2023

A Comparative Study on Critical Thinking Skills of ISEC and Non-ISEC Teachers in Institutions of Higher Education in the North of China

Liqin Tang

A COMPARATIVE STUDY ON CRITICAL THINKING SKILLS OF ISEC AND NON-ISEC TEACHERS IN INSTITUTIONS OF HIGHER EDUCATION IN THE NORTH OF CHINA

By

LIQIN TANG

Master of English Linguistics and Literature, Jilin Normal University, Jilin, China, 2006

Bachelor of English Education, Siping Teachers' College, Jilin, China, 1999

Dissertation

Presented in partial fulfillment of the requirements

For the degree of Doctor of Philosophy

In International Educational Leadership

The University of Montana

Missoula, MT

Approved by:

Ashby Kinch, Dean of the Graduate School

Graduate School

John Matt, Chair
Department of Educational Leadership

William P. McCaw
Department of Educational Leadership

Erica Allen
Department of Educational Leadership

Patty Kero
Department of Educational Leadership

Zhen Cao
Department of World Languages and Cultures

Tang, Liqin Ph.D, Spring 2023 Educational Leadership

A Comparative Study on Critical Thinking Skills of ISEC and Non-ISEC Teachers in Institutions of Higher Education in the North of China

Chairperson: John Matt, Ed.D

Critical thinking (CT) has become a main focus in the higher education and is viewed as one of the essential skills for students to succeed in the 21st century. Many studies focus on Chinese students and their CT skills. There is a scarcity of research targeting teachers' CT. However, teachers are the key to successful education and they play a crucial role in any education reform. Teachers' perception, attitude, and experience impact the educational practice. Therefore, it is imperative to examine teachers' CT.

This study utilized a non-experimental causal-comparative methodology with an explanatory mixed methods research design. The purpose of this study was to explore the status quo (current situation) of Chinese teachers' (including ISEC and non-ISEC teachers) CT, as well as the perception, attitude, and practice of CT among them in institutions of higher education in the north of China. There were 102 participants took the California Critical Thinking Skills Test (CCTST). The results from the quantitative research showed the CT skills of Chinese teachers fell into the upper range of moderate level. There were no significant differences or relationships in CT skills for ISEC and non-ISEC teachers based on the variables: gender, professional rank, educational background, discipline they taught, age, and years of teaching. Twelve participants were interviewed. The core phenomenon or theory emerged from the qualitative data: Chinese teachers advocated and supported CT instruction, and they had a varied and fragmented perception about CT. Although they held a positive attitude towards CT and CT instruction, they applied limited CT teaching strategies in their practice. All participants displayed a strong desire to participate in the CT training programs.

The findings from the qualitative paradigm supported, complemented, and deepened the findings from the quantitative paradigm, which offered a panoramic view of Chinese teachers' CT in institutions of higher education in the north of China. Since there is a scarcity of literature focusing on teachers' CT, this explanatory mixed methods research design filled the gap in this field of the literature. The results of this non-experimental causal-comparative study added new knowledge to the literature on teachers' CT, especially Chinese teachers' CT in institutions of higher education. Future studies should include classroom observations to offer a more authentic picture of how teachers teach students CT, or replicate this study with a larger sample from a wider scope, not only in the north and/or south of China, but also in other countries, in order to generalize the findings to a larger population.

ACKNOWLIEDGEMENTS

I would like to express my sincere thanks to Dr. John Matt, the chair of the Department of Educational Leadership, for providing me with invaluable feedback and advice throughout this endeavor. I am so lucky to have him as my chair because he is a great leader, an accomplished scholar, and yet a very humble person. Without his support, encouragement, and assistance, I could not have completed this journey in such a smooth way. Next, I would like to thank each member of my dissertation committee – Dr. William P. McCaw, Dr. Erica Allen, Dr. Patty Kero and Dr. Zhen Cao. All of my committee members spent time out of their incredibly busy schedules assisting me with this project. They are wonderful scholars, full of solid knowledge base, unique perspectives, and impressive expertise. I have learned a lot from them.

I also need to express my gratitude to my husband, Qingguo Chen, for his constant support for my study in the United States. I am also very thankful to my son, Liangfan Chen, for his understanding, company, and encouragement during the whole journey of my doctoral program. Meanwhile, I would like to give special recognition to my mother and my sisters and brother. My mother is illiterate and did not have the opportunity to go to school, but she knows the importance of education and has always supported me to further my education. My two sisters and one brother never get the chance to go to college, but they are kind, diligent, and brave people, with a great personality. They take good care of my mother, especially when she was in hospital because of the cerebral hemorrhage. I am so grateful to my family members for their understanding, support, and consideration. Without them, I do not think I could put all my heart and soul into my doctoral program and complete my dissertation as expected. Last but not least, I am so grateful to all participants in my study for their kindness and support.

Table of Contents

List of Tables

List of Figures

Chapter One: Introduction to the Study

The origin of critical thinking (CT) can be traced back to Socrates's teaching technique, which occurred over 2,400 years ago (Fasko & Fair, 2021). Socrates' teaching technique best known as the Socratic Method, is delineated in dialogues by Plato as in the *Euthyphro*, the *Apology*, and the *Republic* (Fasko & Fair, 2021). In 1910, Dewey "introduced the term 'critical thinking' as the name of an educational goal" (Hitchcock, 2018, para. 2), and he named this goal as 'reflective thought' or 'reflective thinking' (Hitchcock, 2018, para. 2). In 1946, Black published his book, *Critical Thinking: An Introduction to Logic and Scientific Method*, which is one of the first books with "critical thinking" as the title (Fasko & Fair, 2021)**.** Later, Ennis' (1962) influential article, *A Concept of Critical Thinking*, was published in the Harvard Educational Review (Fasko & Fair, 2021)**.** However, the critical thinking movement did not gain momentum with researchers from psychology, philosophy, and education until the early 1980s (Ennis, 2015; Fasko & Fair, 2021).

Today, critical thinking (CT) is a worldwide concern and the development of critical thinking skills has become the main focus in the context of higher education (Davies & Barnett, 2015). In the past decades or even centuries, China, along with other Asian countries, did not pay much attention to CT (Davies & Barnett, 2015). However, with the development of the critical thinking movement in the Western countries, non-Western countries, especially China, began to focus on CT, and Chinese educators, scholars and researchers started to conduct research on CT. "Attempts to apply the ideas to higher education in China did not begin until the mid-1990s" (Dong, 2015, p. 351), and the growth and progress of the critical thinking movement in China has been made since the late 1990s (Dong, 2015).

Historical Background and Context of CT Education in China

Chinese education can be traced back to Confucius' time in the 6th century B.C.E. "The civil service exam (605–1905) dominated the educational landscape of China for 1300 years before it was replaced by a Western-based school system at the turn of the 20th century" (Tan & Hairon, 2016, p. 315). Therefore, Chinese education has been strongly affected by Confucianism in that "the value of education, textual transmission, academic excellence, meritocracy, and the respect for teachers" are especially stressed (Tan & Hairon, 2016, p. 315). Pang (2012) concluded both traditional Confucius doctrines and foreign pedagogies (first Russia's, and then America's) impacted Chinese educational pedagogies to a large extent.

China has carried out a series of educational reforms since the founding of the People's Republic of China in 1949 (Tan, 2016). All of these reforms have been directed by the important documents of the Communist Party of China (CPC). In 1985, the CPC issued a document, entitled "Decision of the Central Committee of the Communist Party of China on the Reform of the Educational System (中共中央关于教育体制改革的决定)." This document stressed that Chinese colleges and universities need to borrow "curricula design and teaching content from developed countries" (Li, 2017, p. 42) to improve their educational outcomes and qualities. In 1993, the State Council of China released another document, called "The Guidelines for the Reform and Development of Education in China (中华人民共和国教育改革和发展纲要)", and emphasized international educational cooperation between Chinese schools and Western counterparts (Li, 2017). In 2001, the Eighth National Curriculum Reform (第八次全国课程课程改革), again laid stress on changing the traditional Chinese teaching and learning (teacher-centered, exam-oriented, and rote memorization) and promoting all-round development of quality education (Guan & Meng, 2007).

The Ministry of Education (MOE) stressed the imperative of developing critical thinking skills and put CT into its policy document "The College English Curriculum Requirements (2007) (大学英语课程教学要求)" (shortened to The Requirements) as one of the goals for cultivating college students. The Requirements clearly pointed out the importance and urgency for university students to improve their critical thinking skills. (Ministry of Education, 2007). "The Medium-and Long-Term Education Reform and Development Plan (2010-2020) (国家中长期教育改革和发展规划纲要)" put forward the guiding ideology--- "deepening the reform of higher education through the internationalization of Education" (Ju, 2017, p.1). In 2019, the Chinese State Council published a significant plan to drive and advance continued reform in China's education sector, drawing on the range of preceding reforms since China's opening up in 1978 (Zhu, 2019). This document was named "China's Education Modernization 2035 Plan (2035 Plan) (中国教育现代化 2035)" (The State Council, 2019). This plan is aimed to substantially modernize China's education system by 2035, and it is the year that China is determined to realize socialist modernization and become an education powerhouse (The State Council, 2019).

Because of the advocacy and intervention of the Chinese government, internationalization, modernization of higher education, and cultivating innovative talents became the main trend in Chinese colleges and universities (Ryan, 2011; Wei, 2003; Xiao, 2005; Zhu, 2019). Since then, internationalization of higher education, especially critical thinking (CT) education, has taken "center stage in China's curriculum reform" (Tan, 2020, p. 331).

Although Chinese colleges and universities have introduced critical thinking into their school curriculum and made plans for international education development accordingly, they are faced with many challenges and difficulties in the process of CT education and

internationalization. The challenges and difficulties include: (a) how to implement effective CT curriculum development; (b) how to solve the problem of resource shortage; (c) how to further the professional training and development program of faculty and staff; (d) how to advance and deepen the process of CT education and internationalization successfully, etc. (Li, 2017). Under such circumstances, the China Scholarship Council (CSC) employs the rich and extensive resources and channels of international educational exchanges, to develop an international cooperation and exchange project, namely, the International Scholarly Exchange Curriculum (Undergraduate) (ISEC) program, in 2012 (Ju, 2017). The ISEC program mainly focuses on implementing "China's Education Modernization 2035 Plan" (2035 Plan), with the mission of advancing the reform of local or provincial institutions of higher education by means of internationalization and modernization of education (ISEC office, 2022).

As part of the educational reform agenda, one of the main targets of the ISEC program is the ISEC teachers. The curriculum and teaching reform of local colleges and universities will be promoted by improving teachers' overall capabilities and qualities (ISEC office, 2022). The ISEC program builds a platform for local colleges and universities to advance and further their campus internationalization to all-round development, via international curricula design, international faculty and staff training, and international educational resources. etc. (Ju, 2017). As of 2022, the ISEC program has more than 30 member schools. They are located in different cities in different provinces, including Jilin Normal University, Bohai University, Hebei University of Economics and Business, Inner Mongolia Normal University, etc. (ISEC office, 2022).

These ISEC member schools are not part of "Project 211", "Project 985", "C9 League (China 9 University League)", or "Double First Class", which involve only the leading and key

universities in China. According to the data from the Ministry of Education (MOE) of China, there are 2,914 colleges and universities in mainland China, and these universities are grouped under three initiatives (Chiu, 2020). The initiative of Project 211 started in 1995, and this project has funded 112 universities. In 1998, Project 985 initiative was launched, which includes 39 universities. Some universities are included in both Projects 211 and 985, because of their growing prestige and influence in the world. The C9 League was formed out of the initial 9 prestigious universities from Project 985 in 2009. The aim of the C9 League is to construct an elite group of universities in China and attract more talented and intelligent students from around the world (Chiu, 2020). The C9 League universities gain substantial support and funding from the national and local governments of China. They host 3% of research scholars, obtain 10% of national research budget, and produce 20% of cited academic papers in China (Chiu, 2020). The C9 League universities have priority in grants and enjoy special privileges. This league has been "dubbed the Ivy League of China" (Allen, 2017, p. 396), equivalent to the Ivy League in the United States or the Russell Group in the United Kingdom (Yue & Zhu, 2009; Fang et al., 2013). With the goal of comprehensively developing universities and their faculties into world-class and globally-ranked universities by the year 2050, the Double First Class University Plan was implemented in 2015. It includes 42 Double First Class Universities and 465 first class disciplines distributed among 140 schools (95 Double First Class Disciplines Universities) (Chiu, 2020).

Different from the leading and key universities aforementioned, ISEC member schools are from the majority of the ordinary or common colleges and universities in China. They lack educational funding, resources, and support from the national and local governments. There are challenges, risks, and difficulties for these local schools to advance, and deepen the process of

internationalization. This is where the ISEC program comes into play. The ISEC program acts as a lighthouse. It points out a new direction and renders a new momentum for these ordinary schools to go forward. The support and assistance of the ISEC program helps ISEC member schools increase in energy, vigor, and confidence, as they work to internationalize their campuses, and cultivate innovative talents with CT skills.

One of the advantages of the ISEC program is the international advanced teaching ideologies has been adopted, which include student-centeredness, general education, critical thinking (CT), formative assessment, and bilingual teaching (Ju, 2017). The ISEC program advocates and stresses the importance of CT in Chinese higher education, because an overwhelming amount of evidence showed that Chinese students are lacking in CT skills, problem-solving skills, as well as creativity and innovation (Atkinson, 1997; Ballard & Clanchy, 1991; Canagarajah, 2002; Pennycook, 1996; Wan, 1998). In order to help member schools cultivate innovative talents, the ISEC program is focused on fostering students' CT skills and emphasizes this focus in its cultivating goal and curriculum design (Ju, 2017). The ISEC program requires that ISEC teachers change their traditional teacher-centered teaching model to the student-centered model, stimulate students' initiative to study actively, and encourage students to vigorously participate in class activities (Ju, 2017). Students should become life-long learners with CT skills, and the creators of knowledge, rather than the slaves of knowledge (Ju, 2017). As for the curriculum design, the ISEC program offers, *An Introduction to Critical Thinking*, a course as part of the compulsory curricula in general education for ISEC students (Ju, 2017). Other courses in general education provided by the ISEC program also concentrate on cultivating students' CT skills, problem-solving skills, and creativity (Ju, 2017).

There are 3,775 ISEC teachers in total as of August 2022 (ISEC office, 2022). The ISEC program enacts strict rules and regulations in recruiting ISEC teachers. First, ISEC teachers must have obtained professional ranks, such as instructor or lecturer, associate professor or full professor, as well as a certain number of years of teaching experiences, which is variable depending on the teachers' degrees, rank, and area in which they teach. Second, teachers who voluntarily join the ISEC program, must be qualified for the related requirements of academic credentials and professional qualification. Third, if teachers pass the academic credentials and professional qualification screening, they must participate in the pre-service training of the ISEC program to gain the qualification to teach ISEC students. Finally, ISEC teachers must participate in the standard assessment annually (Ju, 2017). The results from the standard assessment were expected to improve each year, or the ISEC teachers will face the risk of losing the qualification to teach ISEC students (Ju, 2017).

Problem Statement

Critical thinking (CT) has become a hot issue in higher education and is regarded as one of the key and essential skills for students to succeed in the 21st century (Halpern, 2003; Adelman et al. 2014; Li, 2016; Roohr et al., 2019). However, CT cultivation in Chinese higher education has been missing for decades. The majority of Chinese teachers spend most of their time giving lectures, and students take notes and learn facts through repetition and rote memorization. The cultivation of students' CT is totally ignored, leading to students' deficiency of CT skills (Li, 2016). According to Li (2016), Chinese students are reproductive learners rather than analytical and speculative learners, and they are passive, unquestioning, and lacking in critical thinking.

Given the fact that Chinese students are lacking in CT skills, innovation, and creativity, the Ministry of Education (MOE) issued documents to advocate, motivate, and advance the CT cultivation in the students in China. Chinese colleges and universities either set up an independent course for critical thinking, or integrate CT into the existing curriculum (Zhai, 2015). However, when they teach CT in schools, teachers are faced with many obstacles and challenges at the personal, socio-cultural, and institutional levels. These obstacles and challenges include shortage of CT ideology, knowledge, and experiences; the great influence of Confucianism; the Chinese tradition "Doctrine of the mean" (中庸), which represents the "middle way between two extremes" (Chen, 2017, p. 532); inadequate educational resources and support; large class size; and the predominantly exam-oriented educational system, etc. Findings from research show that "much of the mandated high-stakes testing have resulted in teachers' over-concentration on lower-order thinking skills" (Zhang & Kim, 2018, p. 160).

According to Statistics Times (2021a), the population of China is 4.35 times greater than that of the US, "with China home to about 1.44 billion people and the United States to 331 million in 2020" (para. 1, United States vs China by Population, January 10, 2021a). Although the population in China is more than four times the US population, "the Per capita income of the United States is 5.78 and 3.61 times higher than that of China in nominal and [Purchasing Power Parity] PPP terms, respectively" (Statistics Times, para 3. Comparing United States and China by Economy, May 15, 2021b). One of the reasons is China is short of innovative and creative talents, which makes China lose its competitive edge in global affairs, economy, and events. To change this situation, the Chinese government is pressing colleges and universities to cultivate a new generation of highly skilled workers and produce innovation and creativity in science and

technology to serve as a recipe for promoting the economic growth and development. Critical thinking is a key component to this recipe.

Purpose of the Study

Critical thinking is at the heart of higher education, and has become a hot topic in both Western and non-Western academia (Wilson, 2016). With the advocacy and intervention of the Chinese Ministry of Education (MOE), colleges and universities across China have introduced CT into their schools, either as a stand-alone course or by infusing it into the existing curriculum (Zhai, 2015). Accordingly, Chinese teachers, both ISEC and non-ISEC teachers, are urged to either integrate CT into their instruction, or teach the course of CT directly, in order to nurture students' critical thinking, innovation, problem-solving, and other necessary competencies in the 21st century (Tan & Hairon, 2016). However, teaching CT effectively in China is faced with many obstacles and challenges, and the result of teaching CT can be far from satisfying.

At present, there is increasing interest in the research on Chinese students and their CT skills. This is only half of the story concerning CT education in China, for teachers' attitude towards the reform and their teaching practice counts for much more in education (Dai et al., 2011). However, there is a dearth of research on Chinese teachers and their CT skills both inside and outside of China. Examining Chinese Teachers' CT skills is fundamental, because teachers are the main agent for school changes and they play a vital role in any successful reform effort (Darling-Hammond & Bransford, 2005). Therefore, the target of this study is ISEC teachers and non-ISEC teachers from local or provincial colleges and universities in the north of China. The reason for choosing these two groups of teachers is as follows: The ISEC program offers guidance, support and assistance to those local colleges and universities in China in reforming their schools and cultivating more innovative talents. The ISEC teachers have more opportunities

to access the advanced educational ideologies, such as critical thinking (CT), student-centeredness, formative assessment, etc. (Ju, 2017). One of the assumptions in this study is that these ISEC teachers have higher levels of CT skills than the non-ISEC teachers in local colleges and universities in the north of China. It is widely accepted that the Qinling-Huaihu Line is used to divide China into 2 regions geographically: the north and the south (Tang et al, 2020). There are a world of differences between the north and south of China, with regard to climate, environment, economy, culture, education, to name a few. This study focuses solely on the institutions of higher education in the north of China, because these colleges and universities share more similarities or commonalities.

By shifting the research focus from students' CT to teachers' CT, the researcher employed the explanatory mixed methods design, the QUAN-Qual Model (Gay, et al., 2006), to explore the status quo (current situation), perception, understanding and practice of CT among the ISEC and non-ISEC teachers in institutions of higher education. Understanding the status quo of teachers' CT skills and their perception and practice of CT becomes critical because it helps construct a panoramic view of CT education in China. The quantitative method was dominant in this study. The researcher used a questionnaire to investigate the status quo of Chinese teachers' CT skills, and whether there was any significant difference in CT skills between the ISEC and non-ISEC teachers to test the third hypothesis of the study. Teacher's instruction, experience, training, to name a few, are the key factors that affect students' learning outcome and CT cultivation (Torff, 2005; Mangena & Chabeli, 2005). Data were analyzed to determine whether there was any significant relationship between CT skills and such variables as age and years of teaching, as well as whether there was any significant difference in CT skills and variables, such as the professional rank, educational background, and discipline. Finally, the follow-up

interviews were conducted to explore how the ISEC and non-ISEC teachers perceived CT, and how they integrated CT into their teaching practice. Findings from this mixed methods study offered an in-depth understanding of the theory and practice of CT education in the north of China, as well as Chinese teachers' perception, understanding, and practice of CT. The following section deals with research questions of this study.

General Research Questions

The basic assumption of a mixed methods research design is that the use of both quantitative and qualitative methods, in combination, provides a better understanding of the research problem and question than either method by itself (Creswell, 2012). In this non-experimental causal-comparative study, the explanatory mixed methods design was used. This explanatory mixed methods design followed the process of first collecting quantitative data and then collecting qualitative data to help explain or elaborate on quantitative results (Creswell, 2012). Answers was sought to the following nine research questions, gaining an overview of CT skills of Chinese teachers in institutions of higher education in the north of China. The researcher tried to find out whether there was any difference in CT skills between the ISEC teachers and non-ISEC teachers. The researcher also examined whether there was any relationship between CT and age, between CT and years of teaching, as well as whether there was any difference in CT skills between the ISEC teachers and non-ISEC teachers, based on gender, professional rank, educational background, and disciplines.

The research questions, alternative hypotheses, and null hypotheses were as follows:

R_1 What is the level of CT skills of overall Chinese teachers in institutions of higher education in the north of China?

R_2 What is the level of CT skills of the ISEC and non-ISEC teachers in institutions of higher education in the north of China, respectively?

R_3 Is there a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers in the north of China?

- H_1 There is a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers in the north of China.
- H_0 There is no statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers in the north of China.

R_4 Is there a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers, based on gender identifying as male and female?

- H_1 There is a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers, based on gender identifying as male and female.
- H_0 There is no statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers, based on gender identifying as male and female.

R_5 Is there a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers based on professional rank?

- H_1 There is a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers based on professional rank.
- H_0 There is no statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers based on professional rank.

R_6 Is there a statistically significant relationship between CT skills and age?

- H_1 There is a statistically significant relationship between CT skills and age.
- H_0 There is no statistically significant relationship between CT skills and age.

R_7 Is there a statistically significant relationship between CT skills and years of teaching?

- H_1 There is a statistically significant relationship between CT skills and years of teaching.
- H_0 There is no statistically significant relationship between CT skills and years of teaching.

R_8 Is there a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers based on the educational background?

- H_1 There is a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers based on the educational background.
- H_0 There is no statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers based on the educational background.

R_9 Is there a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers based on the discipline?

- H_1 There is a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers based on the discipline.
- H_0 There is no statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers based on the discipline.

The Central Question

According to Creswell and Poth (2018), the intent of qualitative central question is to "narrow the purpose to several questions that will be addressed in the study" (p. 137). The central question provides "an opportunity to encode and foreshadow an approach to inquiry" (Creswell & Poth, 2018, p. 137). In this explanatory mixed methods study, the purpose of the central question was to explore how the ISEC and non-ISEC teachers perceived CT, what attitude they held towards CT, and what experience they had regarding CT instruction. The

findings from the qualitative research supplemented and deepened the research results obtained from the quantitative design. The central question was as follows:

1. What is the perception, attitude, and practice regarding CT among the ISEC and non-ISEC teachers in institutions of higher education in the north of China?

Definition of Terms

For the purpose of this explanatory mixed methods study, the following terms were used, followed by their definitions.

Critical thinking (CT). Critical thinking is essential for inquiry, referring to "purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based" (Facione, 1990a, p. 2). CT can be conceptualized into two dimensions: "cognitive skills and affective dispositions" (Facione, 1990a, p. 2).

Critical thinking skills. CT skills are cognitive skills, including "(1) interpretation, (2) analysis, (3) evaluation, (4) inference, (5) explanation and (6) self-regulation. Each of these six is at the core of CT" (Facione, 1990a, p. 4). CT skills can be classified into different subskills. (1) Interpretation covers such subskills as categorization, decoding significance, and clarifying meaning. (2) Analysis involves subskills, like examining ideas, identifying arguments, analyzing arguments. (3) Evaluation includes assessing claims, assessing arguments. (4) Inference consists of querying evidence, conjecturing alternatives, and drawing conclusions. (5) Explanation is composed of stating results, justifying procedures, and presenting arguments. (6) Self-regulation is made up of self-examination and self-correction (Facione, 1990a, p. 6).

Critical thinking dispositions. CT disposition refers to "the consistent internal motivation to engage problems and make decisions by using thinking" (Giancarlo & Facione, 2001, p. 31). It is "an implicit component that can either enhance or hinder critical thinking" (Thomas & Lok, 2015, p. 99), involving the following components of CT:

> inquisitiveness with regard to a wide range of issues; concern to become and remain generally well-informed; alertness to opportunities to use CT; trust in the processes of reasoned inquiry; self-confidence in one's own ability to reason; open-mindedness regarding divergent world views; flexibility in considering alternatives and opinions; understanding of the opinions of other people; fair-mindedness in appraising reasoning; honesty in facing one's own biases, prejudices, stereotypes, egocentric or sociocentric tendencies; prudence in suspending, making or altering judgments; willingness to reconsider and revise views where honest reflection suggests that change is warranted (Facione, 1990a, p. 13).

