Summer 8-2021

A Study of Critical Thinking Dispositions of Undergraduates in Foreign Language Discipline at One Private College in Mainland China

Yuanyuan Zheng
Chapman University, 53306115@qq.com

A Study of Critical Thinking Dispositions of Undergraduates in Foreign Language Discipline at One Private College in Mainland China

A Dissertation by

Yuanyuan Zheng

Chapman University

Orange, CA

Attallah College of Educational Studies

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Education

August 2021

Committee in charge:

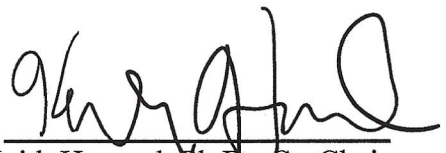
Keith Howard, Ph.D., Co-Chair

Margaret Grogan, Ph.D., Co-Chair

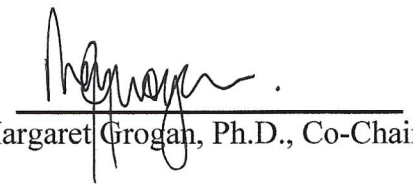
Maria L Martinez, Ph.D.

CHAPMAN UNIVERSITY | Attallah College of Educational Studies

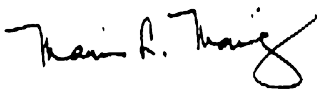
The dissertation of Yuanyuan Zheng is approved.



Keith Howard, Ph.D., Co-Chair



Margaret Grogan, Ph.D., Co-Chair



Maria L Martinez, Ph.D.

July 2021

A Study of Critical Thinking Dispositions of Undergraduates in Foreign Language Discipline at One Private College in Mainland China

ACKNOWLEDGEMENTS

I would like to take this opportunity to extend my gratitude to all who helped me during my PhD studies. First, I extend my millions of thanks to my amazing committee, my co-chairs Drs. Howard and Grogan, and Dr. Martinez, who have always been enlightening, insightful, supportive, and considerate.

I must thank my advisor Dr. Howard, someone with a great sense of humor, an expert at quantitative research, and the mentor who opened the door of SPSS for me. He has given me professional guidance and emotional support that effectively relieved my pressure and anxiety and gave me power and confidence to ensure sustainable growth on my way to completing my dissertation.

I am also grateful to Dr. Grogan, a knowledgeable, capable, and elegant person and professor who has set up a shining example for me in my future work. Thank you for every insightful comment or suggestion you made while reviewing the drafts of my dissertation. I am still touched by your caring, consideration, and warm comfort when I fell down the stairs and was injured last summer.

Many thanks to Dr. Martinez, for your willingness to work on my committee and for your valuable suggestions and precious time, patience, and understanding during the whole process. I can feel your strong interest in China and Chinese higher education when working with you. Welcome to Shanghai!

I also want to say thank you to many other professors at Chapman: To Drs. Dawn Hunter and Kelly Kennedy dedicated so much to work and supporting us. I still remember in the early January of 2020 when you came to Shanghai; unfortunately, we all suffered a lot from high fevers and bad colds. However, in retrospect, I was lucky those were not symptoms of COVID-19. How about you? Anyway, unforgettable memory of our shared experience!

To Dr. Allen Ryan, who accompanied us all the way since we stepped onto Chapman's campus: Without you, Research Questions 4 and 5 would not have been in such good shape and I am so thankful for that.

To Dr. Whitney Mclntyre Miller, whose enlightenment and inspiration in and out of class have still exerted a strong influence on me. Do you still remember the envelopes you reserved for us? I am curious about the message you left us in the mysterious envelopes!

Actually, I want to mention too many professors at Chapman but I cannot name them all here: You have brought me into a new world of scientific research, taught me what a PhD student in education means, and how much the PhD title weighs. Thank you, Chapman!

Next, I would like to thank my dearest classmates. We became a family because of this PhD program. Luna is my roommate at Chapman Grand, my cooking teacher, my SPSS tutor, my . . . many titles for her.

To Li and Helen, capable career women, excellent students, affectionate wives and mommies, my dear companions, and exceptional mutlitaskers. Do you still remember the happy time when we went shopping at Walmart and El Super by Uber and Lyft in the United States?

Becky, the youngest sister in our family, once spared no effort to share her valuable experience at data cleaning and processing with me. Lily Xing, who has always been encouraging and caring during my dissertation writing and defense preparation.

Moreover, Qin's perseverance, Julia's passion, Nancy's gentleness, Echo's intelligence, Rachel's eloquence, Jacky's diligence, and Edy's deliberation all impressed me a lot. I want to say you are all my cherished family.

I am also grateful to Sara, Sabrena, Vivienne, and Lindsey, my editors in the excellent team of Heartful Editor. I highly appreciate and respect your professionalism in reading my drafts millions of times and offering me pertinent suggestions to improve my writing.

I must thank Shanghai municipal education commission, Shanghai Normal University, and Xianda College of Economics and Humanities Shanghai International Studies University, which gave me a golden chance to work and study in this outstanding PhD program at this brilliant university and offered support to help us complete the program.

Last but not least, I must thank my family: my 91-year-old grandmother, my parents and parent-in-law, my husband, my sister, my brother-in-law, and my daughter, Grace. Without your love, support, and understanding, I never would have reached this milestone. Thank you all!

ABSTRACT

A Study of Critical Thinking Dispositions of Undergraduates in Foreign Language Discipline at One Private College in Mainland China

by Yuanyuan Zheng

Critical thinking (CT) has been one of the most important goals of higher education since the 1990s. However, foreign language majors in China have been criticized for the lack of CT and considered inferior to other majors. Despite studies conducted to disprove this belief, less research has explored the CT of undergraduates at private universities. This study investigated the status of undergraduates' CT dispositions at a private college in China, particularly among foreign language majors representing various demographic or academic groupings and with different learning experiences. An online survey was distributed to 5,000 undergraduates at a private college; 1,642 completed the survey (33% response rate). Descriptive statistics showed the undergraduates at this private college showed ambiguity toward CT. The distribution of percentages for this data set were more heavily concentrated in the Ambivalent and Negative category (over 50%) than the Positive category (less than 50%) for the Systematicity, Self-Confidence, and Truth-Seeking scales. Effect size and *t* tests identified no significant differences in the CT between foreign language majors and other liberal arts majors, between the Western and Eastern foreign language majors, or between international joint programs and regular programs; however, differences were significant for the overall scale between student leaders

and nonstudent leaders, for the Academic and Social Dimensions between male and female students, and for the Academic Dimension between students taking and not taking CT courses. Linear and incremental increases were found from freshmen to juniors; however, one-way ANOVA and effect size identified no significant differences in the total scores for the three class levels, but there was a significant difference for the Systematicity scale. Correlational analyses identified that GPA was not correlated with CT for the overall scale, but for sophomores, there was a positive correlation between GPA and CT; most instruction modes were also significantly correlated with CT. The findings suggest private universities should attach importance to nurturing students' stronger inclinations toward CT from the perspectives of curriculum design, faculty professional development, classroom teaching, assessment and evaluation, and extracurricular activities.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1: INTRODUCTION

Since the 1940s, critical thinking has been one of the topics of Western educational reforms. Critical thinking (CT), is described as higher-order thinking (Moore & Parker, 2012; R. Paul & Elder, 2006), and purposeful reflective judgment (Facione, 1990, 2011; Wen et al., 2009), involving cognitive skills and affective dispositions conducive to decision making and problem solving (Facione, 1990; R. Paul et al., 1995; Wen et al., 2009). Cognitive skills include analysis, interpretation, inference, explanation, evaluation, and self-regulation to form and improve judgment (Facione, 1990; Giancarlo & Facione, 2001). On the other hand, CT disposition, as the dimension of personality, is described as "the likelihood that one will approach problem framing or problem solving by using reasoning" (Giancarlo & Facione, 2001, pp. 30–31).

In Western countries like the United States, the cultivation of CT has been considered one of the most important goals of higher education since 1993 (Pithers & Soden, 2000; Yue, 2000). Particularly for liberal education, educators must commit to nurturing students' cognitive skills and their disposition toward CT, which can, in turn, "facilitate students' recognition of opportunity to use thinking to resolve problems and incline students toward doing so" (Giancarlo & Facione, 2001, p. 31). Similarly, in China, educational documents or reports have been used to emphasize the development of students' innovative thinking and the cultivation of students' ability to analyze and solve problems (Wen, 2012).

The Outline of National Medium-and-Long-Term Education Reform and Development Program (2010-2020) issued by China also highlighted the combination of learning and thinking, which advocated for heuristic, inquisitive, discussion, and participatory teaching to help students to learn (The Central People's Government of the People's Republic of China, 2010). Therefore, there is no marked difference between China and Western countries in the understanding of the importance of cultivating students' CT ability (Wen, 2012).

Statement of the Problem

Foreign language majors in mainland China, however, have constantly been criticized for the lack of CT since the late 1990s (Y. H. Gao, 1999; Q. X. He et al., 1999; Y. S. Huang, 1998, 2010; Wen & Zhou, 2006;). English language researchers and teachers have strived to help their students overcome this weakness. Along with this greater importance attached to critical thinking, a belief has emerged that foreign language majors are inferior to other majors because of their poor CT (Project Group of Entry into the WTO and Education of Foreign Language Disciplines, 2001; Wen & Zhou, 2006). Project Group of Entry into the WTO and Education of Foreign Language Disciplines (2001) criticized that foreign language disciplines in China had not attached great importance to the students' thinking ability training, which led to their poorer ability in logical thinking, organization, analysis, synthesis, judgment, and reasoning, compared with other liberal arts majors. Wen and Zhou (2006) argued this assertion by means of logical reasoning and empirical data. They proposed the inadequacy of foreign language proficiency restrained their thinking in the foreign language and college education in foreign language

discipline did not facilitate the advancement of CT the same way as other disciplines. They also collected English essays written by 1st-year English majors to confirm their writings unraveled the problems in CT. In spite of the attempts that some researchers have made to disprove the belief that foreign language majors are inferior to other majors because of their poor CT in China's public universities, there is less research to explore the students' CT, in particular, CT dispositions at China's private universities.

Moreover, *the Outline of National Medium- and Long-Term Education Reform and Development Program (2010-2020)* clearly states that China should expand the opening up of education, strengthen international exchanges and cooperation, and improve the internationalization of China's education. Therefore, China's private colleges have recently promoted more international joint programs leading to foreign degrees in which undergraduate students spend 2–3 years in China and the rest of the time in Western universities. With awareness of the belief that foreign language learners are poor at CT, most international joint programs, committed to creating a strong international language and cultural atmosphere, aim to enhance students' foreign language level and develop students' innovative ability, interpersonal communication ability and critical thinking ability (Xu et al., 2019). However, no studies have been found on the level of CT disposition of foreign language majors and the difference in the CT dispositions between students in the international joint programs and the regular programs.

Research Purpose

This study investigated the status of undergraduates' CT at a private college in mainland China, in particular, CT dispositions among foreign language majors across various demographic or academic groupings and with different training or learning experiences. California Critical Thinking Disposition Inventory (CCTDI), revised by Wen et al. (2011), was used to explore the difference in CT between foreign language majors and other liberal arts majors and the differences in CT of undergraduates in the same foreign language discipline but under different programs at a private university in Shanghai, China. Through this study, I investigated existing problems in teaching to inform the improvement of the integration of CT and foreign language teaching to further promote the cultivation of CT in Chinese undergraduate students.

Research Questions

Research questions are listed as follows:

Question 1: What is the current level of the critical thinking dispositions of undergraduate students, particularly the undergraduates in foreign language discipline, at a private college in Shanghai, China?

Question 2: Are there any significant differences in the CT dispositions between foreign language majors and other liberal arts majors (e.g., International Trade, Law, Business Management, Early Childhood Education, etc.) at a private college in China?

Question 3: Are there any significant differences in the CT dispositions between Western foreign language majors (i.e., English, German, French, and Spanish) and Eastern foreign language majors (i.e., Japanese, Korean, and Arabic)?

Question 4: Are there any significant differences in the CT dispositions among foreign language majors representing different demographic or academic groupings (i.e., gender, class level, GPA, leadership position)?

Question 5: Are there any significant differences in the CT dispositions among foreign language majors with different training/learning experiences (i.e., different programs, different courses, different modes of instruction)?

Study Significance

The focus of the research study is to explore whether there are significant differences in critical thinking dispositions among undergraduate students in foreign language discipline across different demographic or academic groupings and those with different training or learning experiences at a private college in Shanghai, China. Through the findings, I hope to diagnose the existing problems in cultivating CT and identify modes of teaching conducive to developing CT at a Chinese private college to offer suggestions with regard to the integration of critical thinking and foreign language teaching to further promote the development of critical thinking in Chinese undergraduate students. This research sought to provide insights for private university administrators and faculty into what academic features and learning experiences are helpful for the cultivation or the development of students' CT dispositions. Furthermore, the study also

sought to find evidence for the effectiveness of international joint programs in cultivation of critical thinkers.

Definitions of Terms

In this section, the key terms relevant to this study are defined.

Critical Thinking

In the Delphi Project, critical thinking is described as purposeful and self-regulatory judgment, involving cognitive skills and affective dispositions (Facione, 1990). Critical thinking is defined by Wen et al. (2009) as the ability to make purposeful, reasoned judgments about things or opinions in accordance with standards.

Critical Thinking Dispositions

Critical thinking disposition is defined as the likelihood that one will engage problems and make decisions by using reasoning (Giancarlo & Facione, 2001), which include inquisitiveness, open-mindedness, systematicity, analyticity, truth-seeking, self-confidence, and cognitive maturity（Facione, 1990; Facione & Facione, 1992). Wen et al. (2011) proposed that CT dispositions include academic dimensions and social dimensions. The former is related to learning activities, and the latter is related to interpersonal activities. The Academic Dimension can be subdivided into the scales of analyticity, inquisitiveness, systematicity, and self-confidence, and the Social Dimension can be subdivided into truth-seeking, cognitive maturity, open-mindedness, and justice-orientedness.

Critical Thinking Skills

Facione (1990) proposed that a core set of cognitive skills consists of six parts: interpretation, analysis, evaluation, inference, explanation and self-regulation. A critical thinker uses CT skills to form a judgement about what to believe or what to do in a certain context (Giancarlo & Facione, 2001).

The Delphi Project

Under the sponsorship of the American Philosophical Association, a panel of experts, using a qualitative Delphi method, conducted Delphi research to work toward a consensus on the role of critical thinking in educational assessment and instruction (Facione, 1990). The Delphi method required the interactive panel of experts to participate in several rounds of questions and share their expertise and reasoned opinions to work toward "a consensus resolution of matters of opinions" (Facione, 1990, p. 3). The research findings and recommendations were presented in the Delphi project.

Frustration Education

Frustration education means that educators consciously use the existing setbacks or set up frustration situations to guide students to treat various difficulties and setbacks in life positively, to help the students to improve their mental endurance, overcome setbacks, and enhance self-esteem and self-confidence through their individual efforts (Tan, 2014).

College students' frustration education mainly consists of: (a) theoretical guidance, including frustration theory, cultural theory, and social theory, which theoretically enables

college students to know the connotation of frustration, deepen their understanding of frustration, and then improve their ability to resist frustration; (b) situational exercise, including thinking training, role play, and simulation practice, which practically enables college students to deepen their understanding of frustration (Fang & Li, 2014).

International Joint Programs

International joint programs are based on partnership agreements between the Chinese institutions and their foreign partners. Certificates or diplomas from the Chinese universities and their foreign partners are awarded to students after they finish these joint programs (F. Huang, 2003). For undergraduates, international joint programs usually last 4 years, typically consisting of 2- or 3-years' study in China and the remaining 2- or 1-year's study in the foreign partner's institution. Students are expected to complete basic courses in the first 2–3 years in local Chinese institutions. In the 3rd or 4th academic years, some students will enter the foreign partner's institutions for degrees.

Public Universities and Colleges

Public universities or colleges, in contrast with private universities, refer to those institutions supported by public subsidies from the national and provincial governments. In their investigation into the CT of undergraduates in Chinese public universities, Zhang and Shen (2018a, 2018b, 2018c) classified public universities into first-class university, key university, and regular university from highest to lowest in academic ranking order.

Private Universities and Colleges

Private universities or colleges (*Minban* in Chinese) refer to nongovernmental institutions, which are not supported by any public subsidy and operate largely on tuition and fees. Private universities rank lower in academic ranking order compared to public universities, and students admitted to private universities usually score lower on the entrance examination.

Chapter Summary

In this chapter, I first introduced the background of the study. Then I presented the research purpose, research questions, and the significance of the study. Last, I defined the relevant terms of the study. In the next chapter, I review recent literature on undergraduates' CT in China.

CHAPTER 2: LITERATURE REVIEW

In this chapter, I review the literature on the critical thinking (CT) of university students, particularly foreign language majors, in mainland China from 2010 to 2019. I first introduce the conceptualizations of CT, and then move on to the method, describing the process of this review. Next, I report results of the synthesis of empirical studies on CT. Finally, I present the discussion and arrive at a conclusion to shed light on the contribution of my study.

Conceptualizations of Critical Thinking

Critical thinking is defined in different ways involving cognitive skills and affective dispositions conducive to decision making and problem solving. In 1989, Richard Paul, president of the National High-Level Council on Critical Thinking, defined critical thinking as the process of actively and skillfully interpreting, applying, analyzing, synthesizing, and evaluating information that governs beliefs and behaviors (R. Paul, 1989). Under the sponsorship of the American Philosophical Association, a team of experts used a qualitative Delphi method to work toward a consensus on the role of CT in educational assessment and instruction. Their findings were presented in the Delphi Project (Facione, 1990), where CT was described as purposeful and self-regulatory judgment, involving cognitive skills and affective dispositions (Facione, 1990). Facione's (1990) two-dimensional model shows that cognitive skills consist of six parts: interpretation, analysis, evaluation, inference, explanation, and self-regulation (Facione, 1990), while affective dispositions include inquisitiveness, open-mindedness, systematicity, analyticity, truth-seeking, self-confidence, and cognitive maturity（Facione, 1990;

Facione & Facione, 1992). According to R. Paul et al. (1995), CT consists of 26 skills and nine dispositions that guide instructors to design in-class activities to cultivate the CT of students. Moreover, Paul and Elder (2006) proposed a three-dimensional structural model, believing that CT ability includes standards, elements, and intellectual traits. The center of the model is the eight elements. They proposed that thinking should have the elements of the purpose, problem, information, concept, assumption, perspective, reasoning, and revelation, which formed a circular chain of thinking. Each element of thinking should be measured or tested by the 10 standards—clarity, accuracy, correlation, logic, breadth, precision, significance, integrity, argument, and depth. Intellectual virtue, including perseverance, independence, confidence, compassion, etc., must be developed with the thinking ability.

In China, Wen Qiufang, an English professor at Beijing Foreign Studies University, defined CT, on the basis of the Delphi project, as the ability to make purposeful, reasoned judgments about things or opinions in accordance with standards (Wen et al., 2009). Making use of Facione's two-dimensional framework (Facione, 1990) and taking in the standards in Paul and Elder's model (2006), Wen et al. (2009) proposed the hierarchical model that divided critical thinking into two levels: meta-critical thinking and critical thinking. The first level of meta-critical thinking refers to the ability to plan, examine, regulate, and evaluate one's own critical thinking. The second level includes skills and standards in relation to cognition and affective dispositions. In the 2-level hierarchical model, researchers see the importance of one's subjective initiative in critical thinking. A visual of the hierarchical model can be found in Table 1. With

the hierarchical model as the theoretical base, Wen, Liu, et al. (2010) designed new instruments for CT skills, revised the existing instrument of California Critical Thinking Disposition Inventory (CCTDI; Wen et al., 2011), and conducted a pilot test with 750 students from three public universities in 2008 (Wen, Wang, et al., 2010).

Table 1

Critical Thinking Hierarchical Model

Level 1: Meta-Critical Thinking (Self-Regulation)		
Level 2: Critical Thinking		
Cognition		Affective Dispositions
Skills	Standards	
Analyticity	Clarity	Inquisitiveness
Inference	Relevancy	Open-mindedness
Evaluation	Logic	Self-confidence
	Depth	Justice-orientedness
	Flexibility	Systematicity

Note. Adapted from *A Study on Critical Thinking of Chinese University Students,* (p. 32), by Q. F. Wen, 2012, Foreign Language Teaching and Research Press.

The large-scale study by Wen, Wang, et al. (2010) was rooted in the importance attached to the cultivation of CT abilities in China and the lack of empirical studies on the CT of university students. The study was intended to call for more follow-up empirical studies to attain more information and guidance for teaching practice. Wen, Zhang, and Sun (2014) were also concerned about whether the perception was true that foreign language majors are inferior to

other majors because of their poor critical thinking (He et al., 1999; Y. S. Huang, 2010; Wen & Zhou, 2006).

H. M. Han et al. (2011) conducted a review of literature from 2002 to 2010 on the CT of foreign language learners. They found that only a few empirical studies were conducted on the application of CT to classroom teaching (S. J. Han & Wang, 2009; Mi & Yuan, 2009; Q. L. Peng, 2000;) and on the assessment of CT of English majors (Wen, Wang, et al., 2010; Wen & Zhou, 2006). Moreover, the cultivation of CT was more connected with English reading and English writing courses than other courses. H. M. Han et al. (2011) confirmed Wen, Wang, et al.'s (2010) concern about the lack of empirical studies on the CT of undergraduate students majoring in English and indicated that Wen, Liu, et al.'s design of a CT instrument (Wen, Liu, et al., 2010) made a big stride in the measurement of CT in the Chinese context.

Therefore, in this chapter, I attempt to give a close look at the recent development by reviewing the empirical studies on the CT of university students, in particular, the CT of foreign language majors—university/college students majoring in a specific foreign language—in mainland China from 2010 to 2019. Followed by Wen, Liu, et al.'s design of the CT instrument in the Chinese context (Wen, Liu, et al., 2010) and the researchers' urgent call for related empirical studies in this field, more scholars have engaged in the study of CT.

Method

The electronic database China National Knowledge Infrastructure (CNKI) was used to search the literature. I used the keywords of critical thinking, Chinese university students, and/or

foreign language majors in Chinese as search terms to find articles for this review. Only peer-reviewed empirical articles in journals on the Chinese Core Journals Index and the Chinese Social Science Citation Index (CSSCI) published from 2010 to 2019 were included in this review.

I sorted through the abstracts to collect articles reporting on the CT of university students in mainland China. Articles that were not related to undergraduates in higher education were excluded from the review. Literature review articles and dissertations were not included in this review. Moreover, articles without "critical thinking" included in the abstracts or keywords were also excluded. The selection criteria are shown in Table 2.

Table 2

Selection Criteria

Inclusion	Exclusion
Peer-reviewed studies	Studies not conducted with undergraduates in higher education
Empirical studies	Articles not including *critical thinking* in the abstracts or keywords
Published from 2010 to 2019	Literature review articles
	Dissertations

After I gathered appropriate articles from the databases, I also searched through the reference lists and used Baidu Scholar to find similar studies that may have been overlooked in the initial search. Again, I read abstracts to find relevant research. Eventually, 28 peer-reviewed

empirical studies were included in this review. Among them, there were 21 quantitative and seven mixed methods studies.

The results are divided into three sections. In the first section, the use or the development of CT instruments with university students is described. In the second, the cultivation of CT in Chinese universities is reported. In the third, the assessment of CT of university students in China is summarized.

Results

In several studies included in this review, various CT instruments were used either to demonstrate the effect of teaching approaches on CT in the process of cultivation or to measure the current CT level of university students in China. Therefore, I present the results in three sections. The use or development of CT instruments, the cultivation of CT, and the assessment of CT are specifically reported as they are included in each piece.

Use or Development of Critical Thinking Instruments

Researchers in the West began to develop instruments to assess the CT of university students in the 1980s (Facione, 1990, 2000a, 2000b; Facione & Facione, 1992; Watson & Glaser, 1980, 1994). The most frequently used CT instruments by the Chinese scholars include the Watson-Glaser Critical Thinking Appraisal (WGCTA; Watson & Glaser, 1980, 1994), the CCTDI (Facione & Facione, 1992; Facione et al., 2001), and the California Critical Thinking Skills Test (CCTST; Facione, 1990, 2000a).

The WGCTA uses 80 questions to assess the five skills: (a) inference, (b) recognition of assumptions, (c) deduction, (d) interpretation, and (e) evaluation of arguments. In China, Han translated WGCTA into Chinese, and Zhu (2002) used it to test 160 students from 4 medical colleges resulting in high reliability and validity. In the new decade, Y. R. Wu (2014) used it in a study of the cultivation of CT. The preexperimental and postexperimental Cronbach's alpha indices of the WGCTA were .77 and .76.

The CCTDI has 75 questions in seven dimensions: (a) truth-seeking, (b) open-mindedness, (c) analyticity, (d) systematicity, (e) self-confidence, (f) inquisitiveness, and (g) cognitive maturity. There are 34 questions in the CCTST, which are used to test logical thinking skills such as analysis, evaluation, inference, deduction, inductive reasoning, and interpretation. Z. Y. Gao (2013) used the CCTDI-2000 (Facione et al., 2001) to assess Chinese students' CT. In the pilot study, the Cronbach's alpha index of the full scale was .87, and the Cronbach's alpha indices of the subscales were .53, .67, .61, .64, .71, and .64. Q. X. Luo and Yang (2001) translated the CCTDI in 2001 and the CCTST in 2002 and then used the translated Chinese version to test 382 university students, which showed a high reliability. The overall Cronbach's alpha index for Luo and Yang's (2001) full CCTDI was .86 and the subscale alphas ranged from .46 to .83. For the Chinese version of the CCTST, the correlation coefficient of retest after 1 month was .63 and the *p* value was smaller than .01. Moreover, the correlation coefficients were .75 and .8, and the *p* values were smaller than .01. Liu and Jin (2012), X. X. Wu (2015),

and Y. J. Wu et al. (2015) used Luo and Yang's (2001) CCTST in their studies of cultivating CT. Y. X. Dong et al. (2010) used the Chinese version of the CCTST-2000 (Facione, 2000a).

