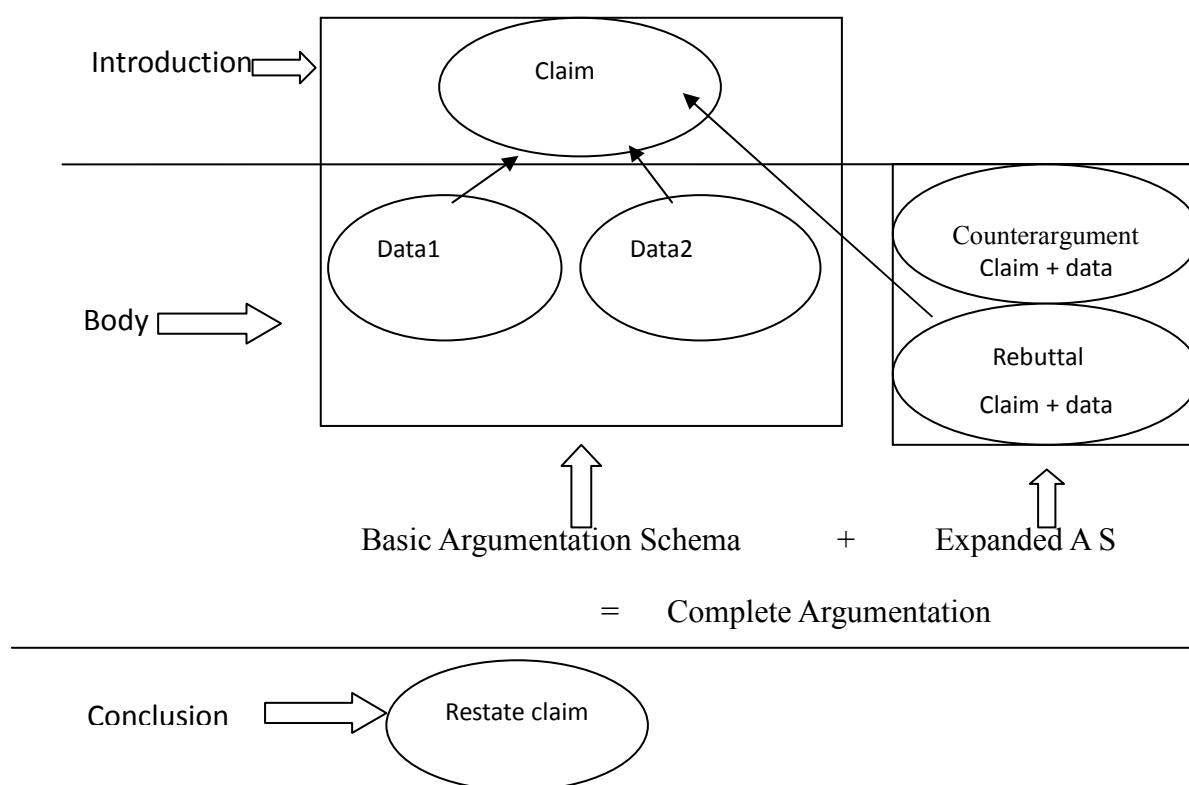
Appendix 12: Topics for Argumentative Writing

1. Many parents think that academic qualifications can give their children the best chance of success in life. How far do you agree?
2. Should celebrities have greater protection from the media?
3. Is stricter punishment the answer to rising juvenile crime?

4. It is said that downloading music without permission is an example of theft and is immoral. In your opinion, is downloading music without permission from the internet wrong?

5. Do you think that the construction of an underground system can solve the traffic congestion in Nanchang? Why or why not?

6. Some people comment that the present Chinese educational system is too examination-oriented and not very successful in developing students' creativity. What do you think of it?

7. "Governments should legislate to prevent monopolies becoming too powerful". Do you agree?

8. Do you think that developers should be permitted to build hotels and resort complexes in the most beautiful places in your hometown?

9. Do you think that people with better educational background should take up more social responsibilities?

10. Are there many things that the old can teach the young or are old people hopelessly out of touch when they reach a certain age?

11. Do you think that filial piety is still one of the most important virtues in the present Chinese society?

12. More and more Chinese people are living with substantial pressure which results in various mental problems. It is also the case in universities. In this condition, is it necessary to provide psychological consultation on the campus of universities?

Appendix 13: An example of coded essays in posttest

Should young people spend more efforts to preserve traditional Chinese culture?

In the force of globalization and fast economic growth, there is a tendency that traditional culture is marginalized, especially among youngsters. I think it's a very dangerous signal and young people should make greater efforts to preserve traditional Chinese culture (claim).

First of all, traditional Chinese culture is our nation's root and treasure (data 1). Without traditional Chinese culture, we can't form our great national spirit – being patriotic, unification, loving peace, being hard-working. Without Confucianism, we can't spread the thought of kindness and live in a harmonious society. Thus, we can see that traditional Chinese culture has brought us a lot of precious ideas and practical social pattern. If it's marginalized, we can't have a stable national root.

Next, young people are a nation's future and should pay more attention to stick to traditional Chinese culture (data 2). generally speaking, young people are energetic and open-minded to accept various culture. after experiencing various culture, young people should form their own the value of thought. Then, there is no doubt that they can't lose their national identity and belonging (data 3). So, determined to preserve traditional culture, youngsters can learn more about traditional culture and apply the treasure to building our nation.

However, some people hold the view that traditional Chinese culture is out-of-date and can't keep the pace with times, so they should be eliminated (counterargument). It's no denying that some parts of traditional Chinese culture is old-fashioned and superstitious, and those should be abandoned. But there are large parts of valuable traditional Chinese culture, such as thrift, respecting the old and caring children, kindness, and so on (rebuttal 1). We should be clear about that nothing is perfect, which applies to culture. just because there exist a little flaws in an excellent culture, then marginalizing it is stupid and ridiculous (rebuttal 2). Only sticking to the valuable parts of traditional culture, can we still keep the pace with times and even promote the social development (rebuttal 3).

In conclusion, youngsters should preserve traditional Chinese culture and learn its valuable parts with greater efforts. Then our society will be more stable and our future will be more wonderful.

claim 1
data 3
Counterargument 1
Rebuttal 3
Score 13

Appendix 14: Definitions and examples of argumentative elements

Argumentative elements	definitions	examples
claim	An assertion in response to a contentious topic or problem.	From my perspective, urbanization is certainly making our life worse.
data	Evidence to support a claim. It can take various forms such as facts, logical explanations, suppositions, statistics, anecdotes, research studies, expert opinions, definitions and analogies.	Urbanization brings about serious social problems as well as environmental problems
counterargument	The possible opposing views and evidence that can challenge the validity of a writer's claim	Some people may argue that urbanization is making our life better because urbanization means higher living standards.
rebuttal	Statements and evidence in which the writer respond to a counterargument	I totally disagree with those people. Despite the convenience and comfort in cities, urbanization is harmful to nature in the long run.