***ISEC program*.** International Scholarly Exchange Curriculum (Undergraduate) program is acronymized as ISEC program, affiliated to the China Scholarship Council (CSC). It is an international education program, based on the strategies of "China's Education Modernization 2035 Plan". The targeted members of ISEC program are local or provincial colleges and universities in China. Its mission is to offer guidance, service and support for local schools in faculty and staff training, teaching management, teaching quality assurance, etc. The aim of the ISEC program is to help local colleges and universities to reform their curricula and teaching, and cultivate innovative talents (ISEC office, 2022).

China Scholarship Council (CSC). China Scholarship Council (CSC), founded in 1996, is a non-profit organization affiliated to the Ministry of Education's (MOE). The CSC offers

support for international academic exchange with China and is the primary vehicle through which the Chinese government awards scholarships. The CSC provides both funding for Chinese citizens and residents to study abroad, and for foreign students and scholars to study in China (China Scholarship Council, 2022).

ISEC member schools. Different from Project 211, Project 985, the C9 League and Double First-Class universities, ISEC member schools are from local or provincial colleges and universities that choose to join in ISEC programs. As of August 2022, there are more than 30 ISEC member schools (ISEC office, 2022). In this study, the local or provincial colleges and universities are also named ordinary or common colleges and universities, different from key and leading universities. They are exchangeable in this study (ISEC office, 2022).

ISEC teachers. Teachers from local or provincial colleges or universities, who have participated in pre-service training of the ISEC program and obtained the qualification of teaching ISEC students, get the title of ISEC teachers. ISEC teachers must attend regular training to keep their knowledge and skills updated. ISEC teachers must be assessed annually, as well. If their assessment is not kept up to the standard, ISEC teachers will lose their qualification of teaching ISEC students (ISEC office, 2022).

Non-ISEC teachers. As opposed to the teachers from Project 211, Project 985, the C9 League and Double First Class universities, teachers who come from ordinary or common colleges and universities (local or provincial colleges and universities) are named non-ISEC teachers. (ISEC office, 2022).

The north of China. The demarcation line or geographical dividing line between the north and south of China is the Qinling-Huaihe Line, or Qin Mountain and its eastern extension to Huai River in central and eastern China (Gao et al., 2011; Song et al., 2017; Zhou et al., 2016).

The north of China includes “15 provinces, autonomous regions and municipalities: Heilongjiang, Jilin, Liaoning, Inner Mongolia, Beijing, Tianjin, Shandong, He’nan, Hebei, Shanxi, Shaanxi, Ningxia, Gansu, Qinghai, and Xinjiang” (Zhou et al., 2016, p. 519). The south of China consists of 17 “provinces, autonomous regions and municipalities, and 2 Special Administrative Regions: Jiangsu, Shanghai, Anhui, Hubei, Hunan, Chongqing, Sichuan, Tibet, Yunnan, Guizhou, Guangxi, Guangdong, Jiangxi, Fujian, Zhejiang, Hainan, Hong Kong and Macao, [Taiwan]” (Zhou et al., 2016, p. 519).

Delimitations

This study was confined to several delimitations. Firstly, the study investigated the ISEC teachers and non-ISEC teachers in institutions of higher education in the north of China, rather than the south of China. Secondly, these teachers came from local or provincial colleges and universities, rather than Project 211, Project 985, the C9 League, and Double First Class universities. Thirdly, the researcher of this study surveyed the CT skills of the ISEC and non-ISEC teachers during a particular time in a specific semester. The study was delimited to the time frame because of the research time constrictions and possible issues with attrition. This study was designed to examine only those critical thinking skills that were measured by the instrument--- California Critical Thinking Skills Test (CCTST). This study was aimed to explore Chinese teachers’ CT. Finally, there are various factors that contribute to or affect teachers’ CT skills, such as experiences of CT training, workplace culture, school policies, etc. This study was only focused on the factors, such as age, years of teaching, educational background, and discipline. The limitations of this study are discussed as follows.

Limitations

There were several limitations that cannot be disregarded in this study. With regard to the section of the quantitative research design, the researcher of the study recruited 102 participants (52 ISEC teachers and 50 non-ISEC teachers) to participate in this survey. This study used a nonprobability sample. The sample size was comparatively small. The small sample size cannot represent all features of the entire teacher population in local or provincial colleges and universities in the north of China. The generalizability of this study was limited, due to the nonprobability sample and small sample size. The main target of this study was the ISEC and non-ISEC teachers in local or provincial colleges and universities. The teachers in Project 211, Project 985, the C9 League, and Double First Class universities need to be investigated in the future CT research studies, in order to have an in-depth understanding about Chinese teachers' CT and CT education in China.

Qualitatively, the interview was conducted to explore how Chinese teachers perceived, understood and integrated CT into their teaching. With the influence of social desirability, participants may give a popular answer that the researcher agrees with rather than their true opinion. During the interview, the participants may not tell the exact truth. Therefore, truthfulness was one limiting factor in the qualitative research. The information provided by the participants is based on their perceptions and memory, which may become an issue. These limitations can influence the qualitative outcome of the study. An additional limitation was, participants were dealing with COVID-19 during the time of this study.

Significance of the Study

Because of the important role that CT plays in higher education, teachers around the world are encouraged to integrate CT into their instructions (Li, 2016). Research suggested that teachers' resistance or unwillingness to make changes may block students from thinking

critically (Torff, 2005), and teachers' beliefs have an impact on their approaches to and success in promoting students' critical thinking (Dike, et al., 2006). Lee (1993) also stressed that the reasoning ability of teachers is vital for promoting students' reasoning. According to Feucht and Bendixen (2010), there are two types of teachers--- absolutist teachers and evaluativist teachers. They hold opposite beliefs in their teaching: Absolutist teachers may suppress the CT development of students via utilizing a transmission teaching approach, while evaluativist teachers may foster students' CT via fostering their interest in new things, and encouraging them to construct knowledge. McBride et al. (2002) summarized if CT is "a valued educational outcome, its teachers should subscribe to it" (p. 133).

Since there is a close link between teachers' pedagogical beliefs and their instructional practices (Borg, 2006, Farrell & Kun, 2008), it is of great significance to investigate teachers CT skills and their perception, attitude and practice of CT via an explanatory mixed methods research design. The significance of this study is as follows. Firstly, teachers play an essential role in students' learning process. Teachers need to be equipped with knowledge and skills of CT to help students become critical and innovative thinkers (Li, 2016). Researching the CT skills of the ISEC and non-ISEC teachers in Chinese local or provincial colleges and universities, as well as how they perceive CT and integrate CT into their teaching practice is imperative. It helps "researchers, policy-makers and teachers themselves identify the guiding principles in relation to their classroom work and go beyond description towards an understanding and explanation of teacher actions" (Li, 2016, p. 274).

Secondly, there is a large volume of research that focuses on students and their CT skills, however, there is a scarcity of research targeting teachers and their CT skills. Especially, it is difficult to find any literature inside or outside of China that concerns CT skills of ISEC and non-

ISEC teachers in institutions of higher education in China. This explanatory mixed methods study addresses this deficit in the literature by specifically concentrating on the status quo (current situation) of Chinese ISEC and non-ISEC teachers' CT skills, as well as how Chinese teachers perceive, understand, and integrate CT into their classrooms. The research findings will offer some hints and clues for those educational leaders, educators, and researchers who are interested in the teaching and learning of CT and those who show great interest in teachers' CT skills. The findings will also provide some implications and recommendations for policy makers to improve CT education. The most important significance of this explanatory mixed methods research is that it offers "base-line data for one group of stakeholders that will allow educators and policy makers to answer questions about whether it is worth the time, expense and effort" to develop and implement CT-based curricula (Li, 2016, p. 274). In summary, the research findings from this explanatory mixed methods study may bridge the long-standing gap between theory and practice in CT and Chinese teachers, and bridge the gap in the literature on ISEC and non-ISEC teachers' CT skills in China.

Summary

Teaching CT skills is vital for student learning and their future development, and thus cultivating global citizens with creativity and innovative spirit (MacDonald, 2005). Some research indicated that Chinese students are obedient and passive, lacking in CT skills in their learning process. Chinese teachers stress "rote learning and knowledge acquiring and retrieving, rather than knowledge construction and creation" (Li, 2016, p. 274), neglecting the cultivation of students' CT skills. Since 2001, the Chinese government has carried out several outstanding educational reforms to change "receptive learning, rote learning, and mechanical drilling and to advocate learner participation, exploration, information collection and comprehension, problem-

solving, negotiation and collaboration" (Li, 2016, p. 274). The ISEC program, as one of the reform agendas of Chinese higher education, plays a leading role for local colleges and universities to implement their education reform on curriculum design and teaching ideology. Among them, cultivating innovative talents and lifelong learners with CT skills is especially emphasized in Chinese education reforms.

A large number of studies focus on Chinese students' learning and their CT skills, however, there is a lack of research targeting Chinese teachers' CT skills. Particularly, there is no literature that concerns the ISEC and non-ISEC teachers in local or provincial colleges and universities and their CT skills. Therefore, this explanatory mixed methods research was designed to investigate the status quo of the ISEC and non-ISEC teachers in local colleges and universities in China, explore the relationship and difference between CT skills and variables (age, years of teaching, educational background and discipline), and examine Chinese teachers' perception, understanding and practice of CT in the context of local colleges and universities. The findings may offer some source and reference for educational leaders, educators, and teachers to advance Chinese higher education reform and especially CT education. The next chapter deals with Review of Related Literature.

Chapter Two: Review of Related Literature

There still exists some controversy on issues concerning CT, such as "whether or not it can be defined and measured, and whether or not it is possible to teach it in the Asian L2 context" (Lin, 2014, p. 16), though the importance of CT is well accepted. The main focus of this chapter is on the theoretical and empirical research on critical thinking (CT). It involves the conceptualization of CT, assessment of CT, CT studies in the West, and CT studies in China. At the end of the chapter, comes the summary of literature review around CT.

Conceptualization of Critical Thinking (CT)

Liu, Frankel and Roohr (2014) claimed that "one of the most debatable features about critical thinking is what constitutes critical thinking---its definition" (p. 2). There is "a notable lack of consensus regarding the definition of critical thinking" (Lai, 2011, p. 4), in spite of its importance in education and the workplace. Researchers and scholars focus on different aspects of critical thinking. Some focus on the "reasoning process specific to critical thinking, while others emphasize the outcomes of critical thinking…" (Liu et al., 2014, p. 2). According to Huitt (1998), cognitive and behavioral psychologists, philosophers and content specialists all contributed to people's understanding of CT from their own perspectives. Other researchers, such as Lewis and Smith (1993), especially, stressed the perspectives of philosophers and psychologists towards critical thinking. Sternberg (1986), in "Critical Thinking: Its Nature, Measurement, and Improvement", elaborated three perspectives of theorizing thought: philosophical, psychological, and educational. These separate academic perspectives or dimensions have offered various approaches to the definition of CT. Each of them reflects their respective thoughts and concerns. These three perspectives---philosophical, psychological, and educational, are explored in detail as follows.

The philosophical perspective

Socrates (470-399 BCE), Plato (427-347 BCE), and Aristotle (384-324 BCE) have been regarded as the founders of critical thinking in ancient times (Atabaki et al., 2015; Ennis, 201; Sternberg, 1986). Socrates "related critical thinking to philosophy by his exploratory dialogues" (Atabaki, 2015, p. 94). In Socrates' point of view, CT referred to discovery of the truth. Then Socrates' student, Plato, as well as Plato's student, Aristotle, continued to conduct research on knowledge and thinking. According to Plato, CT, "as logic, is the tool that will help us find the answer or solution to our confusion and problems" (Thayer-Bacon, 2000, p. 22). Aristotle invented a logical method, called the *syllogism*, to help knowers test out their ideas and even create new knowledge (Thayer-Bacon, 2000). The difference between Plato and Aristotle is that Plato thought, "people are discovering the knowledge while Aristotle believed that intellectual talent [knowledge itself] is one of the most important features of people" (Atabaki, 2015, p. 94). After Socrates, Plato, and Aristotle, American philosopher, psychologist, and educational reformer, Dewey (1859-1952) has been viewed as the modern founder of the CT movement (Sternberg, 1986), or the "true harbinger of critical thinking in this century [twentieth]" (Lipman, 2003, p.35). Dewey (1910) focused on the link among thinking, experience, doing, and the consequences of action in his work, *How We Think* (Thayer-Bacon, 2000). He proposed the model of inquiry or reflective thinking (critical thinking). For Dewey (1910),

> the essence of critical thinking is suspended judgment; and the essence of this suspense is inquiry to determine the nature of the problem before proceeding to attempts at its solution. This, more than any other thing, transforms mere inference into tested inference, suggested conclusions into proof. (p. 74)

By following Dewey's idea, CT is "motivated by a problem" (Tanner, 1988, p. 471).

In more recent times, Ennis, McPeck, Lipman, Paul, Facione, etc., are the important philosophers and scholars who "devoted their attention to understanding the basis of critical thinking" (Sternberg, 1986, p. 3). According to Tian (2008), Ennis is generally credited as an outstanding philosopher that popularized CT in tertiary education in the1960s. Ennis referred to CT as "reasonable reflective thinking that is focused on deciding what is to believe or do" (1987, p. 10; 1996, p. 166). "The emphasis is on reasonableness, reflection and the process of making decisions" (Ennis, 1996, p. 166). In Ennis' definition, an intentional and motivational aspect of CT is stressed, and this is named disposition of CT by other scholars, such as Facione (1990a) and Halpern (1998). According to Ennis (1987), the CT disposition is "a will of thinking critically and reflexively about one's own claims and other people's claims" (Li, 2017, p. 30). There are three key CT dispositions: (a) care that one's beliefs are true and decisions are justified; (b) care to represent one's positions honestly and understand others' positions clearly; and (c) care about everyone's feelings and welfare (Ennis, 1996). The first two are identified as

> *enabling* dispositions or tendencies that must be present within the thinker if critical thinking is to take place, the third is merely a *facilitating* deposition: a tendency of mind that leads to critical thinking being conducted well wherever it happens to take place. (Robinson, 2011, p. 276)

Ennis classified CT skills or abilities under five main categories, and they were further subdivided (Sternberg, 1986). These five categories include "elementary clarification, basic support, inference, advanced clarification, and strategy and tactics" (Sternberg, 1986, p. 9). Ennis is a staunch advocate of the CT transferability argument, which means he holds a strong belief that CT is universal and generic (Moore, 2004) and can be transferred.

McPeck (1981), another seminal philosopher, regarded CT as "the propensity and skill to engage in an activity with reflective skepticism" (p. 8). He stressed CT is the correct usage of reflective skepticism and is "necessarily lined with specific areas of expertise and knowledge" (p. 19). McPeck argued CT skills, "in general, are parasitic upon detailed knowledge of, and experience in, parent fields and problem areas" (p. 10), and "rationality includes critical thinking as a particular aspect (or subset) of itself" (p. 12). To put it another way, CT is a dimension of rationality. This clearly indicates that McPeck emphasized the importance of knowledge and experience, as well as the relationship between CT and rationality. Similar to Ennis' idea, McPeck also believed that "being willing to doubt and being able to doubt are two aspects of CT (i.e. having a disposition to CT)" (Li, 2017, p. 30). They both believed training regarding CT skills is not sufficient to create a critical thinker, and people must develop dispositions to employ their skills. (Ennis, 1987, 1996; McPeck, 1981). According to Moore (2004), however, McPeck is the leading protagonist who strongly argued that CT is subject- or discipline-specific in the generic vs discipline-specific debate. While Ennis held the opposite point of view: CT is generic. As a generalist, Ennis (1987) believed CT has a universal and generic quality, and can be taught as an independent study or course (Moore, 2011). While McPeck, as a specifist, insisted that CT is "not some universal quality, but one that exhibits a good deal of variation and which is shaped irredeemably by the particular problem area under consideration" (Moore, 2011, p. 263). According to McPeck, the useful thinking skills are limited to specific domains or narrower areas of application. Thus, the implication for teaching, from McPeck's specifist perspective, is that "the development of students' critical abilities is always best pursued within the context of their study within the disciplines" (Moore, 2011, p. 263), and CT instruction "must differ from

discipline to discipline" (Lipman, 1987, p. 11). Therefore, McPeck is criticized by many scholars because he insists that CT is not transferable (Atabaki et al., 2015; Li, 2017).

Another philosopher, Lipman, created an approach or a method, named *philosophy for children* (P4C) to develop children's critical thinking through philosophical dialogue in the 1970s (Karadağ & Demirtaş, 2018). Lipman was viewed as "the founder of the movement of philosophy for children" (Karadağ & Demirtaş, 2018, p. 2), because he initially developed the P4C Program, which symbolized one method to introduce CT skills (Lewis & Smith, 1993; Vansieleghem & Kennedy; 2011; Karadağ & Demirtaş; 2018). According to Lipman (1987), CT includes "analyzing, judging, hypothesizing, explaining and many other cognitive activities besides deciding and problem-solving" (p. 5). CT is thus delineated as "a process that occasionally results in decisions or solutions, but the process is not to be defined solely by those occasional consequences" (Lipman, 1987, p. 5). Lipman (1987) stressed the functional definition of CT must consider three characteristics: "(1) It is *self-corrective* thinking; (2) it is thinking *with criteria*; and (3) it is thinking that is *sensitive to context*" (p. 5). Therefore, "thinking that is sensitive to context becomes critical thinking when it is self-correcting and when it makes use of criteria" (Lipman, 1987, p. 5). Lipman also held that CT is "much easier to develop in relations with peers than to be taught in a technical way" (Karadağ & Demirtaş, 2018). In terms of effective educational approaches to CT, Lipman (1987) pointed out, even though there is merit to both propositions of the generalist (Ennis) and the specifist (McPeck), the ideas of the specifist are too narrow. It is similar to the fact that, although science is not the same in physics, chemistry, or mathematics, a general course in the scientific method is still useful and significant (Lipman, 1987). So the same is true of CT. Later, Lipman (1997) proposed The Complex Thinking or Higher Order Thinking Scheme, which represents more cognitive characteristics

(Pacheco & Herrera, 2021). Lipman (1997) believed "complex thinking emerges from the fusion between critical and creative thinking" (Pacheco & Herrera, 2021, p. 2).

Paul (1992), a leading proponent of critical thinking, argued that one definition of CT is too narrow. Instead, it is wise to "retain a hose of definitions" (p. 46), in order to maintain insight into the different dimensions of CT and avoid the limitations of each definition (Paul, 1992). According to Paul (1992), critical thinking is:

> a) the art of thinking about your thinking while you're thinking so as to make your thinking more clear, precise, accurate, relevant, consistent and fair; b) the art of constructive skepticism; c) the art of identifying and removing bias, prejudice, and one-sidedness of thought; d) thinking that rationally certifies what we know and makes clear wherein we are ignorant. (p.47)

Then, Paul (1992) set out another definition of CT --- "disciplined, self-directed thinking which exemplifies the perfections of thinking appropriate to a particular mode or domain of thinking" (p.48). Paul (2005) regarded CT as "the art of thinking about thinking in an intellectually disciplined manner, …focus[ing] on thinking in three interrelated phases" (p. 28). Creative thinking belongs to the third phase, substituting strong thinking for weak thinking, or substituting stronger thinking for strong thinking (Paul, 2005). Creative thinking is a by-product of CT (Paul, 2005). It clearly indicates that Paul believed critical thinking has different degrees (Merrifield, 2018). Paul (1992) proposed two forms of CT: fair-minded or strong sense CT and sophistic or weak sense CT. For Paul (1992), sophistic or weak CT is "the use of critical thinking skills to serve the interest of a particular group of individuals without taking others into consideration" (Merrifield, 2018, p. 28). In contrast, the fair-minded or strong sense of CT, is the use of thinking skills to "take into account the interests of diverse persons or groups" (Paul, 1992, p.48).

Therefore, the strong sense of CT, "critiquing one's own point of view, is related to thinking respectively about the knowledge we had accepted, about what we had believed, in order to update it with better knowledge" (Li, 2017, p. 31). According to Paul (2005), most faculty members lack a substantive concept of CT. What is missing is the "coherence, connection, and depth of understanding that accompanies systematic critical thinking" (Paul, 2005, p. 37). Success in introducing a substantive concept of CT to faculty members demands "well planned, long-term professional development based explicitly in the multiple dimensions of a substantive concept of critical thinking" (Paul, 2005, p. 27).

Facione, another seminal philosopher, "spearheaded the effort to identify a consensus definition of critical thinking using the Delphi approach, an expert consensus approach [in 1990]" (Liu et al., 2014, p. 2). In the Delphi Report Project, 46 panelists, who come from philosophy, education, social sciences, or physical sciences, with experience and expertise in CT, proposed "a list of mental skills and habits of mind [with regards to CT]" (Facione, 2011, p. 5). The cognitive skills are at the very core of CT, including "interpretation, analysis, evaluation, inference, explanation, and self-regulation" (Facione, 1990a, p. 6; Facione, 2011, p. 5). It is clear "a person does not have to be proficient at every skill to be considered a critical thinker" (Liu et al., 2014, p. 2). Forty-six panelists or experts also reached a consensus on the affective or dispositional elements of CT. For instance, "inquisitiveness with regard to a wide range of issues", "concern to become and remain generally well-informed", "alertness to opportunities to use CT", etc. (Facione, 1990a, p. 13). According to Facione (2011), CT "is skeptical without being cynical. It is open-minded without being wishy-washy. It is analytical without being nitpicky. Critical thinking can be decisive without being stubborn, evaluative without being

judgmental, and forceful without being opinionated" (p. 25). Therefore, ideal critical thinkers are:

> habitually inquisitive, well-informed, trustful of reason, open-minded, flexible, fair-mined in evaluation, honest in facing personal biases, prudent in making judgements, willing to reconsider, clear about issues, orderly in complex matters, diligent in seeking relevant information, reasonable in the selection of criteria, focused in inquiry, and persistent in seeking results which are as precise as the subject and the circumstances of inquiry permit. (Facione, 1990a, p. 2; Facione, 2011, p. 27)

According to Liu et al. (2014), the approach that the American Association of Colleges and Universities (AAC & U) employed to define CT is heavily affected by the Delphi Report definitions of CT.

The psychological perspective

Not only philosophers, but also psychologists are interested in the nature of CT (Huitt, 1998; Lai, 2011; Lewis & Smith, 1993; Sternberg, 1986). Different from philosophers, psychologists have elucidated their ideas and thoughts about CT distinctively from research in cognitive and developmental psychology and intelligence theories (Bransford et al., 1987; Halpern, 2003; Sternberg, 1986). The prominent psychologists include Bransford, Bruner, Feuerstein, Sternberg, Halpern, etc. (Sternberg, 1986; Lai, 2011). Psychologists are mainly concerned with "characterizing critical thinking as it is performed under the limitations of the person and environment" (Sternberg, 1986, p. 5). According to Lai (2011), the psychologists differ in two ways from philosophers in terms of CT. First, psychologists who emerged in the behaviorist tradition and experimental research, concentrate on people's actual or true thinking versus the way they think under ideal situations. Second, psychologists try to define CT by the

types of behaviors critical thinkers can do, instead of pointing to features or characteristics of ideal critical thinkers or "standards of good thought" (Lai, 2011, p.7). Sternberg and Halpern are very popular and influential psychologists in modern times, so their ideas about CT are addressed in the following paragraphs, respectively.

Sternberg (1985) chose a psychological perspective, rather than logical perspective to analyze CT. Based on his "componential" description and explanation of thought, Sternberg proposed three types of CT skills. They are "metacomponents, performance components, and knowledge-acquisition components" (Sternberg, 1986, p. 9). As higher order executive processes, metacomponents are utilized for "executive planning and decision making in problem solving" (Clarke, 1986, p. 206). Performance components, belonging to lower order and nonexecutive processes, are employed to "execute the instructions of the metacomponents, and provide feedback to them" (Sternberg, 1986, p. 7), or "used in the execution of problem-solving strategies" (Clarke, 1986, p. 206). Knowledge-acquisition components are used to acquire new information (Clark, 1986). According to Sternberg and Davidson (1983), insight skills consist of three types. They are: (a) selective encoding, "by which relevant information in a given context is sifted from irrelevant information" (Sternberg & Davidson, 1983, p. 51); (b) selective combination, "by which relevant information is combined in a novel and productive way" (Sternberg & Davidson, 1983, p. 51); and (c) selective comparison, "by which new information is related in a novel way to old information" (Sternberg & Davidson, 1983, p. 51). In terms of how to teach critical thinking, Sternberg (1985), in "Teaching Critical Thinking. Part I: Are We Making Critical Mistakes?", argued the first thing and the premise in problem-solving and decision-making, is the recognition of valuable problems in the everyday world. To put it another way, "students need to be taught not only how to solve problems, but also how to find problems

that are worth solving" (Sternberg, 1985, p. 196). However, life does not provide any predictability of critical thinking problems, and people need to learn how to solve real problems and make wise decisions in everyday life (Sternberg, 1985). Later, Atabaki et al (2015) summarized Sternberg's definition of critical thinking in this way: critical thinking mainly deals with solving problems and making decisions. According to Sternberg (1985), the solutions to everyday problems rely on not only informal knowledge, but also formal knowledge. Sternberg (1985) concluded approaches to teaching CT need to "deal adequately with the demand of critical thinking in everyday life" (p. 198). Branford and Stein (1993), in their book *The Ideal Problem Solver: A Guide for Improving Thinking, Learning and Creativity*, also stressed all of the skills or techniques for problem solving should come from everyday examples. These concrete, everyday examples, or instances are beneficial for problem solving and decision making in everyday life.