Taking Chinese cultural context into account, M. C. Peng et al. (2004) revised the CCTDI into a Chinese Version of Critical Thinking Disposition Inventory (CTDI-CV) and used the CTDI-CV to test 1829 nursing students from six universities. Content validity index (CVI) is most often used to measure content validity which addresses how well the items are developed to operationalize a construct to provide an adequate and representative sample of all the items that might measure the construct of interest (Davis, 1992; Kimberlin & Winsterstein, 2008). The overall CVI was .89, with the subscale CVI ranging from .6 to 1. The overall alpha was .90. Subscale alphas ranged between .54 and .77. These readings show satisfactory content validity and internal consistency (M. C. Peng et al., 2004). R. Ma and Qin (2016), Tian et al. (2018), and Y. R. Wu (2014) all used M. C. Peng et al.'s (2004) CTDI-CV. Similarly, X. Y. Huang (2008) translated the CCTDI and used it to test the nursing staff, which was then used by Mu (2016) to assess students' CT.

Wen, Liu, et al. (2010) designed a new instrument to test the CT skills of university students in China based on other Western CT instruments. Moreover, Wen et al. (2011) revised Luo and Yang's (2001) CCTDI by analyzing the relationship among the different dimensions in the instrument to make it more suitable for Chinese university students. The new skill test consisted of 40 questions in 10 categories, aiming to test students' skills of analyticity, reasoning, and evaluation. The overall difficulty index was .61, discriminating power was .31,

and internal consistency reliability was .70. The revised disposition scale was a 6-point Likert scale including 54 statements within eight dimensions. The overall Cronbach's alpha was .87, and the subscale alphas ranged from .50 to .73. Scholars used them to test the CT skills and dispositions of 2677 university students majoring in English and other liberal arts from 11 public universities in 2008 (Wen, 2012; Wen, Wang, et al., 2010; Wen, Zhang, & Sun, 2014) and 18,825 students consisting of English majors, German majors, some other liberal arts majors and science majors from 51 public universities in 2009 (Wen, 2012; Wen, Zhang, & Sun, 2014). More scholars tended to use Wen et al.'s (2011) instruments in their studies in the recent decade (X. D. Chen, 2013; Liang, 2017; H. Liu & Jin, 2012; Wen et al., 2018; Wen & Zhang, 2016; Q. L. Yu et al., 2016; J. S. Zhang & Fu, 2018).

In addition to Wen et al.'s (2011) new instrument, Y. Dong (2012) also designed an instrument to test the CT skills of university students, which included 33 questions with six dimensions. The subdimensions are analytical argumentation structure, meaning clarification, analytical evaluation argumentation and reasoning, deductive meaning by evaluating information narration, evaluation of information reliability, and identification of implicit hypothesis. The Cronbach's alpha was .62, indicating a satisfactory reliability. Scholars used Dong's (2012) instrument to assess the CT skills of first-class or top-level key universities in China (Q. G. Zhang & Shen, 2018a, 2018b, 2018c).

Moreover, standardized tests in the United States are also used to test university students' achievements (Zhao et al., 2015), one of which is the Educational Testing Service (ETS)

Proficiency Profile (EPP) developed by the ETS. The EPP is mainly used to measure the general ability of university students, including reading, critical thinking, mathematics, and writing abilities. In 2012, the EPP test was translated and revised by Chinese scholars at Beijing University of Aeronautics and Astronautics (Zhao et al., 2015). They only focused on the 27 questions in the section of CT. Therefore, the Chinese version of the EPP CT ability test includes seven aspects: (a) analyze and measure the causal explanation of the contradiction; (b) evaluate the consistency of the assumptions and the given facts; (c) judge the relevance of information to analyze an argument or conclusion; (d) judge whether a perceptual understanding is supported by the evidence of materials; (e) determine whether the procedure of investigating a causal relationship is appropriate; (f) evaluate the consistency between the information and the given facts, and between the assumptions and the method; and (g) identify the flaws and contradictions in the argument (Zhao et al., 2015). The Cronbach's alpha of the Chinese version of EPP was .75, close to .78, which was the alpha for the section of CT in the EPP. Zhao et al. (2015) used EPP (Chinese version) to assess the CT of 2,023 students at a first-class university.

Scholars also used the International Critical Thinking Test (ICTT) to measure students' CT ability (Lu et al., 2018). The ICTT consists of two parts—Part 1 measures students' ability to identify elements of thinking, and Part 2 measures students' ability to measure elements of thinking based on the criteria (ICTT, 2018).

Table 3 shows the use of the previously mentioned CT instruments in the empirical studies on the CT abilities of Chinese university students from 2010 to 2019. It should be noted

that four studies in this review are missing from the table, because instead of using the CT instruments or tests, they studied CT levels of students by evaluating their writing or spoken materials in English (Y. P. Chen, 2016; Pei et al., 2018) or analyzing their online learning activities and their posts in the forum (Leng et al., 2018).

Table 3

Use of CT Instruments in Studies From 2010 to 2019

Instruments	Studies (The Cultivation of CT)	Studies (The Assessment of CT)
WGCTA (Han)	Y. R. Wu (2014)	
CCTST -2000 (Chinese version)		Y. X. Dong et al. (2010)
CCTDI -2000 (English version)		Z. Y. Gao (2013)
CCTST (Q. X. Luo & Yang, 2001)	Liu & Jin (2012) Y. J. Wu et al. (2015) X. X. Wu (2015)	
CTDI-CV (M. C. Peng et al., 2004)	Y. R. Wu (2014) Tian et al. (2018)	Ma & Qin (2016)
CCTDI (X. Y. Huang, 2008)		Mu (2016)
CT Skills Test (Wen, Liu, et al., 2010; Wen, Zhao, et al., 2010)		Wen, Wang, et al. (2010) Wen, Zhang, & Sun (2014) Wen et al. (2018) Liang (2017) Zhang & Fu (2018)
CCTDI (Wen et al., 2011)	Liu & Jin (2012) X. D. Chen (2013)	Wen, Zhang, & Sun (2014) Wen & Zhang (2016) Q. L. Yu et al. (2016)
CCTDI (Revised by authors)		M. Zhang et al. (2016) N. Li et al. (2019)

Instruments	Studies (The Cultivation of CT)	Studies (The Assessment of CT)
CT Skills Test (Y. Dong, 2012)		Zhang & Shen (2018a) Zhang & Shen (2018b) Zhang & Shen (2018c)
EPP (2012)		Zhao et al. (2015)
ICTT	Lu et al. (2018)	

In the next section, I report the cultivation of CT and assessment of CT in Chinese universities by use of the instruments as summarized in this section.

Cultivation of Critical Thinking in Chinese Universities

In the first decade of the 21st century, only a few empirical studies were conducted in the areas of nursing education (S. Z. Liu et al., 2005; W. L. Wang et al., 2006) and English teaching (S. J. Han & Wang, 2009; Mi & Yuan, 2009; Q. L. Peng, 2000) to explore the cultivation of CT in specific courses in China. In the new era, scholars tend to apply different approaches to classroom teaching to improve the students' CT in different courses. This synthesis includes 11 empirical studies from 2010 to 2019, six of which use a mixed methods approach and five of which use a quantitative approach. Particularly, more than half of these studies are conducted with students majoring in English.

Blended Learning

Blended learning, a technical term in educational technology, aims to combine the advantages of traditional learning methods with the advantages of e-learning. It is the emerging teaching strategy based on the online context (R. H. Huang et al., 2009). Singh and Reed (2001)

proposed that blended learning is effective in that it enables the learners to solve real-world examples and to gain guidance on how to solve problems. This decade has witnessed a tendency to study how to teach CT in the online environment, and studies indicate that blended learning by way of blending traditional instruction with online learning is popular in today's online context (Leng et al., 2018; Lu et al., 2018; Y. J. Wu et al., 2015; Y. R. Wu, 2014). The new era has put forward higher demands for blended learning that it should target at the development of thinking (Leng et al., 2018).

Y. R. Wu (2014) used a mixed methods approach to investigate whether blended learning can promote CT skills and CT dispositions of 3rd-year students majoring in education technology at a key university. The *t*-test results showed the CT skills and dispositions of the experimental group were significantly different from those of the control group. The experimental group scored higher in both CT skills and CT dispositions than the control group. Similarly, Leng et al. (2018) designed a series of online learning activities to foster the critical thinking of 55 3rd-year students majoring in educational technology at a first-class university and used the visual approach to clearly demonstrate their online interactions and critical thinking behaviors in online discourse. Research results showed the designed online activities could significantly improve the depth and the level of students' CT, which provided pedagogical references for effectively promoting CT of university students.

Y. J. Wu et al. (2015) used a quantitative approach to verify the effectiveness of the new teaching model of developing CT of 3rd-year students majoring in mathematics in the online

learning environment at a key university. The positive results showed the teaching model could display a general trend for mathematics students to utilize the CT process. In addition, they found topic selection, content, and time arrangement of the activities all influenced the overall cultivation of students' CT in the information-based teaching course.

Although no research was conducted to study blended learning and CT of foreign language majors, Lu et al. (2018) adopted a quantitative approach to investigate whether the constructed blended learning model with orientation of CT could improve the CT and English writing abilities of nonEnglish majors using an English writing course at a first-class normal university as a case. Results showed the experimental group outperformed the control group in CT. However, concerning their English writing abilities, data from the two groups did not show a significant difference.

Regarding the research methods, Lu et al. (2018), Y. J. Wu et al. (2015), and Y. R. Wu (2014) conducted a pretest and a posttest and used *t* tests. In terms of CT instruments, Y. R. Wu (2014) used the WGCTA and the Chinese Version of Critical Thinking Disposition Inventory (CTDI-CV). Y. J. Wu et al. (2015) used the CCTST. Lu et al. (2018) used the International Critical Thinking Test; furthermore, different from the other two studies, Y. J. Wu et al. (2015) used graded essays written by students to measure the CT of students. However, not conducting a pretest or posttest or using any instruments, Leng et al. (2018) used a social network to display the distribution of the students' centrality and activity degree and adopted content analysis to analyze 666 posts from students in the forum.

Problem-Based Learning

In addition to blended learning, problem-based learning (PBL), based on collaborative learning, self-directed learning, and hands-on inquiry-based learning, is another approach to promote the CT of university students (X. D. Chen, 2013; Tian et al., 2018). Scholars took a mixed methods approach to investigate the development of CT of 1st-year engineering students in the basic circuit theory course at a first-class university (Tian et al., 2018) and 1st-year English majors in the integrated English course at a regular public university (X. D. Chen, 2013) through PBL. They conducted a pretest and a posttest on both experimental and control groups and used a *t* test to find the difference between the pretest and posttest results of the two groups. Although Tian et al. (2018) used the Chinese Version of the Critical Thinking Disposition Inventory (CTDI-CV) and X. D. Chen (2013) used the Chinese version of the CCTDI revised by Wen, Zhang, et al. (2014), they all found that PBL played an important role in improving students' CT dispositions.

In addition to the quantitative data, scholars collected different qualitative data. Tian et al. (2018) used the observation form of students' behaviors in the classroom and document analysis on the experimental group, while X. D. Chen (2013) conducted interviews with teachers and students. X. D. Chen (2013) used a College English Test (CET) taken in June and December 2015 and found PBL could significantly promote not only the development of English majors' critical thinking but also their second-language-acquisition ability, which enabled students to develop their ability while constructing new knowledge.

Questioning

Other scholars believed the CT process was the process of asking questions (R. W. Paul, 1995; Seker & Kömür, 2008), and CT cultivation was a process of starting with problems, moving through information processing to find the solution (Q. X. Luo, 2001). Therefore, as early as 2008, Seker and Kömür (2008) conducted a study on the relationship between CT skills and in-class questioning behaviors. Researchers have continued to use recent quantitative studies to explore how teachers' or students' questions influence the CT development of students majoring in English (Y. P. Chen, 2016; Pei et al., 2018; X. X. Wu, 2015).

Y. P. Chen (2016) explored the effect of a teacher's questions on the development of 2nd-year English majors' CT in an argumentative writing course at a first-class university in Beijing. Pei et al. (2018) and X. X. Wu (2015) believed Bloom's taxonomy of educational objectives (Bloom, 1956) was the hierarchical basis for questions at different cognitive levels in class. Pei et al. (2018) investigated whether CT of 2nd-year English majors in the course of integrated English at a first-class university in Nanjing could be fostered through group discussions initiated by questions at different cognitive levels. X. X. Wu (2015) explored whether the sheltered instruction observation protocol (SIOP) model, using questioning to stimulate students' thinking at different cognitive levels, could develop 1st-year English majors' CT at a regular university.

Regarding the instruments, only X. X. Wu (2015) used a questioning ability test and the CCTST (Q. X. Luo & Yang, 2001) on two groups of 30 students and found the SIOP model

could effectively cultivate students' ability to raise high-level cognitive questions and thoroughly improve their critical thinking ability. Not using the CT test, Y. P. Chen (2016) used in-class interaction and essays written by 39 2nd-year students to evaluate whether the teachers' questioning could improve one specific CT disposition—the ability to perceive self-centered points of view. The findings indicated that teachers' questions facilitated students' development of CT. In addition, Pei et al. (2018) recorded and transcribed the group discussions of 61 sophomores from two classes into the corpus and then evaluated the corpus in accordance with the self-designed rubrics based on previous studies (Mu, 2016; Wen, 2012; Wen & Hu, 2010). The application of empirical data and corpus analysis showed that higher level cognitive questions of application, analysis, synthesis, and evaluation better fostered the CT than lower-cognitive questions about knowledge and comprehension.

Concerning research methods, Pei et al. (2018) used a *t* test to compare the different effects of higher level cognitive questions and lower level cognitive questions on students' CT. Without using inferential statistics, both Y.P. Chen (2016) and X. X. Wu (2015) used descriptive statistics to demonstrate the difference in the results from the pretest and posttest.

Content-Based Instruction

Content-based instruction (CBI) is meant to teach foreign languages based on professional knowledge (Mohan, 1986). D. X. Yang and Zhao (2011) adopted a mixed methods approach and argued that English majors' lack of critical thinking may be addressed with CBI at a regular university, which is lower in academic ranking than first-class universities and

secondary key universities. The study conducted the quasi-experiment with an experimental group of 42 3rd-year English majors taking the international trade practice course and a control group of 43 3rd-year English majors taking the advanced English viewing, listening, and speaking course. In accordance with Wen et al.'s (2009) five criteria in the hierarchical model, they assessed the CT skills by reviewing students' essays written in their native language. In addition, the questionnaire was used to survey the effect of CBI on students' metacognitive CT skills and dispositions, and interviews were conducted with students to investigate their perceptions. The results confirmed that CBI could foster the improvement of English majors' CT skills and had a positive influence on their critical cognitive ability. However, results also showed English majors still had a big problem in using different dimensions of CT skills in a skilled, flexible way.

Debate

H. Liu and Jin (2012) used a mixed methods approach and investigated whether the English debate course for 2nd-year English majors at a first-class university in Beijing could help develop students' CT ability. The researchers conducted a pretest and a posttest, using the CCTST and the CCTDI on a class of 23 students and semistructured interviews with 11 of them. The data revealed that students' skills in evaluation, inference, deductive reasoning and inductive reasoning had observed statistically significant improvement after 7 months of training, evaluation being the most significant; analytical skills and critical thinking disposition of the students did not see much improvement. Additionally, the examination-oriented education

system, insufficient knowledge and capability, and learning approaches of students were found to be the underlying factors for the insignificant development of analytical skills and critical thinking disposition.

In summary, using a mixed methods or a quantitative approach, Chinese scholars have tried different teaching approaches to cultivate and develop the CT of students of different disciplines and in different grades at different public universities, mostly at first-class key universities or secondary key universities. With the increasing importance attached to the cultivation of CT, a growing number of scholars are committed to assessing the current CT level of students in Chinese universities.

Assessment of the Critical Thinking of University Students in China

From 2010 to 2019, 17 peer-reviewed papers assessed university students' CT skills or dispositions in China, nine of which target English majors. Nine studies examined the CT skills, seven examined CT dispositions, and only one study examined both CT skills and dispositions.

Many scholars compared and contrasted the CT of students from different levels of public universities (for example, first-class key universities, secondary key universities, and regular public universities; Z. Y. Gao, 2013; Q. G. Zhang & Shen, 2018a, 2018b; Zhao et al., 2015), and different disciplines (Z. Y. Gao, 2013; N. Li et al., 2019; Q. G. Zhang & Shen, 2018b; Zhao et al., 2015). Most studies took into consideration different grades (Y. X. Dong et al., 2010; Z. Y. Gao, 2013; Wen et al., 2018; Wen & Zhang, 2016; Q. L. Yu et al., 2016; J. S. Zhang & Fu, 2018; M. Zhang et al., 2016; Q. G. Zhang & Shen, 2018a, 2018b, 2018c; Zhao et al., 2015). A

few scholars explored the differences in critical thinking ability between the university students who were the only child in the family and those who were not (Q. G. Zhang & Shen, 2018c).

Critical Thinking Tests With Students at Different Levels of Universities

Using the CT skills inventory designed by Y. Dong (2012), Zhang and Shen (2018a, 2018b) measured the CT skills of 8,173 1st-year and 5399 4th-year students from 83 universities consisting of 16 first-class universities, 12 national key universities listed in Project 211 initiated in 1995 by the Ministry of Education of the People's Republic of China, and 55 other regular public universities. Zhang and Shen (2018a) showed undergraduates have a greater ability of CT and more of its added value at first-class universities than in other universities. In fact, students from first-class universities had advantages in CT both as 1st- year students and as graduates. Zhang and Shen (2018b) studied the CT of the undergraduates in the first-class key universities, secondary key universities, regular 4-year universities and regular 4-year colleges, ranking from highest to lowest in academics. They found that (a) the education of higher level universities had a greater value-added effect on students' CT than lower level universities; (b) there was a significant disciplinary difference in the value-added effects of university education, and the value-added effects in liberal arts, medicine, science and engineering are descending sequentially; (c) there was diminishing marginal effect in the improvement of the CT, meaning the higher initial CT skills scores, the less future improvement. The third study by Zhang and Shen (2018c), using the same instrument and sample, indicated significant differences in the CT

skills between an only child and a nononly child in both freshmen and senior groups, but the gap showed a gradual decrease with the passing of time.

Zhao et al. (2015) used EPP (Chinese Version) to test the CT of 1,009 undergraduates from a regular local 4-year university and 1,014 students from a first-class research-oriented university. They found students at the first-class research-oriented university scored significantly higher than those from the local 4-year universities. There is a significant difference in the CT of the students in different grades and disciplines in the local 4-year university. To be specific, students made significant progress in their CT skills in a linear and incremental way and liberal arts majors scored lowest among the students in the four disciplines of liberal arts, social science, natural science, and business.

Using the CCTST-2000 (Chinese version), Y. X. Dong et al. (2010) assessed the CT skills of a stratified sample of 25 4th-year students in different disciplines, genders, and GPAs at a first-class university in Hubei Province and argued the CT skills of the first-class university were satisfactory but left room for improvement. However, due to the limitation of the small sample size, the results are less than convincing. Different findings were revealed in M. Zhang et al. (2016), by use of their revised CCTDI. This study investigated 566 students in three disciplines of liberal arts, science, and engineering at a first-class university, including 188 sophomores, 184 juniors, and 194 seniors. Results showed the CT dispositions of the students at this first-class university were weak overall, and there were no significant differences among students in different grades. Similar to M. Zhang et al. (2016), Z. Y. Gao (2013) used the

CCTDI-2000 to assess the CT dispositions of 489 undergraduates from a regular application-oriented public college. The data showed the 4th-year students scored highest and 3rd-year students scored lowest. However, the dispositions of 94.68% of students fell into the range of ambivalence, indicating their dispositions are not clear. Significant differences were not observed in students of different genders and disciplines. It should also be noted the testing language was English in Z. Y. Gao's (2013) study, and the results might be affected by the language proficiency of different students.

Critical Thinking Tests With University Students of Different Disciplines

Using the CCTDI-2000, N. Li et al. (2019) investigated the CT dispositions of 1075 students from five public universities or colleges in Shandong Province. They found that 84.09% of the students showed ambivalent dispositions and there was a significant difference in the CT of students in different disciplines. Specifically, liberal arts students ranked highest and engineering majors lowest, with natural science students and arts students in between. However, it should be noted the percentage of ambivalent dispositions was unusually high, which was not found by other studies.

Scholars show interest in the study of the CT ability of English majors. Some scholars tested the reliability of the existing instrument (R. Ma & Qin, 2016), examined the CT dispositions of English majors, and explored the relationship between CT dispositions and English writing at a key university (R. Ma & Qin, 2016) and a regular university (Mu, 2016).

One of the most representative research teams is led by Professor Wen Qiufang, which aims to investigate the CT abilities of foreign language majors in China. Wen and her colleagues developed and designed the instruments to measure Chinese university students' CT skills and dispositions (J. Q. Wang & Wen, 2011; Wen, 2012; Wen et al., 2009; Wen, Li, et al., 2014). They used their instruments to test the CT skills and dispositions of 2,677 university students majoring in English and other liberal arts from 11 public universities in 2008 (Wen, 2012; Wen, Wang, et al., 2010; Wen, Zhang, & Sun, 2014) and 18,825 students consisting of English majors, German majors, liberal arts majors, and science majors from 51 public universities in 2009 (Wen, 2012; Wen, Zhang, & Sun, 2014). As a result, they found CT abilities of English majors were higher than other liberal arts or science majors (Wen, Wang, et al., 2010; Wen, Zhang, & Sun, 2014). Studies also showed CT dispositions of English and German majors were positive in general, but decreased significantly through the study of the 1st year and kept stable in the 2nd and 3rd years at two key universities (Wen & Zhang, 2016). English and German majors made significant progress in their CT skills, but not necessarily in a linear and incremental way at two first-class universities (Wen et al., 2018).

A few scholars attempted to use the CT instruments newly designed by Wen, Zhao, et al. (2010) to assess the CT skills of students majoring in English (Liang, 2017; Yu et al., 2016; J. S. Zhang & Fu, 2018). However, their findings differed from previous studies. Liang (2017) found the CT skills of English, law and construction equipment majors were not significantly different and the CT skills of English majors stood in the intermediate level, lower than law majors and

higher than construction equipment majors at a key university. Zhang and Fu (2018) found the foreign language majors were superior to international politics majors in the CT skills at a key university but failed to show a marked improvement in their CT skill development throughout their 4-year university study. Q. L. Yu et al. (2016) was the only study on the CT dispositions of the 281 students majoring in English at a private college in Shandong Province. They found the CT dispositions of English majors were higher in social dimensions (e.g., truth-seeking, maturity, open-mindedness, justice) than in academic dimensions (e.g., analyticity, inquisitiveness, determination) like the previous study (Wen, Zhang, & Sun, 2014). However, no significant difference was found in the CT dispositions of English majors in different grades.

Discussion and Conclusion

This review has described literature on the CT of university students in mainland China. In this section, I summarize the results of this review and describe the implications for further research on the CT of foreign language majors.

Use of Critical Thinking Instruments With University Students in China

Regarding the use of instruments, Table 3 shows seven studies used the various versions of the CCTDI (not including Wen's version), four studies used various versions of the CCTST, one study used the WGTCA, one study used the EPP, and one used the ICTT. After Chinese scholars designed new instruments (Y. Dong, 2012; Wen et al., 2011; Wen, Liu, et al., 2010), five studies used Wen, Liu, et al.'s (2010) CT skills instrument, five studies used Wen et al.'s (2011) revised CCTDI, and three studies used Y. Dong's (2012) CT skills inventory. Noticeably,

the CCTDI is more frequently used than the CCTST in China, and Chinese scholars tend to design new CT skills instruments to cater to Chinese university students.

Regarding the research on foreign language majors, three studies did not use the instruments but the graded writing samples or spoken materials in English to measure the CT level (Y. P. Chen, 2016; Pei et al., 2018; D. X. Yang & Zhao, 2011); three other studies respectively used Luo and Yang's (2001) CCTST (X. X. Wu, 2015), Peng et al.'s (2004) CTDI-CV (R. Ma & Qin, 2016), and Huang's (2008) CCTDI (Mu, 2016); and the remaining nine studies used Wen, Zhao, et al.'s (2010) and Wen et al.'s (2011) instruments (X. D. Chen, 2013; Liang, 2017; Liu & Jin, 2012; Wen, Wang, et al., 2010; Wen et al., 2018; Wen & Zhang, 2016; Wen, Zhang, & Sun, 2014; Q. L. Yu et al., 2016; J. S. Zhang & Fu, 2018). The research team represented by Professor Wen has done a lot of work to enrich the empirical studies of the CT of public university students, particularly foreign language majors. Therefore, it was helpful to use Wen et al.'s CT instruments to assess the CT of students in different majors and at different levels of universities, especially those majoring in foreign languages other than English and German and at private universities, which is limited in the current research literature.

Cultivation of Critical Thinking in Chinese Universities

This review found an increase of empirical studies on the cultivation of the CT of university students in China. Scholars agree that integrating CT's cultivation into the instruction of specific courses is an effective way to comprehensively enhance students' CT skills and dispositions (Wen et al., 2009). Ten of the 11 studies reviewed are concerned with the various

courses in disciplines of mathematics (Y. J. Wu et al., 2015), engineering (Tian et al., 2018), education (Leng et al., 2018; Wu, 2014), and English (Chen, 2013, 2016; H. Liu & Jin, 2012; Pei et al., 2018; X. X. Wu, 2015; D. X. Yang & Zhao, 2011), and one study targeted nonEnglish majors (Lu et al., 2018).

Teaching approaches like blended learning (Leng et al., 2018; Lu et al., 2018; Y. J. Wu et al., 2015; Y. R. Wu, 2014;), PBL (X. D. Chen, 2013; Tian et al., 2018), and CBI (D. X. Yang & Zhao, 2011) can be well applied to the teaching process to positively improve the students' CT abilities. Moreover, the practice of in-class teachers' and students' questioning (Y. P. Chen, 2016; Pei et al., 2018; X. X. Wu, 2015) and the training of debate (H. Liu & Jin, 2012) can also help to enhance students' CT abilities to a certain extent. Some research shows higher order cognitive questions of application, analysis, synthesis, and evaluation better foster CT than lower cognitive questions about knowledge and comprehension (Pei et al., 2018) and SIPO model is an effective way to cultivate students' abilities to raise high order cognitive questions (X. X. Wu, 2015). Other research shows that debate significantly improves students' skills in evaluation, inference, deductive reasoning, and inductive reasoning, but brings no significant difference to students' analytical skills and CT disposition (H. Liu & Jin, 2012).

In the sector of English teaching, the courses to incorporate the teaching of CT are extended from English reading and English writing courses (H. M. Han et al., 2011) to the courses of international trade practice (D. X. Yang & Zhao, 2011) and English debate (H. Liu & Jin, 2012). Scholars are actively engaged in the experiments of the application of PBL (X. D.

Chen, 2013), CBI (D. X. Yang & Zhao, 2011), and SIOP model (X. X. Wu, 2015) in the classroom, and the practice or training of questioning (Y. P. Chen, 2016; Pei et al., 2018), group discussion (Pei et al., 2018), and debate (H. Liu & Jin, 2012) in the teaching process. Different ways of cultivation have been proven to positively influence the CT level of students, so the CT of the students should be studied under different programs (Wen et al., 2018). Therefore, it can be a good start to study the difference in the CT of students enrolled in regular programs and those enrolled in international joint programs at universities.

It is noticeable that some studies did not test the CT of English majors by means of instruments, but rather evaluated their CT by grading their writing or spoken materials based upon specified criteria or rubrics (Y. P. Chen, 2016; Lu et al., 2018; Pei et al., 2018; D. X. Yang & Zhao, 2011). As a result, the subjectivity of the evaluators must not be ignored.

Another notable observation is that only three of the 11 studies were conducted at regular universities (X. D. Chen, 2013; X. X. Wu, 2015; Yang & Zhao, 2011) rather than first-class key universities or secondary key universities, and no research was done on the CT cultivation at private universities. Statistics seemed to indicate that higher level universities tried more teaching approaches to cultivate and develop the students' CT. Therefore, it is worth looking at the CT level of students in different levels of universities.

Assessment of the Critical Thinking of University Students in China

Similar to the studies on the cultivation of CT, research on the assessment of CT was done either to compare the CT level of students at different levels of universities (N. Li et al.,

2019; Wen, Wang, et al., 2010; Wen, Zhang, & Sun, 2014; Q. G. Zhang & Shen, 2018a, 2018b, 2018c) or to test the level of CT of students at first-class or key universities (Y. X. Dong et al., 2010; Liang, 2017; Ma & Qin, 2016; Wen et al., 2018; Wen & Zhang, 2016; M. Zhang et al., 2016; J. S. Zhang & Fu, 2018; Zhao et al., 2015).

Studies indicated the CT of higher level university students is satisfactory and better than lower level university students (Y. X. Dong et al., 2010; Wen et al., 2018; Q. G. Zhang & Shen, 2018a, 2018b; Zhao et al., 2015) and the education at higher level universities has a greater value-added effect on students' CT than lower level universities (Q. G. Zhang & Shen, 2018b). It is worth considering the question as to whether the higher level universities attract better critical thinkers to begin with, rather than assuming they provide value-added educational experiences.

However, there are inconsistent findings about the relationship between CT development and disciplines. Zhang and Shen (2018b) proposed that value-added effects of undergraduate education on the CT were significantly different among disciplines, and the value-added effects in liberal arts, medicine, science, and engineering were descending sequentially. Zhao et al. (2015) found a significant difference in the CT of students in different disciplines of liberal arts, social science, natural science, and business, with liberal arts majors scoring lowest. Although N. Li et al. (2019) found a significant difference in the CT of students in different disciplines of liberal arts, arts, natural science, and engineering, they also found liberal arts majors scored highest and the CT of students majoring in liberal arts, natural science, arts, and engineering decreased sequentially. Wen, Wang, et al. (2010), Wen, Zhang, and Sun (2014), and Zhang and

Fu (2018) found the CT abilities of foreign language majors were higher than other liberal arts or science majors, although there were no significant differences. However, Liang (2017) found no significant difference in the CT skills of English, law, and construction equipment majors, though the CT skills of English majors stood in the intermediate level, lower than law majors and higher than construction equipment majors.

Similarly, no consistent findings are presented in the studies concerning the relation between CT development and grades. Zhang and Shen (2018a, 2018b, 2018c) found the higher the CT in the 1st year, the lower the improvements of CT in the 4th year. Zhao et al. (2015) found that students in different grades made significant progress in their CT skills in a linear and incremental way during the 4 years. Wen and Zhang (2016) showed the CT dispositions of English and German majors decreased significantly through the study of the 1st year and kept fairly stable in the 2nd and 3rd years. Wen et al. (2018) indicated English and German majors made significant progress in their CT skills, but not necessarily in a linear and incremental way. However, J. S. Zhang and Fu (2018) failed to find a marked improvement in their CT skill development throughout their 4-year university study. Moreover, M. Zhang et al. (2016) and Q. L. Yu et al. (2016) found no significant difference among students in different grades.

These inconsistencies may be caused by differences in the instruments, use of the instruments, demographics of samples, and the sample size. For example, studies using the CCTDI-2000 with English as the testing language and the revised CCTDI without the reexamination of validity and reliability, indicated weak or ambivalent student dispositions (Z.

Y. Gao, 2013; M. Zhang et al., 2016). It should be noted that participants' lack of English proficiency and the inadequacy in rigorous scholarship, to some extent, might affect the results.

In particular, despite a growing number of studies of foreign language majors, only two studies were conducted respectively at a regular public university (Mu, 2016) and a private university (Q. L. Yu et al., 2016); moreover, the current research was limited to English and German majors. As the only study conducted at a private university, Q. L. Yu et al. (2016) used a sample of 281 English majors to assess their CT dispositions. Therefore, more research is needed on the CT of foreign language majors at private universities and the difference in their CT from public universities. Moreover, students majoring in other foreign languages need to be included in a future study.

Summary and Implications of Literature Review

In this chapter, I reviewed the literature on the CT of university students in China from 2010 to 2019. The synthesis includes the use of CT instruments, the cultivation of CT, and the assessment of CT, particularly regarding foreign language majors. Based on this review, future studies can be extended to examine the level of CT skills and dispositions of students at private universities in China. Wen et al.'s (2010, 2011) instruments can be used to explore the difference in CT between foreign language majors and other majors and under different teaching programs at private universities.

CHAPTER 3: METHODOLOGY

This study aims to investigate the status of undergraduates' critical thinking (CT) at private universities in mainland China; in particular, CT dispositions among foreign language majors across various demographic or academic groupings and with different training or learning experiences. Wen et al.'s (2011) revised California Critical Thinking Disposition Inventory (CCTDI) was used to explore the difference in CT between foreign language majors and other liberal arts majors and differences in CT of undergraduates in the same foreign language discipline but under various teaching programs at a private university in Shanghai, China. Through my study, I intended to investigate the existing problems and challenges in teaching and then inform the improvement of the integration of CT and foreign language teaching to further promote the cultivation of CT in Chinese undergraduate students. Research questions are as follows.

Question 1: What is the current level of the critical thinking dispositions of undergraduate students at a private college in Shanghai, China?

Question 2: Are there any significant differences in the CT dispositions between foreign language majors and other liberal arts majors (e.g., International Trade, Law, Business Management, Early Childhood Education) at a private college in China?

Question 3: Are there any significant differences in the CT dispositions between Western foreign language majors (i.e., English, German, French, and Spanish) and Eastern foreign language majors (i.e., Japanese, Korean, and Arabic)?

Question 4: Are there any significant differences in the CT dispositions among foreign language majors representing different demographic or academic groupings (i.e., gender, class level, GPA, leadership position)?

Question 5: Are there any significant differences in the CT dispositions among foreign language majors with different training/learning experiences (i.e., different programs, different courses, different modes of instruction)?

Research Design

Creswell and Creswell (2018) proposed three research approaches: (a) qualitative, (b) quantitative, and (c) mixed methods. Qualitative research is an inductive approach to knowledge building aimed at generating meaning (Leavy, 2017). The values include the importance of people's subjective experiences and meaning-making processes and acquiring a depth of understanding. However, quantitative research is a deductive approach for proving, disproving, or lending credence to existing theories (Leavy, 2017). The values underlying quantitative research include neutrality, objectivity, and the acquisition of a sizeable scope of knowledge. Mixed methods research (MMR) is an approach for collecting and integrating both quantitative and qualitative data (Leavy, 2017). Creswell and Creswell (2018) defined quantitative research as:

> an approach for testing objective theories by examining the relationship among variables, which can be measured on instruments so that numbered data can be analyzed using statistical procedures. Like qualitative researchers, those quantitative researchers have

> assumptions about testing theories deductively, building in protections against bias, controlling for alternative or counterfactual explanations, and being able to generalize and replicate the findings. (p. 46)

In the definition, I see the approach to research involves philosophical assumptions and distinct methods or procedures. As Creswell and Creswell (2018) proposed, one's philosophical ideas inform the practice of research. The philosophical ideas help to clarify why the researcher chooses the specific approaches for his research. Therefore, next, I address my philosophical ideas and how they shape my research.

The philosophical ideas are termed worldview, which means "a basic set of beliefs that guide action" (Guba, 1990, p. 17). These ideas are also called perspectives that are mainly comprised of basic topics in philosophy: ontology, epistemology, methodology, and values (J. L. Paul, 2005). Ontology is concerned with the study of what is reality. As a quantitative researcher, I believe existing natural laws or social phenomena govern the world, and these need to be tested or verified and refined to understand the world. Epistemology is to study the nature of knowledge. In my understanding, knowledge is conjectural, meaning absolute truth can never be found. Data, evidence, and rational considerations shape knowledge. Knowledge is based on careful observation and measurement of the objective reality that exists out there in the world. Methodology is to study how knowledge is acquired. In practice, as a researcher, I tend to collect information on instruments based on measures completed by participants or by observations recorded by the researcher. Values concern the role of our aesthetic sensibilities and our beliefs

about what is worthy in our judgements of what is true. I value scientific objectivity, researcher neutrality, and replication (Hesse-Biber & Leavy, 2011). In this study, of particular importance to me is to be objective and examine methods and conclusions for bias. The reliability of the instruments are reported.

Therefore, postpositivism most closely aligned with my research goal of investigating the status of undergraduates' CT at private universities in mainland China, in particular, CT dispositions among foreign language majors across different demographic or academic groupings and with different training or learning experiences. This existing status can be measured by use of the instrument.

Also, certain types of social research problems call for specific approaches (Creswell & Creswell, 2018). My research problems call for describing the current level of CT dispositions, uncovering the differences in CT dispositions between foreign language majors and other liberal arts majors, and differences in the CT dispositions among foreign language majors with different training or learning experiences. Creswell and Creswell (2018) proposed that quantitative nonexperimental research can describe something or uncover relationships between two or more factors. Therefore, a quantitative nonexperimental research design is best to choose. Survey research is one of the quantitative nonexperimental designs. The present investigation includes survey distribution and analysis. According to Fowler (2014), survey research offers a numeric description of trends, attitudes, or opinions of a population by studying a sample of that

population. It uses questionnaires for data collection with the intent of generalizing from a sample to a population.

Participants and Sampling

The quantitative study was conducted at a private university in Shanghai, China, with a population of undergraduate students attending China's private universities. I distributed the survey at the private university where I work and thus employed convenience sampling. Site entry emails were sent to the president of the university and then the deans of the four schools (i.e., School of Foreign languages, School of International Exchange, Business School, Management School, School of Education) at the university. To obtain their approval and support, I introduced the study as a survey of undergraduate students' attitudes and personalities at private universities, which aims to improve the way of teaching based on the students' characteristics. I assured them the research approved by the institutional review board (IRB) had minimal risk to the participants and mitigated potential breaches of confidentiality. Upon approval, participants for survey distribution were recruited by sending the message to the WeChat groups of the diverse groups of students in the same disciplines at the same class levels at the private university. The disciplines included all the foreign languages such as English, German, French, Spanish, Japanese, Korean, and Arabic, and other liberal arts majors such as international trade, finance, accounting, business management, event management, hotel management, law, and early childhood education. Class levels range from freshman, sophomore, and junior. Seniors were not considered in this study considering the fact that a large proportion

of seniors were away from campus due to studying abroad or completing internships away from campus. Five thousand students who met the recruitment criteria were informed by message I sent to the WeChat groups to consider whether to participate in the survey research. The response rate was estimated to be 40%.

Data Collection

Once IRB approved the research proposal, the data collection started at the end of December 2020 and lasted approximately 1 month. I used an online survey because it is considered to have the advantages of reduced cost and time, easy access, quick responses, and ease of entering responses into a database (McMillan, 2016). The message sent to the WeChat groups in Week 1 to request their participation in the survey included an overview of the study, the researcher's contact information, an informed consent form, and a hyperlink to the online survey. The study was introduced as a survey of undergraduate students' attitudes and personalities. The expression "critical thinking" was not used, nor did it appear on the CT instrument. I collected anonymous data and did not record the internet protocol addresses. At the end of Weeks 2 and 3, I sent another message to thank participants who completed the survey and reminded participants of the time remaining before the survey closing. In addition to responses on the 54 items of the CCTDI revised by Wen et al. (2011), the questionnaires also asked for gender, year in school, major, self-reported GPAs, program, and leadership position in various student organizations.

Survey Instrument

The most common measures of critical thinking are CT skill tests and CT disposition inventory. Critics have argued CT skill tests are context-based within particular disciplines and are confounded with the test-takers' existing knowledge in a discipline (Ennis, 1989, 1990; McPeck, 1990). CT disposition inventory is an attitudinal measure which is a relatively stable and enduring aspect of one's personality, which can be used to measure the context-neutral critical thinking (Giancarlo & Facione, 2001).

Therefore, this study used the CCTDI (Facione & Facione, 1992) revised by Wen, Zhao, et al. (2011). Facione and Facione's (1992) CCTDI is a 75-item attitudinal measure, intended for college students. Items were derived from the description of the ideal critical thinker by the American Philosophical Associated sponsored Delphi Project (Facione, 1990) and validated to create the CCTDI. It uses a 6-point Likert scale, ranging from 1 (*strongly agree*) to 6 (*strongly disagree*). The CCTDI reports scores on seven scales: truth-seeking, open-mindedness, analyticity, systematicity, critical thinking, self-confidence, inquisitiveness, and maturity of judgment. The seven scales and corresponding Cronbach's alpha indices are shown in Table 4.

Although the CCTDI has satisfactory reliability, researchers are faced with issues of survey translation when conducting research with diverse cultural groups. Culturally and linguistically appropriate instruments to measure various concepts should be used because an instrument that is not properly adapted can have severe ramifications for study conclusions （McGorry, 2000). Taking this into consideration, Wen et al.'s revision (2011) involved the

revision of the scales and the items based on Luo and Yang's (2001) Chinese version of the CCTDI. They added justice-orientedness to the original seven scales, intended to measure the attitude toward injustice, breach of law, and immorality. Wen et al. (2011) believed justice-orientedness is one of the essential dispositions needed to cultivate in undergraduate students. Without justice-orientedness, one may lose moral autonomy. If one has a high level of CT, one is liable to commit more hidden crimes and thus cause greater harm to society (Wen et al., 2011). In addition, they reduced the original 75 items to 54 items by deleting the items that are not in compliance with Chinese culture and designing six new items in the scale of justice-orientedness. They also improved the items that were inappropriately translated and were ambiguous in meaning. It takes approximately 15 minutes to complete the revised CCTDI.

Table 4

Scales and Their Cronbach's Alpha Indices of CCTDI

Scale	Cronbach's alpha indices
Truth-seeking	.72
Open-mindedness	.73
Analyticity	.72
Systematicity	.74
Self-confidence	.78
Inquisitiveness	.80
Cognitive maturity	.75
Full scale	.90

It should be noted that Wen and colleagues (2011) classified the eight scales into four levels: intellectualness, emotionality, rationality, and morality. Furthermore, intellectualness and emotionality are included in the Academic Dimension, and rationality and morality are included in the Social Dimension. The Academic Dimension is related to the study, and the Social Dimension is related to interpersonal activities. Intellectualness is related to the knowledge level in the learning activities, while emotionality is related to the emotion that is involved in the learning. Meanwhile, rationality is related to the quality or the state of being agreeable to reason in interpersonal activities, while morality is related to the norms of social behavior in interpersonal communication.

The classification is supported by factor analysis, which is used to determine how well the items in one construct go together (Urdan, 2017). The indices of the eight scales classified into the two dimensions range from .56 to .82. The factor analysis results show that such a classification of different scales is helpful to examine the relationship among the eight scales and identify the existing problems in undergraduates' disposition to offer more effective training and education (Wen et al., 2011).

Survey Reliability

An overall score on the revised CCTDI is computed by summing the eight scale scores. Cronbach's alpha is a statistic for investigating the internal consistency of a questionnaire (Bland & Altman, 1997; Cronbach, 1951). Cronbach's alpha internal reliability indices of the eight scales that make up the revised CCTDI range from .50 to .73 (see Table 5). Cronbach's alpha

indices of intellectualness, emotionality, rationality, and morality are respectively .74, .79, .70, and .64. Cronbach's alpha indices of academic and social dimensions are .84 and .80. The alpha of the full scale is .87. For research purposes, alpha should be more than 0.7 to 0.8 (Bland & Altman, 1997). Therefore, the statistics indicated the revised CCTDI is internally consistent and can be used in research (Wen, 2012). However, it should be noted that the low end of Cronbach's alpha internal reliability indices of some subscales is below 0.6, which should be regarded as a limitation upon conclusions.

Scholars used this revised CCTDI to measure the CT dispositions of 2,677 university students majoring in English and other liberal arts from 11 public universities in 2008 (Wen, 2012; Wen, Wang, et al., 2010; Wen, Zhang, & Sun, 2014) and 18,825 students consisting of English majors, German majors, some other liberal arts majors and science majors from 51 public universities in 2009 (Wen, 2012; Wen, Zhang, & Sun, 2014). The results indicated the revised instrument is reliable and can be used extensively to measure Chinese university students' disposition toward critical thinking.

Table 5

Cronbach's Alpha Indices of Wen et al.'s Four-Level CT Dispositions Instrument

Full scale	CCTDI revised by Wen et al. (2011)							
α	*.87*							
Dimen-sion	Academic Dimension				Social Dimension			
α	*.84*				*.80*			
Level	Intellectualness		Emotionality		Rationality		Morality	
α	*.74*		*.79*		*.70*		*.64*	
Scale	Analyti-city	Inquisitive-ness	Systemati-city	Self-confiden-ce	Truth-seekin-g	Cogniti-ve maturit-y	Open-minded-ness	Justice-oriented-ness
α	*.58*	*.62*	*.70*	*.73*	*.50*	*.58*	*.55*	*.50*

Data Analysis

Data were imported into Microsoft Excel and analyzed using Statistical Package for Social Sciences (SPSS, version 25.0). A separate codebook was created for analyzed variables and labeled in the statistical software analysis.

Descriptive and inferential statistics were used to determine the level of CT disposition and analyze all closed-ended items. A *t* test is commonly described as the comparison of two means to see whether there are significant differences between them (Urdan, 2017). Therefore, a *t* test was used to measure: (a) the differences in the CT dispositions between foreign language majors and other liberal arts majors, (b) the differences between male and female students, (c) the differences in the CT dispositions between student-leaders and nonstudent-leaders majoring

in foreign languages, (d) the differences in the CT dispositions between foreign language majors who have taken CT-related courses or training and those who have not, and (e) the difference in the CT dispositions among foreign language majors under different teaching programs. Moreover, Urdan (2017) suggested effect size and statistical significance should be taken into consideration to eliminate the effect of the sample size. Therefore, I also examined Cohen's *d,* a common effect size formula used with *t* tests.

One-way ANOVA, which can compare the means of more than two independent groups to find whether they are significantly different among one another (Urdan, 2017), was used to determine the differences across three class levels. Correlational analysis was used for the relationship between GPA and the disposition toward CT. Correlational analysis was also used for the relationship between the mode of instruction in the language courses and the CT dispositions of foreign language majors.

Ethical Considerations

McMillan (2016) stated that most research in the educational setting has some degree of small risk. IRBs, committed to protecting the rights and welfare of human subjects, have the responsibility to review these so-called "minimal" risk studies. I believe the process is essential as it not only safeguards participants but also enhances the quality of research (McMillan, 2016). Therefore, I did not conduct my research until IRB reviewed and approved my research proposal. I was cognizant of ethical issues and actively worked to minimize these ethical risks.

I conducted research at the university where I work and sent site entry emails to the president of the university and the deans of the five schools. To obtain their approval and support, I introduced the study as a survey of undergraduate students' attitudes and personalities at private universities, which aims to improve the methods of teaching based on the students' characteristics. I also made it clear the risk to the participants was minimized and potential breaches of confidentiality were mitigated. In the process of recruitment, I explained in a letter to potential participants how the information would be kept confidential and secured and explain that the purpose of the study is to understand the attitude of the undergraduate toward some issues, problems, or events and not to evaluate the performance of the participants. Coercion of participation was avoided, and anonymous data were collected. Students' self-reported GPAs were collected. Analyzed data have been secured and will be maintained by the researcher for a minimum of 5 years and will be appropriately discarded thereafter (Creswell, 2008).

Expected Outcomes

The research questions and analyses are listed next.

Question 1: What is the current level of the critical thinking dispositions of undergraduate students at a private college in Shanghai, China?

Hypothesis 1: The undergraduates' dispositions toward CT are positive at a private college in Shanghai, China. Descriptive statistics were calculated to examine the positive, negative, or ambivalent dispositions.

Question 2: Are there any significant differences in the CT dispositions between Western foreign language majors (i.e., English, German, French, and Spanish) and Eastern foreign language majors (i.e., Japanese, Korean, and Arabic)?

Hypothesis 2: The null hypothesis is that there are no significant differences in the CT dispositions between Western foreign language majors and Eastern foreign language majors. The alternate hypothesis is that Western foreign language majors score different from Eastern foreign language majors. A *t* test was used to measure the differences in the CT dispositions between the two groups.

Question 3: Are there any significant differences in the CT dispositions between foreign language majors and other liberal arts majors (e.g., International Trade, Law, Business Management, Early Childhood Education, etc.) at a private university in China?

Hypothesis 3: The null hypothesis is that there are no significant differences in the CT dispositions between foreign language majors and other liberal arts majors at a private university in China. The alternate hypothesis is that foreign language majors score different from other liberal arts majors. A *t* test measured the differences in the CT dispositions between the two groups.

Question 4: Are there any significant differences in the CT dispositions among foreign language majors representing different demographic or academic groupings (i.e., gender, class level, GPA, leadership position)?

Hypothesis 4: The null hypothesis is that there are no significant differences in the CT dispositions among foreign language majors across the demographic or academic variables. The alternate hypothesis is there are significant differences in the CT dispositions among foreign language majors across the demographic variables. A *t* test was used to measure the differences in the CT dispositions between male and female students, and between student-leaders and nonstudent-leaders majoring in foreign languages. One-way ANOVA was used to determine the differences across three class levels. Correlational analysis was used for the relationship between GPA and the disposition toward CT.

Question 5: Are there any significant differences in the CT dispositions among foreign language majors with different training/learning experiences (i.e., different programs, different courses, different modes of instruction)?

Hypothesis 5: The null hypothesis is that there are no significant differences in the CT dispositions among foreign language majors with different training or learning experiences. The alternate hypothesis is that there are significant differences in the CT dispositions among foreign language majors with different training/learning experiences. A *t* test and Cohen's *d* effect size were used to measure the difference in the CT dispositions between foreign language majors under international joint programs and regular programs. Cohen's *d* effect size is used to eliminate the effect of sample size. The other *t* test was used to measure the difference in the CT dispositions between foreign language majors who have taken CT-related courses or training and those who have not. Correlational analysis was used for the relationship between the mode of

instruction in the language courses (i.e., blended teaching, problem-based learning [PBL], CBI, instructors' questioning, and group discussions) and the CT dispositions of foreign language majors.

CHAPTER 4: RESULTS

In the previous chapters, I have introduced the study, reviewed the current literature, and presented the research methodology. In this chapter, I restate the research questions, describe how data were analyzed, and report the results of the survey based on the research questions.

Restatement of Research Questions

This study investigated the status of undergraduates' critical thinking (CT) at a private college in mainland China, in particular, CT dispositions among foreign language majors across various demographic or academic groupings and with different training or learning experiences. Wen et al.'s (2011) revised California Critical Thinking Disposition Inventory (CCTDI) was used to explore the difference in CT between foreign language majors and other liberal arts majors, and the differences in CT of undergraduates in the same foreign language discipline but under different programs at a private college in Shanghai, China. Through my study, I investigated the existing problems in teaching to help identify how to improve the integration of CT and foreign language teaching, which could further promote the cultivation of CT in Chinese undergraduate students.

The research questions were:

Question 1: What is the current level of the critical thinking dispositions of undergraduate students, particularly the undergraduates in the foreign language discipline, at a private college in Shanghai, China?

Question 2: Are there any significant differences in the CT dispositions between foreign language majors and other liberal arts majors (e.g., International Trade, Law, Business Management, Early Childhood Education) at a private college in China?

Question 3: Are there any significant differences in the CT dispositions between Western foreign language majors (i.e., English, German, French, and Spanish) and Eastern foreign language majors (i.e., Japanese, Korean, and Arabic)?

Question 4: Are there any significant differences in the CT dispositions among foreign language majors representing different demographic or academic groupings (i.e., gender, class level, leadership position, GPA)?

Question 5: Are there any significant differences in the CT dispositions among foreign language majors with different training/learning experiences (i.e., different programs, different courses, different modes of instruction)?

Data Analysis Description

Upon the collection of data, I analyzed the data using Statistical Package for Social Sciences (SPSS, version 25.0). Descriptive and inferential statistics were used to determine the level of CT disposition and analyze all closed-ended items. A *t* test was used to measure: (a) the differences in the CT dispositions between foreign language majors and other liberal arts majors, (b) the differences between male and female students, (c) the differences in the CT dispositions between student-leaders and nonstudent-leaders majoring in foreign languages, (d) the differences in the CT dispositions between foreign language majors who have taken CT-related

courses or training and those who have not, and (e) the difference in the CT dispositions among foreign language majors under different teaching programs. Cohen's *d* was also used with *t* tests to eliminate the effect of the sample size.

A one-way ANOVA was used to determine the differences across three class levels—freshmen, sophomores, and juniors. Correlational analysis was used to examine the relationship between GPA and the disposition toward CT. Correlational analysis was also used to examine the relationship between the mode of instruction in the language courses and the CT dispositions of foreign language majors.