Halpern, another seminal psychologist, argued that many psychologists proposed definitions of CT, and those definitions were inclined to be similar in content. According to Fischer and Spiker (2000), most definitions of CT included "reasoning/logic, judgment, metacognition, reflection, questioning and mental processes" (Halpern, 2003, p. 6). Halpern (1998) argued CT is "purposeful, reasoned and goal-directed" (p. 450), and it is thinking that includes problem solving, inference formulating, likelihood calculating, and decision making. Later, Halpern (2003) defined critical thinking as

> the use of those cognitive skills or strategies that increase the probability of a desirable outcome. It is used to describe thinking that is purposeful, reasoned, and goal directed---the kind of thinking involved in solving problems, formulating inferences, calculating

> likelihood, and making decisions, when the thinker is using skills that are thoughtful and effective for the particular context and type of thinking task. (p. 6)

Halpern (2003) emphasized, CT is thinking not only concerning making judgments and solving problems, but also using skills and strategies that lead to desirable outcomes. In terms of how to teach CT, Halpern (1998) believed a successful pedagogy "can serve as a basis for the enhancement of thinking" (p. 451). Halpern (1998) posed a model, aimed to guide in teaching thinking skills and promoting transfer to new contexts. This model consists of four parts:

> (a) a dispositional component to prepare learners for effortful cognitive work, (b) instruction in the skills of critical thinking, (c) training in the structural aspects of problems and arguments to promote transcontextual transfer of critical-thinking skills, and (d) a metacognitive component that includes checking for accuracy and monitoring progress toward the goal. (Halpern, 1998, p. 449)

According to Atabaki et al. (2015), Halpern (1998) argued her four-part empirically based CT teaching model "is common to any field" (p. 96). "The question of transferability is important to the discussion of how critical thinking skills should be taught" (Atabaki et al., 2015, p. 96). Lai (2011) pointed out Halpern (2001) contended "instruction in general thinking skills, taught as a broad-based, cross-disciplinary course, is the most effective way of teaching critical thinking" (p. 30). Halpern (2003) stressed there is considerable evidence to show "thinking skills courses and thinking skills instruction that is embedded in other courses can have positive effects that are transferable to many situations" (p. 10). In particular, CT instruction needs to concentrate "overtly and self-consciously on the improvement of thinking, and the learning experience needs to include multiple examples across domains in order to maximize transfer" (Halpern, 2003, p. 13).

The educational perspective

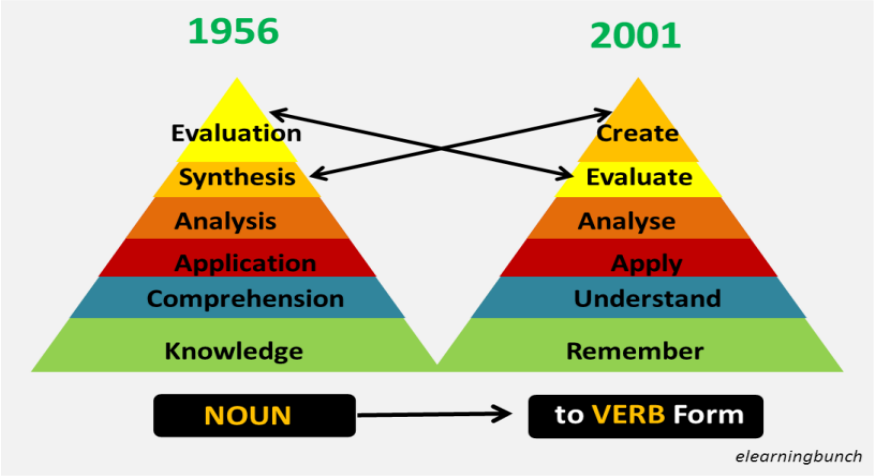
In "Critical Thinking: Its Nature, Measurement, and Improvement", Sternberg (1986) proposed three perspectives with respect to the study of CT. Besides the philosophers and psychologists, the educators or educational theorists cannot be disregarded. The leading figures of educational theorists are Bloom et al. (1956), Gagné (1965), De Bono (1967, 1969), Perkins (1981), and Renzulli (1976) (Sternberg, 1986). For these theorists, "theorizing [critical thinking] seems directly responsive to the skills needed by children in the classroom for problem solving, decision making, and concept learning" (Sternberg, 1983, p. 7). The educational theorists rely heavily on "classroom observation, test analysis, and process analysis of thinking in the classroom to guide their thinking about critical thinking" (Sternberg, 1983, p. 7). Considering their significant influence and contribution, Bloom's, Gagné's, and De Bono's (critical) thinking ideas and thoughts are elucidated in the following paragraphs.

Bloom et al's (1956) taxonomy of cognitive skills is widely used in academic research, classroom contexts, and even textbooks. According to Bloom et al. (1956), teachers should set up educational goals and outcomes in the cognitive area, including such activities as remembering, recalling knowledge, thinking, problem solving, and creating, when teachers build their curricula. The purpose of Bloom et al's taxonomy is to "develop a codification system whereby educators could design learning objectives that have a hierarchical organization" (Marzano & Kendall, 2007, p. 1). Bloom et al's Taxonomy is a "multi-tiered model of classifying thinking according to six cognitive levels of complexity" (Forehand, 2010, p. 48). These six cognitive skills are ranked on a hierarchy: knowledge, comprehension, application, analysis, synthesis and evaluation. Bloom et al. later revised their Taxonomy from noun to verb forms, and they are: remembering, understanding, applying, analyzing, evaluating and creating (Forehand, 2010). The

first three skills (remembering, understanding, applying) belong to the lower order thinking skills, while the other three skills (analyzing, evaluating, & creating) are viewed as higher order thinking skills, which have been taken as critical thinking skills (Scriven & Paul, 1987). Some researchers (Davies & Barnett, 2015; Johnson et al., 2011), used Bloom's Taxonomy of thinking skills as their research framework from the cognitive perspective (Chen, 2017, p. 21). Bloom's Taxonomy is highly valued by scholars, researchers, and educators. Forehand (2010) argued Bloom's Taxonomy has provided "the measurement tool for thinking, [and] an even more powerful tool to help design their lesson plans" (p. 50). According to Marzano and Kendall (2007), Bloom's Taxonomy has had a significant influence on educational theory and practice. However, Bloom's taxonomy has also received some criticism. For Furst (1994), one of the common criticisms is that the taxonomy oversimplifies "the nature of thought and its relationship to learning" (Marzano & Kendall, 2007, p. 8). Additionally, Airasian (1994) pointed out "Bloom's Taxonomy was ultimately replaced by Gagné's (1977) framework as the conceptual organizer for programmed instruction. Although Gagne's framework was less hierarchical than Bloom's Taxonomy, it was more easily translated into instructional practice" (Marzano and Kendall, 2007, p. 2).

Figure 1

Bloom's Taxonomy (Original and Revised)



Note: Taken from https://elearningbunch.wordpress.com/2013/02/20/revised-bloom-taxonomy/

Gagné's well-known theory of instruction "has been seen widespread application in classroom situations and even textbook creation" (Sternberg, 1983, p. 7). Gagné (1972) elucidated effective instruction needs to be designed to take full account of the differences within the domains of learning process. For Gagné (1972),

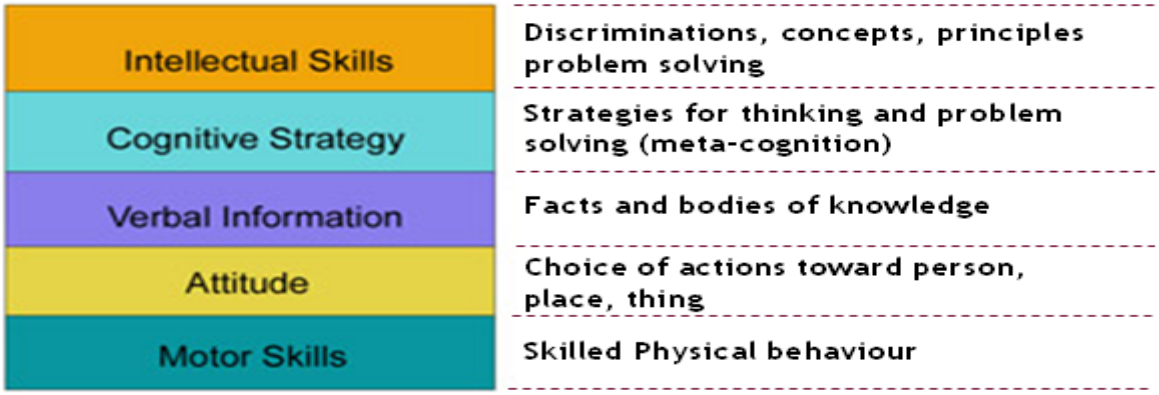
> …different types of learning outcomes call for different types of instruction. There is no best way to teach everything, and the conditions for learning that are appropriate to the type of outcomes we desire will affect our thinking about the design of learning activities and materials (p. 3).

In *Learning Outcomes and Their Effects: Useful Categories of Human Performance,* Gagné (1984), proposed five dimensions of learning outcomes or five categories of human performance. They are: "(1) intellectual skills (procedural knowledge), (2) verbal information (declarative knowledge), (3) cognitive strategies (executive control processes), (4) motor skills, and (5) attitude" (p. 377). Gagné (1984) emphasized "each of these categories may be seen to encompass a broad variety of human activities" (p. 377). According to Connerley and Pedersen (2005), Gagne's theory of instruction "examines the kinds of things people learn and how they

learn them" (p. 65). The theory also implied "there is no one best way to learn everything" (Connerley & Pedersen, 2005, p. 65). DeSimone, Werner, and Harris (2002) argued each of Gagné's five categories of human performance required "a different set of conditions to maximize learning, retention, and transfer" (Connerley & Pedersen, 2005, p. 65). Gagné (1980) believed the important aim of education is to cultivate students to become good thinkers and problems solvers. There are three types of human capabilities involved in good thinking and problem solving: "intellectual skills, verbal knowledge, and cognitive strategies" (Gagné, 1980, p. 86). With regard to how to become good thinkers and problem solvers, Gagné (1980) stressed having three capabilities (intellectual skills, verbal knowledge, and cognitive strategies) play a vital role.

Figure 2

Gagne's Five Categories of Learning Outcomes



Note: Taken from http://gramconsulting.com/2009/02/fun-with-learning-taxonomies/

De Bono, a "pioneering researcher in the field of creative thinking" (Stewart & Krivan, 2021, p. 1) made a great contribution to the understanding of the human brain, mind, and thinking. In his book, *The Use of Lateral Thinking*, de Bono (1967) coined the term *lateral thinking*, which was defined as "employing unconventional approaches to problem solving" (Stewart & Krivan, 2021, p. 1). His other book, *the Mechanism of Mind* (1969), is viewed as

"breaking the mould, and his elegant and simple technique of lateral thinking and provocation as devices to break thinking processes free from the shackles of backwards-forwards logical thinking" (Johnston, 2000, p. 152). De Bono (1985) defined thinking as "the operating skill with which intelligence acts upon experience" (p. 1). For de Bono (2002), argument is crude and highly ineffective during the process of exploring the subject. The Six Hat method, developed by de Bono (2002), is

> much more efficient and can reduce meeting times to one fifth or less. Each of the Six Hats indicated a mode of thinking which everyone follows in parallel. This makes use of changes in brain chemicals. The method is now widely in use in major corporations and also with children in schools. (de Bono, 2002, p. 10)

De Bono's Six Hats method design offers the Western thinking with a constructive idiom instead of adversarial argument for the first time (de Bono & Zimbalist, 1970). According to Sternberg (1986), de Bono put forward "a series of techniques to improve people's critical thinking". An instructional program on thinking skills that is called CoRT (Cognitive Research Trust) has shown to be effective (de Bono, 1985). The CoRT program "deliberately focuses on 'tools' that can be transferred" (de Bono, 1985, p. 5). One of the most famous of these is called the PMI, which is the acronym of three words: plus, minus, and interesting (Sternberg, 1986). For de Bono (1985), the PMI is a scanning tool, rather than a judgement tool. It is designed to avoid the point-to-point thinking. The person first looks at the Plus (good points), and then at the Minus (bad points), and finally at the Interesting (things that might be worth noting) (De Bono, 1985). According to Sternberg (1986), de Bono's opinion is

> usual for its stress on the evaluation of the interest as well as the positive and negative features of each solution. Getting people to think in this way encourages them to develop

> their ability to see both familiar and unfamiliar problems in novel and potentially interesting ways. (p. 13)

In summary, critical thinking has become an important and inseparable part of education. Given the complicated nature of CT, Beyer described CT as "one of the most abused terms in our thinking skills vocabulary. Generally, it means whatever its users stipulate it to mean" (p. 32, as cited in French & Rhoder, 1992, p. 184). The study of CT "is of particular interest because of its confluence of three traditions of thought --- the educational, the philosophical, and the psychological" (Sternberg, 1986, p. 3). These three academic strands have "developed different approaches to defining critical thinking that reflect their respective concerns" (Lai, 2011, p. 4). The philosophy field developed from discourse and argumentation, while psychology grew from experiment and research (Lewis & David, 1993). Educational theories benefit from their close tie or link with classroom observation and experience (Sternberg, 1986). For Sternberg (1986), education theorists or educators concentrate on the requirements of CT in the classroom; philosophers focus on the requirements of formal logical systems; and psychologists put their attention on the performance and behavior of humans in laboratory settings. Philosophers show their interest in

> …logical reasoning and perfections of thinking to decide what to believe and do, [while] psychologists are more concerned with the thinking process and how this process can help people make sense out of their experience by construction meaning and imposing structure. (Lewis & David, 1993, p. 132)

Educators have mixed philosophers' idea of specifying what people can do with psychologists' idea of specifying what people actually do, "with the nature and proportions of the mix less than clearly specified" (Sternberg, 1986, p. 7).

Table 1

Summary of CT Definitions and Related Concepts

Scholar	**Definition**
Socrates	Relate critical thinking to philosophy by his exploratory dialogues.
Plato	CT, as logic, is the tool that will help us find the answer or solution to our confusion and problems.
Aristotle	Invent a logical method, called the syllogism, to help knowers test out their ideas and even create new knowledge.
Dewey	The essence of critical thinking is suspended judgment; and the essence of this suspense is inquiry to determine the nature of the problem before proceeding to attempts at its solution. This, more than any other thing, transforms mere inference into tested inference, suggested conclusions into proof.
Ennis	Reasonable reflective thinking that is focused on deciding what is to believe or do.
McPeck	The propensity and skill to engage in an activity with reflective skepticism.
Lipman	A process that occasionally results in decisions or solutions, but the process is not to be defined solely by those occasional consequences.
Paul	Disciplined, self-directed thinking which exemplifies the perfections of thinking appropriate to a particular mode or domain of thinking
Facione	CT is skeptical without being cynical. It is open-minded without being wishy-washy. It is analytical without being nitpicky. Critical thinking can be decisive without being stubborn, evaluative without being judgmental, and forceful without being opinionated.
Sternberg	Propose three types of CT skills: metacomponents, performance components, and knowledge-acquisition components.
Halpern	The use of those cognitive skills or strategies that increase the probability of a desirable outcome. It is used to describe thinking that is purposeful, reasoned, and goal directed---the kind of thinking

	involved in solving problems, formulating inferences, calculating likelihood, and making decisions, when the thinker is using skills that are thoughtful and effective for the particular context and type of thinking task.
Bloom	Propose six cognitive skills, which are ranked on a hierarchy: knowledge, comprehension, application, analysis, synthesis and evaluation. The later revised Bloom's Taxonomy is: remembering, understanding, applying, analyzing, evaluating and creating.
Gagné	Propose five dimensions of learning outcomes or five categories of human performance: (1) intellectual skills (procedural knowledge), (2) verbal information (declarative knowledge), (3) cognitive strategies (executive control processes), (4) motor skills, and (5) attitude.
de Bono	Define thinking as the operating skill with which intelligence acts upon experience.

Assessment of Critical Thinking (CT)

Similar to the diversity of opinions on the definition of CT, there is no agreement among scholars about how to assess or measure critical thinking (Ku, 2009; Merrifield, 2018). To a large extent, the lack of consensus is the "result of the ongoing debate over how critical thinking should be conceptualized" (Merrifield, 2018, p. 58). Halpern (2001) also argued any assessment of CT needs to be based on an operational definition of CT. According to Ku (2009), "the conceptualization and assessment of critical thinking are interdependent issues that must be discussed together: how critical thinking is defined determines how it is best measured" (p. 71). Liu et al. (2014) contended the major challenge in developing an instrument for measuring CT is "to strike a balance between the assessment's authenticity and its psychometric quality" (p. 8). Although there are many challenges in assessing or measuring CT skills and dispositions, scholars developed a series of CT tests or instruments, some of which are very popular (Lai,

2011; Liu et al., 2014). For instance, the California Critical Thinking Skills Test (CCTST; Facione, 1990a) , the California Critical Thinking Disposition Inventory (CCTDI; Facione & Facione, 1997), the Cornell Critical Thinking Test (CCTT; Ennis, Millman, & Tomko, 2005), the Watson-Glaser Critical Thinking Appraisal (WGCTA; Watson & Glaser, 2008a), the Ennis-Weir Critical Thinking Essay Test (EWCTET; Ennis & Weir, 1985), and the Halpern Critical Thinking Assessment Using Everyday Situations (HCTAES; Halpern, 2010). These tests or instruments have been grouped into three categories: multiple-choice tests of CT, open-ended or constructed-response tests of CT, and a combination of the multiple-choice and open-ended or constructed-response test of CT (Ku, 2009; Liu, et al., 2014).

Multiple-choice tests of CT

Using the multiple-choice response format, the California Critical Thinking Skills Test (CCTST; Facione, 1990a) is based on the American Psychological Association (APA) Delphi definition of CT. The CCTST has been used in the USA and in nearly 70 countries worldwide with graduate student populations, executive level adult populations, and undergraduate students in all fields (Insight Assessment, 2021). It has been "characterized as the best commercially available CT skills assessment instrument" (Facione, 1991, p. 3), "specifically designed to assess selected, core critical thinking skills" (Facione, 1990b, p. 2). It is a discipline-neutral measure of reasoning skills (Insight Assessment, 2021). The CCTST has been translated into about 20 languages, including simplified and traditional Chinese, Arabic, Dutch, French Canadian, German, Hebrew, Italian, Japanese, Korean, Portuguese, Spanish, Turkish, Thai, and Vietnamese (Insight Assessment, 2021). The CCTST is composed of two delivery formats: online or paper/pencil, and the time frame to complete is 45 minutes. The CCTST contains 34 items

(vignette based), which are divided into 5 scales of CT. They are: analysis, evaluation, inference, induction, and deduction (Liu et al., 2014).

According to Facione (1991), the CCTST reports six scores: an overall score on CT cognitive or reasoning skills and five subscores of "analysis, evaluation, inference, deductive reasoning and inductive reasoning" (p. 2). The overall score indicates "overall strength using reasoning to form reflective judgments about what to believe and what to do" (Insight Assessment, 2021, p. 6). Analysis or analytical skills are employed to "identify assumptions, reasons, themes, and the evidence used in making arguments or offering explanations" (Insight Assessment, 2021, p. 7). Evaluation skills are utilized to evaluate "the credibility of the claims people make or post, and to assess the quality of the reasoning people display when they make arguments or give explanations" (Insight Assessment, 2021, p. 7). Inference skills "enable us to draw conclusions from reason, evidence, observations, experiences, or our values and beliefs" (Insight Assessment, 2021, p. 7). Induction or inductive reasoning counts on "estimating likely outcomes. Decision-making in contexts of uncertainty relies on inductive reasoning. Inductive decisions can be based on analogies, case studies, prior experience, statistical analysis, simulations, hypotheticals, trusted testimony…" (Insight Assessment, 2021, p. 8). Deduction skills are employed when "we determine the precise logical consequences of a given set of rules, conditions, beliefs, values, policies, principles, procedures, or terminology" (Insight Assessment, 2021, p. 8). According to Merrifield (2018), "the CCTST manual (Facione & Facione, 2002) reports 'high correlations with standardized tests of college-level preparedness in higher-order reasoning' (p. 46)" (p. 67), high predictive value to predict educational success and high internal consistency.

The California Critical Thinking Disposition Inventory (CCTDI), which "derives its conceptualization of the disposition toward CT from the APA Delphi Report" (Facione et al., 1995, p. 5), is the first tool to measure the disposition of CT. The CCTDI has two delivery formats: online or paper/pencil, lasting 30 minutes (Liu et al., 2014). It consists of 75 items using a six-point Likert scale, which has been grouped into seven aspects or scales: "Inquisitiveness, Open-mindedness, Systematicity, Analyticity, Truth-seeking, CT Self-confidence, and Maturity" (Facione et al., 1995, p. 5). These seven CCTDI dispositional scales are discipline neutral, rather than discipline specific (Facione et al., 1995). Each scale can be "readily interpreted within the liberal arts and science as well as professional discipline" (Facione et al., 1995, p. 6).

According to Facione et al. (1995), the Inquisitiveness on CCTDI scale assesses people's "intellectual curiosity and one's desire for learning even when the application of the knowledge is not readily apparent" (Facione et al., 1995, p. 6). The Open-mindedness scale "addresses being tolerant of divergent views and sensitive to the possibility of one's own bias" (Facione et al., 1995, p. 6). The Systematicity scale "measures being organized, orderly, focused, and diligent in inquiry" (Facione et al., 1995, p. 7). The Analyticity scale "targets prizing the application of reasoning and the use of evidence to resolve problems, anticipating potential conceptual or practical difficulties, and consistently being alert to the need to intervene" (Facione et al., 1995, p. 7). The Truth-seeking scale assesses "being eager to seek the best knowledge in a given context, courageous about asking questions, and honest and objective about pursuing inquiry even if the findings do not support one's self-interests or one's preconceived opinions" (Facione et al., 1995, p. 8). The CT Self-Confidence scale "measures the trust one places in one's own reasoning processes" (Facione et al., 1995, p. 8). The Maturity scale "targets the disposition to be judicious in one's decision making" (Facione et al., 1995, p. 9). For Facione and Facione (1992),

if a person gets a score of 30 and below on any scale, it indicates his or her given attribute or characteristic is consistently opposite or weak (Facione et al., 1995). The score of 40 shows "minimal endorsement on average" (Facione et al., 1995, p. 5). The score of 50 and above demonstrates "consistent endorsement or strength of the given characteristics" (Facione et al., 1995, p. 5).

As an example of utilizing multiple-choice response format, the Cornell Critical Thinking Test (CCTT; Ennis, Millman, & Tomko, 2005) consists of two forms: CCTT Level X and CCTT Level Z. Level X is suitable for assessing 4-14-grade students' CT skills, and Level Z is aimed at high school students, college students, graduate students, and other adults (Ennis et al., 2005). CCTT Level X includes 71 items, which is intended to be taken within 50 minutes (Ennis et al., 2005; Liu et al., 2014). CCTT Level Z involves 52 items, with the length of 50 minutes. (Ennis et al., 2005; Liu et al., 2014). "Each item on each test has three choices and one keyed answer" (Ennis et al., 2005, p. 1). According to Ennis et al. (2005), Level X measures such skills as induction, deduction, observation, credibility, and identification of assumptions. Level Z assesses the following skills: induction, deduction, observation, credibility, identification of assumptions, and meaning (Ennis et al., 2005). According to Merrifield (2018), the CCTT manual claims that the CCTT Level Z has substantial validity given standard conditions. Frisby (1992) challenged the CCTT's statement that "the heterogeneity of critical thinking naturally reduces the internal reliability of test items" (Merrifield, 2018, p. 65). Although this statement illustrates low internal consistency across all items, it is still likely to test the psychometric properties of each test section (Merrifield, 2018). Opposed to Frisby (1992), Possin (2008) gave the CCTT a positive comment that it is "well-constructed and has a well-documented history" (p. 218).

The Watson-Glaser Critical Thinking Appraisal (WGCTA; Watson & Glaser, 1980), the other example of multiple-choice response format, is the standard test for assessing CT skills and decision making among tertiary students, graduate students, and professionals. There are several types of forms: Form A, Form B, short form, and Watson-Glaser II. Form A and B belong to the standard form, including 80 items with a time limit of 40-60 minutes (Liu et al., 2014). The short form consists of 40 items, with 30 minutes, if timed, and Watson-Glaser II consists of 40 items with a time limit of 40 minutes (Liu et al., 2014). The WGCTA is made up of five subscales or tests: inferences, recognition of assumptions, deduction, interpretation, and evaluation of arguments (Liu et al., 2014). Inference requires "discrimination among degrees of validity of inferences drawn from given data" (Sternberg, 1986, p. 17). Recognition of assumptions demands "recognition of assertion" (Sternberg, 1986, p. 17). Deduction requires "determination of given statements of premises" (Sternberg, 1986, p. 17). Interpretation includes "weighing of evidence and deciding if generalizations or conclusions based on the given data are warranted" (Sternberg, 1986, p. 17). Evaluation of arguments demands "distinguishing between arguments that are strong and relevant and those that are weak or irrelevant to a particular question at issue" (Sternberg, 1986, p. 17). Accordingly, each of these five tests involves neural and controversial passages and scenarios taking place at work, in the classroom, or in the media. Finally, only the total score is reported (Watson & Glaser, 2008a, 2008b). According to Watson & Glaser (2010), "measures and provides interpretable subscores for three critical thinking skill domains that are both contemporary and business relevant including the ability to: (a) recognize assumptions, (b) evaluate arguments, and (c) draw conclusions" (Liu et al., 2014, p.7). For Hassan & Madhum (2007), the WGCTA is designed to measure "different, though interdependent, aspects of critical thinking through its subsets..." (pp. 363-364).