General Summary of Sample Demographics

Participants for survey distribution included diverse groups of students in various disciplines at different class levels at a private college. The disciplines included all the foreign languages such as English, German, French, Spanish, Japanese, Korean, and Arabic, and other liberal arts majors such as international trade, finance, accounting, business management, event management, hotel management, early childhood education, law, and journalism. Class levels included freshmen, sophomores, and juniors. The survey link was sent to approximately 5,000 students. The response rate reached 40%. A total of 1,901 undergraduates across various demographic or academic groupings participated in the study and responded to the survey. When data collection was completed, I imported the data into Microsoft Excel and started data cleaning. During the data cleaning process, I removed the cases with missing or mistaken data,

which included those who opted *Below 18 years*, *Senior*, and *Other majors*, and those who did not fill in their GPA. The final sample size was 1,642.

A total of 1,642 students—400 males and 1,242 females—participated in the present study. Table 6 presents the demographic characteristics of the sample. Among the subjects for the sample, there were 694 freshmen, 536 sophomores, and 412 juniors majoring in liberal arts fields. Of the total 1,642, 764 were foreign language majors, which consisted of 565 Western and 199 Eastern foreign language majors, and 878 were other liberal arts majors. Among the 565 Western foreign language majors, 136 participated in international joint programs, and 429 students were not in the international joint programs. Of the overall sample, 756 were student leaders, and 886 were nonstudent leaders; moreover, 582 had taken CT-related courses or training and 1,060 had not.

Table 6

Sample Demographics: Sample Sizes by Category

Demographic		Frequency	%	Valid %	Cumulative %
Gender	Male	400	24.4	24.4	24.4
	Female	1,242	75.6	75.6	100.0
	Total	1,642	100.0	100.0	
Class level	Junior	412	25.1	25.1	25.1
	Sophomore	536	32.6	32.6	57.7
	Freshman	694	42.3	42.3	100.0
	Total	1,642	100.0	100.0	
Leadership position	Student leader	756	46.0	46.0	46.0
	Nonstudent leader	886	54.0	54.0	100.0
	Total	1,642	100.0	100.0	
CT-related courses/training	Taking CT courses/training	582	35.4	35.4	35.4
	Not taking CT courses/training	1,060	64.6	64.6	100.0
	Total	1,642	100.0	100.0	
Disciplines	Foreign language majors	764	46.5	46.5	46.5
	Other liberal arts majors	878	53.5	53.5	100.0
	Total	1,642	100.0	100.0	
Foreign language majors	Western	565	34.4	74.0	74.0
	Eastern	199	12.1	26.0	100.0
	Total	764	46.5	100.0	
Missing	System	878	53.5		
Total		1642	100.0		
Foreign language majors in different programs	International joint programs	136	8.3	24.1	24.1
	Regular programs	429	26.1	75.9	100.0
	Total	565	34.4	100.0	
Missing	System	1,077	65.6		
Total		1642	100.0		

Results

Question 1: What is the current level of the critical thinking dispositions of undergraduate students, particularly the undergraduates in foreign language discipline, at a private college in Shanghai, China?

As mentioned in Chapter 3, the revised CCTDI reports scores on eight scales. The four-level model of CT dispositions is comprised of the top level of CT disposition; the second level of the academic and social dimensions; the third level of intellectualness, emotionality, rationality, and morality; and the bottom level of the eight scales: analyticity, inquisitiveness, systematicity, self-confidence, truth-seeking, cognitive maturity, open-mindedness, and justice-orientedness. According to Giancarlo and Facione (2001) and Wen (2012), for each of the eight scales on the bottom level, a person's score on the revised CCTDI may range from a minimum of 10 points to a maximum of 60 points. Scores are interpreted in accordance with the following guidelines: (a) a score of 40 points or higher indicates a positive inclination or affirmation of the characteristic; (b) a score of 30 or less indicates opposition, disinclination, or hostility toward that same characteristic; and (c) a score in the range of 31–39 points indicates ambiguity or ambivalence toward the characteristic. For the third level, a score of 80 points or higher indicates a positive inclination or affirmation of the characteristic, and a score of 60 or less indicates opposition, disinclination, or hostility toward that same characteristic. For the second level, (a) a score of 160 points or higher indicates a positive inclination or affirmation of the characteristic; (b) a score of 120 or less indicates opposition, disinclination, or hostility toward that same

characteristic; and (c) a score in the range of 121–159 points indicates ambiguity or ambivalence toward the characteristic (Giancarlo & Facione, 2001).

An overall score on the revised CCTDI is computed by summing the eight scale scores. Overall CCTDI scores may range from a minimum of 80 points to a maximum of 480 points. Similar interpretative guidelines are used when looking at overall scores: a total score of 320 points or higher indicates a positive overall disposition toward critical thinking; whereas, a total score of 240 or lower indicates a negative disposition toward critical thinking (Giancarlo & Facione, 2001).

The results are presented based on the four-level model and be reported from the top level to the bottom. The mean scores for participants (N = 1,642) in this data set are presented in Table 7. Mean scores were 318.79 points for the overall scale, indicating ambiguity or ambivalence toward critical thinking. On the second level, 160.16 points indicate a positive inclination or affirmation of the Academic Dimension and 158.62 points indicate ambiguity or ambivalence toward the Social Dimension. On the third level, participants were positively disposed toward Intellectualness and Morality (above 80 points); although, they were ambivalent toward Emotionality and Rationality (below 80 points). It is interesting to find students scored highest on Intellectualness and lowest on Emotionality on the third level, both of which belong to the Academic Dimension on the second level. Correspondingly, on the bottom level, they were positively disposed toward Analyticity, Inquisitiveness, Open-Mindedness, and Justice-

Orientedness (above 40 points) and were ambivalent toward Systematicity, Self-Confidence, Truth-Seeking, and Cognitive Maturity (below 40 points).

Table 7

Descriptive Statistics of Overall Sample

CT Disposition (318.79* A)							
Academic Dimension (160.16* P)				Social Dimension (158.62* A)			
Intellectualness (84.51* P)		Emotionality (75.65* A)		Rationality (77.02* A)		Morality (81.60* P)	
Analyticity (41.83* P)	Inquisitiveness (42.68* P)	Systematicity (36.67* A)	Self-Confidence (38.98* A)	Truth-Seeking (37.42* A)	Cognitive Maturity (39.60* A)	Open-Mindedness (40.43* P)	Justice-Orientedness (41.17* P)

Note. N = 1,642. A = Ambivalent. P = Positive. * = Mean.

Table 8 offers a visual profile of the rankings of the scale scores. Students scored highest on Inquisitiveness and lowest on Systematicity. The other scales—Analyticity, Justice-Orientedness, Open-Mindedness, Cognitive Maturity, Self-Confidence, and Truth-Seeking—respectively, ranked 2–7.

Table 8

Rankings of Scale Scores

Scale	Mean	CT disposition	Rankings
Analyticity	41.83	P	2
Inquisitiveness	42.68	P	1
Systematicity	36.67	A	8
Self-Confidence	38.98	A	6
Truth-Seeking	37.42	A	7
Cognitive Maturity	39.60	A	5
Open-Mindedness	40.43	P	4
Justice-Orientedness	41.17	P	3

Table 9 shows the scores of foreign language majors for the scales on the four levels. The mean scores for each of the CT dispositions scales of foreign language majors were slightly lower than those of the overall sample. A noticeable difference lies in that the mean scores of all the students in the sample showed a positive tendency toward the Academic Dimension, and those of foreign language majors showed ambivalence toward the Academic Dimension. No other differences were found in the two groups in terms of the mean score, rankings of scores, and the disposition toward each characteristic on the four levels (see Table 10).

Table 9

Descriptive Statistics of Foreign Language Majors

CT Disposition (317.17* A)							
Academic Dimension (159.38* A)				Social Dimension (157.79* A)			
Intellectualness (84.15* P)		Emotionality (75.23* A)		Rationality (76.75* A)		Morality (81.03* P)	
Analyti-city (41.59* P)	Inquisi-tiveness (42.56* P)	System-acity (36.32* A)	Self-Confidence (38.90* A)	Truth-Seeking (37.29* A)	Cognitive Maturity (39.47* A)	Open-Minded-ness (40.06* P)	Justice-Orientedness (40.97* P)

Note. n = 764. A = Ambivalent. P = Positive. * = Mean.

Table 10

Rankings of Scale Scores: A Comparison

Scale	Mean		CT disposition	Rankings
	Overall sample	Foreign language majors		
Analyticity	41.83	41.59	P	2
Inquisitiveness	42.68	42.56	P	1
Systematicity	36.67	36.33	A	8
Self-Confidence	38.98	38.90	A	6
Truth-Seeking	37.42	37.29	A	7
Cognitive Maturity	39.60	39.47	A	5
Open-Mindedness	40.43	40.06	P	4
Justice-Orientedness	41.17	40.97	P	3

Question 2: Are there any significant differences in the CT dispositions between foreign language majors and other liberal arts majors (e.g., International Trade, Law, Business Management, Early Childhood Education) at a private college in China?

In terms of the mean scores on all four levels, I found the mean scores for foreign language majors (n = 764) and other liberal arts majors (n = 878) in this data set were 317.17 and 320.20 points for the overall scale, respectively. The statistics indicated foreign language majors were ambiguous or ambivalent toward critical thinking, and other liberal arts majors showed a positive overall disposition toward critical thinking. The mean scores for both groups were below 160 for the Social Dimension. However, the mean scores for foreign language majors were below 160 for the Academic Dimension but those for other liberal arts majors were above 160 for this dimension. The mean scores for both groups were above 80 for Intellectualness and Morality, but below 80 for Emotionality and Rationality. For the eight scales, the mean scores for both groups were above 40 for Analyticity, Inquisitiveness, Open-Mindedness, and Justice-Orientedness but below 40 for the other scales (i.e., Systematicity, Self-Confidence, Truth-Seeking, and Cognitive Maturity). No substantial differences were found in the two groups in terms of the mean score, ranking of scores, and the disposition toward each of the characteristics on the four levels.

An independent samples t test was conducted to compare the scores of the overall scale for foreign language majors and other liberal arts majors (see Table 11). There was no significant difference in the scores for foreign language majors (M = 317.17, SD = 39.61) and other liberal arts majors (M = 320.20, SD = 40.83; t (1640) = -1.52, p = .128, two-tailed). The magnitude of the differences in the means (mean difference = -3.03, 95% CI [-6.94 to -.88]) was very small. The effect size, calculated using Cohen's d, was .01.

Table 11

Group Statistics and Independent Samples Test: Foreign Language Majors vs. Other Liberal Arts Majors

Scale	Disciplines	*M*	*SD*	*t*	*df*	Sig. (2-tailed)
Total Score	Foreign language majors	317.17	39.61	-1.52	1640	0.128
	Other liberal arts majors	320.20	40.83			
Academic Dimension	Foreign language majors	159.38	26.10	-1.12	1640	0.263
	Other liberal arts majors	160.85	26.76			
Social Dimension	Foreign language majors	157.79	28.29	-1.14	1640	0.253
	Other liberal arts majors	159.35	27.14			
Intellectualness	Foreign language majors	84.15	16.63	-0.81	1640	0.418
	Other liberal arts majors	84.82	16.79			
Emotionality	Foreign language majors	75.23	13.16	-1.23	1640	0.219
	Other liberal arts majors	76.02	13.05			
Rationality	Foreign language majors	76.75	16.63	-0.64	1640	0.523
	Other liberal arts majors	77.26	15.51			
Morality	Foreign language majors	81.03	13.47	-1.57	1640	0.116
	Other liberal arts majors	82.09	13.73			
Analyticity	Foreign language majors	41.59	8.89	-1.01	1640	0.314
	Other liberal arts majors	42.04	8.87			
Inquisitiveness	Foreign language majors	42.56	8.75	-0.52	1640	0.603
	Other liberal arts majors	42.79	8.92			
Systematicity	Foreign language majors	36.33	8.07	-1.61	1640	0.108
	Other liberal arts majors	36.97	8.00			
Self Confidence	Foreign language majors	38.90	7.85	-0.41	1640	0.683
	Other liberal arts majors	39.05	7.79			
Truth Seeking	Foreign language majors	37.29	8.36	-0.63	1640	0.531
	Other liberal arts majors	37.54	7.98			
Cognitive Maturity	Foreign language majors	39.47	9.15	-0.58	1640	0.562
	Other liberal arts majors	39.72	8.59			
Open Mindedness	Foreign language majors	40.06	8.63	-1.63	1640	0.104
	Other liberal arts majors	40.75	8.49			
Justice Orientedness	Foreign language majors	40.97	6.71	-1.10	1640	0.270
	Other liberal arts majors	41.34	6.80			

Note. n = 764 (Foreign language majors). n = 878 (Other liberal arts majors).

I conducted *t* tests to compare scales on the other three levels for both groups. No significant differences were found in the scores in these areas for foreign language majors and other liberal arts majors.

To the extent that universities would like to increase the disposition toward critical thinking among all of its students, the data of greatest interest are data that represent the attitudes of portions of the student body (Giancarlo & Facione, 2001). Therefore, scores on the bottom level scales were converted to categories of Positive, Ambivalent, or Negative, using the same cut-off score criteria as explained previously in the discussion of the results of Research Question 1. Table 12 shows the percentages of students in each of the three categories, displayed separately for foreign language majors and other liberal arts majors. No substantial difference in the distribution of percentages in the three categories was found between the two disciplines. However, a notable finding was that the distributions of percentages for both groups were more heavily concentrated in the Ambivalent and Negative category (greater than 50%) than the Positive category (less than 50%), for the Systematicity, Self-Confidence, and Truth-Seeking scales. Educators are more concerned with the ambivalent or negatively disposed students because they are supposed to offer specific ways to change or adjust students' dispositions so students will show a positive shift toward CT (Giancarlo & Facione, 2001).

Table 12

Percentages of Students in Each Category of Endorsement on the Eight Scales Categorized by Discipline

Scale	Foreign language majors			Other liberal arts majors		
	Positive	Ambivalent	Negative	Positive	Ambivalent	Negative
Analyticity	62.2	25.4	12.4	62.5	25.4	12.1
Inquisitiveness	66.5	24.1	9.4	67.0	22.8	10.3
Systematicity	33.8	43.3	22.9	32.9	47.2	19.9
Self-Confidence	42.3	46.5	11.3	42.8	45.4	11.7
Truth-Seeking	44.4	37.3	18.3	42.6	40.4	17.0
Cognitive Maturity	56.8	27.5	15.7	57.2	30.0	12.9
Open-Mindedness	51.3	37.7	11.0	52.1	39.6	8.3
Justice-Orientedness	58.6	34.8	6.5	60.8	33.8	5.4

Question 3: Are there any significant differences in the CT dispositions between Western foreign language majors (i.e., English, German, French, and Spanish) and Eastern foreign language majors (i.e., Japanese, Korean, and Arabic)?

Regarding the mean score for the overall scale, although Western foreign language majors (n = 565) scored slightly higher than Eastern foreign language majors (n = 199), both groups showed ambiguity toward critical thinking. The mean scores for both groups were below 160 for both the Academic and Social Dimensions. The mean scores for both groups were above 80 for Intellectualness and Morality, but below 80 for Emotionality and Rationality. In terms of

mean score of the eight scales, despite similarities, Western foreign language majors tended in the positive direction for Open-Mindedness categorized under Morality, and Eastern foreign language majors were ambivalent toward this characteristic.

An independent samples *t* test was conducted to compare the scores of the overall scale for Western and Eastern foreign language majors. There was no significant difference in the scores for Western foreign language majors (M = 317.29, SD = 40.42) and Eastern foreign language majors (M = 316.81, SD = 37.31; t (762) = .146, p = .884, two-tailed; see Table 13). The magnitude of the differences in the means (mean difference = -.476, 95% CI[-5.94 to 6.89]) was very small (Cohen's d = .01).

I conducted *t* tests to compare the scale scores on the other three levels for both groups. No significant difference was found in the scores in these areas for Western and Eastern foreign language majors.

Table 13

Group Statistics and Independent Samples Test: Western Foreign Language Majors Versus Eastern Foreign Language Majors

Scale	Foreign language major	*M*	*SD*	*t*	*df*	Sig. (2-tailed)
Total Score	Western	317.29	40.42	.15	762	.884
	Eastern	316.81	37.31			
Academic Dimension	Western	159.18	25.82	-.35	762	.725
	Eastern	159.94	26.94			
Social Dimension	Western	158.11	28.01	.53	762	.597
	Eastern	156.87	29.12			
Intellectualness	Western	84.21	15.99	.16	762	.871
	Eastern	83.99	18.35			
Emotionality	Western	74.97	13.30	-.91	762	.366
	Eastern	75.95	12.74			
Rationality	Western	76.84	16.41	.24	762	.812
	Eastern	76.51	17.28			
Morality	Western	81.27	13.53	.82	762	.414
	Eastern	80.36	13.33			
Analyticity	Western	41.50	8.58	-.51	762	.612
	Eastern	41.87	9.73			
Inquisitiveness	Western	42.71	8.45	.82	762	.410
	Eastern	42.12	9.57			
Systematicity	Western	36.28	8.18	-.27	762	.788
	Eastern	36.46	7.75			
Self Confidence	Western	38.69	7.72	-1.24	762	.215
	Eastern	39.49	8.20			
Truth Seeking	Western	37.38	8.19	.52	762	.602
	Eastern	37.02	8.83			
Cognitive Maturity	Western	39.46	9.07	-.04	762	.965
	Eastern	39.49	9.38			
Open Mindedness	Western	40.13	8.64	.38	762	.702
	Eastern	39.86	8.60			
Justice Orientedness	Western	41.14	6.76	1.15	762	.250
	Eastern	40.50	6.55			

Note. n = 565 (Western foreign language majors). n = 199 (Eastern foreign language majors).

Table 14 shows the percentages of students in each of the three categories, displayed separately for Western and Eastern foreign language majors. Though not to the extent of being statistically significant, higher percentages of Western foreign language majors were positively inclined to CT compared to Eastern foreign language majors for Analyticity, Inquisitiveness, Systematicity, Truth-Seeking, and Justice-Orientation scales. For Self-Confidence, Cognitive Maturity, and Open-Mindedness scales, Western foreign language majors were less positive than Eastern foreign language majors. The distribution of percentages for both groups was more heavily concentrated in the Ambivalent and negative category (over 50%) than the positive category (less than 50%) for the Systematicity, Self-Confidence, and Truth-Seeking scales.

Table 14

Percentages of Students in Each Category of Endorsement on the Eight Scales Categorized by Western and Eastern Foreign Language Majors

	Western			Eastern		
Scale	Positive	Ambivalent	Negative	Positive	Ambivalent	Negative
Analyticity	62.8	24.8	12.4	60.3	27.1	12.6
Inquisitiveness	67.6	23.5	8.8	63.3	25.6	11.1
Systematicity	34.0	42.8	23.2	33.2	44.7	22.1
Self-Confidence	41.9	46.2	11.9	43.2	47.2	9.5
Truth-Seeking	44.8	36.6	18.6	43.2	39.2	17.6
Cognitive Maturity	55.8	28.8	15.4	59.8	23.6	16.6
Open-Mindedness	50.6	38.6	10.8	53.3	35.2	11.6
Justice-Orientedness	59.5	34.3	6.2	56.3	36.2	7.5

Question 4: Are there any significant differences in the CT dispositions among foreign language majors representing different demographic or academic groupings (i.e., gender, class level, leadership position, GPA)?

To answer this question, the first variable examined was gender. Table 15 offers a visual image of the mean scores of both groups for the 4-level scales. The mean scores for both males (n = 186) and females (n = 578) in this data set were below 320 for the overall scale. The mean scores for males were above 160 for the Academic Dimension but below 160 for the Social Dimension; however, the mean scores for females were below 160 for the Academic Dimension but above 160 for the Social Dimension. The mean scores for both males (n = 186) and females (n = 578) in this data set were above 80 for Intellectualness but below 80 for Emotionality and Rationality. In terms of Morality, the mean scores of males were below 80 points and those of females were above 80 points. For the eight scales, the mean scores for males and females were above 40 for both Analyticity and Inquisitiveness, and below 40 for the Systematicity and Truth-Seeking scales. For the Self-Confidence scale, the mean score of males was above 40 points but the mean score of females was below 40. However, for the other three scales (i.e., Cognitive Maturity, Open-Mindedness, and Justice-Orientedness), the mean score of males was below 40 points but the mean score of females was above 40.

Table 15

Group Statistics and Independent Samples Test: Males Versus Females

Scale	Gender	*N*	*M*	*SD*	*t*	*df*	Sig. (2-tailed)
Total Score	Male	186	315.23	41.73	-.77	762	.445
	Female	578	317.79	38.92			
Academic Dimension	Male	186	164.96	27.99	3.38	762	.001**
	Female	578	157.58	25.23			
Social Dimension	Male	186	150.27	34.05	-4.21	762	.000***
	Female	578	160.20	25.74			
Intellectualness	Male	186	86.84	18.62	2.54	762	.011**
	Female	578	83.29	15.86			
Emotionality	Male	186	78.12	12.96	3.48	762	.001**
	Female	578	74.29	13.09			
Rationality	Male	186	72.68	20.49	-3.88	762	.000***
	Female	578	78.06	14.96			
Morality	Male	186	77.59	15.22	-4.05	762	.000***
	Female	578	82.14	12.67			
Analyticity	Male	186	43.19	9.73	2.83	762	.005**
	Female	578	41.08	8.54			
Inquisitiveness	Male	186	43.65	9.75	1.95	762	.051
	Female	578	42.21	8.39			
Systematicity	Male	186	37.12	7.93	1.53	762	.126
	Female	578	36.08	8.10			
Self-Confidence	Male	186	41.01	8.39	4.26	762	.000***
	Female	578	38.22	7.55			
Truth-Seeking	Male	186	35.68	10.28	-3.04	762	.002**
	Female	578	37.81	7.57			
Cognitive Maturity	Male	186	37.00	10.95	-4.27	762	.000***
	Female	578	40.26	8.34			
Open-Mindedness	Male	186	37.93	9.53	-3.91	762	.000***
	Female	578	40.74	8.21			
Justice-Orientedness	Male	186	39.66	7.21	-3.09	762	.002**
	Female	578	41.40	6.49			

Note. n = 186 (Male). n = 578 (Female). * for <.05. ** for <.01. *** for <.001.

Statistically significant gender differences were found for Analyticity, t (762) = 2.834, p = .005; Self-Confidence, t (762) = 4.259, p = .000; Truth-Seeking, t (762) = -3.038, p = .002; Cognitive Maturity, t (762) = -4.271, p = .000; Open-Mindedness, t (762) = -3.906, p = .000; Justice-Orientedness, t (762) = -3.088, p = .002; but not for Inquisitiveness and Systematicity. The magnitude of the differences in the means was small (Cohen's d = .2 or .3). Significant difference was also found for all the scales on the second and third level. Males were found to score significantly higher than females on the Academic Dimension, t(762) = 3.378, p = .001, but significantly lower than females on the Social Dimension, t(762) = -4.212, p = .000. However, no significant difference was found for the overall scale.

To gauge the practical significance of the difference found, scores on the scales on the bottom level were again converted to categories of Positive, Ambivalent, or Negative. Table 16 shows the percentages of students in each of the three categories, displayed separately for males and females. Despite highly similar mean scores for each of the CCTDI scales, substantially more male students strongly endorsed Analyticity and Self-Confidence (70.4%, 53.8%, respectively) than female students (59.5%, 38.6%, respectively). Substantially fewer male students strongly endorsed Cognitive Maturity and Open-Mindedness (47.8%, 41.9%, respectively) than female students (59.7%, 54.3%, respectively). For the Truth-Seeking and Justice-Orientedness scales, males were less positive (41.9%, 53.2%, respectively) and less ambivalent (28.5%, 33.9%, respectively) than females (45.2%, 60.4% Positive, respectively, and 40.1%, 35.1% Ambivalent, respectively). These differences were statistically significant.

Likewise, though not to the extent of being statistically significant, there was a higher percentage of positively inclined male than female students for the Inquisitiveness scale and a lower percentage of positively inclined male than female students for the Systematicity scale.

Table 16

Percentages of Students in Each Category of Endorsement on the Eight Scales Categorized by Gender

Scale	Male			Female		
	Positive	Ambivalent	Negative	Positive	Ambivalent	Negative
Analyticity	70.4	18.8	10.8	59.5	27.5	13
Inquisitiveness	71.5	18.3	10.2	64.9	26.0	9.2
Systematicity	33.3	48.4	18.3	33.9	41.7	24.4
Self-Confidence	53.8	39.2	7.0	38.6	48.8	12.6
Truth-Seeking	41.9	28.5	29.6	45.2	40.1	14.7
Cognitive Maturity	47.8	26.9	25.3	59.7	27.7	12.6
Open-Mindedness	41.9	40.9	17.2	54.3	36.7	9.0
Justice-Orientedness	53.2	33.9	12.9	60.4	35.1	4.5

The second variable investigated was leadership position. Table 17 offers a visual image of the mean scores of student leaders and nonstudent leaders for the 4-level scales. For the overall scale, the mean scores for student leaders (n = 361) were above 320 points but those for nonstudent leaders (n = 403) were below 320. The mean scores for student leaders (n = 361)

were above 160 points but those for nonstudent leaders (n = 403) were below 160 for the Academic Dimension. The mean scores for both groups in this data set were below 160 for the Social Dimension. The mean scores for both groups were above 80 for Intellectualness, but below 80 for Emotionality and Rationality. For Morality, the mean scores of student leaders were above 80, but those of nonstudent leaders were below 80. The mean scores for both groups were above 40 for Analyticity, Inquisitiveness, and Justice-Orientedness, but below 40 for Systematicity, Truth-Seeking, and Cognitive Maturity. For Self-Confidence and Open-Mindedness, the mean scores for student leaders were above 40, but those for nonstudent leaders were below 40.

The difference between the two groups can be seen in Table 17. Statistically significant differences were found for the overall scale, t (762)= 3.379, p = .001, and the Academic Dimension, t (762) = 3.224, p = .001, on the second level; Intellectualness, t (762) = 2.246, p = .003, Emotionality, t (762) = 3.552, p = .000, Morality, t (762) = 2.424, p = .016, on the third level; Inquisitiveness, t (762) = 2.572, p = .010, Systematicity, t (762) = 2.151, p = .032, Self-Confidence, t (762) = 3.733, p = .002, and Open-Mindedness, t (762) = 2.703, p = .007, on the 4th level. The magnitude of the differences in the means was small (Cohen's d = .2 or .3). No statistically significant differences were found for the other scales.