Open-ended tests of CT

The Ennis-Weir Critical Thinking Essay Test (EWCTET; Ennis & Weir, 1985) is a typical example of open-ended instrument of CT assessment. It is a "general test of critical thinking skills in an essay format" (Werner, 1991, p. 494), with a time limit of 40 minutes. The EWCTET is aimed at "grades 7 through college. [It is] also intended to be used as a teaching material" (Ennis, 1993, p. 183). There is only one delivery format: paper/pencil. The EWCTET measures CT in the context of argumentation and takes "the form of a letter to the editor of a fictional newspaper. In the letter, the writer makes a proposal and offers a variety of arguments in support of it" (Ennis & Weir, 1985, p. 2). The EWCTET is composed of a nine-paragraph essay/letter. This assessment measures the following CT skills or competences: "(a) getting the point, (b) seeing reasons and assumptions, (c) stating one's point, (d) offering good reasons, (e) seeing other possibilities, and (f) responding appropriately to and /or avoiding argument" (Liu et al., 2014, p. 6). Ku (2009) stressed the EWCTET mainly assesses test-takers' ability to "analyze and respond to arguments and debates in authentic situations" (p. 73). According to Werner (1991), the test manual offers "detailed criteria for analyzing and scoring each paragraph of the response, and the grader awards points for how well the response reflects these criteria" (p. 494). Ennis and Weir do not offer any information about the content validity of the EWCTET. They only claim that "the situation that the test presents to examinees is a common type of situation in which skill at appraising and formulating arguments is manifested, …" (Ennis & Weir, 1985, p. 3). Predictive and concurrent validity are not examined either, since "there is no established criterion for the ability the test was designed to measure" (Ennis & Weir, 1985, p. 3). The reliability estimates of EWCTET are .86 and .82, built upon interrater comparisons (Werner, 1991).

Combination of multiple-choice and open-ended tests of CT

The Halpern Critical Thinking Assessment Using Everyday Situations (HCTAES; Halpern, 2007) is a typical example that combines both multiple-choice and open-ended response forms into one instrument or test. It has two formats: forced choice (multiple choice, ranking, or rating of alternatives) and opened-ended (Liu et al., 2014). Compared to the EWCTET, the HCTAES is "less structured and presents more life-like situations" (Ku, 2009, p. 73). The HCTAES is a standardized, computer-based test, consisting of 25 everyday scenarios that test-takers have to analyze and critique (Butler et al., 2012). The test-taker first answers "open-ended questions (recall-based) related to the scenario and then responds to forced-choice questions (recognition-based) related to scenario. The scenarios involve thinking in a variety of domains including health, education, work, social policy, and others" (Butler et al., 2012, p. 113). The multiple-choice part of each question measures "recognition of correct responses from a list of alternatives, whereas the open-ended portion assesses strategic use of thinking skills as well as the ability to self-construct solutions without hints" (Ku, 2009, pp. 73-74). According to Halpern (2010), the HCTAES assesses five CT subskills: (a) verbal reasoning skills (e.g., recognizing the use of pervasive or misleading language), (b) argument and analysis skills (e.g., recognizing reasons and conclusions in arguments), (c) skills in thinking as hypothesis testing (e.g., understanding sample size, generalizations), (d) using likelihood and uncertainty (e.g., applying relevant principles of probability, base rates), and (e) decision-making and problem-solving skills (e.g., identifying the problem goal, generating and selecting solutions among alternatives) (Butler et al., 2012; Ku, 2009; Liu et al., 2014). For Halpern (2007), the open-ended part of the HCTAES "attempts to reveal more of the dispositional component of thinking…" (Ku, 2009, p. 74). According to Butler et al. (2012), many studies have established the reliability

and validity of the HCTAES. The validity of the HCTAES has been evaluated by many methodologies and the high face validity makes this test easily transfer to a general audience. Evidence of both criterion validity and content validity has been established as well (Butler et al., 2012).

In summary, there are a number of CT tests published. All these CT tests or instruments have imperfections. Put another way, "there is no single 'best' means of assessing critical thinking" (Halpern, 2003, p. 365). Each has its own strengths and weaknesses. According to Ennis (1993), most of these tests or instruments are multiple-choice tests. One of the advantages of multiple-choice tests is they are efficient and economical. While the disadvantage is they are not comprehensive (Ennis, 1993). Norris (1989) expressed a similar idea in his article, "Can We Test Validly for Critical Thinking?". Norris (1989) stressed multiple choice tests "can provide evidence on some fundamental aspects of critical thinking, such as the ability to make credibility judgments [or deductive reasoning]. [However, they] cannot test all of critical thinking" (p. 26). Liu et al (2014) offered a more exhaustive summary with respect to the strengths and weaknesses of current CT tests. The advantages of multiple-choice tests "lie in their objectivity, efficiency, high reliability, and low cost" (p. 10). The disadvantage of such tests is "[multiple-choice] items alone will not be able to meet the psychometric standards due to their low internal consistency, one type of reliability" (Liu et al., 2014, p. 11). In contrast, open-ended tests are more expensive to score and are time consuming, compared to multiple-choice tests (Ennis, 1993; Liu et al., 2014). Butler (2012), Halpern, (2010), and Ku (2009) call for multiple-item tests or a combination of multiple-choice and open-ended tests concerning CT assessments. According to Liu et al. (2014), "a combination of multiple item formats offers the potential for an authentic and psychometrically sound assessment" (p. 11).

Table 2

Summary of CT Instruments or Tests

Test	**Format**	**Length**	**Forms and Items**	**Themes/ Topics**
California Critical thinking Skills Test (CCTST)	Multiple-choice (MC)	45 min	34 items (vignette based)	The CCTST returns scores on the following scales: (a) analysis, (b) evaluation, (c) inferences, (d) deduction, (e) induction, and (f) overall reasoning skills
California Critical Thinking Disposition Inventory (CCTDI)	Selected-response (Likert scale--- extent to which students agree or disagree)	30 min	75 items (seven scales: 9-12 items per scale)	This test contains seven scales of CT: (a) truth-seeking, (b) open-mindedness, (c) analyticity, (d) systematicity, (e) confidence in reasoning, (f) inquisitiveness, and (g) maturity of judgment (Facione, Facione, & Sanchez, 1994)
Cornell Critical Thinking Test (CCTT)	MC	50 min (can also be administered untimed)	Level X: 71 items	Level X is intended for students in Grades 5-12+ and measures the following skills: (a) induction, (b) deduction, (c) credibility, and (d) identification of

				assumptions (The Critical Thinking Co., 2014)
			Level Z: 52 items	Level Z is intended for students in Grades 11-12+ and measures the following skills: (a) induction, (b) deduction, (c) credibility, (d) identification of assumptions, (e) semantics, (f) definition, and (g) prediction in planning experiments (The Critical Thinking Co., 2014)
Watson-Glaser Critical Thinking Appraisal tool (WGCTA)	MC	Standard: 04-60 min (Forms A and B) if timed	80 items	The WGCTA is composed of five tests: (a) inference, (b) recognition of assumptions, (c) deduction, (d) interpretation, and (e) evaluation of arguments. (Watson & Glaser, 2008a, 2008b)
		Short form: 30 min if timed	40 items	
		Watson-Glaser II: 40 min if timed	40 items	Measures and provides interpretable subscores for three critical thinking skill domains that are

				both contemporary and business relevant, including the ability to: (a) recognize assumptions, (b) evaluate arguments, and (c) draw conclusions (Watson & Glaser, 2010)
The Ennis-Weir Critical Thinking Essay Test (EWCTET)	Essay	40 min	Nine-paragraph essay/letter	This assessment measures the following areas of the critical thinking competence: (a) getting the point, (b) seeing reasons and assumptions, (c) stating one's point, (d) offering good reasons, (e) seeing other possibilities, and (f) responding appropriately to and/or avoiding argument weaknesses (Ennis & Weir, 1985)
The Halpern Critical Thinking Assessment Using Everyday Situations (HCTAES)	Forced choice (MC, ranking, or rating of alternatives) and open-ended	60-80 min, but test is untimed (Form S1) 20 min, but test is untimed (Form S2)	25 scenarios of everyday events (five per subcategory) S1: Both open-ended and forced choice items S2: All forced choice items	This test measures five critical thinking subskills: (a) verbal reasoning skills, (b) argument and analysis skills, (c) skills in thinking as hypothesis testing, (d) using likelihood and uncertainty, and (e)

				decision-making and problem-solving skills (Halper, 2010)

Note: This table is taken and adapted from Liu et al., 2014, pp. 5-7.

CT Studies in the West

The roots of the Euro-Western view of rational thought can be traced back to the great Greek philosophers: Socrates, Plato, and Aristotle (Thayer-Bacon, 2000). One American pragmatist, Dewey, made great contributions to the development of the Euro-Western CT theory (Thayer-Bacon, 2000). The CT movement has gained momentum at all levels of education in the globalized world (Paul, 1992). The epicenter of the CT movement "is in North America but its influence is being felt in Europe and beyond" (Paul, 1992, p. 33). According to Paul (1992), the CT movement has been manifested in many aspects, such as, a variety of academic publications and research projects, educational mandates, curriculum articulations, and school restructuring. Soeherman (2010) claimed a large number of empirical studies have been conducted on critical thinking since the 1980s. This section of the review is mainly targeted on empirical studies on CT, and these studies can be divided into three themes: cultivating critical thinking, critical thinking and culture, and critical thinking instruction.

Cultivating Critical Thinking

Researchers and scholars have reached a consensus that CT consists of skills and dispositions (McPeck, 1981; Ennis, 1987; Siegel, 1988; Lipman, 1988; Facione, 1990a; Paul, 1992; Halpern, 1998; Bailin et al., 1999; Soeherman, 2010). Critical thinking "is not inherent in humans from birth" (Hidayati & Sinaga, 2019), but can be cultivated after birth (Soeherman, 2010). Different aspects of CT have been studied and implications for cultivating CT have been presented since Dewey introduced the term of critical thinking in 1910. Some researchers,

scholars and educators "correlate CT performance with either cognitive ability or personality disposition separately" (Clifford et al., 2004, p. 170). Other researchers concentrate on both CT skills and dispositions. According to Behar-Horenstein and Niu's (2011) literature review of empirical studies (from the year of 1994 to 2009) on CT, the standardized tests, such as CCTST and WGCTA, have been used more frequently than the CCTT. This subsection is limited to those empirical studies that discussed promoting students' CT skills through instructional intervention and assessed students' CT skills via employing one of the three CT tests (CCTST, WGCTA, or CCTT).

Yang et al. (2005) investigated the efficacy of utilizing Socratic questioning to improve students' CT skills in asynchronous discussion forums (ADF) in distance learning courses. Yang et al. employed an experimental research design. The independent variable was the teaching and modeling of Socratic questioning within the ADF. The dependent variables included students' CT skills, assessed by CCTST, and class discussions on the ADF. The research results showed: "(a) teaching and modeling of Socratic questioning helped students demonstrate a higher level of CT skills, and (b) students maintained their CT skills after exposure to and modeling of Socratic questioning in the ADF" (Yang et al., 2005, p. 163). In Yang et al.'s (2005) study, the CCTST was used because it is "aimed at college/graduate students and adult professionals. The test reports an overall score on one's CT skills and five subscales: analysis, evaluation, inference, deductive reasoning, and inductive reasoning" (Yang et al., 2005, p. 166). The findings indicated structured Web-Based Bulletin Boards (WBBs) improved students' CT skills and attitudes toward learning (Yang et al., 2005).

Quitadamo and Kurtz (2007) examined the effects of writing on CT performance in general education biology courses, using a quasi-experimental pretest/posttest control group

design. The pretest/ posttest control group design was used to "minimize internal and external validity threats and maximize the ability to determine the effects of writing on student critical thinking performance" (Quitadamo & Kurtz, 2007, p. 146). In this study, CCTST was utilized because "only the CCTST measures cognitive and meta-cognitive skills associated with critical thinking… and has been evaluated for validity and reliability for measuring critical thinking at the college level" (Quitadamo & Kurtz, 2007, p. 145). Quitadamo and Kurtz (2007) collected students' demographic information, such as gender, ethnicity, age, class standing, academic term, and time of day, as well as raw scores for a total CT skill, and raw scores for analysis, inference, and evaluation, in order to compare CT "performance of students who experienced a laboratory writing treatment with those who experienced traditional quiz-based laboratory in a general education biology course" (p. 140). Several statistical tests were conducted to examine the effects of writing on CT performance. An analysis of co-variance (ANCOVA) test was used to investigate whether there are differences in overall CT performance between the writing and non-writing group. Then, a multivariate analysis of covariance (MANCOVA) test was employed to examine "changes in particular component critical thinking skills (analysis, inference, and evaluation)…" (Quitadamo & Kurtz, 2007, p. 146). The research results demonstrated that the writing group improved their CT skills, while the nonwriting group did not. Analysis and inference skills "increased significantly in the writing group but not the nonwriting group. Writing students also showed greater gains in evaluation skills; however, these are not significant" (Quitadamo & Kurtz, 2007, p. 140). Besides writing, other factors, such as prior CT skill and the teacher, also influenced students' CT performance (Quitadamo & Kurtz, 2007).

Scott et al. (1998) examined changes in CT skills between their entry and near the end of the third year for 68 medical school students, and evaluated "the predictive ability of a test of

critical thinking skills, and assessed the concurrent validity of clerkship components and final grade" (p. 14). The WGCTA was utilized to determine whether the changes do exist. The research findings indicated medical students' CT skills "improved modestly but significantly from entry to medical school to near the end of year 3" (Scott et al., 1998, p. 14). Scott et al. (1998) "attributed the gains [or changes] to the medical education process" (Behar-Horenstein & Niu, 2011, p. 29). According to Scott et al. (1998),

> The ability of a critical thinking test to predict clerkship performance was limited; the correlation between WGCTA total score at entry and the components and final grade of five major clerkships ranged from near 0 to 0.34. The concurrent validity of clerkship components and final grade was also limited; correlations with WGCTA total score near the end of year 3 ranged between 0.08 and 0.49. The correlation between WGCTA total score and United States Medical Licensing Examination Step 2 was higher at year 3 than at medical school entry. (p. 14)

Finally, Scott et al. (1998) cautioned that the single test scores of CT cannot provide all-round information or value in predicting who will succeed in clinical clerkships.

Iwaoka et al. (2010) examined whether there were gains by students in Food Science and Human Nutrition Courses after carrying out problem-based learning activities. The aim of this longitudinal study was to determine whether "the conventional CCTT, Level Z (Ennis & others, 1985) administered as a pretest and posttest could measure any gains in CT among students in a Food Science Experimental Foods Course offered every fall semester during the period 2001-2008" (Iwaoka et al., 2010, p. 69). In this study, the CCTT test scores of the students in an Experimental Foods class reported over an eight years period was similar to other students "from a wide variety of classes at different universities as reported by Ennis and others (1985)"

(Iwaoka et al., 2010, p. 71). The statistical analysis showed there were significant gains in deduction and assumption (*p* values 0.024 and 0.045 in 2002, and 0.020 and 0.036 in 2004, respectively) in the year of 2002 and 2004, but not in the other aspects, such as, meaning, observation/credibility, and induction (Iwaoka et al., 2010). The research results also indicated "gains in CT scores occurred both in students who had initially lower as well as higher CT scores" (Iwaoka et al., 2010, p. 74). According to Iwaoka et al. (2010), the CCTT is only one way to document gains in general CT skills of students as developed by the Problem Based Learning (PBL) method. If the CCTT, a more traditional CT test, cannot assess gains in all CT skills area, "the collective evidence from student journal feedback and other PBL exercises might be more useful and meaningful, and accurate in assessing development of CT skills than depending only on scores obtained from administering a pre- and post-CCTT multiple-choice test" (Iwaoka et al., 2010, p. 74).

Some researchers, scholars and educators showed great interest in CT dispositions, especially the relationships or associations between CT disposition, academic achievement, gender, age, and educational experiences. Bers et al. (1996), investigated the disposition level of 224 community college students via the CCTDI instrument. Bers et al. (1996) explored whether there were statistically significant differences in subscale or total scales of CCTDI of community college students, based on their gender, age, or prior education. Secondly, whether students' CT disposition changed over one semester was examined via pre- and post-tests. Thirdly, whether there were relationships between course enrollment (CT or non-CT section), gender or age, and changes in the CT disposition over one semester were investigated. Finally, whether there were relationships between academic achievement and the CT disposition were explored. CCTDI offered mean scores of pre-and post-test for the 7 subscales (truth-seeking, open-mindedness,

inquisitiveness, systematicity, maturity, self-confidence, and analyticity) and total score of the disposition, "as well as the percent of students testing 40 or more on a subscale or 281 and above on the total score" (Bers et al., 1996, p. 204). The research results showed females' CT disposition was much stronger than males, older students had more CT disposition than younger students, and "students with more education had higher scores than those with less education" (Bers et al., 1996, p. 205). There were statistically significant differences in "mean scores for male and female, for older and younger students, and for those with varying levels of prior education" (Bers et al., 1996, p. 204). After re-administering the CCTDI at the end of the semester, the research findings indicated that mean subscale and total score changes were small. After conducting a chi-square test of relationship between the "nominal categories and enrollment in a critical or non-critical thinking course section" (Bers et al., 1996, p. 209), there was no statistically significant difference in "whether or not students' strengthened, weakened, or remained the same in their CT dispositions that were associated to the enrollment in the critical or non-critical thinking classes" (Bers et al., 1996, p. 209). The results finally demonstrated there were "no significant correlations for any subscale or total score and percent of fall courses successfully completed" (Bers et al., 1996, p. 212). However, there were significant correlations between the subscale and total scores, and term or cumulative GAPs. All these correlations were positive, indicating "a stronger subscale or total disposition was associated with a higher term or cumulative GPA" (Bers et al., 1996, p. 212). Overall, students showed stronger CT disposition on Truth-seeking, Open-mindedness, Inquisitiveness, Maturity, and Analyticity (Bers et al., 1996). Bers et al. (1996) concluded there were many factors which contributed to academic success and CT disposition was one of them.

Walsh and Hardy (1999) explored dispositional differences among university majors and across gender, using an exploratory study. In this study, students' disposition was measured by Facione's CCTDI. There were 334 participants, and they were third-year baccalaureate undergraduates, including 121 males and 213 females. They came from 6 majors belonging to either practice disciplines (nursing, education, business), or nonpractice disciplines (English, history, psychology). Walsh and Hardy (1999) conducted a "2×6 factorial MANCOVA between gender (male and female) and major (English, history, psychology, business, education, and nursing) on the overall score and the seven subscale scores of the CCTDI, using GPA as a covariate" (p. 153). The MANCOVA indicated the mean score of the English majors was the highest, followed by psychology, nursing, history, education, and business. There was a significate main effect for the student's major on the overall score and the subscale scores (Wilks F [40, 1372] = 3.16, $p \leq .001$). There were no interactions of major and gender (Walsh & Hardy, 1999). Walsh and Hardy (1999) divided the six majors into practice and nonpractice, and a 2×2 factorial MANCOVA between gender and major was calculated on the overall score and the CCTDI subscale scores, taking GPA as a covariate. The MANCOVA showed there were "differences for major on the overall CCTDI score and subscale scores (Wilks F = 3.5, [8, 32], p = .01). There were no interactions between practice or nonpractice major and gender on the CCTDI subscales" (p. 153). Walsh and Hardy (1999) summarized the research findings of their study, and their findings supported the idea that "there are differences among college students in different majors in the disposition toward critical thinking" (p. 154). According to Walsh and Hardy (1999), "the nature of the differences among majors in disposition toward critical thinking needs further exploration, [and] additional demographic questions concerning the eventual career

goal of participants may aid in the discrimination of difference between practice and nonpractice majors" (p. 155).

Redding (2001) conducted a descriptive longitudinal study to examine the relationship between freshman nursing students' CT disposition assessed by the CCTDI, "their composite scores on the American College Testing (ACT) examination, their high school percentile rank (HSPR), and their cumulative grade point average (GPA) in nursing education by the end of the sophomore year" (p. 125). This study also explored the difference in CT disposition between students who persist in nursing education and those who leave. The results of this study indicated there was a positive relationship between GPA and a combination of CT disposition, ACT and HSPR (Redding, 2001). However, there was no association between CT disposition and GPA, which did not support the previous research result: there was a positive relationship between CT disposition and academic performance in nursing education. According to Redding (2001), the main reason for this difference was the earlier research involved only one institution, but her research included seven. After examining the institutions individually, instead of collectively, "a relationship was found between CT disposition and GPA in two institutions, in one case positive and one negative" (Redding, 2001, p. 125). Another possible reason was that the different characteristics of the study sample may lead to unique research results. Contrary to the previous studies, mostly targeting at seniors, Redding's study was focused on the developmental state of students from freshmen year to their sophomore year. Redding argued senior students demonstrated "the multiple level of intellectual development and elementary critical thinking skills" (p. 125), compared to the novice students: the freshman and sophomore. According to Redding, only "about one third of the variance in GPA is explained by a combination of CCTDI score, HSPR, and ACT score, and only a small portion of that variance is explained by CCTDI

score" (p. 126). Therefore, academic achievement can be influenced by other factors, such as curricular, co-curricular, and extracurricular events. These factors need to be explored in the future studies, as well (Redding, 2001). In addition, the mean CCTDI score of those who persisted was slightly lower than the mean of the entire group. The group who left school had a narrower range of scores, including the highest score of the entire group. The mean CCTDI score of those who left was higher than the mean of the entire group. It indicated that those who left school had a strong CT disposition. Redding (2001) suggested further research needs to be done to figure out why students with strong CT dispositions tend to drop out of school.

Researchers and scholars have begun to focus on both CT skills and dispositions in order to get a comprehensive understanding of students' CT. Colucciello (1997) undertook "a cross-sectional, descriptive, comparative, and correlational study" (p. 236) to determine whether there was significant difference in students at different academic levels, and whether there was a relationship between nursing students' CT skills and CT dispositions. In Colucciello's study, two research instruments were utilized. One was the CCTST Form A, and the other was CCTDI. According to Colucciello (1997), the CCTST and CCTDI together provided "the mechanism to assess reasoning ability and attitudes essential for professional nurses to possess" (p. 239). A convenience sample of 328 nursing students participated in this study. They represented five academic levels: Sophomore II ($n = 94$), Junior I ($n = 65$), Junior II ($n = 64$), Senior I ($n = 59$), and Senior II ($n = 46$). The research findings showed there was a statistically significant difference in CT skills among the academic levels ($F = 6.63$, $p < .05$). After running the analysis of variance (ANOVA), Junior I students had the highest critical thinking mean scores, and the sophomore II students attained the lowest scores, "reflecting a significant difference ($p \leq .05$)" (Colucciello, 1997, p. 236). An ANOVA also indicated there was a significant difference in the

total mean scores of CT disposition "between students at the junior I and senior I and II levels and those at the sophomore II level ($p \leq .0000$)" (Colucciello, 1997, p. 236). The research results also showed that the truth-seeking disposition among students at all levels was weak. Pearson Product Moment Correlation Coefficient "yielded a significant positive relationship between critical thinking skills and critical thinking dispositions ($p \leq .01$)" (Colucciello, 1997, p. 236).

A similar non-experimental study on nursing students' CT skills and dispositions was undertaken by Profetto-McGrath (2003) to explore the CT skills and CT dispositions of students in a 4-year baccalaureate program at a university in Canada. Profetto-McGrath employed a cross-sectional design with a volunteer sample. The research instruments included a background/demographic questionnaire, the CCTST, and the CCTDI. The reliability of the CCTST and CCTDI was established "using the Kuder Richardson 20 and Cronbach Alpha respectively" (Profetto-McGrath, 2003, p. 569). According to Profetto-McGrath, the CCTST and CCTDI were the up-to-date instruments "available with acceptable levels of reliability and validity" (p. 572). These two instruments were easy to administer and "deemed effective in measuring CTS [critical thinking skills] and CTD [critical thinking dispositions] of baccalaureate nursing students" (Profetto-McGrath, 2003, p. 572). A total of 228 nursing students have participated in the survey, including 38 freshmen, 53 sophomores, 57 juniors, and 80 seniors. The research results indicated the students' level of CT skills fell into the adequate category. The mean scores from the CCTST increased from Years 1 through 4 with the exception of Year 3. There was no statistically significant difference among the four student groups ($F = 1.234$, $p = 0.295$). Meanwhile, there were no statistically significant differences in critical thinking disposition scores. However, students' CT disposition scores "differed significantly on the systematicity subscale [$F = 0.822$, $p = 0.483$]" (Profetto-McGrath, 2003, p. 573). There was a

statistically significant relationship between students' overall CT skills and CT dispositions scores (Profetto-McGrath, 2003). Profetto-McGrath concluded that 38% of students had adequate levels of CT skills and 85.5% had adequate levels of CT dispositions. According to Profetto-McGrath, it was imperative to improve nursing students' CT skills and dispositions.

Critical Thinking and Culture

The concepts of CT have been developed mainly in the Western countries. Some scholars, such as Atkinson (1997), Pennycook (1996), and Canagarajah (2002), argued CT is a "unique western product and incompatible with Asian collectivist traditions" (Tian & Low, 2011, p. 61). Others, like Paton (2005), claimed CT is not only a Western product, but also "belongs to all existing successful cultures in the world" (Tian & Low, 2011, p. 62). According to Atkinson (1997), CT is a culture-based concept and CT is regarded as a social practice in which "an individual is automatically immersed by virtue of being raised in a particular cultural milieu…" (p. 73). For Atkinson (1997), CT is "cultural thinking" (p. 89). Questions have been posed, such as what CT means in the non-Western regions, for instance, Asia, the Middle East and Africa, or whether culture affects the development of CT in those contexts. Some research and studies have been undertaken to identify the impact of culture on CT development. This subsection focuses on the studies regarding CT and culture from three regions: Asia, the Middle East, and Africa.

Western researchers and scholars performed studies on CT and Chinese students from the perspective of culture. They discovered the Chinese students in the Western universities were inclined to avoid debating in the classroom. The arguments these students made were weak when they were presented with a statement containing fallacies, which indicated a lack of CT (Atkinson, 1997; Cortazzi & Jin, 1997). In studies of why Chinese students are not proficient at CT, some researchers focused on the cultural constraints, and claimed that Chinese students are

weak in CT because CT is not compatible with the Chinese culture, which advocates collectivism, conformity, and respecting authority (Atkinson, 1997; Canagarajah, 2002; Hu, 2002; Liu, 1998; Pennycook, 1996; Wan, 1999). Some research offered evidence that the cultural background of Chinese students affected their reasoning (Tian & Low, 2011). For instance, Durkin (2008a) conducted qualitative research on the East Asian master students' perception of critical argumentation in U.K. universities. Participants consisted of 23 Chinese, 7 Taiwanese, 7 Thai, 3 Japanese, 1 Korean, and 1 Indonesian. In the same year, Durkin (2008b) undertook qualitative research, with the aim at exploring "the adaptation experiences of East Asian master students in the UK in dealing with western academic norms of critical thinking and debate" (Durkin, 2008b, p. 15). Forty-one East Asian students were interviewed, including 23 Chinese, 6 Taiwanese, 3 Japanese, 7 Thai, and 2 Indonesian students. These two studies conducted by Durkin (2008a, 2008b) achieved similar results: East Asian students developed their own way of arguing that "blend[ed] western CT (allowing the possibility of alternative) views and Asian elements (conciliatory reasoning, avoiding direct challenge and disagreement)" (Tian & Low, 2011, p. 66). For Durkin (2008b), East Asian students "opted for a 'middle way' which synergizes their own cultural approach to critical thinking with those aspects of western-style critical thinking and debate that are culturally acceptable to them" (p. 15). However, there is a controversy about the view that Chinese students' lack of CT resulted from the Chinese culture. Paton (2005) especially argued against this idea in his article "Is Critical Analysis Foreign to Chinese Students?". Paton (2005) pointed out the key problem of Chinese students' poor CT performance stems from lacking special training in CT, instead of from culture barriers.

Other researchers focused less on the cultural dimension and stressed that Chinese culture was not the only reason Chinese students performed worse in CT. Rather, these researchers

concentrated more on psychological dimensions. According to Tian and Low (2011), there was limited psychometric research on Chinese students' CT, and they were mostly restricted to testing the CT dispositions of students by using the California Critical Thinking Dispositions Inventory (CCTDI). For example, McBride et al. (2002) compared the dispositions towards CT between 218 American students and 234 Chinese students using the CCTDI instrument. All participants were junior or senior volunteers from physical education secondary teaching methods classes. Cronbach alpha coefficients were calculated for the seven subscales and the total score to determine internal consistency or reliability of the CCTDI. Because of the low reliability of the Chinese version of the CCTDI, three subscales (Analyticity, Systemacity and Open-mindedness) were "dropped from further analysis" for Chinese students (McBride et al., 2002, p. 136). A 2 ×2 MANOVA was calculated in order to examine "whether culture and gender influence CT dispositions on the seven subscales. Follow-up univariate tests were conducted for significant main effects" (McBride et al., 2002, p. 136). The research findings indicated the American students outscored the Chinese students on maturity and self-confidence, while the two groups achieved similar scores on truth-seeking and inquisitiveness. According to McBride et al, two factors attributed to the difference between American and Chinese students on maturity and self-confidence: one was the United States embraces an individualist culture and China follows collectivistic tradition; the other was different teaching and learning styles in China and the United States (Tian & Low, 2011). Owing to the purposive sampling, McBride et al. admitted that their findings cannot be generalized to the larger Chinese population.

Tiwari et al. (2003) conducted cross-sectional and descriptive research to compare the CT dispositions between Hong Kong Chinese and Australian nursing students. There were 384 participants in this study. They came from two universities: one was in Hong Kong, and the other

was in Australia. The English CCTDI was administered to the Australian students, and the Chinese CCTDI was delivered to the Hong Kong students (Tiwari et al., 2003). The research results indicated the Hong Kong Chinese nursing students showed a negative disposition towards CT, with a total mean score of 268.36, based on the criterion that a total score of 280 or higher showed a positive disposition towards CT (Facione & Facione, 1997). The total mean score of the Australian students was 287.73, which indicated they had a positive disposition towards CT. The one-way analysis of variance (ANOVA) was calculated to examine whether "there were differences between the Hong Kong Chinese and Australian students in total and subscale mean scores" (Tiwari et al., 2003, p. 302). The ANOVA showed there were statistically significant differences between the Hong Kong Chinese and Australian students for the subscale scores of Truthseeking, Open-mindedness, Inquisitiveness, Systematicity, Maturity, and the CCTDI total scores (Tiwari et al., 2003). The similarities between the two groups of students lied in the fact that the subscale scores of Truthseeking and Systematicity were both low. Tiwari et al. attributed the similarities between Chinese and Australian students to the institutional traditions and teaching and learning practices in the two countries (Tian & Low, 2011). For Tiwari et al., the low scores of Openmindedness and Maturity among Hong Kong Chinese students "may be a product of cultural norms rather than proof of their diminished intellectual capacity" (p. 305).

Researchers and scholars showed interest in the CT movement not only in Asian countries, but also in the countries of the Middle East. For Egege and Kutieleh (2004), CT is "valued differently in various cultures" (Bali, 2015, p. 317). According to Bali (2015), the ideas and practices of CT do exist in the Egyptian Islam culture. Similar to Paton's (2005) idea, Bali (2015) also believed CT is not "an exclusively Western notion, but one ingrained in Islamic scholarship and informal Egyptian culture" (p. 318). Bali made use of the evidence from the

American University in Cairo (AUC) to explore the cultural issues concerning CT development in the Egyptian context, because of his in-depth teaching experience and research there. School curricula in the Arab region mainly "encourage submission, obedience, subordination and compliance, rather than free critical thinking" (UNDP 2003, iv, as cited in Bali, 2015, p. 319). For Nurullah (2006), Arab and Muslim societies are "less likely to apply critical and questioning approaches to Islamic scholarship" (as cited in Bali, 2015, p. 319). Similar to Chinese education, Egyptian education also stresses memorizing and avoiding criticism. Bali admitted a lack of CT in Egyptian formal educational contexts does create challenges for teachers.

Bali (2013) conducted qualitative research on CT development at the American University in Cairo (AUC) in an Arab/Muslim country --- Egypt. Bali (2013) carried out semi-structured interviews with students and faculty at the AUC, "integrating contextual factors within AUC itself, and the wider Egyptian and global sociopolitical environment to understand how different[ly] students develop CT" (p. 18). The use of triangulation/crystallization, thick description, and prolonged engagement with the field of study contributed to the authenticity of this qualitative study (Bali, 2013). One theme that emerged from student interviews was diversity and it exerted great influence on students' CT development. Diversity here referred to "exposure to diverse disciplines within AUC's liberal arts education, exposure to diverse people/cultures at university, exposure to diverse professors and ways of teaching, diverse readings, diverse media sources, or exposure to diverse viewpoints in a classroom or friendly discussion" (Bali, 2013, p. 141). Bali categorized the factors that influenced students' CT into three types:

1. External [factors] to AUC altogether such as high school, parents and friends;

2. Extra-curricular factors within (but some also before/outside) AUC such as intercultural interaction and experiential learning situated in authentic contexts [note: some intercultural/authentic experiences were also found in academic courses];

3. Academic/curricular factors such as the core curriculum, rhetoric and composition courses and particular professors. (pp. 141-142)

Bali finally summarized exposure to diverse worldviews and interacting with these different views was an important and effective way to develop students' CT.

Researchers and scholars also displayed much interest in CT development in Africa, especially, in South Africa. Statistics South Africa (2001) reported that South Africa consists of many diverse cultures (Grosser & Lombard, 2008). The majority (75%) of the population is black Africans, including Xhosa, Zulu, Ndebele, Venda, Sotho and Tswana (Grosser & Lombard, 2008). Although urbanization brings about political, economic, and social changes, the African people still embrace their traditional African culture and lifestyle, in which "the philosophy 'ubuntu' is emphasized" (Grosser & Lombard, 2008, p. 1368). Ubuntu stands for "collective personhood and collective morality" (Grosser & Lombard, 2008, p. 1368), which may be elaborated by the Xhosa proverb, "umuntu ngumuntu ngabantu" (I am because we are) (Mbigi, 1997, p. 2, as cited in Grosser & Lombard, 2008, p. 1368). According to Mbigi (1997), ubuntu stresses collective solidarity, rather than individual self-sufficiency (as cited in Grosser & Lombard, 2008). Ubuntu plays an important role in building harmony and reconciliation in South Africa (Grosser & Lombard, 2008). According to Mbigi and Maree (1995), a derivation from Ubuntu is "the Afrocentric view. The landmark of the Afrocentric philosophy is about being a good community member" (as cited in Grosser & Lombard, 2008, p. 1368). African traditional

education focuses on cultivating young people and turning them into "collective social, economic, spiritual and political stewardship" (Grosser & Lombard, 2008, p. 1368).

Grosser and Lombard (2008) undertook exploratory research to investigate the CT "abilities of a mixed cultural group of prospective teachers, and to provide insight into the relationship between critical thinking abilities of the groups of prospective teachers and their various cultural environments" (p. 1369). One hundred and fourteen first-year students participated in this study, including 46 Black African students (representing African culture, Sesotho speaking) and 68 White Afrikaans (representing the Western culture, English speaking). The Watson -Glaser Critical Thinking Appraisal (WGCTA, UK version) was used to measure CT skills in this study. The research results indicated the White Afrikaans students representing the Western culture performed better than those Black African students representing African culture. Grosser and Lombard (2008) argued that

> the response to the appraisal that cognitive process in the Western culture tend to focus on the utilization of critical thinking abilities, in contrast to the African culture where there is an absence of utilizing cognitive processes in which critical thinking plays an important role. (p. 1371)

Grosser and Lombard (2008) claimed the South African classrooms are still teacher-centered without offering students opportunities to think critically. Therefore, it is imperative for African teachers to implement key teaching strategies and "create a learning environment which encourage the development of thinking strategies…" (p. 1372).

Critical Thinking Instruction

Researchers and scholars have reached an agreement on the importance of teaching CT in higher education (Behar-Horenstein & Niu, 2011). However, there is a debate on how CT can be

promoted through instruction (Tsui, 2002). Some scholars, McPeck (1981), for example, claimed CT instructions become effective "when it is integrated in teaching subject specific knowledge and skills" (Behar-Horenstein & Niu, 2011, p. 25). Other scholars, such as Ennis (1989), argued CT skills are "a generalized subset of skills that should be taught separately" (Behar-Horenstein & Niu, 2011, p. 25).

A series of studies were carried out to examine the effectiveness of different CT teaching strategies and interventions, aiming at improving students' CT. Behar-Horenstein and Niu (2011) grouped CT instructional interventions into two categories: "programmatic, pertaining to the whole curriculum of a degree program, and instructional, pertaining to specific instructional approaches" (p. 29). The characteristics of the studies utilizing programmatic approach are either to examine students' CT at the beginning or at the end of the program, or to undertake a pretest and posttest over a period of time to investigate the difference (Behar-Horenstein & Niu, 2011). Bartlett and Cox (2002) conducted an exploratory study to examine the change in CT dispositions and skills of physical therapy students over the middle year of the program, using a one-group repeated measures design. Twenty-eight physical therapy students (26 females and 2 males) participated in this research over a one-year period. The CCTDI and CCTST instruments were used before and after the academic year, as well as after their clinical placements in this study.

> A one way repeated measures analysis of variance [ANOVA] and Tukey post hoc tests were conducted on the subscales and total scores of the CCTDI and the CCTST. Multiple regression was used to determine the correlations between the descriptive variables and change in total CCTDI and CCTST. (Bartlett & Cox, 2002, p. 65)

The results from the ANOVA and Tukey post hoc tests demonstrated statistically significant gains in all subscale scores and the total scores for both the CCTDI and CCTST. The greatest change on the CCTDI occurred in the subscales of truth-seeking ($p < 0.001$) and self-confidence ($p < 0.001$), and "the greatest change in the total score occurred in the academic portion of the program" (Bartlett & Cox, 2002, p. 66). The greatest improvement on the CCTST was found in the deductive dimension ($p < 0.001$), and "the greatest change in the total score occurred in the clinical portion of the program" (Bartlett & Cox, 2002, p. 66). According to Bartlett and Cox (2002), changes in CT skills were more moderate than changes in CT dispositions. The descriptive variables, such as, age, gender, years of completed postsecondary education, highest level of education completed, were not correlated with change in the total score of CCTDI. There was a negative relationship between age and change on CCTST. Bartlett and Cox (2002) claimed the CCTDI can be utilized to "monitor individual change and to evaluate the effectiveness of a program in enhancing CT ability" (p. 64). Behar-Horenstein and Niu (2011) argued Bartlett and Cox (2002) attributed the significant gains to their medical education process.

Magnussen et al. (2000) undertook a study to investigate the impact of the inquiry-based learning (IBL) on the development of CT. Compared to problem-based learning (PBL), IBL is "more holistic and flexible" (Magnussen et al., 2000, p. 360), despite some similarities between the two methodologies. Magnussen et al. (2000) pointed out two major differences between PBL and IBL: "(1) PBL framed learning issues in the context of client problems (IBL focused discussion to include the client's strengths), and (2) in PBL, classroom sessions were always conducted as tutorials" (p. 361). The WGCTA was used as the instrument in the four-year period program. Form A was administered in the first semester as a pretest and Form B during the last semester of the program as a posttest. Two hundred and twenty-eight students participated at the

beginning of the program, and 257 students at the end of the program. Among them, 150 students had "paired-scores collected from testing the same students on admission and at graduation. There was no statistical difference between these scores" (Magnussen et al., 2000, p. 363). The participants were also placed into three groups of approximately equal size: the high level (48 students), the medium level (53 students), and the low level (49 students). After conducting the paired *t*-test, a statistically significant difference was found in the low level group between pretest and posttest, with an increase in mean score of 2.23 ($t = 2.76$, $p < .01$). There was no difference in the medium level group. A significant decrease was shown for the high level group, with a mean score of 4.79 ($t = - 4.81$, $p < .001$) (Magnussen et al., 2000). Magnussen et al. summarized there were no significant differences between scores at the beginning of the program and at graduation. Some students' scores were even lower in the posttest. According to Magnussen et al., the possible reason for this decline in the final semester before graduation was because students lost the motivation to achieve higher performances, which was consistent with the research findings of Vaughan-Wrobel et al. (1997). Magnussen et al. concluded that the low level students in the pretest benefited the most from IBL, compared to the medium and high level students. Further research needs to be conducted to validate the research findings of this study (Magnussen et al., 2000).

> Ennis (1989) proposed four types of instructional approaches, and they are:
>
> critical thinking can be taught 'separately' (the "general" approach), be infused in instruction in existing subject-matter areas (the "infusion" approach), result from a student's immersion in the subject matter (the "immersion" approach), or … be taught as a combination of the general approach with infusion or immersion. (Ennis, 1989, p. 4)

Behar-Horenstein and Niu (2011) listed the specific instructional methods investigated, which include concept mapping, scenario-based course exercises, active learning techniques, problem-based learning, inquiry-based learning, question approach, guided practice, computer-assisted instruction, structured web-based bulletin boards, and online instruction.

Wheeler and Collins (2003) carried out a quasi-experimental study to investigate the effect of concept mapping in developing CT skills of the baccalaureate nursing students, using a pretest and posttest design. Wheeler and Collins used a convenience sampling strategy, which included 76 volunteer participants. The participants were "randomly assigned to experimental group (n = 44) and control group (n = 32)" (Wheeler & Collins, 2003, p. 339). The experimental group was trained to "use concept mapping of patient information to prepare for clinical experiences. The control group was taught to use traditional nursing care plans" (Wheeler & Collins, 2003, p. 339). As a popular and proper instrument, the CCTST was used to measure nursing students' CT skills. Form A of the CCTST was used for the pretest. Form B was used as the posttest at the end of the fall semester of students' junior year. The pretest scores of the experimental group and the control group did not differ significantly. The ANCOVA was calculated "on the mean difference between pretest and posttest scores on the overall CCTST and the subscales, with pretest scores used as a covariant…" (Wheeler & Collins, 2003, p. 343). The post hoc tests (Fisher LSD) showed the mean score of the experimental group on the posttest differed significantly from the mean score of its pretest on the overall CCTST ($p = .02$), but there was no significant difference between the experiment group and control group on the overall CCTST ($p = .52$). The scores of the experimental group improved significantly on the overall score of CCTST and its subscales of analysis and evaluation. The control group only significantly improved the evaluation score but declined significantly on the subscale of

inference (Wheeler & Collins, 2003). Wheeler and Collins (2003) summarized that the significant difference between pretest and posttest scores of CCTST and suggested the students' CT skills were improved after one semester of training in concept mapping methodology. These authors noted that a longitudinal study needs to be conducted to examine "the long term effect of concept mapping as well as how it takes students to master the method" (Wheeler & Collins, 2003, p. 344).

Yuan et al. (2008) conducted quasi-experimental, two-group pretest-posttest research to investigate the effectiveness of problem-based learning (PBL) on the CT skills of nursing students in China. There were 46 participants in this study. They were all sophomores who registered for the course "Introduction to Nursing" at the university in Shanghai, China. In this study, participants were randomly grouped into the experimental group (n = 23) and the control group (n = 23). The Chinese Taiwan Version of CCTST Form A was utilized. The CCTST Form A was administered to the Chinese students in both the PBL and the lecture groups as a pretest and later a posttest during a one-semester nursing course (Yuan et al., 2008). The PBL approach was conducted for the experiment group, and the lecture approach was used for the control group. Independent sample t-test was "performed to compare the mean scores and change scores of the critical thinking skills by the PBL and lecture groups at the significance level of 0.05" (Yuan et al., 2008, p. 72). The results showed no statistically significant difference in age concerning CT skills between the PBL group and lecture group ($t = 1.577$, $p = 0.123$). The overall mean scores of CCTST in the pretest was lower than those in the posttest. There was "no significant difference in critical thinking skills at pretest ($p = 0.429$), whereas, significant differences in critical thinking skills existed between the PBL and lecture groups for the posttest ($p = 0.040$)" (Yuan et al., 2008, p. 72). Yuan et al. (2008) summarized there was a significant

improvement on the overall CCTST, analysis, and induction scores for the PBL group, compared to the lecture group. Therefore, PBL is an effective strategy to develop students' CT skills. Yuan et al. (2008) suggested that a randomized controlled trial with a larger sample be utilized in order to achieve a more conclusive result.

Naber and Wyatt (2014) undertook an experimental, pretest-posttest study to investigate the effect of reflective writing intervention on the CT skills and dispositions of baccalaureate nursing students at universities in the southern United States. The convenience sampling strategy was used, with 70 fourth-semester students in the baccalaureate nursing program. This is a level-III pretest-posttest design. Students are randomly assigned to an experimental group (n = 34) and a control group (n = 36). The CCTST and the CCTDI were used to measure nursing students' CT skills and dispositions. The pre-intervention CCCTST and CCTDI were administered to the students at the beginning of the semester. After six reflective writing assignments were completed, the post-intervention CCTST and CCTDI were administered to the students. Naber and Wyatt performed an "independent sample t-test to compare base-line CCTST and CCTDI scores. A MANCOVA was performed to determine if there were group differences with regard to the average change in CCTST and CCTDI subscales while controlling for experience" (p. 69). The results from the Mann-Whitney U, a non-parametric test, showed no significant difference between the control and experimental groups in age or months of experience regarding CT skills. There were also no "significant pre-intervention differences between the control and experimental groups on total CCTST, total CCTDI, or any of the subscales" (Naber and Wyatt, 2014, p. 70). The MANCOVA was employed to test "total scores and each set of subscales between the groups with regard to the average change in total scores and subscores while controlling for experience" (Naber and Wyatt, 2014, p. 70). The test results indicated a

significant difference in the subscale of truthseeking, which means the reflective writing intervention was effective and it has improved experimental students' CT. There was no significant difference in the CCTDI total scores between the experimental and control groups. There was also no significant group difference with CCTST subscales (Naber and Wyatt, 2014). Although there was no group difference in CCTST subscales, "the experimental group's scores increased on four of the five subscales" (Naber and Wyatt, 2014, p. 70), which are much "higher than control group's on three CCTST subscales" (Naber and Wyatt, 2014, p. 67). Naber and Wyatt (2014) concluded it was "imperative for nursing schools to consider including reflective writing---especial assignments based on Paul's (1993) model---in nursing courses" (p.67)

In summary, this section mainly concentrated on empirical studies on CT from three perspectives: cultivating critical thinking, critical thinking and culture, and critical thinking instruction. Scholars, researchers, and educators reached enough agreement to "pursue research on strategies for developing critical thinking skills and dispositions" (Reed, 1998, p. 45). Most theorists claimed CT is composed of both skills and dispositions (Bailin & Siegel, 2003; Ennis, 1987; Facione & Facione, 1997; Halpern, 1998; Norris & Ennis, 1989). Scholars, researchers, and educators either focused on students' CT skills or CT dispositions, or concentrated on the relationship between CT skills and CT dispositions, using standardized tests, such as the CCTST, CCTDI, WGCTA, CCTT, EWCTET, and HCTAES. With regards to the influence of culture on the CT, many studies indicated students from non-Western contexts have exhibited lower CT skills and dispositions, compared with their counterparts from the West (Soeherman, 2010). Some scholars argued the lower level of students' CT from non-Western countries mainly results from their cultures (McBride et al., 2002; Grosser & Lombard, 2008; Yang & Chou, 2008). Others believed that culture is not the only factor that affects students' CT, and that other factors,

such as insufficient knowledge, target language deficiency and insufficient training, influence students' CT, as well (Bali, 2015; Egege and Kutieleh, 2004; Paton; 2005). Regarding the CT instruction, scholars, researchers, and educators developed a variety of programs, models, and approaches to teaching CT and conducted various empirical studies to examine the effect of these methodologies. However, the research results were "mixed; no single instructional method was shown to be either always effective or ineffective" (Behar-Horenstein, 2011, p. 30).

CT Studies in China

To date, CT is valued in both the Western educational system and the Chinese educational system (Li, 2017). Since the late 1990s, the CT movement in China has begun to grow and develop (Dong, 2015). Similar to the research findings of the Western scholars and researchers, some Chinese scholars also discover that Chinese students have more difficulties in constructing their own ideas and they are less likely to challenge authorities, compared to the Western students (Liu, 1998; Hu, 2002). The evidence shows Chinese students lack CT, which is a concern for both Chinese students who study abroad, and those who study at home (Tian, 2008). Partly thanks to these research findings, Chinese policy makers, such as, the Ministry of Education (MOE) of China, as well as researchers, scholars, and educators started to realize the importance of thinking. They hope that "introducing CT to Chinese students will enable them to become innovative and creative thinkers and entrepreneurs, and introducing CT into Chinese HE [higher education] will help realize the comparative success of Western HE" (Li, 2017, p. 43). This section focuses on CT studies in China, and is limited to two dimensions: theoretical and empirical studies on CT.

Theoretical Studies on CT in China

The National People's Congress released the "Higher Education Law of the People's Republic of China (中华人民共和国高等教育法)" in 1998, which claimed cultivating innovative talents was one of the main aims of Chinese higher education (HE) (Li, 2017). Since then, internationalization of the Chinese HE and fostering innovative talents with CT have become the main focus of Chinese universities and colleges (Li, 2017). More Chinese scholars began to explore CT and its instructional approaches. However, compared to the Western scholars' research, Chinese scholars' studies on CT are not deep enough and lack a systematic inquiry. Many of these studies remain more at the theoretical level of CT, i.e., its definition, its assessment, introduction of CT development and CT instruction, lacking theoretical creation and innovation (Zhu, 2017).

Zhixian Zhu (1985), in his article, "A Review of Modern Cognitive Psychology", claimed CT is the intellectual quality of being good at strictly estimating thinking materials, and carefully examining thinking processes in thinking activities. CT is the degree of independent analysis and CT activities. CT should be regarded as an integral part of problem solving and creative thinking (Huang, 2019). According to his classification of CT theory, Qingxu Luo (2002) defined critical thinking as a personality quality in which individuals make self-regulatory judgment on the correctness of the knowledge generation process, theory, method, background, evidence and criteria for evaluating knowledge. Qingxu Luo (2002) believed CT includes the personality tendency of critical thinking and the psychological characteristics of personality. Xisheng Wang (2006) argued that with the purpose of achieving a rational grasp of the world, people take the initiative to reflect on the rationality of others or their own thoughts to arrive at a rational judgment. Therefore, CT is a type of thinking that consciously evaluates the rationality of certain beliefs and behaviors (Huang, 2019). Zhenzhi Gu & Zhuanghu Liu (2015) referred to

CT, in a broad sense, as the ability to develop and improve people's world view and apply it to all aspects of life effectively. Specifically, CT is the ability to make rational decisions about what to believe or do (Huang, 2019).