Table 17

Group Statistics and Independent Samples Test: Student Leaders Versus Nonstudent Leaders

Scale	Student leader in various organizations at college?	*M*	*SD*	*t*	*df*	Sig. (2-tailed)
Total Score	Student leader	322.25	38.42	3.38	762	.001**
	Nonstudent leader	312.61	40.15			
Academic Dimension	Student leader	162.58	25.41	3.22	762	.001**
	Nonstudent leader	156.52	26.41			
Social Dimension	Student leader	159.67	29.48	1.75	762	.081
	Nonstudent leader	156.10	27.10			
Intellectualness	Student leader	85.58	16.52	2.25	762	.025*
	Nonstudent leader	82.88	16.64			
Emotionality	Student leader	77.00	12.61	3.55	762	.000***
	Nonstudent leader	73.64	13.44			
Rationality	Student leader	77.39	17.46	1.01	762	.313
	Nonstudent leader	76.18	15.84			
Morality	Student leader	82.28	13.91	2.42	762	.016*
	Nonstudent leader	79.92	12.98			
Analyticity	Student leader	42.16	8.70	1.67	762	.096
	Nonstudent leader	41.09	9.04			
Inquisitiveness	Student leader	43.42	8.86	2.57	762	.010*
	Nonstudent leader	41.79	8.60			
Systematicity	Student leader	36.99	7.78	2.15	762	.032*
	Nonstudent leader	35.74	8.28			
Self-Confidence	Student leader	40.01	7.94	3.73	762	.000***
	Nonstudent leader	37.90	7.65			
Truth-Seeking	Student leader	37.58	8.81	.91	762	.362
	Nonstudent leader	37.03	7.94			
Cognitive Maturity	Student leader	39.82	9.49	1.00	762	.317
	Nonstudent leader	39.15	8.83			
Open-Mindedness	Student leader	40.95	9.08	2.70	762	.007**
	Nonstudent leader	39.26	8.13			
Justice-Orientedness	Student leader	41.33	6.63	1.39	762	.164
	Nonstudent leader	40.65	6.77			

Note. n = 403 (Nonstudent leaders). n = 361 (Student leaders). * for <.05. ** for <.01. *** for < .001.

Table 18 shows the percentages of students in each of the three categories, displayed separately for student leaders and Nonstudent leaders. Student leaders were more positive than nonstudent leaders for all the scales. The distributions of percentages for both groups were more heavily concentrated in the ambivalent and negative category (greater than 50%) than the positive category (less than 50%) for the Systematicity, Self-Confidence, and Truth-Seeking scales. For the Open-Mindedness scale, the distributions of percentages for nonstudent leaders were more heavily concentrated in the Ambivalent and Negative category (greater than 50%) as compared to less than 50% for student leaders.

Table 18

Percentages of Students in Each Category of Endorsement on Eight Scales Categorized by Leadership Positions

Scale	Student leader			Nonstudent leader		
	Positive	Ambivalent	Negative	Positive	Ambivalent	Negative
Analyticity	63.7	26.3	10.0	60.8	24.6	14.6
Inquisitiveness	70.1	21.9	8.0	63.3	26.1	10.7
Systematicity	36.8	42.1	21.1	31.0	44.4	24.6
Self-Confidence	47.9	33.8	18.3	36.2	50.4	13.4
Truth-Seeking	41.9	28.5	29.6	41.2	40.4	18.4
Cognitive Maturity	59.8	25.8	14.4	54.1	29.0	16.9
Open-Mindedness	58.2	30.7	11.1	45.2	43.9	10.9
Justice-Orientedness	62.0	31.0	6.9	55.6	38.2	6.2

The third variable examined was class level. The mean scores of juniors (n = 170) were above 320 points, but those of sophomores (n = 273) and freshmen (n = 321) were below 320. Juniors scored above 160 for both the Academic and Social Dimension, while sophomores and freshmen scored below 160 for both. For Intellectualness and Morality, students in the three class levels scored above 80; for Emotionality and Rationality, the three groups all scored below 80. For the eight scales, students in the three class levels scored above 40 points for the Analyticity, Inquisitiveness, and Justice-Orientedness scales, but below 40 for the Systematicity, Self-Confidence, and Truth-Seeking scales. For the Cognitive Maturity scale, only juniors scored above 40 points. For the Open-Mindedness scale, both juniors and sophomores scored above 40.

For universities to increase the disposition toward critical thinking among all of its students, one of the greatest concerns about the data is how the attitudes of students change over time (Giancarlo & Facione, 2001). Figures 1–15 show the tendency of mean scores for each scale across the three class levels. Linear increases were found from freshmen to juniors for the overall scale, as well as the Academic and Social Dimensions; Analyticity, Truth-Seeking, Cognitive Maturity, and Open-Mindedness. For Inquisitiveness and Self-Confidence, the mean scores of sophomores were higher than both juniors and freshmen. A sharp increase was found from freshmen to sophomores for the two scales. A sharp drop was found for Self-Confidence from sophomores to juniors. For the Systematicity and Justice-Orientedness scales, sophomores scored lowest, with a sharp drop from freshmen to sophomores and a sharp increase from sophomores to juniors.

Figure 1

Line Chart of Mean Scores Across Class Levels for Total Score

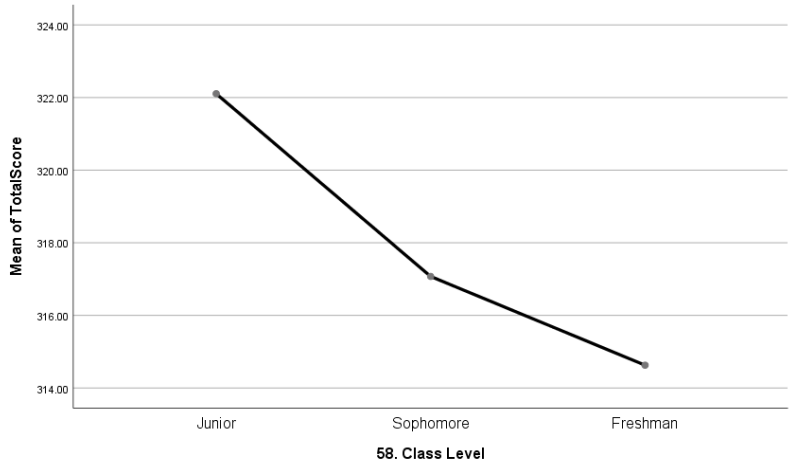


Figure 2

Line Chart of Mean Scores Across Class Levels for Academic Dimension

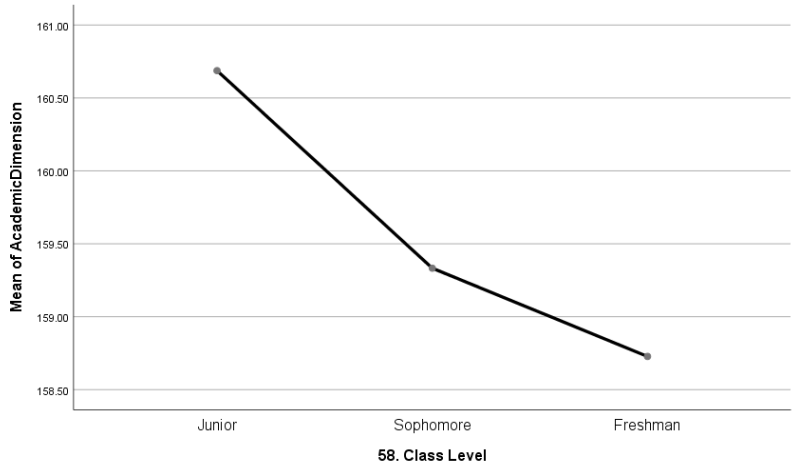


Figure 3

Line Chart of Mean Scores Across Class Levels for Social Dimension

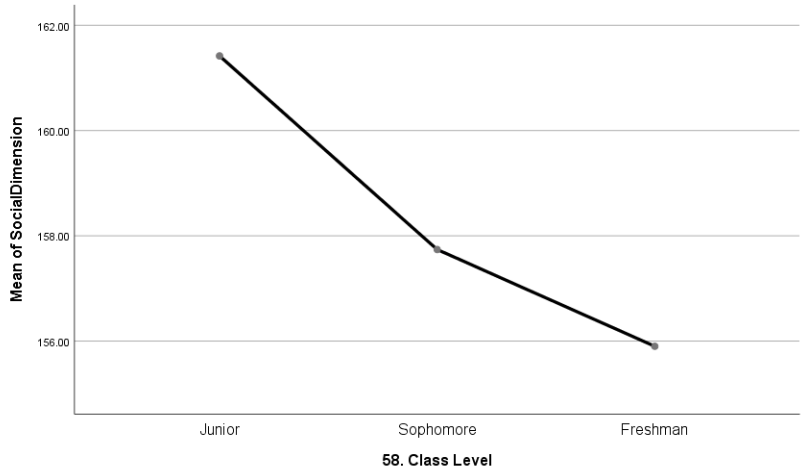


Figure 4

Line Chart of Mean Scores Across Class Levels for Intellectualness

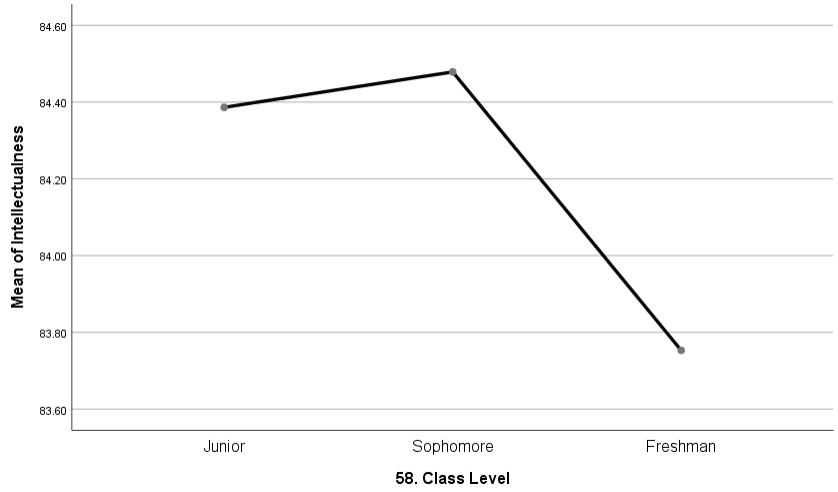


Figure 5

Line Chart of Mean Scores Across Class Levels for Line Chart of Mean Scores Across Class Levels for Emotionality

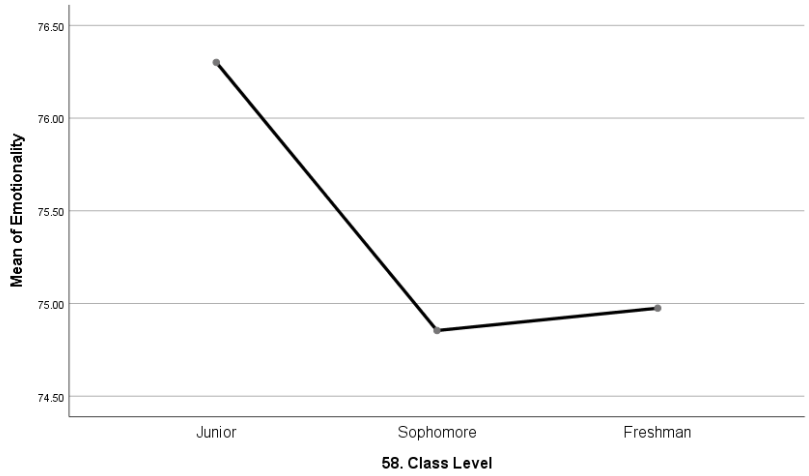


Figure 6

Line Chart of Mean Scores Across Class Levels for Rationality

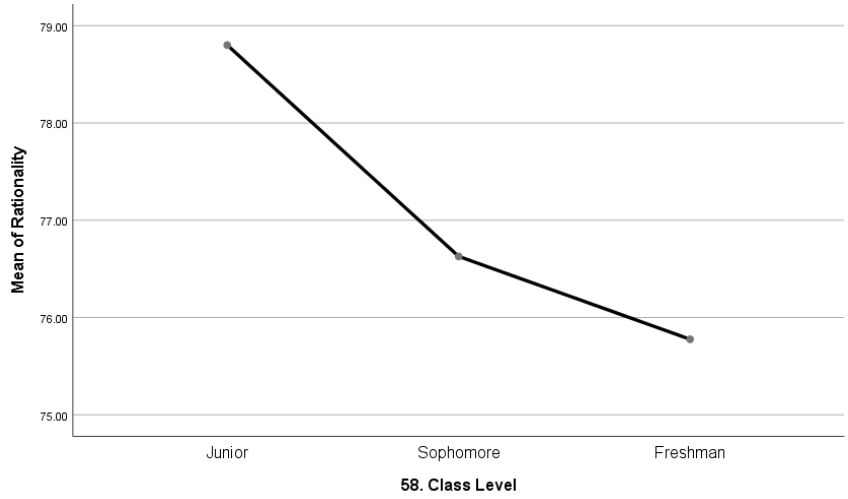


Figure 7

Line Chart of Mean Scores Across Class Levels for Morality

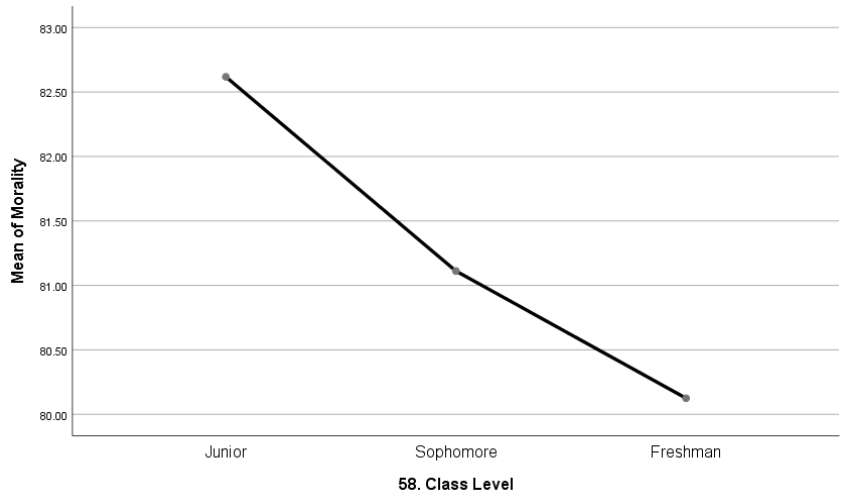


Figure 8

Line Chart of Mean Scores Across Class Levels for Analyticity

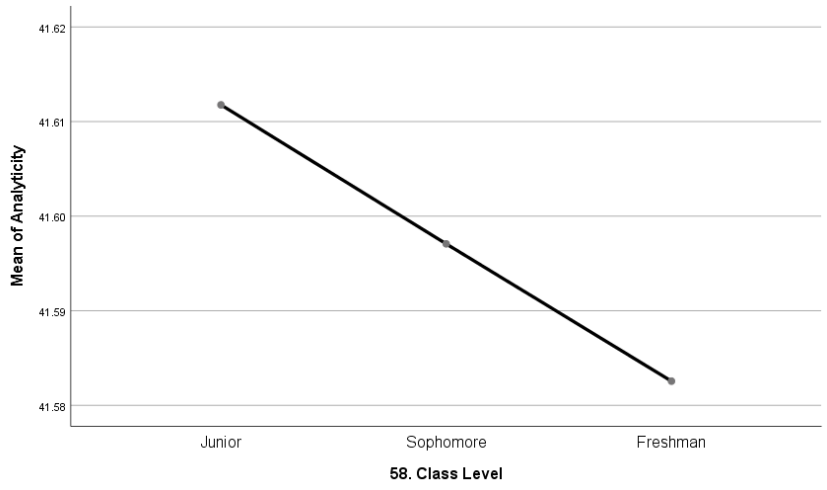


Figure 9

Line Chart of Mean Scores Across Class Levels for Inquisitiveness

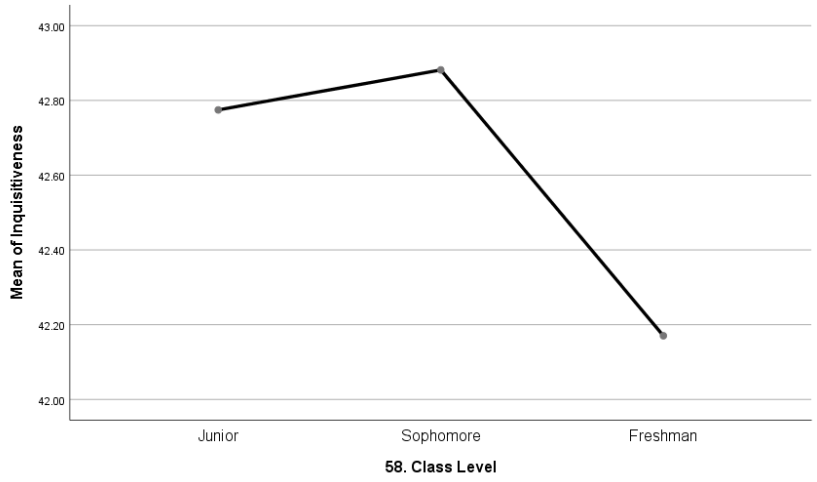


Figure 10

Line Chart of Mean Scores Across Class Levels for Systematicity

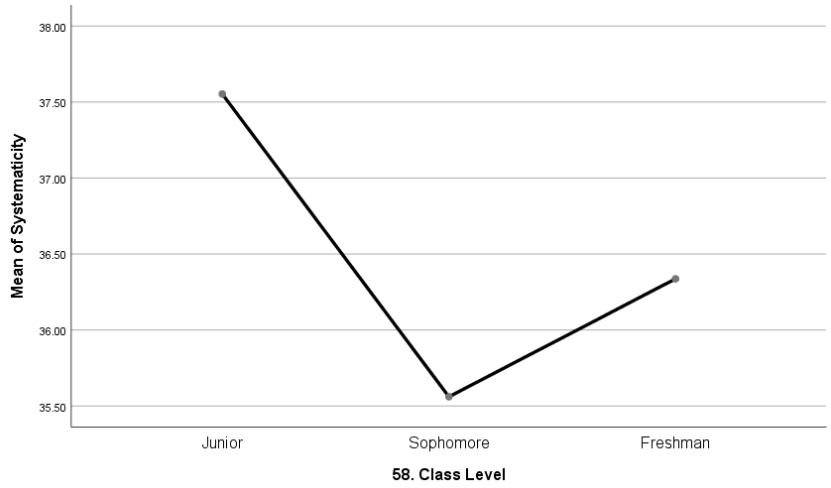


Figure 11

Line Chart of Mean Scores Across Class Levels for Self-Confidence

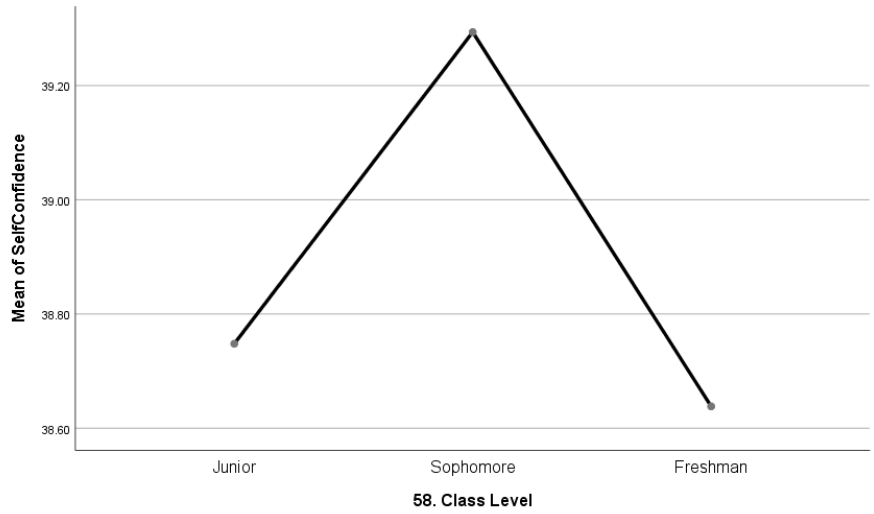


Figure 12

Line Chart of Mean Scores Across Class Levels for Truth-Seeking

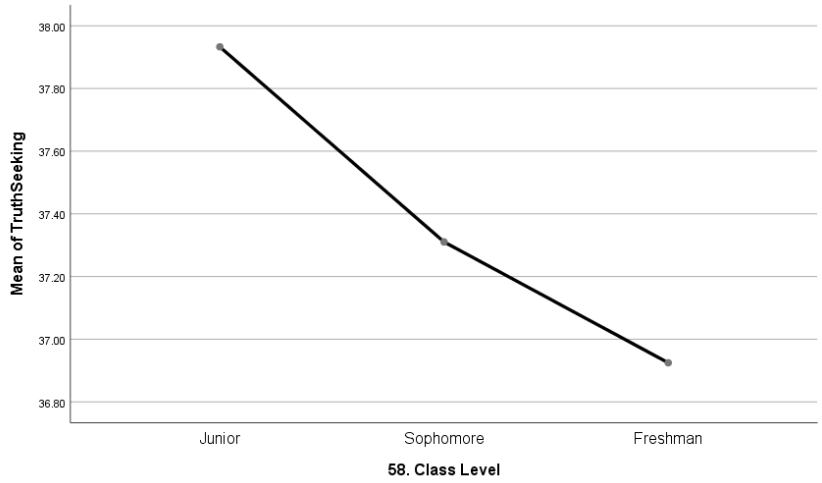


Figure 13

Line Chart of Mean Scores Across Class Levels for Cognitive Maturity

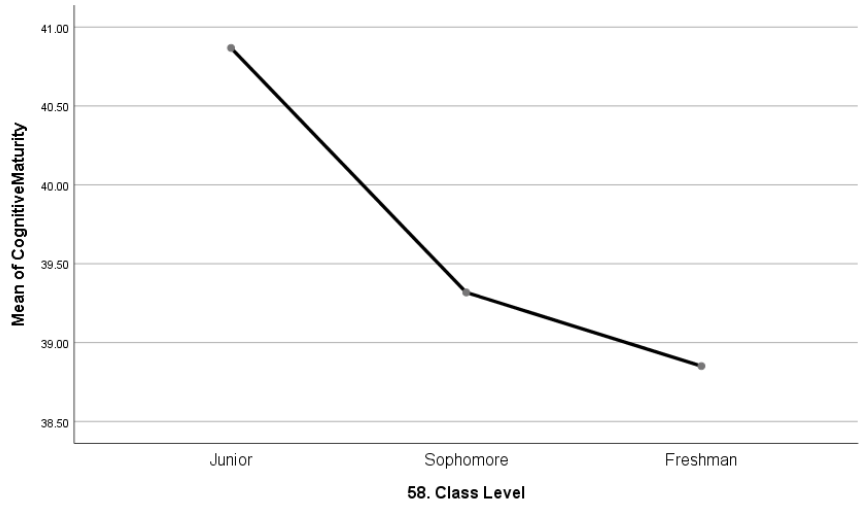


Figure 14

*Line Chart of Mean Scores Across Class Levels for Op**e**n-Mindedness*

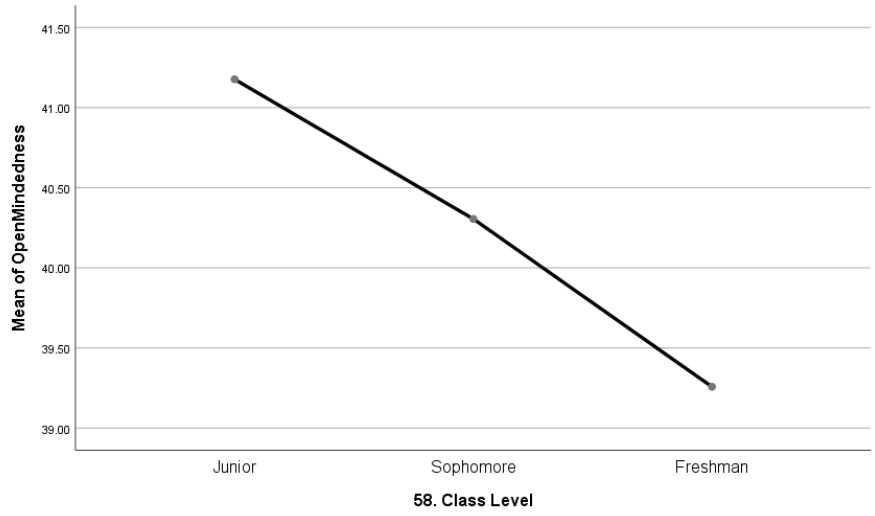
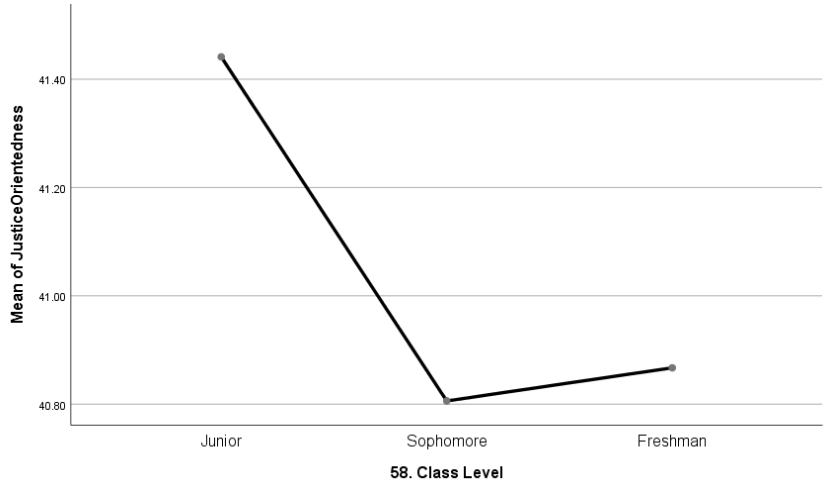


Figure 15

Line Chart of Mean Scores Across Class Levels for Justice-Orientedness



A one-way between-groups analysis of variance was conducted to explore the impact of class level on Total Scores for the overall scale, as measured by the revised CCTDI. Participants were divided into three groups according to their class level (Group 1: Juniors; Group 2: Sophomores; Group 3: Freshmen). There was no statistically significant difference at the $p < .05$ level in scores for the three groups: F (2, 761) = 1.986, p = .138.

A one-way between-groups analysis of variance was conducted to explore the impact of class level on scale scores on the bottom level. As displayed in Table 19, for the Systematicity scale, there was a statistically significant difference at the p < .05 level in scores for the three class level groups: F (2, 761) = 3.214, p = .041. Despite reaching statistical significance, the actual difference in mean scores between the groups was quite small. The effect size, calculated using eta squared, was .01. Post-hoc comparisons using the Tukey HSD test indicated the mean

score for Group 1 (*M* = 37.55, *SD* = 8.36) was significantly different from Group 2 (*M* = 35.56, *SD* = 8.23). Group 3 (*M* = 36.34, *SD* = 7.70) did not differ significantly from either Group 1 or 2.

Table 19

One-Way ANOVA

Dependent variable		Sum of squares	*df*	Mean square	*F*	Sig.
Systematicity	Between Groups	415.940	2	207.970	3.214	.041
	Within Groups	49244.940	761	64.711		
	Total	49660.880	763			
TotalScore	Between Groups	6217.145	2	3108.572	1.986	.138
	Within Groups	1191073.166	761	1565.142		
	Total	1197290.310	763			

Dependent variable	(I) 58. Class Level	(J) 58. Class level	Mean difference (I-J)	Std. Error	Sig.	95% CI Lower bound	95% CI Upper bound
Systematicity	Junior	Sophomore	1.99250*	.78593	.031	.1469	3.8381
		Freshman	1.21649	.76305	.249	-.5754	3.0084
	Sophomore	Junior	-1.99250*	.78593	.031	-3.8381	-.1469
		Freshman	-.77601	.66229	.470	-2.3313	.7792
	Freshman	Junior	-1.21649	.76305	.249	-3.0084	.5754
		Sophomore	.77601	.66229	.470	-.7792	2.3313
TotalScore	Junior	Sophomore	5.03336	3.86521	.394	-4.0433	14.1100
		Freshman	7.47705	3.75267	.115	-1.3354	16.2895
	Sophomore	Junior	-5.03336	3.86521	.394	-14.1100	4.0433
		Freshman	2.44369	3.25714	.734	-5.2051	10.0924
	Freshman	Junior	-7.47705	3.75267	.115	-16.2895	1.3354
		Sophomore	-2.44369	3.25714	.734	-10.0924	5.2051

Note. * The mean difference is significant at the 0.05 level. CI = confidence interval.

Table 20 shows the percentages of students in each of the three categories, displayed separately for juniors, sophomores, and freshmen. Juniors were less positive and more ambivalent than sophomores and freshmen for the Analyticity scale. Juniors were more positive and less ambivalent than sophomores and freshmen for the Systematicity, Open-Mindedness, and Justice-Orientedness scales. For the Self-Confidence scale, juniors were less positive and more ambivalent than sophomores but more positive and less ambivalent than freshmen. For the Truth-Seeking scale, juniors were less positive and more ambivalent than sophomores but more positive and more ambivalent than freshmen. Juniors were more positive and more ambivalent than sophomores but more positive and less ambivalent than freshmen for the Cognitive Maturity scale. No substantial difference was found for the Inquisitiveness scale. The distribution of percentages for all the three class levels was more heavily concentrated in the Ambivalent and Negative category (greater than 50%) as compared to the Positive category (less than 50%) for the Systematicity, Self-Confidence, and Truth-Seeking scales. Moreover, the distribution of percentages for freshmen was more heavily concentrated in the Ambivalent and Negative category (greater than 50%) than the Positive category (less than 50%) for the Open-Mindedness scale.

Table 20

Percentages of Students in Each Category of Endorsement on the Eight Scales Categorized by Class Level

Scale	Junior			Sophomore			Freshman		
	P	A	N	P	A	N	P	A	N
Analyticity	58.8	28.8	12.4	64.1	22.0	13.9	62.3	26.5	11.2
Inquisitiveness	66.5	22.9	10.6	66.7	23.8	9.5	66.4	24.9	8.7
Systematicity	42.9	40.0	17.1	30.8	42.9	26.4	31.5	45.5	23.1
Self-Confidence	42.9	45.3	11.8	46.5	41.0	12.5	38.3	51.7	10.0
Truth-Seeking	44.1	41.8	14.1	45.4	36.3	18.3	43.6	35.8	20.6
Cognitive Maturity	59.4	27.6	12.9	57.1	27.1	15.8	55.1	27.7	17.1
Open-Mindedness	57.1	35.3	7.6	52.4	38.5	9.2	47.4	38.3	14.3
Justice-Orientedness	63.5	28.8	7.6	58.2	34.8	7.0	56.4	38.0	5.6

Note. P = positive; A = ambivalent; N = negative.

The fourth variable to be examined was GPA. Pearson product-moment correlation coefficient found no statistical correlation between GPA and CT for juniors and freshmen (see Table 21). The relationship between GPA and Total Score for the overall scale in the group of sophomores was also investigated using Pearson product-moment correlation coefficient. Preliminary analyses were performed to ensure no violation of the assumptions of normality, linearity, and homoscedasticity. There was a small, positive correlation between the two

variables (r = .16, n = 273, p =.009 < .01), with higher levels of GPA associated with Total Score. GPA helped to explain 2% of the variance in respondents' scores on the Overall Scale.

Then I looked into the correlation between GPA and the scores for the scales of the other three levels for sophomores (see Tables 22–24). There was a small, positive correlation between GPA and the Academic Dimension (r = .13, n = 273, p =.003 < .01), with higher levels of GPA associated with the Academic Dimension. GPA helped to explain nearly 2% of the variance in respondents' scores on the Academic Dimension on the second level.

There was a small, positive correlation between GPA and Emotionality, (r = .19, n = 273, p =.002 < .01), with higher levels of GPA associated with Emotionality. GPA helped to explain nearly 4% of the variance in respondents' scores on Emotionality on the third level. There was a small, positive correlation between GPA and Systematicity, (r = .22, n = 273, p = .000 < .01), with higher levels of GPA associated with Systematicity. GPA helped to explain nearly 5% of the variance in respondents' scores on Systematicity on the fourth level.

Table 21

Correlation Between GPA and CT Disposition

Class level			Academic achievement	Totalscore
Junior	Academic achievement	Pearson Correlation	1	.053
		Sig. (2-tailed)		.493
		N	170	170
	Totalscore	Pearson Correlation	.053	1
		Sig. (2-tailed)	.493	
		N	170	170
Sophomore	Academic achievement	Pearson Correlation	1	.158**
		Sig. (2-tailed)		.009
		N	273	273
	Totalscore	Pearson Correlation	.158**	1
		Sig. (2-tailed)	.009	
		N	273	273
Freshman	Academic achievement	Pearson Correlation	1	-.058
		Sig. (2-tailed)		.303
		N	321	321
	Totalscore	Pearson Correlation	-.058	1
		Sig. (2-tailed)	.303	
		N	321	321

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

Table 22

Correlation Between GPA and Academic Dimension

Class level			Academic achievement	Academic dimension
Junior	Academic achievement	Pearson correlation	1	-.018
		Sig. (2-tailed)		.811
		N	170	170
	Academic dimension	Pearson correlation	-.018	1
		Sig. (2-tailed)	.811	
		N	170	170
Sophomore	Academic achievement	Pearson correlation	1	.129*
		Sig. (2-tailed)		.033
		N	273	273
	Academic dimension	Pearson correlation	.129*	1
		Sig. (2-tailed)	.033	
		N	273	273
Freshman	Academic achievement	Pearson correlation	1	-.087
		Sig. (2-tailed)		.119
		N	321	321
	Academic dimension	Pearson correlation	-.087	1
		Sig. (2-tailed)	.119	
		N	321	321

Note. * Correlation is significant at the 0.05 level (2-tailed).

Table 23

Correlation Between GPA and Emotionality

Class level			Academic achievement	Emotionality
Junior	Academic achievement	Pearson correlation	1	-.012
		Sig. (2-tailed)		.879
		N	170	170
	Emotionality	Pearson correlation	-.012	1
		Sig. (2-tailed)	.879	
		N	170	170
Sophomore	Academic achievement	Pearson correlation	1	.188**
		Sig. (2-tailed)		.002
		N	273	273
	Emotionality	Pearson correlation	.188**	1
		Sig. (2-tailed)	.002	
		N	273	273
Freshman	Academic achievement	Pearson correlation	1	-.061
		Sig. (2-tailed)		.279
		N	321	321
	Emotionality	Pearson correlation	-.061	1
		Sig. (2-tailed)	.279	
		N	321	321

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

Table 24

Correlation Between GPA and Systematicity

Class level			Academic achievement	Systematicity
Junior	Academic achievement	Pearson correlation	1	-.007
		Sig. (2-tailed)		.927
		N	170	170
	Systematicity	Pearson correlation	-.007	1
		Sig. (2-tailed)	.927	
		N	170	170
Sophomore	Academic achievement	Pearson correlation	1	.221**
		Sig. (2-tailed)		.000
		N	273	273
	Systematicity	Pearson correlation	.221**	1
		Sig. (2-tailed)	.000	
		N	273	273
Freshman	Academic achievement	Pearson correlation	1	-.022
		Sig. (2-tailed)		.696
		N	321	321
	Systematicity	Pearson correlation	-.022	1
		Sig. (2-tailed)	.696	
		N	321	321

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

Question 5: Are there any significant differences in the CT dispositions among foreign language majors with different training/learning experiences (i.e., different programs, different courses, different modes of instruction)?

Three variables were examined to answer this question. The first variable examined was different programs. Table 25 offers a visual image of the mean scores of both international joint programs (n = 136) and noninternational joint programs (n = 429) for the 4-level scales. The

mean score for students in international joint programs was above 320, but for noninternational joint programs, it was below 320 for the overall scale. The mean score for international joint programs was above 160 for the Academic Dimension, but for noninternational joint programs, it was below 160 for the Academic Dimension; however, mean scores for both groups were below 160 for the Social Dimension. The mean scores for both groups were above 80 for Intellectualness and Morality, but below 80 for Emotionality and Rationality. For the eight scales, the mean scores for both groups were above 40 for Analyticity, Inquisitiveness, and Justice-Orientedness, but below 40 for Systematicity, Self-Confidence, Truth-Seeking, and Cognitive Maturity. For Open-Mindedness, the mean score of international joint programs was above 40 points, but the mean score of noninternational joint programs was below 40.

Statistically significant differences were found for Intellectualness, t (563) = 2.118, p = .035; Analyticity, t (563) = 1.982, p = .048; and Inquisitiveness, t (563) = 1.996, p = .046. The magnitude of the differences in the means was small (Cohen's d = .2).

Table 25

Group Statistics and Independent Samples Test: International Joint Programs Versus Regular Programs

Parameter	Foreign language majors Different Programs	*M*	*SD*	*t*	*df*	*Sig. (2-tailed)*
Totalscore	International joint programs	320.43	42.61	1.04	563	.300
	Regular programs	316.29	39.71			
Academic Dimension	International joint programs	161.36	26.09	1.13	563	.260
	Regular programs	158.49	25.72			
Social Dimension	International joint programs	159.07	29.14	.46	563	.646
	Regular programs	157.80	27.67			
Intellectualness	International joint programs	86.74	15.80	2.12	563	.035*
	Regular programs	83.41	15.99			
Emotionality	International joint programs	74.62	14.59	-.35	563	.726
	Regular programs	75.08	12.88			
Rationality	International joint programs	76.85	17.18	.01	563	.992
	Regular programs	76.83	16.18			
Morality	International joint programs	82.22	14.12	.94	563	.348
	Regular programs	80.97	13.33			
Analyticity	International joint programs	42.76	8.77	1.98	563	.048*
	Regular programs	41.10	8.49			
Inquisitiveness	International joint programs	43.97	8.07	2.00	563	.046*
	Regular programs	42.32	8.54			

Parameter	Foreign language majors Different Programs	*M*	*SD*	*t*	*df*	*Sig. (2-tailed)*
Systematicity	International joint programs	36.03	8.91	-.42	563	.678
	Regular programs	36.36	7.95			
Self-Confidence	International joint programs	38.59	8.12	-.17	563	.869
	Regular programs	38.72	7.60			
Truth-seeking	International joint programs	37.08	8.69	-.49	563	.624
	Regular programs	37.48	8.04			
Cognitive Maturity	International joint programs	39.77	9.62	.46	563	.646
	Regular programs	39.36	8.90			
Open-Mindedness	International joint programs	41.23	8.85	1.70	563	.090
	Regular programs	39.78	8.55			
Justice-Orientedness	International joint programs	40.99	7.19	-.29	563	.773
	Regular programs	41.18	6.62			

Note. n = 429 (Regular programs). n =136 (International joint programs). * for < .05.

Table 26 shows the percentages of students in each of the three categories, displayed separately for Students in International Joint Programs and Regular Programs. For the Analyticity, Inquisitiveness, Self-Confidence, Open-Mindedness, and Justice-Orientedness scales, students in International Joint Programs were more positive than students in Regular Programs. For the Systematicity, Truth-Seeking, and Cognitive Maturity scales, students in International Joint Programs were less positive than students in regular programs. For the Open-Mindedness scale, the distributions of percentages for students in international joint programs

were less heavily concentrated in the Ambivalent and Negative category (41.9%) as compared to the 51.7% for students in regular programs. For the Systematicity, Self-Confidence, and Truth-Seeking scales, the distributions of percentages for both groups were more heavily concentrated in the Ambivalent and Negative category (greater than 50%) than the Positive Category (less than 50%).

Table 26

Percentages of Students in Each Category of Endorsement on the Eight Scales Categorized by Program

Scale	International joint programs			Regular programs		
	Positive	Ambivalent	Negative	Positive	Ambivalent	Negative
Analyticity	66.9	25.0	8.1	61.5	24.7	13.8
Inquisitiveness	69.9	24.3	5.9	66.9	23.3	9.8
Systematicity	32.4	41.2	26.5	34.5	43.4	22.1
Self-Confidence	44.9	39.0	16.2	41.0	48.5	10.5
Truth-Seeking	43.4	36.8	19.9	45.2	36.6	18.2
Cognitive Maturity	55.1	30.9	14.0	55.9	28.2	15.9
Open-Mindedness	58.1	33.1	8.8	48.3	40.3	11.4
Justice-Orientedness	61.0	29.4	9.6	59.0	35.9	5.1

The second variable investigated was different courses. Table 27 offers a visual image of the means scores of the students who have taken CT-related courses or training (Group 1, n = 218, M = 321.57, SD = 41.21) and those who have not (Group 2, n = 546, M = 315.41, SD =

38.85) for the 4-level scales. The mean score for Group 1 in this data set was above 320 for the overall scale, but for Group 2, it was below 320 points for the overall scale. The mean score for Group 1 was above 160 for the Academic Dimension, but for Group 2 it was below 160 for the Academic Dimension; however, the mean scores for both groups were below 160 for the Social Dimension. The mean scores for both groups were above 80 for Intellectualness and Morality, but below 80 for Emotionality and Rationality. For the eight scales, mean scores for both groups were above 40 for Analyticity, Inquisitiveness, and Justice-Orientedness, but below 40 for Systematicity, Truth-Seeking, and Cognitive Maturity. For Self-Confidence, the mean score for Group 1 was above 40, but for Group 2, it was below 40; however, for Open-Mindedness, the mean score for Group 1 was below 40, but for Group 2, it was above 40.

Statistically significant differences were found for the Academic Dimension, t (762) = 3.539, p = .000, two-tailed; Intellectualness, t (762) = 2.590, p = .010, two-tailed; Emotionality, t (762) = 3.739, p = .000, two-tailed; Analyticity, t (762) = 2.245, p = .025, two-tailed; Inquisitiveness, t (762) = 2.638, p = .009, two-tailed; Systematicity, t (762) = 2.633, p = .009, two-tailed; and Self-Confidence, t (762) = 3.545, p = .000, two-tailed. The magnitude of the differences in the means was small (Cohen's d = .2 or .3). No significant differences were found for the overall scale and other scales.

Table 27

Group Statistics and Independent Samples Test: CT Courses or No CT Courses

68. Have you taken any courses or training related to thinking?		*M*	*SD*	*t*	*df*	*Sig. (2-tailed)*
TotalScore	Taking CT courses/training	321.57	41.21	1.94	762	.052
	Not taking CT courses/training	315.41	38.85			
Academic Dimension	Taking CT courses/training	164.63	28.25	3.54	762	.000***
	Not taking CT courses/training	157.28	24.91			
Social Dimension	Taking CT courses/training	156.94	31.24	-.52	762	.600
	Not taking CT courses/training	158.12	27.05			
Intellectualness	Taking CT courses/training	86.61	18.44	2.59	762	.010*
	Not taking CT courses/training	83.17	15.76			
Emotionality	Taking CT courses/training	78.02	13.19	3.74	762	.000***
	Not taking CT courses/training	74.11	12.99			
Rationality	Taking CT courses/training	76.22	18.63	-.56	762	.579
	Not taking CT courses/training	76.96	15.77			
Morality	Taking CT courses/training	80.71	14.48	-.42	762	.678
	Not taking CT courses/training	81.16	13.06			
Analyticity	Taking CT courses/training	42.73	9.58	2.25	762	.025*
	Not taking CT courses/training	41.14	8.56			
Inquisitiveness	Taking CT courses/training	43.88	9.70	2.64	762	.009**
	Not taking CT courses/training	42.03	8.30			
Systematicity	Taking CT courses/training	37.54	7.95	2.63	762	.009**
	Not taking CT courses/training	35.85	8.07			
Self-Confidence	Taking CT courses/training	40.48	8.46	3.55	762	.000***
	Not taking CT courses/training	38.27	7.51			
Truth-Seeking	Taking CT courses/training	37.18	9.24	-.22	762	.827
	Not taking CT courses/training	37.33	7.99			
Cognitive Maturity	Taking CT courses/training	39.04	10.07	-.81	762	.418
	Not taking CT courses/training	39.64	8.76			
Open-Mindedness	Taking CT courses/training	39.47	8.98	-1.19	762	.235
	Not taking CT courses/training	40.29	8.48			
Justice-Orientedness	Taking CT courses/training	41.24	6.92	.69	762	.490
	Not taking CT courses/training	40.87	6.63			

Note. n = 218 (Taking CT courses). n = 546 (Not taking CT courses). * for <.05. ** for <.01.

*** for <.001.

Table 28 shows the percentages of students in each of the three categories, displayed separately for students taking and not taking CT-related courses. Students who had taken CT-related courses or training were more Positive than students who had not, for all the scales except the Cognitive Maturity scale. For the Self-Confidence scale, the distribution of percentages for students who had taken CT-related courses or training were less heavily concentrated in the Ambivalent and Negative category (48.6%) as compared to the 61.3% for students who had not. For the Systematicity and Truth-Seeking scales, the distributions of percentages for both groups were more heavily concentrated in the Ambivalent and Negative category (great than 50%) than the Positive Category (less than 50%).

The third variable examined was modes of instruction. The six modes of instruction were ranked in order of frequency according to the classroom teaching the participants experienced at college. Kendall's tau-b correlation coefficient is a nonparametric measure of the strength and direction of association that exists between two variables measured on at least an ordinal scale (Magiya, 2019). Therefore, the relationship between modes of instruction and the overall scale was investigated using Kendall's tau-b correlation coefficient. Significant correlations were found between all the modes of instruction and the overall scale, except content-based instruction (CBI; see Table 29).

Table 28

Percentages of Students in Each Category of Endorsement on the Eight Scales Categorized by Course

Scale	Taking CT courses			Not taking CT courses		
	Positive	Ambivalent	Negative	Positive	Ambivalent	Negative
Analyticity	67.0	22.5	10.6	60.3	26.6	13.2
Inquisitiveness	72.5	19.3	8.3	64.1	26.0	9.9
Systematicity	38.5	42.7	18.8	31.9	43.6	24.5
Self-Confidence	51.4	40.8	7.8	38.6	48.7	12.6
Truth-Seeking	45.4	34.9	19.7	44.0	38.3	17.8
Cognitive Maturity	56.0	27.1	17.0	57.1	27.7	15.2
Open-Mindedness	51.8	33.5	14.7	51.1	39.4	9.5
Justice-Orientedness	61.9	29.8	8.3	57.3	36.8	5.9

There was a small, positive correlation between the two variables of traditional teaching mode and the overall scale (r = .084, n = 764, p = .003) with higher levels of traditional teaching mode associated with total score. Traditional teaching mode helped explain 1% of the variance in respondents' scores on the overall scale.

Likewise, there was a small, positive correlation between the two variables of blended learning and the overall scale (r = .082, n = 764, p = .002) with higher levels of blended learning associated with total score. Blended learning helped to explain 1% of the variance in respondents' scores on the overall scale.

A small, positive correlation was also found between the two variables of problem-based learning (PBL) and the overall scale (r = .089, n = 764, p = .001) with higher levels of PBL associated with total score. PBL helped to explain 1% of the variance in respondents' scores on the overall scale.

There was a small, positive correlation between group discussion and the overall scale (r = .057, n = 764, p = .029) with higher levels of group discussion associated with total score. There was a small, positive correlation between debate and the overall scale (r = .069, n = 764, p = .012) with higher levels of debate associated with total score.

Table 29

Correlation Between Modes of Instruction and CT Disposition

Kendall's tau_b		TotalScore	Traditional Teaching Mode
TotalScore	Correlation Coefficient	1.000	.084**
	Sig. (2-tailed)	.	.003
	N	764	764
Traditional Teaching Mode	Correlation Coefficient	.084**	1.000
	Sig. (2-tailed)	.003	.
	N	764	764
Kendall's tau_b		TotalScore	Blended Learning
TotalScore	Correlation Coefficient	1.000	.082**
	Sig. (2-tailed)	.	.002
	N	764	764
Blended Learning	Correlation Coefficient	.082**	1.000
	Sig. (2-tailed)	.002	.
	N	764	764

	Kendall's tau_b	TotalScore	PBL
TotalScore	Correlation Coefficient	1.000	.089**
	Sig. (2-tailed)	.	.001
	N	764	764
PBL	Correlation Coefficient	.089**	1.000
	Sig. (2-tailed)	.001	.
	N	764	764
	Kendall's tau_b	TotalScore	Content-Based Instruction
TotalScore	Correlation Coefficient	1.000	.051
	Sig. (2-tailed)	.	.054
	N	764	764
Content-Based Instruction	Correlation Coefficient	.051	1.000
	Sig. (2-tailed)	.054	.
	N	764	764
	Kendall's tau_b	TotalScore	Group Discussion
TotalScore	Correlation Coefficient	1.000	.057*
	Sig. (2-tailed)	.	.029
	N	764	764
Group Discussion	Correlation Coefficient	.057*	1.000
	Sig. (2-tailed)	.029	.
	N	764	764
	Kendall's tau_b	TotalScore	Debate
TotalScore	Correlation Coefficient	1.000	.069*
	Sig. (2-tailed)	.	.012
	N	764	764
Debate	Correlation Coefficient	.069*	1.000
	Sig. (2-tailed)	.012	.
	N	764	764

Note. * Correlation is significant at the 0.05 level (2-tailed). **Correlation is significant at the 0.01 level (2-tailed)

Summary

This chapter reviewed the research questions and the rationale of the methods, then reported the survey results in line with each question. Descriptive and inferential statistics were both analyzed to answer the questions. In the next chapter, I interpret and analyze the results in depth.

CHAPTER 5: DISCUSSIONS AND CONCLUSIONS

The purpose of this study was to investigate the status of undergraduates' critical thinking (CT) at a private college in mainland China. In particular, the study focused on CT dispositions among foreign language majors across various demographic or academic groupings and with different training or learning experiences. Wen et al.'s (2011) revised California Critical Thinking Disposition Inventory (CCTDI) was used to explore the difference in CT between foreign language majors and other liberal arts majors and differences in the CT of undergraduates in the same foreign language discipline but under different programs at a private college in Shanghai, China. This chapter provides a summary and discussion of the results to address the research questions, presents implications for policies and practice, discusses the current study limitations, and offers suggestions for future research.

Summary and Interpretation of Results

In this section, I summarize the major results, compare my findings with previous studies, and provide possible explanations for or speculation about my findings.

Current Level of Undergraduate Student CT Dispositions

The undergraduates at this 4-year private college showed ambiguity toward critical thinking. This result is inconsistent with Wen (2012) and Wen, Zhang, et al. (2014), who found the students at Chinese public universities showed a positive inclination toward CT dispositions, and also different from Q. X. Luo and Yang (2001), who found the negative disposition toward CT of Chinese students consisting of 4-year university/college and 3-year junior college

students. The result seemed to suggest the CT of higher level universities' students is satisfactory and better than lower level universities' students, which was consistent with the previous research (Y. X. Dong et al., 2010; Wen et al., 2018; Q. G. Zhang & Shen, 2018a, 2018b; Zhao et al., 2015). Conscious efforts were made at public universities to apply more modes of instruction to the explicit teaching of the students' CT, which has been found to be better than implicit instruction and is conducive to cultivating students' self-reflection and self-monitoring ability (Wen & Sun, 2015). Therefore, another possible explanation is that private college students' ambiguous dispositions toward CT are to some extent related to the implicit instruction of CT, in addition to the possibility that higher level universities might admit better critical thinkers. However, less than 2 points below a positive disposition overall would suggest mixed results between ambivalence and a positive disposition.

Likewise, students in this sample set were positively inclined toward the academic dimension and ambivalent toward the social dimension at the level close to the cutoff point. On the third level, they were positively disposed toward Intellectualness and Morality, while they were ambivalent toward Emotionality and Rationality. On the bottom level, they were positively disposed toward Analyticity, Inquisitiveness, Open-Mindedness, and Justice-Orientedness and were ambivalent toward Systematicity, Self-Confidence, Truth-Seeking, and Cognitive Maturity.