Hongzhi Wu (2004), based on his systematic introduction to the meaning of CT abroad, claimed CT is mainly composed of critical skills, critical knowledge, critical techniques and critical thinking habits or temperaments. Qiufang Wen et al. (2009), based on the existing critical thinking model, proposed a hierarchical model of critical thinking. This model includes two aspects: metacognitive ability and thinking ability. Metacognitive ability refers to self-regulation ability, including the ability to plan, check, adjust, and evaluate one's own thinking. Thinking ability is divided into two dimensions: cognitive and emotional. The cognitive dimension consists of core skills and cognitive criteria. There are three core skills: analysis, reasoning, and evaluation; cognitive criteria include clarity, relevance, logic, depth, and flexibility. The emotional dimension of critical thinking can be summarized as curiosity, openness, confidence, integrity, and perseverance (Wen, 2009). Minghui Xiong (2014) regarded CT as a purposeful reflective judgment and presented the six core cognitive skills of CT, including interpretation, analysis, evaluation, inference, clarification, and self-regulation (Huang, 2019).

Most of the existing CT tests in China are translated directly from the West. For instance, Meici Peng et al (2004) revised the Chinese version of the California Critical Thinking Disposition Inventory (CCTDI) into the Critical Thinking Disposition Inventory Chinese Version (CTDI-CV), making it more suitable for the Chinese context (Wang, 2011, p. 6). Qingxu Luo (2002), Qingxu Luo and Xinhui Yang (2001, 2002) revised the Chinese version of the California Critical Thinking Disposition Inventory (CCTDI) and the California Critical Thinking Skills Test (CCTST). There are self-designed CT assessment tools in China as well. Some domestic scholars

also compiled a CT test scale. For example, Taiwan scholars, Yuzhu Ye and Piling Ye believed identifying hypothesis, reasoning, deduction, explanation, and evaluation are important skills of critical thinking, so they have compiled a Critical Thinking Skills Test for Primary and Secondary School Students. (as cited in Li, 2019). Xingtai Lin and Yucheng Zhang developed the Short-form Roche Advanced Cognitive Ability Test, which focuses on the critical thinking level of students in Grades Four to Six (as cited in Li, 2019). Qiufang Wen (2009), an English education researcher in mainland China, compiled a Measurement Tool of Foreign Language College Students' Critical Thinking Ability, which examines the critical thinking level of foreign language college students. (Li, 2019). Linlu Wang et al. (2011) compiled the Critical Thinking Scales for College Students, including two sections, assessing college students' CT dispositions and CT skills. Chengyan Huang (2015) compiled the Critical Thinking Disposition Scale for College Students, consisting of four dimensions with 22 items, using 5-point Likert scale. For Li (2019), with the increasing emphasis on CT education in China, more CT assessment tools will be designed, created, and developed.

Qingxu Luo (2002), based on the systematic analysis of the extant CT theories and tests, analyzed the current situation and problems of the Western critical thinking assessments in detail. Qingxu Luo (2002) selected two representative Western critical thinking test tools for revision: the Chinese version of the California Critical Thinking Skills Test (CCTST) and the Chinese version of the California Critical Thinking Disposition Inventory (CCTDI). It is the first time that a Western critical thinking assessment tool was revised for use in China, and the revision work was viewed as a new method for the development of critical thinking assessment (Luo, 2002). According to Luo (2002), the CCTST mainly involves logical reasoning problems, the statements of its items are more precise, and the understanding of the items is less affected by

different languages and cultural backgrounds. Luo (2002) sampled 382 college students, including 181 males and 201 females, ranging from 18 to 24 in age. Luo (2012) only examined the test-retest reliability coefficient ($a = 0.63$, $p < 0.01$) and split-half reliability coefficient ($a = 0.75$, $a = 0.80$, $p < 0.01$) of the Chinese version of CCTST. Luo (2002) claimed that the simplified Chinese version of CCTST has good reliability, based on this result. Luo (2002) also examined construct validity and criterion validity of the simplified Chinese version of CCTST. The scores of the participants on the CCTST test have increased significantly, and the results showed the simplified Chinese version of CCTST has good construct validity (Luo, 2002). The statistical results also indicated the scores of the simplified Chinese version of CCTST were significantly correlated with the Raven's standard reasoning test scores and grade point average (GPA) of the participants, but the correlation between the CCTST scores and the GPA of the test pairs was smaller than the correlation between the CCTST scores and the Raven's standard reasoning test, which indicated that the simplified Chinese version of the CCTST has better criterion validity (Luo, 2002).

After testing the simplified Chinese version of CCTDI, Luo (2002) discovered the Cronbach *a* of the Chinese version was lower than the English version. The coefficients of internal consistency of the subscales ranged from 0.46 to 0.83 (Huang, 2015). For Luo (2002), this internal consistency of the CCTDI was still acceptable. However, it suggested the CT disposition measured by CCTDI has obvious cultural differences. Therefore, it is necessary to design a critical thinking assessment tool with the Chinese cultural characteristics (Luo, 2002). Luo (2002) finally noted that due to funding and time constraint, a complete computer-based tool for the CT assessment based on genetic algorithms and neural networks was not completed by the end of his dissertation.

On the basis of the literature review, interviews, and open question questionnaire, Chengyan Huang (2015) put forward the theoretical conception of college students' CT disposition. Through the item analysis and exploratory factor analysis of 457 valid initial questionnaires collected, a formal questionnaire of Critical Thinking Disposition with four dimensions and 22 questions was designed. The four dimensions are: (1) universal skepticism; (2) self-reflection; (3) open-mindedness; (4) careful thinking (Huang, 2015). Universal skepticism is the courage to challenge authority, and the universal attitude of skepticism and reflection towards knowledge. Self-reflection is the ability to expose one's mistakes or deficiencies and self-correction; Open-mindedness is the willingness to consider different beliefs and keep a keen interest in things. Careful thinking refers to the ability to think objectively about the problem and seek appropriate supporting evidence before making a decision (Huang, 2015). Regarding reliability testing, internal consistency and test-retest reliability were included. The validity test included content validity, construct validity and calibration validity. The results showed each reliability and validity index was good. These analyses of the self-designed critical thinking disposition scale for college students have good reliability and validity and can be used as a measurement tool to assess college students' critical thinking disposition (Huang, 2015).

Li (2019) claimed most Chinese scholars conducted research only on the introduction and dissemination of CT and its instruction. For instance, Hongzhi Wu from Yan'an University was engaged in the research and teaching of critical thinking for a long time and carried out a detailed categorization and summary of the basic concepts and teaching of critical thinking (Li, 2019). To date, Hongzhi Wu published 9 books on CT, such as *Critical Thinking*, and *A Preliminary Study of Critical Thinking*, which ranked number one in China (Huang, 2019). Jianwu Zhou from Renmin University of China published 4 books on CT (e.g., *Critical Thinking Course --- Logical*

Reasoning and Argumentation), having the second most books addressing CT, and followed by Yu Dong from Huazhong University of Science and Technology, who published 3 books on CT (e.g., *Principles and Methods of Critical Thinking*) (Huang, 2019).

There were numerous journal articles on the subject of CT. In his "On the Cultivation of Critical Thinking in College Students", Qingxu Luo (2000) analyzed and discussed the nature and structure of critical thinking, the significance and way of developing critical thinking training in Chinese colleges and universities. He also pointed out the problems that need to be addressed in order to obtain better results for critical thinking training. According to Luo (2000), CT is a personality quality in which individuals make self-regulatory judgments on the correctness of the process regarding knowledge generation, theory, method, background, evidence, and evaluation criteria. CT includes two aspects: the personality tendency and the personality psychological characteristics of CT. Personality tendency reflects the critical spirit of the individual, and personality psychological characteristics reflect the critical ability of the individual (Luo, 2000). Luo (2000) proposed three key points regarding the significance to foster Chinese students' CT: (1) Cultivating students' CT is the demand of the 21century. It is imperative to cultivate the innovative spirit and ability of college students because of the fierce competition in the global market. (2) Developing college students' CT can improve their ability to choose, understand, digest, and evaluate knowledge in the era of knowledge explosion, and enhance their ability to distinguish all kinds of knowledge from the internet and mass media, so as to improve the adaptability of social labor force in the new era. (3) Developing students' CT can help them effectively resist superstition and blind faith, and prevent social crisis and unrest caused by blind faith and superstition. Luo (2000) finally summarized three ways of cultivating students' CT abroad: (1) set up a separate critical thinking classroom for CT training; (2)

combine the cultivation of CT with subject instruction and develop students' CT through regular classroom teaching; (3) develop students' CT by developing hidden curriculum (Luo, 2000).

In "Research on Critical Thinking and its Skills", Xisheng Wang (2006) elaborated on two popular definitions: Facione et al's (46 scholars reach a consensus on the definition of CT) and Ennis' (one of the pioneers of the American CT movement). According to Wang (2006), these two exemplary definitions have the following characteristics in common: (1) both of them frame the cognitive object of critical thinking in the perspective of knowledge modification; (2) both emphasize the thinking subject's initiative to judge the cognitive object; (3) both regard critical thinking as a process or activity. These two definitions have their merits, but they also have weaknesses. For example, it is too narrow to frame the objective of CT only in the scope of "knowledge modification" (Wang, 2006). Another weakness is the trait of critical thinking --- "positive and active reflection" is not highlighted as it should be (Wang, 2006). Wang (2006) claimed CT is a type of thinking that consciously evaluated the rationality of certain beliefs and behaviors, and specifically, CT should be the organic combination of "dare to criticize" (spirit) and "good at criticism" (skill) (p. 7). Wang (2006) summarized the methods of CT skill training can be divided into four categories: General Approach, Immersion Approach, Infusion Approach, and Mixed Approach, according to the relationship between the teaching and training of CT and the subject content of college students. Only by combining the CT technical training of college students with daily knowledge learning and the cultivation of CT spirit, can they improve their critical thinking skills and improve their learning effect (Wang, 2006).

Guangfu Xiang (2012), in his article "Review of Critical Thinking Literature Abroad", introduced the definitions of CT proposed by seminal scholars, such as Dewey, Glaser, Fisher, and Facione. Xiang (2012) then described the components of CT: dispositions, skills and

knowledge, based on the ideas of Ennis, Paul, McPeck, and Facione. He also explained the three dimensions of CT: philosophical, psychological and educational. The debate on the transferability of CT was introduced as well. Scholars who systematically studied CT, such as Ennis, proposed CT skills can be taught independently of specific disciplines, so CT is generic and can be transferred (Xiang, 2012). In contrast to Ennis, McPeck held the opposite view and claimed CT is discipline-specific and cannot be transferred between disciplines (Xiang, 2012). Xiang concluded the debate on the transferability of CT results from the different theoretical presuppositions, and stated the reason for it is how the researchers view the problem from a purely logical perspective and make the mistake of either/or. Xiang also pointed out that the lack of consistent understanding of CT due to the different perspectives of researchers limits the development of assessment tools for the comprehensive evaluation of CT.

In summary, the CT definition from Delphi Report spearheaded by Facione (1990a) and Ennis' definition (1987) are more popular and are more frequently cited in China (Wang, 2006). By synthetically comparing different CT research literature in the Western academia, Chinese scholars proposed their own definitions of CT, mainly based on the ideas of the Western scholars. Chinese scholars also believed that various definitions of CT mainly resulted from different perspectives: philosophical, psychological, and educational. Philosophers regard CT as a cognitive ability, a subjective attitude or tendency. Psychologists take CT as a mental activity. While for educators, CT is a practice (Luo, 2002; Huang, 2015; Huang, 2019). Each definition of CT has its own legitimate components, its own advantages and disadvantages, and it is used in a certain context (Huang, 2015). Both the Western scholars and Chinese Scholars reached an agreement that CT includes skills, dispositions, and background knowledge (Huang, 2019). As for the assessment of CT, China lacks localized CT instruments or tests. Most existing CT tests

or assessment tools are borrowed from the West. The instruments of CCTST, CCTDI, and WGCTA are most frequently utilized to assess students' CT in China (Huang, 2019). More scholars and educators have begun to design their own CT skill and disposition instruments which take Chinese culture and Chinese students' characteristics into consideration. Most Chinese research on CT belonged to theoretical dimension, focusing on the definition, the assessment, literature review, or introduction of CT achievements that Western scholars have gained.

Empirical Studies on CT in China

Fortunately, Chinese researchers, scholars, and educators realized the significance of empirical studies on CT and began to explore CT teaching and learning in the recent years (Du et al. 2008; Liu, 2015; Li & Li, 2016; Li, 2019). Chinese scholars mainly target the Chinese students' CT skill and dispositions. However, there is a scarcity of research on Chinese teachers' CT. This section concentrates on reviewing the empirical studies on CT in China, and is limited to two dimensions: empirical studies on Chinese students' CT and empirical studies on Chinese Teachers' CT.

With respect to the empirical studies on Chinese students' CT, Tianjiao Du, Na Yu, and Shuying Guo (2008) explored whether the Problem-based Learning (PBL) teaching model can promote the CT skills of the clinical medical students at the China Medical University. There were 119 participants from 4 classes of clinical medicine in this study. Du et al. (2008) utilized the Chinese version of California Critical Thinking Skills Test (CCTST) and Eysenck Personality Questionnaire (EPQ) to assess Chinese medical students' CT skills and their personality. This study lasted one year from October, 2006 to October, 2007. A comparison was done of CT skills of medical students before and after adopting the PBL teaching model by

paired samples *t* test. It indicated that students' CT scores after the PBL teaching model were significantly higher than their scores before the PBL teaching model ($p < 0.05$). There was a statistically significant difference between the medical students' scores of analysis scale before and after the PBL teaching model. Although the evaluation and inference scores of students after the PBL teaching model were higher than their scores before the PBL teaching model, there were no statistically significant differences ($p > 0.05$). Results from the *t*-test showed there was a significant difference in students with intermediate type of personality after and before the PBL teaching model ($p < 0.001$). However, there were no significant differences in introvertive and extrovertive personality of students after the PBL teaching model and before the PBL teaching model. These research results indicated the PBL teaching model does promote CT skills of the clinical medical students. Students' analysis skills were also improved after the PBL teaching model (Du et al., 2008).

Yang and Chou (2008) employed a pretest and posttest quasi-experimental design with a comparison group to explore the relationship between CT skills and CT dispositions, and the effect of three different levels of instructional strategy (asynchronous online discussions [AODs], CT skill instruction via AODs, and CT skills instruction with CT dispositions cultivation via AODs), in improving students' CT skills and dispositions. The participants included 220 students enrolled in Electricity and Life, a general education course in a large university in Taiwan, China. They were divided into two groups: 145 students in Section A and 128 students in Section B. These students are randomly assigned to 10 discussion groups, and each discussion group consisted of 14 students. Yang and Chou randomly selected five discussion groups in Section A as a comparison group ($n = 75$). "Five discussion groups in Section B served as Treatment I (71 students), and the other five groups in Section B as Treatment II (74 students)" (Yang & Chou,

2008, p. 677). The CCTST and the CCTDI were used to measure students' CT skills and dispositions respectively. Students were required to take the CCTST and CCTDI in class at the beginning and at the end of the semester. The results of Pearson product-moment correlation coefficient (r) indicated the pretest CCTST and CCTDI scores of students of three groups had a significantly positive relationship ($r = 0.63$, $p = 0.00$). Because the coefficient of determination $r^2 = 0.3969$, it suggested "40% of the variance in CCTST scores was potentially attributable to the differences in these students' CCTDI scores, and vice versa" (Yang & Chou, 2008, p. 677). Further analysis showed there were no Chinese students reaching the high level of CCTDI. The significant correlation only existed between students with high CT skills and medium CT dispositions. The results of MANCOVA indicated "the posttest scores of CCTST and CCTDI among three research groups differ significantly" ($p = 0.00$, Yang & Chou, 2008, p. 678). The pared-sample t-tests were used to examine whether there was "an improvement between the pretest and posttest scores in CT skills and CT dispositions in each group" (Yang & Chou, 2008, p. 674). The improvement in CT skills reinforced CT dispositions, but the improvement in CT disposition did not increase the CT skills (Yang & Chou, 2008). Yang and Chou finally recommended the instructional strategy, as well as the CT skill instruction with the CT disposition cultivation should be utilized in order to enhance students' CT skills and dispositions.

Ng et al. (2022) conducted a cross-sectional correlational study on Chinese community college students to assess their CT cognitive skills and CT dispositions in order to identify the associated factors. "This is the first study of its kind to assess the CT abilities of community college students in an Asian context" (Ng et al., 2022, p.1). Ng et al. used a convenient sampling strategy, and 209 students at Hong Kong Community College (HHCC) were recruited in this study. They came from the Division of Science, Engineering and Health Studies (SEHS). This

was a one-year study, starting from March 26, 2020 to March 27, 2021. Three instruments were employed to measure the participants' CT skills, dispositions, and socio-demographic and academic factors. They are: CCTST, CCTDI and the self-developed Socio-Demographic and Academic Characteristics Questionnaire. The research findings indicated the mean score of the community college students' CT skills was 17.82 ± 4.20, falling into the upper end of the moderate range. The mean score of the community college students' CCTDI was 278.81 ± 22.61, indicating this was an inconsistent/ambivalent overall disposition towards CT. However, there are four subscales: open-mindedness, analyticity, confidence in reasoning, and inquisitiveness which showed a positive disposition towards CT. After running a regression analysis, the University Entrance Examination total score (HKDSE total score), CT subject grade level, CCTDI truth-seeking, and CCTDI analyticity were identified "as the four significant factors associated with their CT cognitive skills" (Ng et al., 2022, p. 1). According to Ng et al. (2022), these research results offered "significant implications when reviewing the program design and curriculum as well as the addition of CT elements in a separate course to promote students' CT abilities for sustainable development" (Ng et al., 2022, p. 1).

Although most studies on Chinse students' CT employ quantitative methods, there are a few qualitative studies on students' CT as well. For instance, Chen (2017) conducted a qualitative case study on Chinese students' conceptualization of CT. There were 46 volunteer participants who came from a university in Guangzhou, China. Their majors included Business, Literature, Education, and Engineering. With the purpose of exploring and identifying the patterns of Chinese college students' conceptualization of CT, Chen employed "a multi-step open-coding strategy to identify potential patterns in the data" (p. 143). The research findings indicated most Chinese students do not have difficulty offering definitions of CT, and they can

provide examples from their experience to interpret CT. However, two freshmen refused to offer the definition of CT. After Chen presented several definitions, these two students gave examples based on their understanding. One student gave a definition, but could not support it with an example. Six themes emerged from the Chinese students' definitions of CT: cognitive thinking skills (36 students, 75%), intellectual autonomy (27 students, 56%), omnipresence of the opposite point of view (11 students, 23%), multiple perspective (more than two, and aware of conflicts) (8 students, 17%), revolutionary thoughts in art, science, and other fields (7 students, 15%), and knowledge and skills outside the scope of standard examination (1student, 2%). Chen explained the total percentage of students' themes was over 100%, because "most students mentioned several different themes" (p. 144). According to Chen, *cognitive thinking skills, intellectual autonomy* and *the omnipresence of the opposite point of view* were the most frequently mentioned themes by Chinese students. "The emphasis on ownership of ideas against authority is characterized as *intellectual autonomy*…" (Chen, 2017, p. 148). Chen also discussed the Chinese students' understanding of pros and cons and the significance of harmony. *Intellectual autonomy* and *the omnipresence of the opposite point of view* "reflect the influence of Chinese Indigenous Philosophy and the Western concept of individualism, among many other influences" (Chen, 2017, p. 148). Chen stressed his research results challenged the stereotype that Grimshaw (2007) refuted in his ethnographic study on Chinese college students. Avoiding overemphasizing the differences between the Chinese and Western students concerning CT is a good way to erase the stereotype of Chinese students (Chen, 2017).

Although teachers play a vital role in the development of students' thinking skills, little is known about teachers' conceptions, beliefs, and practice (Li, 2016). Some scholars make attempts to conduct research on Chinese teachers' thinking skills. For example, Li (2016)

conducted a qualitative case study to examine how English as a Foreign Language (EFL) teachers in China "conceptualize thinking skills and how these skills are perceived to be promoted in subject learning" (p. 273). Participants in this study were 473 volunteer teachers, including 398 females and 75 males. Most teachers (n = 366) were aged between 26 and 35. Li (2016) employed a funneling approach by means of a questionnaire, "four focus group interviews with 18 teachers and a further follow-up classroom observation and video-based reflection of three teachers" (p. 273). The data analysis concentrated on bringing all data together to "generate an in-depth understanding about how teachers conceptualize thinking skills and how these skills are perceived to be promoted in subject learning" (Li, 2016, p. 273). The research results indicated EFL teachers in China have difficulty in defining thinking skills. Their concepts about thinking are fragmented and insufficient, though they show positive attitudes toward integrating thinking skills into language teaching. Most of these Chinese teachers do not teach thinking skills in their classrooms, but they believe thinking skills can be taught (Li, 2016). According to Li, there is "a strong case for arguing for immediate teacher training to develop both content and pedagogical knowledge of teaching thinking skills" (p. 273).

The research on Chinese teachers' CT is entitled "Investigating University EFL Teachers' Perceptions of Critical Thinking and Its Teaching: Voices from China" by Zhang et al. (2020). Their study investigated Chinese EFL teachers' perceptions of CT and their teaching practice by means of a mixed methods research design. There were 336 Chinese university EFL teachers recruited. They were from 24 provinces (out of total 27, not including Taiwan, Hong Kong, and Macao) and "four municipalities directly under the central government…" (Zhang et al., 2020, pp. 485-486). The average length of these teachers' service was 16 years with a range from 1 to 39. The quantitative methodology employed a questionnaire with a ten-item Likert

scale, adapted from Stapletons' (2011) study, to examine "the participants' beliefs about the meaning of CT, and its relationship with English language teaching (ELT)" (Zhang et al., 2020, p. 486). After the questionnaire, follow-up interviews were carried out with the purpose of offering more insight on teachers' perceptions of CT and its teaching (Zhang et al., 2020). In the telephone interview, demographic information of the teachers, such as region, gender, teaching experience, and their student types, were collected. In this mixed methods study, means, standard deviation (SD), and percentage were reported to provide "a general picture of Chinese university EFL teachers' perception about CT, the relationship between CT and EFL instruction, and their teachers' perceived need for further training on teaching CT" (Zhang et al., 2020, p. 486). T-tests, ANOVA, and MANOVA tests of statistical significance were performed. The interview data generated from the open-ended questions were input into NVivo 11 for data coding. The research findings demonstrated EFL teachers in China gave a strong endorsement to integrating CT with language teaching ($M = 3.26$). Participants reached an agreement that they need more training about the theory and methods of teaching CT in EFL classrooms and "teaching CT is an important goal in university EFL curriculums ($M = 3.02$)" (Zhang et al., 2020, p. 488).

A one-way ANOVA was conducted to "compare teachers' educational background and their perceptions of CT and its teaching" (Zhang et al., 2020, p. 488). The results of the ANOVA indicated teachers with master's (MA) and doctoral degrees score higher than those with bachelor's (BA) degrees, and these differences were statistically significant ($F = 14.238$, $p < 0.001$). Results from a paired-sample t test indicated those who teach English majors have significantly higher scores than the teachers who do not teach English ($t = 2.715$, $p < 0.01$). The results also demonstrated "the perceptions of integrating CT with language teaching were significantly associated with academic background ($F = 5.256$, $p < 0.05$)" (Zhang et al., 2020, p.

488). Zhang et al. summarized the participants display only a fragmented understanding of CT, which is not consistent with their high scores on the questionnaire. This suggested there is a disparity between teachers' perceptions of CT and "a generally recognized more robust meaning of CT" (Zhang et al., 2020, p. 490). There was also a disparity between teachers' perceptions of CT teaching and their actual classroom teaching, which does not conform with their perception of the importance of CT indicated from the questionnaire. It implied teachers do not help foster students in their classrooms, which "contradicted their professional perceptions" (Zhang et al., 2020, p. 490). Teachers reported they lack professional knowledge of CT and strategies of implementing CT, and they need more learning and training opportunities, as well as resources and contextual support (Zhang et al., 2020).

In summary, this section mainly targeted the empirical studies on CT in China. Two dimensions were the main focus: empirical studies on Chinese students' CT and empirical studies on Chinese Teachers' CT. It is worth noting that there is a large volume of studies on Chinese students' CT, but there is a scarcity of research with regard to Chinese teachers' CT. This review of the literature found on literature on ISEC teachers' CT. The empirical research on Chinese students in both China and the West indicated Chinese students lack critical thinking. They are reticent in class, prefer a reproductive learning approach, and rely on a limited range of learning strategies, like rote learning and memorization (Atkinson, 1999; Ballard, 1996; Flowerdew, 1998; Tsui, 2002). However, some scholars, such as Grimshaw (2007), refuted these ideas and regarded them as a cultural stereotype of Chinese learners. The empirical studies on Chinese teachers' CT indicated EFL teachers in China demonstrate positive attitudes towards CT in the classroom, but they did not have clear conception or perception regarding how to define

CT. Their ideas about thinking are fragmented and insufficient. They need systemic learning and professional training on CT and strategies to integrate CT into their teaching practices.

Summary

This chapter focused on reviewing the related theoretical and empirical research on critical thinking (CT). It consisted of four sections: the conceptualization of CT, assessment of CT, CT studies in the West, and CT studies in China. Scholars and researchers have not reached a consensus on the conceptualization or definition of CT. They defined CT from different perspectives: philosophical, psychological and educational. However, there is an agreement on the components or dimensions of CT: CT is composed of skills, dispositions and knowledge. There is also an ongoing debate on whether CT is general or domain-specific. Most scholars are proponents of generalization of CT, including Ennis (1989), Halpern (2001), Lipman (1988), Gelder (2005), Moore, (2011), and Davies, (2013). The advocates of domain specificity belong to the minority group, involving McPeck (1981), Bailin (1999), and Willingham (2007). The disagreement on the degree of domain specificity leads to the debate of whether CT is transferable to new contexts (Lai, 2011). Most scholars and researchers claim critical thinking skills and abilities can be taught (Lai, 2011).