The students' ambivalent tendency toward Systematicity and Self-Confidence can be interpreted based on their family background and life or learning experiences. Due to the high tuition fees, students enrolled in private universities or colleges usually come from wealthy

families. It is possible that superior family conditions and inadequate life skills training lead to their weak will (Xia, 2011; C. Y. Yu, 2013), lacking psychological resilience, perseverance, and tenacity. Therefore, it is possible they feel frustrated due to different setbacks. Besides, as foreign language learning takes a long time and requires considerable patience, many students cannot see their immediate achievements, gradually sinking in frustration, which affects their learning enthusiasm and self-confidence to a certain extent (J. Y. Yang, 2009). In addition, foreign language learning often requires speaking in public and expressing ideas, which often causes shy and introverted students to feel unconfident (L. N. Sun, 2004).

Students' ambivalent tendency toward Truth-Seeking indicated they lack objective attitudes toward pursuing knowledge. The lack of confidence in self-reflection and questioning may lead them to blindly believe in book knowledge and seldom challenge authorities (J. Y. Ma et al., 2015). One reason is related to the education system. For a long time, students have received examination-oriented education, adapted to a teacher-centered, spoon-feeding teaching method, and learned standard answers by memorization (Ran, 2010). As a result, they often lack the initiative consciousness of independent thinking, develop the inertia of passively receiving knowledge, and lose critical thinking abilities (Xiao, 2015). Besides, the students' ambivalence toward Truth-Seeking is also related to the influence of Confucian educational ideas, which emphasize obedience. Therefore, parents and teachers influence students with this ideological requirement, making students increasingly obedient to authority and unified in their thinking patterns (J. Y. Ma et al., 2015).

It is interesting to find students scored highest on Intellectualness and lowest on Emotionality at the third level under the Academic Dimension. This finding is consistent with Wen (2012). The current study reinforced the contradictory characteristic feature exhibited by the undergraduates in China who were ready to meet the challenge and stay hungry for new knowledge but unconfident in their critical thinking ability and lacked the concentration and willpower to tackle complicated problems.

As Wen (2012) was the most recent large-scale study of the CT at the public universities using the same instrument as this current study, I offer a visual profile of the scale scores' rankings compared with rankings reported by Wen (2012; see Table 30). There is a similarity between undergraduates at public universities and the private college. In this study, rankings of the Inquisitiveness, Open-Mindedness, Self-Confidence, Truth-Seeking, and Systematicity scales of private college students were the same as that of public university students. They both displayed strong curiosity about new knowledge and positive attitudes toward analysis, but were less positive about activities that required hard work, courage, and perseverance. This similarity can be seen from the high rank of Inquisitiveness and Analyticity but the low rank of Systematicity and Truth-Seeking.

Table 30

Rankings of Scale Scores

Scale	Mean	CT disposition	Rankings	Rankings at public universities (Wen, 2012)
Analyticity	41.83	P	2	3
Inquisitiveness	42.68	P	1	1
Systematicity	36.67	A	8	8
Self-Confidence	38.98	A	6	6
Truth-Seeking	37.42	A	7	7
Cognitive Maturity	39.60	A	5	2
Open-Mindedness	40.43	P	4	4
Justice-Orientedness	41.17	P	3	5

However, differences in the rankings existed on the Analyticity, Cognitive Maturity, and Justice-Orientedness scales. Private college students were more intolerant of the injustice, illegal activities, and practices that went against the moral standards than public university students, which were reflected in the different rankings of Justice-Orientedness between private college and public university undergraduates. Moreover, private college students showed a greater tendency toward a conformist, arbitrary, and simplified mentality and a weaker inclination to see the complexity in problems and to desire prudent and timely decision making because of their lower rank of Cognitive Maturity than public university students. The difference in the disposition toward Justice-Orientedness might be attributed to students' superior family conditions, strong social adaptability, and rich social experience, which facilitate their tendency to seek justice (Q. L. Yu et al., 2016). On the other hand, due to their poorer academic achievement and learning autonomy, private college students tended to believe in the authority of

teachers or book knowledge so that they were ambivalent toward Cognitive Maturity (M. Zhang & Yin, 2011).

In terms of foreign language discipline, students were ambiguous toward critical thinking. Different from the overall sample, they showed ambivalence toward the Academic Dimension. No other differences were found in the two groups in terms of the mean score, rankings of scores, and disposition toward each characteristic on the four levels (see Table 31). One finding of foreign language majors at the private college was that mean scores for the Academic Dimension were higher than the social dimension, although not reaching a statistically significant difference. This finding was inconsistent with Wen (2012), who found undergraduates in foreign language discipline at public universities scored significantly higher in the Social Dimension than the Academic Dimension. It is possible the academic study at the private college constantly develops the foreign language majors' dispositions toward Intellectualness (Wen, 2012). However, with a lack of self-confidence and sense of frustration during their language study (J. Y. Yang, 2009), they might be reluctant to take part in the social practice or activities and make human connections, so their social dimension scores were lower than those in the academic dimension in this sample set.

Table 31

Rankings of Scale Scores: A Comparison

Scale	Mean		CT disposition	Rankings	Rankings at public universities (Wen, 2012)
	Overall Sample	Foreign language majors			
Analyticity	41.83	41.59	P	2	3
Inquisitiveness	42.68	42.56	P	1	1
Systematicity	36.67	36.33	A	8	8
Self-Confidence	38.98	38.90	A	6	6
Truth-Seeking	37.42	37.29	A	7	7
Cognitive Maturity	39.60	39.47	A	5	2
Open-Mindedness	40.43	40.06	P	4	4
Justice-Orientedness	41.17	40.97	P	3	5

Differences Between Foreign Language Majors and Other Liberal Arts Majors

For mean scores, foreign language majors scored lower than other liberal arts majors on the overall scale, which was inconsistent with previous studies (Wen, 2012; Wen, Wang, et al., 2010; Wen, Zhang, & Sun, 2014; J. S. Zhang & Fu, 2018), but consistent with Liang (2017). However, no significant differences were found in all the scores on the four levels for foreign language majors and other liberal arts majors, which was consistent with Liang (2017) and Wen (2012).

Furthermore, no substantial differences in the distributions of percentages in the three categories of positive, ambivalent, or negative were found between the two disciplines. However,

the distribution of percentages for both groups was more heavily concentrated in the ambivalent and negative categories than the positive category, for the Systematicity, Self-Confidence, and Truth-Seeking scales. Therefore, nurturing a stronger inclination toward Systematicity, Self-Confidence, and Truth-Seeking might become goals of curricular and extracurricular programs both for foreign language majors and other liberal arts majors. As educators, we are supposed to offer specific ways to change or adjust their opinions or attitudes to show a positive shift toward CT.

Differences Between Western Foreign Language Majors and Eastern Foreign Language Majors

No significant differences were found for the scales on all the levels between the two clusters. Though not to the extent of being statistically significant, Western foreign language majors tended in the positive direction for Open-Mindedness categorized under Morality, while Eastern foreign language majors were ambivalent toward this characteristic. This finding might, to some degree, reflect the different influences of culture on the students. Influenced by the Western culture valuing open-mindedness, Western foreign language majors can keep an open mind and tend to accept exotic things (Song & Kou, 2010). Therefore, intercultural communication and Oriental and Western culture courses should be included in the curriculum design for both Eastern and Western foreign language majors. Students can improve their communication skills and comparative and contrastive skills, which boost their ability to make rational decisions and set up the value system that guides them to lead a positive life.

Differences Among Foreign Language Majors Representing Different Demographic or Academic Groupings

Gender Differences

The current study found no significant difference for the overall scale, but statistically significant differences for Analyticity, Self-Confidence, Truth-Seeking, and Justice-Orientedness, which was inconsistent with Giancarlo and Facione (2001). However, consistent with Giancarlo and Facione (2001), statistically significant gender differences were found for Cognitive Maturity and Open-Mindedness. In addition, statistically significant gender differences were also found for all scales on the second and third levels.

The significant gender differences in the subscales did not drive the gender difference in the overall scale score because men in this sample were significantly higher in the Academic Dimension but lower in the Social Dimension than female students. Men had a stronger analytical ability of logical reasoning and showed greater confidence in tackling difficult problems. However, women exhibited strengths in the social dimension: Compared to male students, they were more open-minded, more tolerant of different opinions, and more willing to resist their own biases and listen to others, with a better sense of justice and a stronger ability to avoid following blindly.

Leadership Position Differences

Statistically significant leadership position differences were found for the overall scale; the Academic Dimension; Intellectualness, Emotionality, Morality; Inquisitiveness,

Systematicity, Self-Confidence, and Open-Mindedness. Student leaders should assume administrative and organizational tasks and stand out as role models. Therefore, they tend to set high and strict demands on themselves. In leadership positions, they gradually develop and enhance their self-confidence, organizational and managerial skills, and problem analysis and problem-solving abilities. To some extent, these experiences help explain the statistically higher scores of student leaders. In comparison, because of the lack of such experiences, nonstudent leaders showed a weaker inclination toward CT. This finding was consistent with Yuan et al. (2019) on the critical thinking ability of undergraduate nursing students in a military medical university.

Class Level Differences

For the overall scale, no statistically significant differences were found among students in different class levels, which was consistent with M. Zhang et al. (2016) and Q. L. Yu et al. (2016), but different from Zhao et al. (2015) and Wen and Zhang (2016). However, the current study found a statistically significant difference only in the scale of Systematicity between sophomores and juniors. Systematicity measures whether students claim to approach problems with planning, focus, and will. Students gain unprecedented freedom after entering the university without the pressure of the college entrance examination and away from the nanny-style supervision and protection of high school teachers. It is possible they live and study at will at the transition phase and become less motivated, hard-working, and self-disciplined than in high school. Thus, they may fail to make plans and study with concentration and strong will. This

possibility can explain the decrease from freshmen to sophomores for the scale of Systematicity, and is consistent with the study, which uncovered the students' learning problems in the goal, self-management, self-monitoring, and regulation at secondary-tertiary transition (Z. R. Wang et al., 2015).

Though not to the extent of reaching statistical difference, linear and incremental increases were found from freshmen to juniors for the overall scale, inconsistent with Wen and Zhang (2016). Moreover, only freshmen were ambivalent toward the scale of Open-Mindedness. Freshmen were experiencing the transition from high school education to higher education. High school graduates had an obvious tendency of examination-oriented education and formed a mode of thinking of "either wrong or right" views on people and things. They were rigidly bound by book knowledge and had solid standardized thinking. They were used to asking, "what is the correct answer" and blindly believed in the authority of teachers and books (M. Zhang & Yin, 2011). Furthermore, high school students generally came from the same city or region, so the possibility of conflicts in their viewpoints is less than that of college students, and the space for cultivating their disposition toward Open-Mindedness is more limited than that of college students. In terms of teaching, universities emphasize the complexity of the questions and the diversity of answers. Therefore, universities were more conducive to the cultivation of positive inclination toward Open-Mindedness. Also, implicit education through rich and colorful extracurricular activities has become a unique feature of quality-oriented education in Chinese colleges and universities (Pang & Huan, 2015). The colorful life at universities and colleges has

freed students from the shackles of the College Entrance Examination, which in turn enables them to become more mature and cautious. After a year of college study, they gradually began to form the habit of independent thinking, avoided absolute and extreme perceptions, and could have a more real, rational evaluation of themselves. This explanation can be seen from the result that only juniors showed positive inclinations toward the scale of Cognitive Maturity.

GPA

There was a statistically positive correlation between GPA and total score for the overall scale, the Academic Dimension, Emotionality, and Systematicity for Sophomores. Positive correlations between GPA and CT indicated the faculty recognize and reward the stronger disposition toward CT with higher grades (Giancarlo & Facione, 2001). Positive correlations between GPA, overall scale, and the scale of Systematicity indicated students' inclination to be organized, focused, diligent, and persevering in inquiry is closely related to excellent academic performance. Examinations at universities and colleges require students to memorize standard answers (Yao, 2012; Q. E. Yu & Wang, 2016), encouraging students' good learning attitudes, focus, diligence, and perseverance.

However, the results of no significant relationship between GPA and most other scales on the instrument showed dispositions toward critical thinking are not key components in the assignment of grades at this college. It is important to remember that being disposed to think critically is different from being skilled at thinking critically. Perhaps the CT skills should be reflected in GPA. However, the examinations ignore students' reasoning, analysis, application,

and creativity (Yao, 2012; Q. E. Yu & Wang, 2016), which seems inconsistent with the teaching objectives of cultivating CT abilities. Therefore, how to implement the combination of teaching and evaluation deserves consideration.

Differences Among Foreign Language Majors With Different Training and Learning Experiences

Different Programs

Statistically significant differences were found for Intellectualness, Analyticity, and Inquisitiveness. International joint programs are usually characterized by faculty with overseas experiences, curriculum design that takes the requirements of the foreign partners' universities into consideration, high tuition and fees, and small class size. Faculty with overseas experience underline the importance of seminar and heuristic instruction, which should have a positive influence on students' dispositions toward CT (L. J. He & Tang, 2006; Y. N. Li & Wang, 2010). Current study results showed international joint programs significantly improved students' inclination toward Analyticity and Inquisitiveness scales. Students in international joint programs were strongly inclined toward approaching problems with curious minds and by the use of reason and evidence. However, it is surprising that international joint programs did not significantly improve students' dispositions toward CT for most scales. This result is partly in line with H. Y. Liu (2018), who found small-class-size teaching experiences were inadequate in improving students' self-confidence and social communication ability, which may be caused by teachers' unclear training objectives and lack of understanding of students' development (H. Y.

Liu, 2018). Moreover, different curricula did not lead to the statistical differences, perhaps because courses required to be completed in 4 years were compressed into 2–3 years in Chinese colleges which affected the effectiveness of the cultivation of CT dispositions.

Additionally, the finding that no significant differences existed for most scales might also be explained by the current teaching reforms prevalent in China. Recent teaching reforms have advocated the transformation from faculty-centered classrooms to student-centered classrooms, from traditional teaching modes to heuristic teaching, and from examination-oriented education to whole-person education. The influential teaching reforms urged all the faculty to make radical changes in their teaching career by learning how foreign universities have implemented teaching to increase their competitive advantage. Therefore, reforms have influenced faculty's modes of instruction and reduced the difference between the faculty with and without overseas experience, possibly benefitting students in both programs.

Different Courses

Statistically significant differences were found for the Academic Dimension; Intellectualness, Emotionality; Analyticity, Inquisitiveness, Systematicity, and Self-Confidence between students taking and not taking CT-related courses. This finding offered sufficient evidence that CT courses played a positive role in nurturing CT dispositions for the Academic Dimension. CT courses aimed to help students develop the skills and dispositions they need to "become an independent, self-directed thinker and learner" (Bassham et al., 2010, p. xi). CT courses significantly improve students' course work and their inclination toward tackling

problems by reasoning and using evidence, their urgent desire to acquire knowledge, the attitude toward solving problems in a planned, focused, and strong-willed way, and the confidence in their own CT ability. However, CT courses found no support in (a) enhancing students' desire to seek truth and express themselves in an objective and truthful way, (b) awareness of following blindly and solving problems in an arbitrary and simplified way, (c) tolerance of different schools of ideas, and (d) alertness to their own biases, and attitudes toward the unjust affairs that violate the moral standards and breach the law. In general, the students' CT dispositions toward the Social Dimensions were not nurtured by CT-related courses.

Different Modes of Instruction

Traditional teaching mode, blended learning, problem-based learning (PBL), group discussion, and debate were significantly positively correlated to CT dispositions. These results confirm Li's (2010) finding, which proposed traditional skills-based language teaching models do not necessarily go against the development of critical thinking; therefore, the question of whether to replace skill-based language courses with content courses still needs further investigation. Consistent results were also found with most literature reviewed in Chapter 2 about blended learning (Leng et al., 2018; Y. J. Wu et al., 2015; Y. R. Wu, 2014), PBL (X. D. Chen, 2013; Tian et al., 2018), group discussion (Pei et al., 2018), and debate (H. Liu & Jin, 2012). However, the current investigation did not find a significant relationship between content-based instruction (CBI) and CT dispositions, inconsistent with D. X. Yang and Zhao (2011). A

possible explanation is that CBI was least frequently used in the classroom teaching at the college to exert a great influence on students' CT dispositions.

Implications for Policies and Practice

The present study found the undergraduate students at this private college showed ambivalence toward the scales of Systematicity, Self-Confidence, Truth-Seeking, and Cognitive Maturity, and the distribution of percentages for the students were more heavily concentrated in the Ambivalent and Negative category than the Positive category, for the Systematicity, Self-Confidence, and Truth-Seeking scales. Therefore, to develop students' CT disposition, private universities and colleges should attach importance to nurturing their stronger inclination toward these scales. The policy makers and administrators, the faculty members, and the students themselves should make a concerted effort to develop students' CT dispositions in the areas of curriculum design, faculty professional development, classroom teaching, assessment and evaluation, and extracurricular activities.

Curriculum Design

The decrease from freshmen to sophomores for the scale of Systematicity supports the recommendation that private universities should design a freshman orientation course or freshman seminar. Due to various problems in basic education in China, students have many defects in learning methods and independent thinking ability (H. X. Zhang, 2009). Therefore, critical thinking and a questioning spirit should be key content in this newly added course or seminar for freshmen transitioning from high school graduates to university undergraduates. It is

also important to strengthen the introduction of the nature of the disciplines and the system and structure of the curriculum. It is important to guide freshmen to understand the university's organization and management to help them understand the overall cultivation plan of 4-year education, the significance of courses offered in each grade, and the logic of their priorities so they will set up learning goals and make short- and long-term plans. Based on the findings that students were ambivalent toward Systematicity and Self-Confidence, this course or seminar could also include (a) reading ability training, (b) time management, (c) how to cope with pressure and maintain physical and mental health, (d) note-taking skills, (e) how to participate in group activities and cooperative learning, and (f) introduction to the school environment and equipment.

Western foreign language majors tended in the positive direction for Open-Mindedness; therefore, the courses of Intercultural Communication and Oriental and Western Culture should be included in the curriculum design for both Eastern and Western foreign language majors. Furthermore, as CT-related courses were found to have a statistically significant relationship with students' CT dispositions toward the Academic Dimensions, CT-related courses should be added to the list of the general education courses across all the disciplines at the private college.

At the same time, for the international joint programs to reduce the course load in the first 2–3 years of study in China, and thus improve the effectiveness of instruction, credit transfer should be further promoted and implemented in the international joint programs. Administrators should remove some courses with similar objectives and content in the foreign partner

universities that can be recognized as equivalent to the courses required by the Chinese universities.

Faculty Professional Development

International joint programs significantly improved students' dispositions toward Analyticity and Inquisitiveness in this study. However, the faculty's skills for nurturing CT should be developed to improve students' CT dispositions toward the overall scale. Therefore, future faculty professional development should underline the improvement of faculty's CT abilities and the skills of nurturing students' CT abilities. Foreign language teachers must have a deep understanding at both the theoretical and operational levels (Y. Z. Sun et al., 2016). The future faculty professional development is warranted as to the following topics: (a) What is critical thinking skills and dispositions? (b) What is the relationship between critical thinking ability and language ability? (c) How to develop critical thinking ability in language skills classroom teaching? (d) How to cultivate the ability of thinking in classroom teaching of professional knowledge? (e) How can the examinations guide and promote the cultivation of critical thinking ability? (f) How to write textbooks to promote the development of critical thinking ability? Obviously, the prominence of the cultivation of critical thinking ability in foreign language education poses a challenge to the development of foreign language teachers. Only when the faculty members consolidate the knowledge of CT and are trained how to effectively combine the cultivation of CT with foreign language teaching, can they apply what they know to practice and significantly improve students' CT.

Classroom Teaching

Critical thinking ability can be enhanced as language skills are developed and fostered so long as the faculty reduce the number of skills training hours, overhaul the syllabus, and adjust teaching methods (L. W. Li, 2010). The current study found significant, positive correlations between all the modes of instruction and the overall scale, except CBI. Based on the current study results, a diversity of instruction modes can be applied to the development of CT dispositions to cultivate students' ability of analysis, synthesis, criticism, and debate, and the ability to discover and solve problems. For example, to improve students' disposition toward Cognitive Maturity, faculty should encourage more group discussions in class where students feel free to express their distinctive viewpoints on a certain topic and more debates where students can learn to defend themselves against opponents.

Additionally, taking student leaders' stronger inclination toward CT into account, faculty should help students achieve their full leadership potential in the teaching process, stimulate the enthusiasm of students to participate in teaching activities, support and encourage students to communicate and cooperate with each other, and provide them with the opportunity to think, create and display their talents, to improve their self-confidence and sense of responsibility.

Furthermore, the integration of CT cultivation in foreign language teaching should be promoted both implicitly and explicitly. For example, at the end of the integration of teaching, the CT skills implicitly taught and practiced in foreign language teaching should be explicitly summarized and consolidated, which can effectively enhance students' CT awareness, and in the

meantime help students to improve their initiatives in enhancing CT. Also, specific training on CT should be complemented with the integration of teaching so that students can have a systematic understanding of the conceptualization of CT and the objectives of learning CT. In the future study, it is worth examining the relationship between the extent to which professors expect CT from the students and professors demonstrate and test CT in the classroom and the students' ability to think critically.

Assessment and Evaluation

The findings that students showed ambivalence toward Truth-Seeking and no significant relationship existed between GPA and the CT dispositions for most scales supports the recommendation that examinations reduce the content of rote memorization and increase the part that tests students' ability to understand, analyze, apply, and judge. Thus, examinations could fully reflect the students' knowledge and appropriately reflect the students' various CT abilities. Considering the importance universities and colleges attach to the cultivation of CT, educators should be aware of the difference between CT dispositions and CT skills. Therefore, perhaps it deserves further study on how students' CT skills and CT dispositions can be reflected in GPA, which might raise students' awareness of improving their CT abilities and their initiative in CT training activities.

Extracurricular Activities

To temper students' will and nurture their disposition toward Systematicity, private institutions should strengthen frustration education that guides students to face setbacks in

studies and in life. In addition, more extracurricular activities or competitions should be organized to offer students more opportunities to display their diverse talents, and enhance their egos and self-confidence. Furthermore, these activities can train the students to communicate and socialize, enhance their abilities to organize and coordinate, and develop their ability to withstand setbacks and courage to face difficulties head on (H. F. Liu, 2013).

For the sake of effective implementation, cooperation between faculty and counselors is encouraged. Based on the characteristics of the course, faculty take the responsibility to design extracurricular activities, such as debate, speech contests, research projects, and campaigns. These activities are designed to allow the students to make full use of theoretical knowledge to solve practical problems and are organized by the counselors out of class time. The students' performances could be graded and considered as a part of a course assessment. By integrating classroom teaching and extracurricular activities, students can exercise their comprehensive ability to apply knowledge to practice.

In addition, faculty and counselors should be fully aware of the gender differences in their dispositions toward CT. Male and female students should be involved in each curricular or extracurricular activity. In this way, students of different genders will have more chances to learn from each other.

Strengths and Limitations of the Study

The current study is a significant contribution to the quantitative research on private college students' CT in China. No other large-scale survey investigation with a large sample size

and demographic diversity had been conducted in Chinese private universities and colleges. Therefore, this study provides new evidence for the undergraduates' CT dispositions at private universities in China, particularly those students in a foreign language discipline. Moreover, new evidence is provided for the effectiveness of international joint programs in cultivating critical thinkers. Most importantly, this study offers new valuable suggestions on the teaching reform projects or policies on higher education.

However, limitations also exist in the current study. First, a limit was convenience sampling, which influenced its generalization. Second, cross-sectional research is limited by examining the change of undergraduates' CT over 4 years at college. Third, self-reported GPA is liable to lack accuracy, and thus, affect the reliability of results. Fourth, although the instrument I used for this study was revised specifically for Chinese undergraduates, the low end of Cronbach's alpha internal reliability indices of some subscales fell below 0.6, which might have affected results. Fifth, researchers should pay attention to avoid statistical results overshadowed by scale cutoff points in the analysis of different groups of students, which should be considered to be guidelines rather than hard breakpoints in CT dispositions.

Recommendations for Future Research

The present study has explored the CT dispositions of undergraduates' in foreign language disciplines across different demographic or academic groups and those with different training or learning experiences at a private college in Shanghai, China. I found foreign language majors showed ambivalence toward CT dispositions, and no significant differences existed

between foreign language majors and other liberal arts majors for this sample set. Gender, class level, GPA, leadership positions, programs, courses, and modes of instruction were investigated to find the difference in the CT dispositions among foreign language majors at the private college.

Future research is warranted due to the limitations. First, despite a large sample size for this study, the sample was limited to only one private college, which might not represent the whole population. Therefore, samples from other private universities or colleges should be included for better generalizability. Second, longitudinal research should be conducted to explore the undergraduates' CT in the development process throughout the 4 years at private universities or colleges. A 4-year longitudinal study can be conducted to find the development of critical thinking dispositions of the foreign language majors. Third, qualitative methods should be used to triangulate quantitative results. Research questions that can be addressed include the students' perception of CT dispositions and the students' lived experiences of developing the CT dispositions at China's private institutions. Fourth, the internal reliability of the subscales should be improved in future research. Last, regression analyses should be applied to determine the effect of each instruction mode while controlling for the effect of the other two modes.

Conclusion

This study examined the CT dispositions of China's private college undergraduates across different academic groupings and with different learning experiences. The independent variables included gender, class levels, GPA, leadership positions, programs, courses, and modes

of instruction. This study contributed to expanding the CT knowledge base of foreign language majors at private universities and colleges. Through this investigation, I uncovered existing problems in nurturing students' CT dispositions and made suggestions on developing CT dispositions in both curricular and extracurricular programs.

The study found undergraduate students' ambivalence toward Systematicity, Self-Confidence, Truth-Seeking, and Cognitive Maturity at this private college. Suggestions were offered from the perspectives of curriculum design, faculty professional development, classroom teaching, and extracurricular activities.

For the policy makers and administrators, a freshmen orientation seminar or course, intercultural communication and oriental and western culture courses, and CT courses should be added to the curriculum. Courses with the equivalent features in foreign partner universities should be removed from the curriculum in international joint programs in the home country. However, faculty professional development should be extended to the field of the conception of critical thinking and include ways to integrate CT cultivation into foreign language teaching.

For faculty members, a diversity of instruction modes can be applied to the development of CT dispositions to cultivate students' abilities of analysis, synthesis, criticism and debate, and to discover and solve problems. Moreover, they should consider to what extent students' CT skills and CT dispositions should be reflected in GPA. In addition, faculty and counselor cooperation is encouraged to effectively organize more extracurricular activities or competitions to display their diverse talents and enhance their egos and self-confidence.

Undergraduates should receive frustration education and learn to face setbacks in life and in their studies to improve their dispositions toward systematicity. Furthermore, they should actively participate in extracurricular and in-class and seize opportunities of exploring and developing their leadership potentials. Male and female students should also be involved in each curricular or extracurricular activity to make it possible to learn from other's strong points to offset one's weaknesses concerning CT dispositions.

REFERENCES

Bassham, G., Irwin, W., Nardone, H., & Wallace, J. M. (2010). *Critical thinking: A student's introduction*. McGraw Hill Higher Education.

Bland, J. M., & Altman, D. G. (1997). Cronbach's alpha. *BMJ Clinical Research, 314*(7080), 572.

Bloom, B. S. (1956). *Taxonomy of educational objectives. Vol. 1: Cognitive domain.* David McKay.

Brown, H. D. (2004). Some practical thoughts about students-sensitive critical pedagogy. *The Language Teacher, 28*(7), 23–27.

The Central People's Government of the People's Republic of China. (2010, July 29). *The Outline of National Medium-and-Long-Term Education Reform and Development Program (2010-2020).* http://www.gov.cn/jrzg/2010-07/29/content_1667143.htm

Chen, X. D. (2013). PBL 教学模式对非英语专业学生批判性思维能力影响的实证研究 [An empirical study on the influence of PBL teaching approach on critical thinking ability of non-English majors]. *Journal of PLA University of Foreign Languages, 36*(4), 68–72, 88.

Chen, Y. P. (2016). 教师提问与学习者批判性思维能力的培养 [Investigation into students' critical thinking development through teachers' questions]. *Foreign Languages and Their Teaching, 2*, 87–96, 146–147. http://doi.org/10.13458/j.cnki.flatt.004245

Creswell, J. W. (2008). *Research design: Qualitative, quantitative, and mixed methods approaches*. SAGE Publications.

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334.

Davis, L. L. (1992). Instrument review: getting the most from a panel of experts. *Applied Nursing Research 5*(4), 194–197.

Dong, Y. (2012). 批判性思维三大误解辨析 [Three fundamental misconceptions of critical thinking]. *Journal of Higher Education, 33*(11), 64–70.

Dong, Y. X., Li, K., & Liu, F. (2010). 大学生的批判性思维技能：评估与培养 [The critical thinking skills of college English students: Assessment and cultivation]. *Computer-Assisted Foreign Language Education, 5*, 33–38.

Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and need research. *Educational Researcher, 18*(3), 4–10.

Ennis, R. H. (1990). The extent to which critical thinking is subject-specific: Further clarification. *Education Researcher, 19*(4), 13–16.

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction — The Delphi report*. California Academic Press.

Facione, P. A. (2000a). *California Critical Thinking Skills Test: CCTST-2000*. California Academic Press.

Facione, P. A. (2000b). The disposition toward critical thinking: Its character, measurement, and relationship to critical thinking skill. *Informal Logic, 20*(1), 61–84.

Facione, P. A. (2011). Critical thinking: What it is and why it counts. *Insight Assessment*, *2007*(1), 1–23.

Facione, P. A., & Facione, N. C. (1992). *The California critical thinking skills test*. California Academic Press.

Facione, P. A., Facione, N. C., & Giancarlo, C. A. F. (2001). *California critical thinking disposition inventory: CCTDI.* California Academic Press.

Fang, H. Z., & Li, C. M. (2014). 大学生挫折教育研究综述 [A review on college students' frustration education]. *China Youth Study*, *6*, 102–105, 110. https://doi.org/10.19633/j.cnki.11-2579/d.2014.06.020

The Foundation for Critical Thinking. (2019). *International Critical Thinking Test.* https://www.criticalthinking.org/pages/international-critical-thinking-test/619

Fowler, F. J. (2014). *Survey research methods* (5th ed.). SAGE Publications.

Gao, Y. H. (1999). 外语学习木桶的"短板" [The weakness of foreign language learning]. *Foreign Language Teaching Abroad, 3,* 6–9, 5.

Gao, Z. Y. (2013). 应用型本科大学生批判性思维倾向现状调查 [A survey on the critical thinking disposition of students in application-oriented four-year university]. *Higher Education Exploration, 2,* 129–133.

Giancarlo, C. A., & Facione, P. A. (2001). A look across four years at the disposition toward critical thinking among undergraduate students. *The Journal of General Education, 50*(1), 29–55.

Guba, E. (Ed.). (1990). The alternative paradigm dialog. In *The paradigm dialog* (pp.17–31). SAGE Publications.

Han, H. M., Ma, Q., & Li, N. (2011). 外语学习者批判性思维能力研究综述 [A review of research on critical thinking of foreign language learners]. *Journal of Hebei University (Philosophy and Social Science), 36*(5), 154–156.

Han, S. J., & Wang, X. Y. (2009). 英语专业精读教学与学生批判性思维能力的培养 [Cultivation of critical thinking in the intensive English reading course]. *Foreign Language Education, 30*(6), 67–70. https://doi.org/10.16362/j.cnki.cn61-1023/h.2009.06.022

He, L. J., & Tang, Y. G. (2006). 研讨式教学:高等学校教学的新走向 [Discussion teaching: A new trend of teaching in colleges and universities]. *Heilongjiang Research on Higher Education, 5,* 145–147.

He, Q. X., Yin, T. S., Huang, Y. S., & Liu, H. P. (1999). 对外语专业本科教育改革的几点建议 [Some suggestions on the reform of undergraduate education for foreign language majors]. *Foreign Language Teaching and Research, 1,* 24–28.

Hesse-Biber, S. N., & Leavy, P. (2011). *The practice of qualitative research* (2nd ed.). SAGE Publications.

Huang, F. (2003). Transnational higher education: A perspective from China. *Higher Education Research & Development, 22*(2), 193–203. https://doi.org/10.1080/07294360304114

Huang, R. H., Ma, D., Zheng, L. Q., & Zhang, H. S. (2009). 基于混合式学习的课程设计理论 [Curriculum design theory based on. blended learning]. *E-education Research, 1*, 9–14.

Huang, X. Y. (2008). 中文版批判思维倾向性测量表于台湾护理人员之测试 [*Using Chinese version of CCTDI to test Taiwan nurses*]. [Unpublished doctoral thesis]. National Cheng Kung University, Taiwan.

Huang, Y. S. (1998). The absence of critical thinking. *Foreign Languages and Their Teaching, 7*, 1, 19.

Huang, Y. S. (2010). 英语专业课程必须彻底改革——再谈“思辨缺席”[The curriculum for English majors must be overhauled—Syndrome of absence of critical thinking reconsidered]. *Foreign Language World, 1*, 11–16.

Kimberlin, C. L., & Winsterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. *American Society of Health-System Pharmacists, 65*, 2276–2284.

Leavy, P. (2017). *Research design: Quantitative, qualitative, mixed methods, arts-based, and community-based participatory research approaches.* Guilford Publications.

Leng, J., Guo, R. F., Hou, Y. R., & Gu, X. J. (2018). 促进大学生批判性思维的在线活动设计研究及可视化分析[Study on online activity design and visual analysis for promoting college students' critical thinking]. *E-Education Research, 39*(10), 77–84. https://doi.org/10.13811/j.cnki.eer.2018.10.011

Li, L. W. (2010). 试析英语专业技能课程与批判性思维能力培养的关系[On the relationship between English language skills training and the development of critical thinking]. *Foreign Languages in China, 6,* 68–73. https://doi.org/10.13564/j.cnki.issn.1672-9382.2010.06.011

Li, N., Han, Q. E., & Zhong, W. X. (2019). 大学生批判性思维素质现状及差异分析——基于山东省的调查研究[The general situation and variation analysis of undergraduates' critical thinking: Based on an investigation in Shandong Province]. *China Higher Education Research, 2,* 49–52. https://doi.org/10.16298/j.cnki.1004-3667.2019.02.09

Li, Y. N., & Wang, D. S. (2010). 创新型人才培养模式研究 [Research on the training mode of innovative talents]. *Shandong Social Sciences, 4,* 174–176. https://doi.org/10.14112/j.cnki.37-1053/c.2010.04.023

Liang, F. J. (2017). 本科高校英语专业大学生思辨能力发展调查［A survey of critical thinking skills of English majors in Chinese universities]. *Foreign Language and Literature Research, 3*(5), 94–102. https://doi.org/10.16651/j.cnki.fllr.2017.0070

Liu, H., & Jin, L. M. (2012). An empirical study on English debate and critical thinking development of college students. *Foreign Languages and Their Teaching, 5*, 24–28. https://doi.org/10.13458/j.cnki.flatt.003825

Liu, H. F. (2013). 大学新生入学适应问题及策略研究[A study on the problems and strategies of freshmen's adaptation to college entrance]. *Education and Vocation, 2*, 78–79. https://doi.org/10.13615/j.cnki.1004-3985.2013.02.035

Liu, H. Y. (2018). 小班研讨课：一流本科教育的核心要素[Small class seminars: The core element of a first-class undergraduate education]. *Jiangsu Higher Education, 7*, 60–65. https://doi.org/10.13236/j.cnki.jshe.2018.07.013

Liu, S. Z., Li, J. P., & Zhu, D. (2005). 不同层次护生批判性思维意向及影响因素的研究 [Critical thinking disposition and the influence factors of nursing students in different educational level]. *Modern Nursing, 7*, 496–499.

Lu, D., Xie, Y. G., Tang, Y. W., & Yang, X. (2018). 批判性思维导向的新型混合学习模式研究——以英语写作教学为例 [Research on the new model of blended leaching with orientation of critical thinking — A case of English writing teaching]. *China Educational Technology, 377*(6), 140–145.

Luo, Q. X. (2001). 批判性思维的结构、培养模式及存在的问题 [The construction, training model of critical thinking and problems in research and training]. *Journal of Guangxi University for Nationalities (Natural Science Edition), 3*, 215–218. https://doi.org/10.16177/j.cnki.gxmzzk.2001.03.017

Luo, Q. X., & Yang, X. H. (2001).《加利福尼亚批判性思维倾向问卷》中文版的初步修订 [Revision for CCTDI (Chinese Version)]. *Psychological Development and Education, 3*, 47–51.

Ma, J. Y., Zhao, Q., Zhang, Y., & Zhu, A. L. (2015). 高校师范生批判性思维倾向的调查研究［Investigation on critical thinking disposition for normal university students]. *Journal of Mathematics Education*, *6*, 21–25.

Ma, R., & Qin, X. Q. (2016). 英语专业大学生的批判性思维倾向特征研究 [Traits of critical thinking disposition of English majors]. *Journal of Xi'an International Studies University*, *24*(4), 60–63. https://doi.org/10.16362/j.cnki.cn61-1457/h.2016.04.014

Magiya, J. (2019). *Kendall rank correlation explained*. https://towardsdatascience.com/kendall-rank-correlation-explained-dee01d99c535

McGorry, S. (2000). Measurement in a cross-cultural environment: Survey translation issues. *Qualitative Market Research*, *3*(2), 74–81.

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

McPeck, J. E. (1990). Critical thinking and subject specificity: A reply to Ennis. *Educational Researcher*, *19*(4), 10–12.

Mi, B. F., & Yuan, P. H. (2009). 高校英语专业读写课程中学生评判性思维能力的培养 [The cultivation of students' critical thinking ability in reading and writing courses for English majors in higher education]. *Continue Education Research, 6*, 58–60.

Mohan, B. A. (1986). *Language and content*. Addison-Wesley.

Moore, B. N., & Parker, M. (2012). *Critical thinking* (10th ed.). McGraw-Hill.

Mu, C. J. (2016). 英语专业学生的写作思辨能力调查[An investigation on the critical writing ability of English majors]. *Modern Foreign Languages, 39*(5), 693–704, 731.

Pang, H. S., & Huan, X. H. (2015). 素质教育与大学教育改革 [Reforming of university education caused by Suzhi Education]. *China Higher Education Research, 9,* 73–78. https://doi.org/10.16298/j.cnki.1004-3667.2015.09.14

Paul, J. L. (Ed.). (2005). *Introduction to the philosophies of research and criticism in education and the social sciences*. Prentice Hall.

Paul, R. (1989). Critical thinking in North America: A new theory of knowledge, learning and literacy. *Argumentation, 3,* 197–235.

Paul, R. W. (1995). *Critical thinking: How to prepare students for a rapidly changing world*. Foundation for Critical Thinking.

Paul, R., Binker, A., Martin, D., & Vetrano, C. (1995). *Critical thinking handbook: 6th – 9th Grades. A guide for remodeling lesson plans in language arts, social studies and science*. Foundation for Critical Thinking.

Paul, R., & Elder, L. (2006). *The miniature guide to understanding the foundations of ethical reasoning*. Foundation for Critical Thinking.

Pei, Z. W., Xing, Y. H., & Xu, L. (2018). 不同认知层级问题驱动下的小组讨论与思辨能力培养 [Fostering students' critical thinking skills through group discussions initiated by questions of different cognitive levels]. *Foreign Languages and Their Teaching, 5*, 118–-128 &150–151. https://doi.org/10.13458/j.cnki.flatt.004532

Peng, M. C., Wang, G. C., Chen, J. L., Chen, M. H., Bai, H. H., Li, S. G., Li, J. P., Cai, Y. F., Wang, J. Q., & Yin, L. (2004). 批判性思维能力测量表的信效度测试研究 [Validity and reliability of the Chinese critical thinking disposition inventory]. *Chinese Journal of Nursing, 9*, 7–10.

Peng, Q. L. (2000). 思辨与创新——口语课堂上的演讲、辩论初探 [Critical thinking and innovation—A preliminary study of speech and debate in the oral English course]. *Foreign Language World, 2*, 39–44.

Pithers, R. T., & Soden, R. (2000). Critical thinking in education: A review. *Educational Research, 42*(3), 237–249.

Project Group of Entry into the WTO and Education of Foreign Language Disciplines. (2001). Some thoughts on the restructuring of China's higher educational system and teaching models in foreign language disciplines: At the era of China's entry into the WTO (1). *Foreign Language World, 5*, 9–15, 65.

Ran, Y. H. (2010). 基础教育应试倾向的深层原因论析 [Analysis on the deep causes of exam-oriented tendency in basic education]. *Journal of Teaching and Management, 28*, 7–9.

Singh, H., & Reed, C. (2001). A white paper: Achieving success with blended learning. *Centra. Software, 1,* 1–11.

Song, Y., & Kou, Q. (2010). 中西方文化差异与道德文化自觉——兼论德育现代化[Cultural differences between China and the West and moral culture consciousness—On moral education modernization]. *Education Exploration, 12*, 6–8.

Sun, L. N. (2004). 情感因素在大学英语教学中的作用[The role of emotional factors in college English teaching]. *Education Exploration*, 9, 69–70.

Sun, Y. Z., Wen, Q. F., Wang, L. F., Feng, Y. H., Gu, P. Y., & Zhang, H. (2016). 准确理解《国标》精神，积极促进教师发展 [Accurately understand the spirit of the National Standards for Teaching Quality of Foreign Language Undergraduate Majors in Colleges and Universities, and actively promote the development of teachers]. *Foreign Language World, 6*, 9–15, 56.

Tan, P. (2014). 大学生挫折教育研究 [Study on college students' frustration education]. *Education and Vocation, 5,* 75–77. https://doi.org/10.13615/j.cnki.1004-3985.2014.05.033

Tian, S. P., Wang, L. J., & Qiu, Y. H. (2018). 问题式教学法对工科大学生批判性思维倾向影响的实证研究［An empirical study on the influence of problem-based learning on critical thinking disposition of engineering students]. *Research in Higher Education of Engineering, 6,* 156–160.

Urdan, T. C. (2017). *Statistics in plain English* (4th ed.). Routledge.

Wang, J. Q., & Wen, Q. F. (2011). 国外思维能力量具评介及启示［Introduction of critical thinking instruments from abroad and some inspirations]. *Journal of Jiangsu Teachers University of Technology*, *17*(7), 38–42, 77.

Wang, W. L., Chen, Y. K., Cheng, S. S., Zhao, M., & Su, P. Y. (2006). 评判性思维教学实践及效果评价 [Practice and effect of teaching of critical thinking for nursing students]. *Chinese Journal of Nursing*, *9*, 834–836.

Wang, Z. R., Zhou, Y. Y., & Wu, H. (2015). 学业过渡期大学生英语自主学习能力的动态研究 [A dynamic study of university EFL learners' autonomous learning ability at secondary-tertiary transition]. *Foreign Language World*, *2*, 25–33.

Watson, G., & Glaser, E. (1980). *Critical thinking appraisal manual*. Harcourt Brace Jovanovich.

Watson, G. B., & Glaser, E. M. (1994). *Watson-Glaser critical thinking appraisal Form S manual*. Harcourt Brace.

Wen, Q. F. (2012). 中国外语类大学生思辨能力现状研究 [*A study of the critical thinking of foreign language majors in China]*. Foreign Language Teaching and Research Press.

Wen, Q. F., & Hu, J. (2010).中国大学生英语口语能力发展的规律与特点 [*Rules and characteristics of the development of Chinese College students' oral English ability].* Foreign Language and Research Press.

Wen, Q. F., Li, Y. J., Sun, M., Zhang, L. L., & Yang, S. (2014). 研制多套难度相似的思辨技能量具: 实践与反思 [Developing parallel tests of critical thinking skills: Practice and reflections]. *Computer-Assisted Foreign Language Education, 4,* 3–7, 13.

Wen, Q. F., Liu, Y. P., Wang, H. M., Wang, J. Q., & Zhao, C. R. (2010). 我国外语类大学生思辨能力量具的修订与信效度检验研究 [Study on the reliability and validity of the revised critical thinking skills test for Chinese university students]. *Foreign Language World, 4,* 19–26, 35.

Wen, Q. F., & Sun, M. (2015). 评述高校外语教学中思辨力培养存在的问题 [On key issues about critical thinking development in college EFL classroom]. *Foreign Language Learning Theory and Practice*, *3,* 6–12, 94.

Wen, Q. F., Sun, M., & Zhang, L. L. (2018). 外语专业大学生思辨技能发展趋势跟踪研究［Development of critical thinking skills: A longitudinal study on foreign language majors]. *Foreign Language World, 6,* 12–19.

Wen, Q. F., Wang, H. M., Wang, J. Q., Zhao, C. R., & Liu, Y. P. (2010). 我国英语专业与其他文科类大学生思辨能力的对比研究 [A comparative study of critical thinking skills between English and other liberal arts majors]. *Foreign Language Teaching and Research, 42*(5), 350–355, 400.

Wen, Q. F., Wang, J. Q., Zhao, C. R., Liu, Y. P., & Wang, H. M. (2009). 构建我国外语类大学生思辨能力量具的理论框架 [Developing a conceptual framework for assessing Chinese university students' critical thinking skills]. *Foreign Language World, 1,* 37–43.

Wen, Q. F., Wang, J. Q., Zhao, C. R., Liu, Y. P., & Wang, H. M. (2011). 对我国大学生思辨倾向量具信度的研究 [Study of the reliability of critical thinking disposition inventory for Chinese college students]. *Computer-Assisted Foreign Language Education, 6,* 19–23.

Wen, Q. F., & Zhang, L. L. (2016). 外语专业大学生思辨倾向变化的跟踪研究 [The development of critical thinking disposition of foreign language majors: A longitudinal study]. T*echnology Enhanced Foreign Language Education, 1*, 3–8, 16.

Wen, Q. F., Zhang, L. L., & Sun, M. (2014). 外语专业学生的思辨能力逊色于其他专业学生吗? [Are foreign language majors inferior to the other majors in critical thinking?] *Modern Foreign Languages, 37*(6), 794–804, 873.

Wen, Q. F., Zhao, C. R., Liu, Y. P., Wang, H. M., & Wang, J. Q. (2010). 我国外语类大学生思辨能力客观性量具构建的先导研究 [A pilot study of a self-designed critical thinking skills test for foreign language majors in China]. *Foreign Language Education, 31*(1), 55–58, 63. https://doi.org/10.16362/j.cnki.cn61-1023/h.2010.01.019

Wen, Q. F., & Zhou, Y. (2006). 评述外语专业学生思维能力的发展 [Comment on the development of thinking ability of foreign language majors]. *Foreign Language Research, 5*, 76–80. https://doi.org/10.16263/j.cnki.23-1071/h.2006.05.018

Wu, X. X. (2015). 美国 SIOP 模式与中国外语专业大学生批判性思维能力发展研究 [Research on American SIOP model and the development of critical thinking ability of Chinese foreign language majors]. *Foreign Language Education, 36*(4), 60–63. https://doi.org/10.16362/j.cnki.cn61-1023/h.2015.04.017

Wu, Y. J., Zhao, H., & Chen, L. (2015). 网络环境下大学生批判性思维培养教学模式的实践［The teaching practice of the teaching model of developing critical thinking of college students in the online learning environment]. *Modern Distance Education Research, 2,* 76–86.

Wu, Y. R. (2014). 混合式学习促进大学生批判性思维能力发展的实证研究［An empirical study on how blended learning promotes the critical thinking ability of college students]. *E-Education Research, 8,* 83–88. https://doi.org/10.13811/j.cnki.eer.2014.08.013

Xia, L. (2011). 当代大学生人文精神的现状与培育 [The present situation and cultivation of humanistic spirit of contemporary college students]. *Continue Education Research, 1,* 117–118.

Xiao, W. W. (2015). 批判性思维缺失的教育反思与培养策略 [Reflection on education and cultivation strategy of lack of critical thinking]. *Journal of the Chinese Society of Education, 1,* 25–29.

Xu, Y. J., Li, G. Q., & Yu, X. (2019). 国际化视角下高等教育信息类人才培养模式分析与探索 [Analysis and exploration of the training mode of the talents of higher education information from the perspective of internationalization]. *Education Modernization, 7, 50*–52. https://doi.org/10.16541/j.cnki.2095-8420.2019.07.016

Yang, D. X., & Zhao, Y. P. (2011). 内容依托式教学对英语专业学生思辨能力的影响 [The influence of content-based teaching on students' critical thinking ability]. *Foreign Language Education, 32*(5), 61–64. https://doi.org/10.16362/j.cnki.cn61-1023/h.2011.05.024

Yang, J. Y. (2009). 大学生外语学习焦虑实证研究及应对策略［An empirical study on college students' foreign language learning anxiety and its coping strategies]. *Gansu Social Sciences, 5,* 235–238. https://doi.org/10.15891/j.cnki.cn62-1093/c.2009.05.035

Yao, Y. (2012). 高校素质教育与大学考试改革探析 [Quality education and examination reform in colleges and universities]. *Education and Vocation, 3, 3*168–169. https://doi.org/10.13615/j.cnki.1004-3985.2012.03.051

Yu, C. Y. (2013). 对大学生心理健康教育的思考［Reflections on college students' mental health education]. *Education Exploration, 10,* 131–132.

Yu, Q. E., & Wang, L. Y. (2016). 创新型人才培养与课程考试改革的实践研究 [A practical research on the training of innovative talents and the reform of examination]. *Education Review, 6,* 87–89.

Yu, Q. L., Xiao, D. F., & Wang, J. (2016). 独立学院英语专业学生思辨人格倾向研究 [Study of critical thinking dispositions of English majors in independent colleges]. *Foreign Languages Research, 33*(4), 53–56, 82. https://doi.org/10.13978/j.cnki.wyyj.2016.04.010

Yuan, Y., Xu, X., & Lu, F. (2019). 某军医大学护理本科生评判性思维能力现状及影响因素 [Status quo and influencing factors of critical thinking ability of undergraduate nursing students in military medical university]. *Chinese Nursing Research, 14,* 2524–2527.

Yue, X. D. (2000). 批判思维的形成与培养: 西方现代教育的实践及其启示 [Formation and cultivation of critical thinking: The practice of western modern education and its enlightenment]. *Educational Research, 8,* 65–69.

Zhang, H. X. (2009). 美国大学的新生研讨课及其启示［Freshman seminars in the U.S. universities and their implications]. *China University Teaching, 11,* 93–96.

Zhang, J. S., & Fu, S. F. (2018). 外语专业学生批判性思维能力测试: 一项对比研究 [A critical thinking abilities test among foreign language majors: A contrastive study]. *Journal of Xi'an International Studies University, 26*(2), 72–76. https://doi.org/10.16362/j.cnki.cn61-1457/h.2018.02.015

Zhang, M., Ru, J. F., & Yin, Y. (2016). 大学生批判性思维现状及成因研究［Study on university students' critical thinking disposition and its causes]. *Journal of Chongqing University (Social Science Edition), 22*(3), 202–207.

Zhang, M., & Yin, Y. (2011). 批判性思维：研究生开启科学创新之门的钥匙 [Critical thinking: The key to scientific innovation for graduate students]. *Academic Degrees & Graduate Education, 9,* 29–32. https://doi.org/10.16750/j.adge.2011.09.008

Zhang, Q. G., & Shen, H. (2018a). 一流大学本科生批判性思维能力水平及其增值——基于对全国 83 所高校本科生能力测评的实证分析 [Undergraduates' ability of critical thinking and its added value in first-class universities—An empirical study based on the assessment of undergraduates in China's 83 universities]. *Educational Research, 39*(12), 109–117.

Zhang, Q. G., & Shen, H. (2018b). 中国大学教育能提高本科生批判性思维能力吗——基于"2016 全国本科生能力测评"的实证研究 [Can China's undergraduate education improve collegiate critical thinking ability? —An empirical analysis based on the 2016 National Assessment of Collegiate Capacity]. *China Higher Education Research, 289*(6), 69–76. https://doi.org/10.16298/j.cnki.1004-3667.2018.06.11

Zhang, Q. G., & Shen, H. (2018c). 独生子女与非独生子女大学生批判性思维能力的差异性分析 [Analysis of the differences in critical thinking ability between the only child and non-only child]. *Fudan Education Forum, 4*, 58–64. https://doi.org/10.13397/j.cnki.fef.2018.04.010

Zhao, T. T., Yang, Y., Liu, O., & Mao, L. Y. (2015). 大学生学习成果评价的新途径——EPP（中国）批判性思维能力试测报告 [A new approach to evaluate college student learning outcomes — Pilot test report of EPP (Chinese version) critical thinking]. *Educational Research, 9*, 64–71, 118.

Zhu, X. L. (2002). WGCTA 和 CCTDI 量表在国内护生中的初步应用 [*Preliminary application of WGCTA and CCTDI scales in nursing students in China*] [Unpublished master's thesis]. Peking Union Medical College, Beijing.