There are three types of CT instruments or tests: multiple-choice, open-ended, and a combination of multiple-choice and open-ended tests. Most frequently used CT instruments are multiple-choice tests, such as the CCTST, the CCTDI, and the WGCTA. The commonly employed open-ended test is the EWCTET with an essay-based format. The HCTAES is a typical example of a combination of the multiple-choice and open-ended test. To date, all CT assessment tools have their own strengths and weaknesses. There are no perfect instruments.

According to Liu et al. (2014), the combination of multiple-choice and open-ended test becomes more popular, because it provides the potential to measure CT authentically and appropriately.

With regard to the CT studies in both China and the West, it is obvious that the Western scholars and researchers gained fruitful achievements on all aspects of CT, despite the disagreement or debate on some of the issues. The CT studies in the West are comparatively mature, compared to the CT research in China, which started in the late 1990s and are still in their infancy. Most Chinese scholars focused on the theoretical research on CT, and translate and borrow CT theories from the West. With the development of globalization and internationalization, more and more Western scholars and Chinese counterparts began to concentrate on Chinese students, especially their CT skills and dispositions. However, there is minimal research on Chinese teachers' CT to date. Especially, there are no studies that target the CT of ISEC teachers and non-ISEC teachers in China. Therefore, it is imperative to do research on teachers' CT, especially ISEC teachers' and non-ISEC teachers' CT, to bridge the gap of literature. The next Chapter will focus on the Methodology of this study.

Chapter Three: Methodology

In order to explore the status quo (current situation), the perception, understanding, and practice of CT among ISEC and non-ISEC teachers in institutions of higher education in the north of China, the researcher used a non-experimental causal-comparative study. In this non-experimental causal-comparative study, the explanatory mixed methods design or the QUAN-Qual Model (Gay et al., 2006), was utilized. According to Gay et al. (2006), the aim of the mixed methods research design is "to build on the synergy and strength that exists between quantitative and qualitative research methods in order to understand a phenomenon more fully than is possible using either quantitative or qualitative methods alone" (p. 490). For the quantitative research paradigm, the computer-based CCTST on the Insight Assessment website was utilized to explore the status quo of CT skills of Chinese ISEC and non-ISEC teachers. For the qualitative research paradigm, interviews via Zoom were employed to collect data with respect to teachers' perception, attitudes, and experiences concerning CT and CT teaching in China. By means of data collection and data analyses, the results and findings offered a better understanding of the research purpose and questions.

Quantitative

Research Design

This was a non-experimental causal-comparative study with an explanatory sequential mixed methods research design. According to Creswell (2012), in the explanatory sequential mixed methods research design, the researcher first collected quantitative data and then collected qualitative data to help explain or elaborate on the quantitative findings. The quantitative research paradigm in this study employed a survey method (Creswell, 2012), and data were collected utilizing the computer-based CCTST on Insight Assessment website. The researcher

sought answers to the following nine research questions to gain an overview of CT skills of Chinese teachers in institutions of higher education. Research questions, alternative hypotheses, and null hypotheses were:

R_1 What is the level of CT skills of overall Chinese teachers in institutions of higher education in the north of China?

R_2 What is the level of CT skills of the ISEC and non-ISEC teachers in institutions of higher education in the north of China, respectively?

R_3 Is there a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers in the north of China?

- H_1 There is a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers in the north of China.
- H_0 There is no statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers in the north of China.

R_4 Is there a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers, based on gender identifying as male and female?

- H_1 There is a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers, based on gender identifying as male and female.
- H_0 There is no statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers, based on gender identifying as male and female.

R_5 Is there a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers based on professional rank?

- H_1 There is a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers based on professional rank.

- H_0 There is no statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers based on professional rank.

R_6 Is there a statistically significant relationship between CT skills and age?

- H_1 There is a statistically significant relationship between CT skills and age.
- H_0 There is no statistically significant relationship between CT skills and age.

R_7 Is there a statistically significant relationship between CT skills and years of teaching?

- H_1 There is a statistically significant relationship between CT skills and years of teaching.
- H_0 There is no statistically significant relationship between CT skills and years of teaching.

R_8 Is there a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers based on the educational background?

- H_1 There is a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers based on the educational background.
- H_0 There is no statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers based on the educational background.

R_9 Is there a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers based on the discipline?

- H_1 There is a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers based on the discipline.
- H_0 There is no statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers based on the discipline.

Population and Sample. The ISEC teachers and non-ISEC teachers from the local or provincial colleges and universities, rather than from Project 211, Project 985, or C9 League, the

key and leading universities in China, comprised the population of this non-experimental causal-comparative study. According to the ISEC Office (2022), there are 3775 ISEC teachers in total as of August 2022. The number of non-ISEC teachers in the local colleges and universities is unknown. Nonprobability sampling was used in this non-experimental causal-comparative study, because "the probability of selecting a single individual is not known" (Salkind, 2000). One hundred and two volunteer participants were recruited, including 52 ISEC teachers and 50 non-ISEC teachers, to take the Chinese version of CCTST. Email and WeChat (a popular Chinese social media) invitations were sent to all the ISEC and non-ISEC teachers that the researcher knew. The researcher asked the participants to recruit other volunteers. These steps were repeated until the needed sample size (102 participants) was reached. This was a nonprobability snowball sampling (Glen, n.d.). After obtaining quantitative data through CCTST, 12 out of 102 participants were selected for interviews because of their willingness to participate, including 6 ISEC and 6 non-ISEC teachers. According to Guest et al. (2006), 12 interviews were the minimum number to reach theoretical saturation.

The researcher followed the ethical standards overseen by the Institutional Review Board (IRB) at the University of Montana (UM). Participants' confidentiality was guaranteed using the following process. The participants were informed about the purpose of the study, importance of a high response rate, and the benefits and potential risks of participation before they formally agreed to participate and signed the consent form. Every participant signed his or her consent form on a voluntary basis. The confidentiality of identities was ensured by using pseudonyms for the qualitative portion of this study. The original data of CT skills obtained from Insight Assessment Company, as well as the recording and transcript of participants' interviews were kept in a safe area, which can only be accessed by the researcher herself. The researcher deleted

the recordings after they were transcribed. Participants were informed they had the right to refuse to participate in this study or withdraw from the study at any time they liked without any negative consequences.

Variables and Levels of Measurement. In this study, both the difference and correlation (or relationship) were explored. With regard to the difference, the score of CT skills obtained from the CCTST was the dependent variable. The other factors---gender, educational background, professional ranks, and discipline---were used as independent variables. In terms of correlation, the CT skill score was the criterion variable, and the other two factors: age and years of teaching were predictor variables. The process of measurement is critical to any quantitative research. To explore the difference and relationship between variables, the researcher needed to know the properties of the variables: the score of CT skill was a ratio level of measurement, and the level (Not Manifested, Moderate, Strong) of CT skills was ordinal data. Both age and years of teaching were the ratio level of measurement. Educational background, professional ranks, and discipline were the nominal levels of measurement.

Collection and Instrument. Two instruments were used in this study with an explanatory mixed methods research design. The first instrument was the Chinese version of the California Critical Thinking Skills Test (CCTST), which consisted of two sections. The first section included the participants' demographic information, such as university name, age, gender, years of teaching, professional rank (assistant, lecturer, associate professor, full professor), educational background (bachelor, master, doctorate, post-doctorate), teaching discipline, whether they participated in any form of CT training, and whether they were ISEC teachers or not, to collect quantitative data. The second section was the Chinese version of

CCTST, aiming at assessing teachers' CT skills. The second instrument was the interview protocol in Chinese, which was explained in detail in the qualitative paradigm section.

The Chinese version of the CCTST measured five core CT skills: analysis, inference, evaluation, deductive reasoning, and inductive reasoning. According to Facione (1990a), analysis skills include determining significance, interpreting meaning, and detecting possible inferential relationships. (2) Inference skills involve bringing together all the various elements needed to draw reasonable conclusions, and forming conjectures and hypotheses. (3) Evaluation skills refer to assessing the credibility of statements and the strength of the arguments. Justifying one's reasoning based on relevant evidence, concepts, methods or standards. (4) Deductive reasoning includes the assumed truth of the set of premises purportedly necessitates the truth of the conclusion. (5) Inductive reasoning involves an argument's conclusion is apparently warranted, but not necessitated, by the assumed truth of its premises. The CCTST consists of 34 multiple-choice questions that are discipline-neutral. It takes participants 45 minutes to complete. The test produces a total score and five subscale cores. One point is obtained for each correct answer, with the maximum score of 34. According to the CCTST manual (2021), there are five levels of CT skills. Table 3 shows the qualitative description of CCTST OVERALL score. Table 4 indicates the qualitative interpretation of five scale scores, including three levels.

Table 3

Qualitative Description of CCTST OVERALL Score

CCTST	Not Manifested	Weak	Moderate	Strong	Superior
CCTST OVERAL Score (34- point version)	0-7	8-12	13-18	19-23	24 or higher

Note: Table is taken and adapted from the CCTST manual (2021), p. 28

Table 4

CCTST 34-Point Scale Score Interpretation

CCTST 34-point Versions	Not Manifested	Moderate	Strong
Analysis	0-2	3-4	5 or more
Inference	0-5	6-11	12 or more
Evaluation	0-3	4-7	8 or more
Induction	0-5	6-11	12 or more
Deduction	0-5	6-11	12 or more

Note: Table is taken and adapted from the CCTST manual (2021), p. 30

Reliability and Validity. According to the CCTST manual (2021), the reliability and validity of the CCTST were established by Insight Assessment. The CCTST meets

> the threshold for strong internal consistency reliability (a minimum alpha of 0.80) and are observed to maintain this performance in all samples of adequate variance. The internal consistency reliability for the individual scales included in any of the mindset measures range from .71 to .80. (the CCTST manual, 2021, p. 47)

According to Facione et al. (2010), the internal consistency reliability KR-20 of CCTST Form 2000 ranged from .78 to .82 (as cited in Soeherman, 2010).

The content validity, construct validity, and criterion (predictive) validity were established respectively. With regard to different language translation versions of CCTST, validity, reliability, and cultural competence are achieved and ensured in all authorized translations by close collaboration with international scholars who are native language speakers (the CCTST manual, 2021). Content validity, construct validity, and criterion (predictive) validity of each of the skills have been supported by the independent and peer-reviewed research publications written by educators, researchers, and doctoral dissertation scholars all over the

world. Soeherman (2010) summarized the content validity of CCTST is established by the consensus of 46 theorists on CT skills sponsored by APA. For criterion validity, Facione et al. (2010) reported the correlation with college GPA as .20, correlation with SAT Verbal score as .55, and correlation with SAT Math score as .44 (Soeherman, 2010). In this non-experimental causal-comparative study, the reliability was checked after the data were collected. Cronbach alpha was calculated to measure reliability, or internal consistency.

Data Analysis

In this study, once the validity and reliability of the instrument were established, parametric tests, such as Student *t* test, Pearson correlation, and ANOVA were used to test the hypotheses proposed in this study. After the participants completed their tests, the researcher downloaded an individual report or a spreadsheet of the group's assessment scores. All data were input into SPSS Statistics 28 to conduct statistical tests. Descriptive statistics were used to determine overall Chinese teachers' CT level (R_1), as well as CT levels of ISEC and non-ISEC teachers respectively (R_2). The independent t-test was used to investigate the difference in the levels of CT between ISEC and non-ISEC teachers (R_3). A two-way ANOVA was used to explore the difference in CT skills between ISEC teachers and non-ISEC teachers, based on gender (R_4). A two-way ANOVA was used to explore differences in CT skills among ISEC teachers and non-ISEC teachers based on their professional titles or ranks (R_5). A Pearson correlation was run to examine the relationship between CT skills and age (R_6). A Pearson correlation was run to examine the relationship between CT skills and years of teaching (R_7). A two-way ANOVA was used to explore the differences in CT skills between ISEC teachers and non-ISEC teachers based on the educational background (bachelor, master, or doctorate) (R_8). A two-way ANOVA will be used to explore the differences in CT skills between ISEC teachers and non-ISEC teachers based on the discipline (R_9). Following a significant two-way ANOVA,

the post hoc test, Tukey's HSD, was employed to determine which pair or pairs of group means significantly differ. According to Privitera (2015), post hoc tests are necessary when there are more than two groups ($k > 2$), "because multiple comparisons are needed" (p. 372). If the ANOVA is not significant, post hoc tests are not needed, because "no group means significantly differ" (Privitera, 2015, p. 372).

***A Priori* Assumptions.** Assumptions need to be made before conducting parametric tests, such as Student t-test, Pearson correlation, and ANOVA**.** There are several *a priori* assumptions for this study:

(1) Levels of Measurement. According to Pallant (2010), the level of measurement for parametric tests should be interval or ratio (continuous), or one dichotomous (categorical) and one continuous dependent variable. In this study, the CT skill scores from the CCTST are dependent variables, belonging to the ratio level of measurement (continuous). Two independent variables of age and years of teaching are both ratio level data (continuous). Other independent variables, such as gender, professional ranks, educational background, and discipline are all perceived as nominal data (categorical /dichotomous).

(2) Linearity. The relationship between the two variables is assumed to be linear.

(3) Random Sampling. An assumption of the ANOVA is random sampling. This study will use a nonprobability sample. This is a limitation of the study.

(4) Independence of Observations. It is assumed that the probability of each measured outcome in this study is independent or equal. In this study, independence of observations will be met through Chinese teachers' independent completing of the online CCTST and individual interviews via Zoom.

(5) Normal Distribution. All ISEC teachers and non-ISEC teachers in Chinese local colleges and universities in the north of China will be covered in this study. Data in the population or population being sampled from are assumed to be normally distributed because of the Central Limit Theorem.

(6) Homogeneity of Variance: This study will use a Levene's test for homogeneity of variance, in order to test whether the variance in scores is the same for each of the levels of teachers' CT skills. Alpha equals .05.

(7) Alpha Level. In this study, an a priori alpha level will be set at .05 to determine statistically significant differences and relationships.

Qualitative

Research Design

This non-experimental causal-comparative study employed an explanatory (sequential) mixed methods design. For the qualitative research paradigm, the grounded theory design was utilized to gather data concerning ISEC and non-ISEC teachers' perception, attitudes, and experiences about CT and CT instruction. The interview protocol, including nine questions, was employed. All the questions in the interview protocol were open-ended. The individual interview was conducted via Zoom, which lasted 40-60 minutes.

The Central Question. According to Creswell and Poth (2018), the qualitative research questions "are open-ended, evolving, and nondirectional" (p. 137). Creswell and Poth (2018) recommended researchers reduce to one central question and several subquestions. The central question in this explanatory mixed methods study was: What is the perception, attitude, and practice regarding CT among the ISEC and non-ISEC teachers in institutions of higher education in the north of China? This central question was explored through three subquestions.

Subquestions. The subquestions "essentially take the central question and break it down into its constituent parts" (Creswell & Poth 2018, p. 140)**.** The three subquestions in this study were: (1) what is the perception of CT among Chinese ISEC and non-ISEC teachers in institutions of higher education in the north of China? (2) What is the attitude toward CT among Chinese ISEC and non-ISEC teachers in institutions of higher education in the north of China? (3) What is the experience of CT teaching among Chinese ISEC and non-ISEC teachers in institutions of higher education in the north of China?

Interview questions. The interview protocol included nine questions. During the interview nine questions were addressed in Chinese (see Appendix D), because all participants were Chinese teachers. These questions were adapted from Xia (2018)'s dissertation.

Q_1 How do you understand critical thinking (CT)? What is your definition of CT?

The first question was designed to gain insights on how ISEC and non-ISEC teachers in China perceive CT and define CT. It was aimed to measure teachers' analysis skills of CT (Facione, 1990a). According to Borg (2003), teacher cognition refers to "the unobservable cognitive dimension of teaching --- what teachers know, believe, and think" (p. 81). As one of the subsets of cognition, perception refers to a higher mental process "by which we extract meaningful information from physical stimulation. It is the way we interpret our sensation" (Sainn & Ugwuegbu, 1980, p. 90). Choy and Cheah (2009) claimed teachers' perceptions affect their behaviors in the classroom. If teachers perceive themselves as disseminators of information, rather than mediators of learning, then "there is little regard for student input and feedback" (Choy & Cheah, 2009, p. 199). Therefore, it is difficult for students to acquire CT skills from such teachers (Choy & Cheah, 2009). Teachers themselves need to understand and define CT in order to help students improve their CT skills (Choy & Cheah, 2009).

Q_2 Do you think CT is helpful or useful for students' learning, exams, and their future development? Why or why not?

This question was developed to learn about ISEC and Non-ISEC teachers' disposition or attitude towards CT cultivation. It was also designed to measure teachers' analysis skills (Facione, 1990a). Holding a positive attitude towards CT is essential for the effective CT cultivation and "a fair and valid assessment of students' critical thinking" (Lloyd & Bahr, 2010, p. 3). According to Stapleton (2011), "any conception of CT must include its dual dimensionality, i.e., skills and disposition …" (p. 16). However, in this mixed methods study, the dominant quantitative paradigm only focused on the CT skills of Chinese ISEC and non-ISEC teachers using CCTST, the CT disposition was not measured. This dispositional or attitudinal question rendered a deeper understanding of CT skills and dispositions.

Q_3 Do you like those students who challenge or question you in your class? How do you deal with such situations in class?

This question was intended to explore the attitude of ISEC and non-ISEC teachers towards their students' CT disposition, together with the ways they handled the situations in class. It was aimed to assess teachers' analysis and inference skills (Facione, 1990a). According to Facione (1990a), questioning, inquisitiveness, and self-confidence regarding a wide range of issues are components of CT disposition. Teachers need to encourage students to pose questions they are curious about or do not know, and thus students' CT disposition is "improved as a result of educational experience" (Bers et al., 1996).

Q_4 What do you think of your students' CT skills, high, medium, or low? How do you assess your students' CT skills?

This question sought to examine how ISEC and non-ISEC teachers evaluated or assessed their students' CT skills. It was designed to measure teachers' evaluation skills (Facione, 1990a). Facione (1990s) elaborated that evaluation skills refer to measuring credibility and the strength of statements or arguments, and justifying one's reasoning built upon relevant evidence, concepts, or standards, etc. Teachers need to assess their students' CT effectively using standardized tests, and thus adjust their CT teaching strategies or interventions accordingly, based on the results of students' CT assessment.

Q_5 What do you think of your own CT skills, high, medium, or low? How do you know?

This question was aimed to understand how ISEC and non-ISEC teachers evaluated or assessed their own CT skills, which was to measure teachers' evaluation skills (Facione, 1990a). According to Darling-Hammond and Bransford (2005), teachers play an important role in the reform effort. Their attitude towards the reform and their instruction affect the results of the reform (Dai et al., 2011). It is necessary for educational leaders and teachers to learn about or understand teachers' CT skills. Based on the assessment results of teachers' CT, leaders need to offer an appropriate CT training program for teachers and teachers themselves need to improve their weak dimensions of CT.

Q_{6a} Have you cultivated your students' CT skills in your course or classroom? If yes, what are your teaching strategies on CT instruction? What difficulties or challenges have you met?

Q_{6b} If no, what are the reasons for it?

This question consisted of two portions, which mainly focused on ISEC and non-ISEC teachers' experiences in their teaching practice. It was aimed to assess teachers' analysis, inference and deductive reasoning skills (Facione, 1990a). Scholars and researchers reach a consensus that CT instruction is vital in the institutions of higher education (Behar-Horenstein &

Niu, 2011). According to Behar-Horenstein and Niu (2011), CT instruction can be implemented from two dimensions: programmatic and instructional. It is imperative to examine the effectiveness of different CT teaching strategies in order to improve students' CT.

Q_7 What do you think of textbooks and so-called standard answers? Recently, there is a popular assignment among senior high school students: Students are required to find the mistakes in their Chinese textbooks, including spellings, punctuation marks, wrong characters, or logical problems in the texts, or what aspects the textbook need to be corrected and improved, etc. What do you think of this assignment? Are you willing to assign such homework to your students? Do you think it is beneficial to students' learning or exams? Why or why not?

This question was designed to understand ISEC and non-ISEC teachers' attitude towards the teaching tools, such as textbooks, assignments. It was to measure teachers' skills of analysis (analyze the scenario), inference (synthesize their ideas), inductive reasoning (draw a conclusion) (Facione, 1990a). Teachers' attitude towards the teaching tools influences the effectiveness of their CT teaching strategies (Asgharheidari & Tahriri, 2015). Teachers need to hold an appropriate attitude in the process of their teaching process.

Q_{8a} Does your college or university emphasize and require the cultivation of students' CT? If yes, how does your school emphasize or require CT instruction or cultivation?

Q_{8b} If no, what recommendations would you like to present to your school?

This question included two parts, which was aimed to examine whether schools of ISEC and non-ISEC teachers value and support CT instruction and cultivation. It was designed to assess teachers' inference and analysis skills (Facione, 1990a). The advocacy and support from schools is a strong motivator for teachers to implement CT instruction.

Q_{9a} Do you have any ideas about how to improve teachers' CT skills?

Q_{9b} Do you think there is a need to participate in the CT training (program)? Why or why not?

This question involved two parts, designed to explore the strategies that ISEC and non-ISEC teachers use to improve their CT skills and their attitude towards CT training. It was used to assess teachers' analysis and inference skills (Facione, 1990a). According to Li (2016) and Zhang et al. (2020), the concept of CT among Chinese EFL teachers is fragmented and insufficient. Many of them do not foster students' CT in their classrooms. Chinese teachers need more CT learning and training (Zhang et al., 2020).

Participants. Purposeful sampling is mainly used in the qualitative study (Creswell, 2012), because such a sample can "purposefully inform an understanding of the research problem and central phenomenon in the study" (Creswell & Poth, 2018, p. 158). The participants in this qualitative design were ISEC and non-ISEC teachers in institutions of higher education in the north of China. According to Guest, et al. (2006), in order to obtain saturation and variability, the minimum number of the participants for the interview should be twelve. In this study, twelve teachers, who were from the 102 volunteer participants in the local or provincial colleges and universities in the north of China, were purposefully selected. The teachers were purposefully selected because they shared some degree of experience of the phenomenon being questioned in this study. Six out of twelve participants were ISEC teachers and the remaining six were non-ISEC teachers.

Data collection

Procedures. The qualitative grounded theory design was used to collect data through interviews via Zoom. According to Creswell and Poth (2018), a grounded theory design "is to move beyond description and to generate or discover a theory" (p. 82). The researcher of this study finally generated a theory of the process of teaching, the teaching action, or their

interactions with students, shaped by Chinese ISEC and non-ISEC teachers. First, the participants were invited to participate in the interview via email and WeChat. If they agreed to be interviewed, the Informed Consent form was delivered to all participants for their signature. The interview was audio and video recorded with the permission of participants. The researcher transcribed all the interviews. During the interview, field notes were taken by the researcher.

Transferability. Transferability was established by showing readers evidence that the results and findings from the research could be transferable or applicable to other contexts or people with similar experiences (Creswell & Poth, 2018). Lincoln and Guba (1985) stressed the researchers themselves could not decide whether the findings can be transferred or not. Instead, it was the readers' decision. According to Creswell and Poth (2018), once trustworthiness was achieved, the researchers may transfer the knowledge gained from the study, if the readers or users had similar experiences and in similar social contexts.

Trustworthiness. In order to establish the trustworthiness of a qualitative study, the data need to be obtained or collected from multiple sources, such as interview, observation, field notes, and focus groups (Creswell & Poth, 2018). In this qualitative grounded theory, data were obtained from the individual interviews, transcriptions, the emerging categories based on the interview transcription, the interviewer's notes, and participants' demographic data.

Accuracy. In this study, accuracy was established by video and audio recording of all interviews. Then interviews were transcribed by the researcher. The researcher reviewed, checked and revised the transcriptions after the initial transcription. Finally, verbatim transcriptions were checked for accuracy and authenticity.

Verification Procedures. Creswell and Poth (2018) put forward three categories of verification: the researcher's lens, the participant's lens, and the reader's lens. In this qualitative

grounded theory design, three verification steps were used. First, the researcher's ideas, experiences, values, assumptions, and bias could influence the research findings. The researcher needed to clarify her own bias first, which was a critical facet of verification for a qualitative study. The researcher of this study came from a northern university in China. She became an ISEC teacher in February, 2017. With 24 years of teaching experiences, she has developed biases and formed opinions about ISEC and non-ISEC teachers in China. For example, before conducting this study, she assumed the CT skills of ISEC teachers must be higher than those of non-ISEC teachers, because ISEC teachers are required to participate in the CT training program by the ISEC office. She understood the challenges and difficulties that ISEC and non-ISEC teachers face and may show empathy towards the participants. The researcher reduced the bias to the minimum through the external audit of her dissertation chair and committee members.

The researcher followed the grounded theory research procedures of Strauss and Corbin (1998) strictly to guarantee the confirmability of the study. Through the stages of data collection and data analysis, rich and thick descriptions were achieved through nine open-ended interview questions, and the participants were encouraged to tell their own stories. During the interview, the researcher asked interviewees for clarification to establish accuracy and authenticity. According to Strauss and Corbin (1998), theoretical saturation is vitally important for qualitative grounded theory methodology. In this qualitative research paradigm, theoretical saturation was achieved when there was no new or relevant data emerging regarding a category, the category was well developed with regard to its properties and dimensions indicating variation, and the relationships among categories were well established and validated (Strauss & Corbin, 1998). Finally, after the researcher developed codes and categories of data, the participants were invited to check and review the codes and categories for consensus, accuracy and authenticity.

Data Analysis

The researcher applied Strauss and Corbin's (1998) procedures for data analysis. During the initial coding stage, open coding was aimed at conceptualizing and categorizing data via two analytic procedures: making constant comparisons and asking questions. This process helped to identify initial categories. In the intermediate stage, axial coding was designed to "begin the process of reassembling data that were fractured during open coding" (Strauss & Corbin, 1998, p. 124). In axial coding, categories were "related to their subcategories to form more precise and complete explanations about the phenomena" (Strauss & Corbin, 1998, p. 124). Major categories or themes emerged from this stage. During the final stage, selective coding was employed to develop and establish central or core categories grounded in data, and thus the theory was generated in a final narrative to explain the detailed process of the phenomenon.

Summary

This was a non-experimental causal-comparative study with an explanatory sequential mixed methods design. This non-experimental causal-comparative study mainly focused on the CT of the ISEC and non-ISEC teachers in China, whether there was a statistically significant difference in the CT level between the ISEC teachers and non-ISEC teachers, as well as how the ISEC and non-ISEC teachers defined, valued, and applied CT in their teaching process. The reliability and validity of the quantitative design were established by Insight Assessment. In this explanatory sequential mixed methods study, Cronbach alpha was calculated to check the reliability, or internal consistency, after the quantitative data were downloaded from the website of Insight Assessment. To guarantee the transferability of the qualitative design, trustworthiness of the data was established through accuracy and verification procedures. These included audio and video recording of all interviews, verbatim transcriptions, member checking, external audit of the dissertation chair and committee. Using these strategies, the results and findings from this

study offered a better understanding of the research problems and questions. The next chapter focuses on Data Analysis.

Chapter Four: Data Analysis

This non-experimental causal-comparative study with an explanatory mixed methods research design was aimed to examine the status quo of CT skills among ISEC and Non-ISEC teachers in institutions of higher education in the north of China, as well as their perception, attitude, and practice of CT in their teaching process. The quantitative method was employed first and was dominant in this study. Then the qualitative grounded theory method was followed. For the quantitative paradigm, the demographic information was collected first. Then the CCTST was employed to explore all participants' overall CT skills and their five core skills of CT: *analysis, inference, evaluation, inductive reasoning,* and *deductive reasoning*. After collecting the quantitative data (Group Score Report) from the account established by Insight Assessment, the researcher of this study initiated its qualitative portion. Twelve interviews were conducted with 6 ISEC and 6 Non-ISEC teachers via Zoom. Both quantitative and qualitative paradigms in this study addressed the research questions and the central question. The research results from both quantitative and qualitative rendered a panoramic view of the Chinese teachers' CT in the north of China, through descriptive statistics, inferential statistics, as well as categories (themes) and the theory generated from the qualitative data. This chapter focused on the data analysis and research findings based on the statistical results and the grounded theory research results. In the following sections, the quantitative data analysis is addressed first, and then followed by the qualitative data analysis.

The Quantitative Data Analysis Procedures

Participants were recruited through a nonprobability snowball sampling, and 102 volunteers, including 52 ISEC and 50 Non-ISEC teachers, participated in the quantitative research. Participants were allowed to complete the Chinese version of CCTST from December

26, 2022 to January 6th, 2023. After 102 participants completed the CCTST, the researcher of this study downloaded the Group Score Report from the Insight Assessment account, on January 6th, 2023. All raw data were input into SPSS Statistics 28 to conduct statistical tests. Both descriptive and inferential statistics were utilized to address all the research questions.

Participants' Demographic Profiles

One hundred and two teachers from 22 colleges and universities in the north of China participated in the research. These teachers provided their demographic information, including name, gender, age, school name, years of teaching, professional rank, educational background, college or department, discipline they teach, whether they have participated in CT training or not, and whether they are ISEC or Non-ISEC teachers. Geographically, 64 participants (62.7%) were from Jilin province, 17 were from Inner Mongolia Autonomous Region (16.7%), 13 were from Shandong province (12.7%), 5 from Liaoning province (4.9%), 2 from Heilongjiang province (2.0%), and 1 from Henan province (1.0%).

There were 37 males (36.3%) and 65 females (63.7%). Among them there were 17 (16.7%) male and 35 (34.3%) female ISEC teachers. Of the Non-ISEC teachers 20 (19.6%) were male and 30 (29.4%) were female. With regard to the professional rank, there were 2 assistants (2.0%), 39 instructors (38.2%), 51 associate professors (50.0%), and 10 professors (9.8%). In terms of their educational background (highest degree earned), there were 2 bachelor's degrees (2.0%), 77 masters (75.5%), 21 doctorates (20.6%), and 2 (2.0%) post doctorates.

Colleges or departments that participants belonged to were divided into 7 groups, because different schools categorized and named their colleges or departments in different ways based on their own characteristics and market demands. These colleges or department were College of Foreign Languages (n = 47, 46.1%) , College of Economics, Management and Law, including

Tourism and Business (n = 24, 23.5%), College of Mathematics and Computer Science, including Information and Technology (n = 11, 10.8%), College of Physical Education (n = 6, 5.9%), College of Liberal Arts, including History and Marxism (n = 6, 5.9%), College of Education and International Education (n = 6, 5.9%), and College of Chemistry, Mathematics and Physics (n = 2, 2.0%).

The disciplines that the participants teach were categorized into 7 groups as well. They were Foreign Languages (n = 47, 46.1%), Economics, Management, Accounting, and Tourism (n = 24, 23.5%), Mathematics, Physics, and Chemistry (n = 8, 7.8%), Computer Science and Technology (n = 6, 5.9%), Physical Education (n = 6, 5.9%), Liberal Arts, including Chinese, Marxism, and Literature (n = 6, 5.9%), Education and International Education (n = 5, 4.9%). The descriptive statistics showed the average age of the participants was 43 (SD = 5.78), and the mean score of years of teaching was 18 (SD = 7.18). Table 5 describes participants' age and years of teaching.

Table 5

Participants' Age and Years of Teaching

	n	Range	Minimum	Maximum	Mean	SD
Age	102	34	26	60	43.01	5.76
Years of Teaching	102	35	1	36	17.81	7.18

The researcher of this study used Cronbach's alpha to check the reliability or internal consistency of the study, because the reliability of a scale "can vary depending on the sample. It is therefore necessary to check that each of your scales is reliable with your particular sample" (Pallant, 2010, p. 97). The value of Cronbach's alpha was .86 in this study, suggesting good

internal consistency or reliability for this group of participants or sample. According to Pallant (2010), values of Cronbach's alpha above .70 are considered acceptable, and values that are above .80 are preferable.

Quantitative Research Questions

In this study, the alpha level, or the level of significance was set at .05. For the quantitative research paradigm, nine research questions were addressed. The researcher first conducted descriptive statistics to address the first two research questions. The independent *t*-test, two-way ANOVA, and Pearson correlation were used to address the remaining seven research questions.

Research Question 1: What is the level of CT skills of overall Chinese teachers in institutions of higher education in the north of China?

The descriptive statistics indicated the mean score of the overall CT skills was 17.34 (SD = 3.66). According to the CCTST user manual (2021), if the overall scores range from 0-7, then they fall into the "Not Manifested" level; if the overall scores are from 8-12, they belong to the "Weak" level; if scores are 13-18, they fall into the "Moderate" level; scores ranging from 19 to 23 belong to the "Strong" level; and 24 or higher falls into the "Superior" level. This criterion showed the overall CT mean scores of Chinese teachers in the north of China fell into the upper range of the "Moderate" level.

The mean scores of the 5 core CT skills---*analysis, inference, evaluation, induction reasoning, and deduction reasoning* were 4.21 (SD = 1.36), 8.62 (SD = 2.44), 4.52 (SD = 1.65), 9.25 (SD = 2.12), 8.10 (SD = 2.35) respectively. Based on the CCTST 34-Point Scale Score Interpretation in the user manual (2021), the mean value of *analysis* skill of Chinese teachers fell into the range between Moderate (3-4) and Strong (5 or more), belonging to the lower range of

the "Strong" level. The *inference* skill of Chinese teachers fell into the middle range of the "Moderate" level (6-11). The *evaluation* skill was in the lower range of the "Moderate" level (4-7). The *induction reasoning* fell into the medium range of "Moderate" level (6-11), and the *deductive reasoning* was in the medium range of "Moderate" level (6-11) as well (see Table 6).

Table 6

Five Core CT Skills of Chinese Teachers

Five Core CT Skills	Analysis	Inference	Evaluation	Induction	Deduction
Mean	4.21	8.62	4.52	9.25	8.10
SD	1.36	2.44	1.65	2.12	2.35
Level of CT skills	Lower range of the "Strong" level	Middle range of the "Moderate" level	Lower range of the "Moderate" level	Medium range of the "Moderate" level	Medium range of the "Moderate" level

Research Question 2: What is the level of CT skills of the ISEC and non-ISEC teachers in institutions of higher education in the north of China, respectively?

The descriptive statistics (Table 7) showed the mean score of the overall CT skills of the ISEC teachers was 17.39 (SD = 3.63), and for the non-ISEC, the mean score was 17.30 (SD = 3.74). The overall score of ISEC teachers was slightly higher than that of non-ISEC teachers (see Table 7). The mean values of ISEC teachers' five core CT skills: *analysis, inference, evaluation, induction* and *deduction*, were 4.17 (SD = 1.32), 8.39 (SD = 2.49), 4.83 (SD = 1.57), 9.46 (SD = 2.11), and 7.92 (SD = 2.22), compared to those of non-ISEC teachers' five skills: 4.24 (SD = 1.41), 8.86 (SD = 2.39), 4.20 (SD = 1.68), 9.02 (SD = 2.13), and 8.28 (SD = 2.48). With regard to *analysis*, *inference*, and *deduction*, the mean scores of non-ISEC teachers were slightly higher than those of ISEC teachers. The other two core skills: *evaluation* and *induction*, the mean scores of ISEC teachers were higher than those of non-ISEC teachers (see Table 8).

Table 7

Overall CT Scores of ISEC and Non-ISEC Teachers

Overall Scores	n	Range	Minimum	Maximum	Mean	SD
ISEC	52	18	9	27	17.39	3.63
Non-ISEC	50	18	8	26	17.30	3.74

Table 8

Scores of Five Core Skills of ISEC and Non-ISEC Teachers

Five Core skills	Group	n	Range	Minimum	Maximum	Mean	SD
Analysis	ISEC	52	6	1	7	4.17	1.32
	Non-ISEC	50	6	1	7	4.24	1.41
Inference	ISEC	52	11	3	14	8.39	2.49
	Non-ISEC	50	9	4	13	8.86	2.39
Evaluation	ISEC	52	7	2	9	4.83	1.57
	Non-ISEC	50	6	1	7	4.20	1.68
Induction	ISEC	52	11	5	16	9.46	2.11
	Non-ISEC	50	9	4	14	9.02	2.13
Deduction	ISEC	52	11	3	14	7.92	2.22
	Non-ISEC	50	11	3	14	8.28	2.48

Research Question 3: Is there a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers in the north of China?

An independent t-test was conducted to compare the overall CT scores and five scores of the core CT skills for the ISEC and non-ISEC teachers. Although the mean overall CT score of the ISEC teachers was higher than that of non-ISEC teachers, there was no statistically significant difference in the overall CT scores for the ISEC teachers (M = 17.39, SD = 3.63) and

non-ISED teachers (M = 17.30, SD = 3.74; $t_{(100)}$ = .12, p = .91, two-tailed). The researcher of this study failed to reject the null hypothesis. The magnitude of the differences in the means (mean difference = .08, 95% CI: -1.36 to 1.53) was very small (Cohen d = .02). The effect size was interpreted based on Cohen's (1988) rule of thumb: a value of 0.2 represents a small effect size; a value of 0.5 represents a medium effect size, and a value of 0.8 represents a large effect size.

In terms of *analysis, inference*, and *deduction,* although the scores of non-ISEC teachers were slightly higher than those of the ISEC teachers, there were no statistically significant differences in scores between ISEC teachers (*analysis*: M = 4.17, SD = 1.32; *inference*: M = 8.39, SD = 2.49; *deduction*: M =7.92, SD = 2.22) and non-ISEC teachers (*analysis*: M = 4.24, SD = 1.41; $t_{(100)}$ = -.25, p = .81, two-tailed; *inference*: M = 8.86, SD = 2.39; $t_{(100)}$ = -.98, p = .33, two-tailed; *deduction*: M =8.28, SD = 2.48; $t_{(100)}$ = -.77, p = .45, two-tailed). The magnitudes of the differences in the means of *analysis* (mean difference = -.07, 95% CI: -.60 to .47), *inference* (mean difference = -.48, 95% CI: -1.44 to .48), and *deduction* (mean difference = -.36, 95% CI: -1.28 to .57) were very small or small (*analysis*: Cohen d = .05; *inference*: Cohen d = .20; *deduction*: Cohen d = .15). For the other two core skills: *evaluation* and *induction*, the scores of the ISEC teachers were higher than those of non-ISEC teachers, but there were no statistically significant differences in scores for the ISEC teachers (*evaluation*: M = 4.83, SD = 1.57; *induction*: M =9.46, SD = 2.11) and non-ISEC teachers (*evaluation*: M = 4.20, SD = 1.68; $t_{(100)}$ = 1.95, p = .054, two-tailed; *induction*: M =9.02, SD = 2.13; $t_{(100)}$ = 1.05, p = .30, two-tailed). The magnitudes of the differences in the means of *evaluation* (mean difference = .63, 95% CI: -.01 to 1.26) and *induction* (mean difference = .44, 95% CI: -.39 to 1.28) were small because *evaluation*: Cohen d = .39 and *induction*: Cohen d = .21. Table 9 contained the results of the independent t-test.

Table 9

Independent Samples t-test of CT Scores for ISEC and Non-ISEC teachers

	Group	Mean	SD	Mean difference	t	Sig.	Cohen d	95% CI Lower	95% CI Upper
Overall	ISEC	17.39	3.63						
	Non-ISEC	17.30	3.74	.08	.12	.91	.02	-.1.36	1.53
Analysis	ISEC	4.17	1.32						
	Non-ISEC	4.24	1.41	-.07	-.25	.81	.05	-.60	.47
Inference	ISEC	8.39	2.49						
	Non-ISEC	8.86	2.39	-.48	-.98	.33	.20	-.1.44	.48
Evaluation	ISEC	4.83	1.57						
	Non-ISEC	4.20	1.68	.63	1.95	.054	.39	-.01	1.26
Induction	ISEC	9.46	2.11						
	Non-ISEC	9.02	2.13	.44	1.05	.30	.21	-.39	1.28
Deduction	ISEC	7.92	2.22						
	Non-ISEC	8.28	2.48	-.36	-.77	.45	.15	-1.28	.57

Research Question 4: Is there a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers, based on gender (identifying as male and female)?

A two-way between-groups analysis of variance (ANOVA) was conducted to explore the difference in CT skills between the ISEC teachers and non-ISEC teachers, based on gender. According to the descriptive statistics (see Table 10), the overall mean score of male teachers (M = 17.11, SD = 3.99, n = 37) was lower than that of female teachers (M = 17.48, SD = 3.49, n = 65). The mean value of male ISEC teachers (M = 17.65, SD = 3.35, n = 17) was higher than that of female ISEC teachers (M = 17.26, SD = 3.79, n = 35). The mean score of male non-ISEC teachers (M = 16.65, SD = 4.50, n = 20) was lower than that of female non-ISEC teachers (M =

17.73, SD = 3.14, n = 30). The mean score of female non-ISEC teachers ranked the highest, and then followed by male ISEC and female ISEC. The male non-ISEC teachers ranked the lowest. The interaction effect between gender and two teacher groups was not statistically significant, $F_{(1, 98)} = .93$, $p = .34$, and effect size was very small (partial eta squared = .01). There was no statistically significant main effect for two groups of teachers: ISEC and non-ISEC, $F_{(1, 98)} = .12$, $p = .73$. The main effect for gender, $F_{(1, 98)} = .21$, $p = .65$, did not reach statistical significance, either (see Table 11). The researcher of this study failed to reject the null hypothesis.

Table 10

Descriptive Statistics of Overall CT Scores for ISEC and Non-ISEC Teachers Based on Gender

Group	Gender	Mean	SD	N
Total	Male	17.11	3.99	37
	Female	17.48	3.49	65
	Total	17.34	3.66	102
ISEC	Male	17.65	3.35	17
	Female	17.26	3.79	35
	Total	17.39	3.63	52
Non-ISEC	Male	16.65	4.50	20
	Female	17.73	3.14	30
	Total	17.30	3.74	50

Table 11

Two-Way ANOVA of Overall CT Scores for ISEC and Non-ISEC Teachers Based on Gender

Source	F	Sig.	Partial Eta Squared
ISECNon-ISEC	.12	.73	.00
Gender	.21	.65	.00
ISECNon-ISEC*Gender	.93	.34	.01

*Indicates interaction effect

Research Question 5: Are there any statistically significant differences in CT skills between the ISEC teachers and non-ISEC teachers based on professional rank?

A two-way ANOVA was performed to examine the differences in CT skills between the ISEC teachers and non-ISEC teachers based on professional rank. Because there were only two assistant ISEC teachers, the rank of assistant was not considered in the two-way ANOVA. With respect to descriptive statistics (see Table 12), the overall mean score of instructors (M = 17.54, SD = 3.93, n = 39) was the highest, and then followed by that of associate professors (M = 17.26, SD = 3.27, n = 51) and professors (M = 16.30, SD = 3.97, n = 10). The mean value of ISEC professors (M =18.00, SD = 2.16) was higher than that of non-ISEC professors (M =15.17, SD = 4.67). Professors were not further considered because there were only 4 ISEC professors and 6 non-ISEC professors. The mean value of ISEC instructors (M = 18.09, SD = 3.91, n = 22) was higher than that of ISEC associate professors (M = 16.33, SD = 3.09, n = 24). The magnitude of the differences in the means (mean difference = 1.76) was medium (Cohen *d* = .5). Therefore, the difference between these groups had some practical significance (Pallant, 2010). The mean value of non-ISEC instructors (M = 16.82, SD = 3.96, n = 17) was lower than that of non-ISEC associate professors (M = 18.07, SD = 3.27, n = 27). The magnitude of the differences in the means (mean difference = 1.25) was small (Cohen *d* = .3). The mean score of ISEC instructors was higher than that of non-ISEC instructors. The magnitude of the differences in the means (mean difference = 1.27) was small (Cohen *d* = .3). While the mean score of ISEC associate professors was lower than that of non-ISEC associate professors. The magnitude of the differences in the means (mean difference = 1.74) was medium (Cohen *d* = .6). The difference between these groups had some practical significance (Pallant, 2010).

Table 13 indicated that the interaction effect between three professional ranks and two teacher groups was not statistically significant, $F_{(2, 94)} = 2.87$, $p = .06$, and effect size was very small (partial eta squared = .01). There was no statistically significant main effect for two groups of teachers: ISEC and non-ISEC, $F_{(1, 94)} = .73$, $p = .39$. The main effect for the professional rank, $F_{(1, 94)} = .24$, $p = .79$, did not reach statistical significance, either (see Table 13). The researcher of this study failed to reject the null hypothesis. Although there were no statistically significant differences in CT skills for the ISEC teachers and non-ISEC teachers based on the professional rank, the practical significance did exist between ISEC associate professors and non-ISEC associate professors, as well as ISEC instructors and ISEC associate professors.

Table 12

Descriptive Statistics of Overall CT Scores for ISEC and Non-ISEC Teachers Based on the Professional Rank

Group	Professional Rank	Mean	SD	N
Total	Instructor	17.54	3.93	39
	Associate prof	17.26	3.27	51
	Professor	16.30	3.97	10
	Total	17.27	3.59	100
ISEC	Instructor	18.09	3.91	22
	Associate prof	16.33	3.09	24
	Professor	18.00	2.16	4
	Total	17.24	3.48	50
Non-ISEC	Instructor	16.82	3.96	17
	Associate prof	18.07	3.27	27
	Professor	15.17	4.47	6
	Total	17.30	3.74	50

Table 13

Two-Way ANOVA of Overall CT Scores for ISEC and Non-ISEC Teachers Based on the Professional Rank

Source	F	Sig.	Partial Eta Squared
ISECNon-ISEC	.73	.39	.01
ProfessionalRank	.24	.79	.01
ISECNon-ISEC* ProfessionalRank	2.87	.06	.06

*Indicates interaction effect

Research Question 6: Is there a statistically significant relationship between CT skills and age?

According to Pallant (2010), before conducting a correlation analysis, it is necessary to generate a scatterplot. The scatterplot helps to check for violation of the assumptions of linearity and homoscedasticity, as well as offering "a better idea of the nature of the relationship between variables" (Pallant, 2010. p. 129). Figure 3 provides the distribution of the two variables: overall CT skill scores and age. The data points were spread far apart and lacked linear consistency, suggesting a very low correlation. The upward or downward direction of this relationship is hard to distinguish based on this scatterplot. Then the relationship between overall CT scores and age was investigated performing Pearson product-moment correlation coefficient. Preliminary analyses were performed to ensure no violation of the assumptions of normality, linearity and homoscedasticity. The Pearson correlation coefficient was r = -.17, n = 102, p = .09. Cohen (1988) suggested such guidelines: small relationship, if r = .10 to .29; medium relationship, if r = .30 to .49; and large if r = .50 to 1.0. Therefore, there was a weak, negative correlation between the two variables: overall CT skill scores and age. According to Privitera (2015), a negative correlation means that the values of two factors change in the opposite direction (as the values of

one factor increase/decrease, the values of the second factor decrease/increase). The older participants generally had lower CT scores and vice versa. There was no statistically significant relationship between overall CT scores and age, because the *p* value was larger than .05. The coefficient of determination r^2 = .03, which meant only 3% of the variance in overall CT scores could be explained by age (See Table 14). The researcher of this study failed to reject the null hypothesis.

Figure 3

Relationship between Overall CT Scores and Age

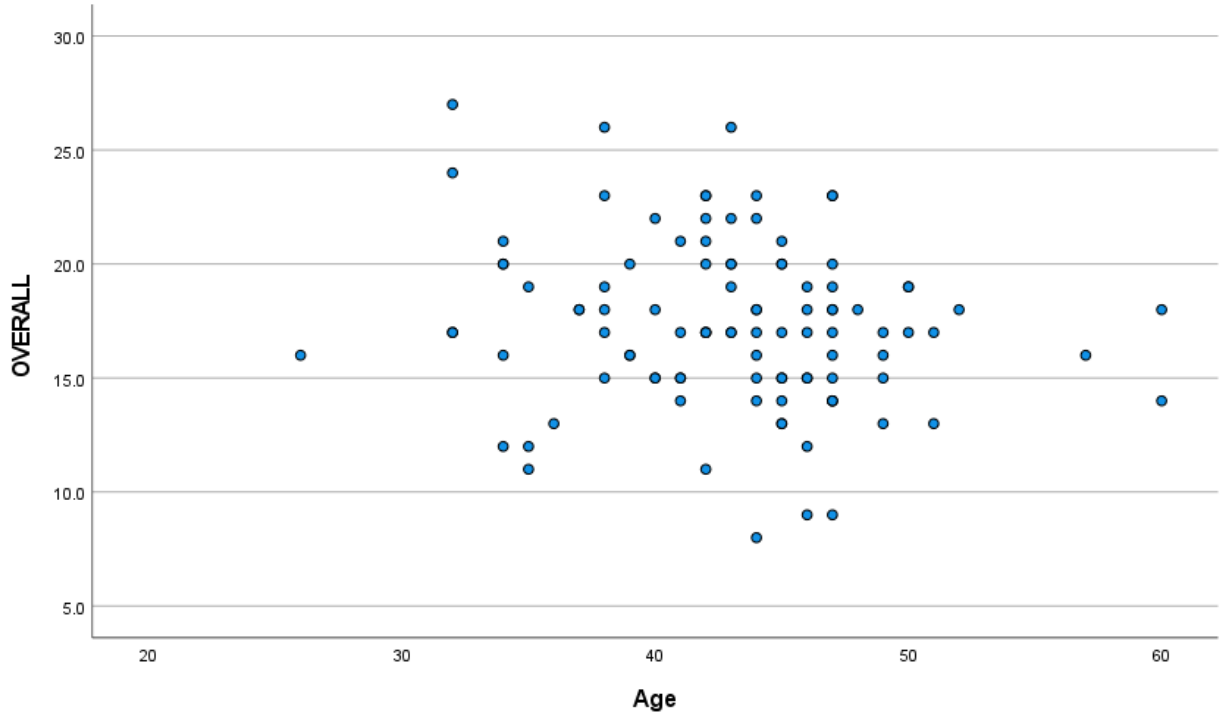


Table 14

Pearson Correlation between Overall CT Scores and Age

		Overall CT Scores	Age
Overall CT Scores	Pearson correlation	1	-.17
	Sig. (2-tailed)		.09
	n	102	102

The relationships between the five core CT skill scores and age were investigated. Table 15 indicated there was a weak, negative correlation between analysis and age (r = -.12, p = .23), inference and age (r = -.10, p = .31), evaluation and age (r = -.13, p = .21), induction and age (r = -.24, p = .02), as well as deduction and age (r = -.05, p = .66). The correlation or relationship between induction and age ($p < .05$) was statistically significant. The coefficient of determination r^2 = .06, which meant only 6% of the variance in induction scores could be explained by age. The remaining correlations were not statistically significant.

Table 15

Pearson Correlations between Five Core CT Skill Scores and Age

		Analysis	Inference	Evaluation	Induction	Deduction
Age	Pearson correlation	-.12	-.10	-.13	-.24	-.05
	Sig. (2-tailed)	.23	.31	.21	.02*	.66
	n	102	102	102	102	102

* Correlation is significant at the .05 level (two-tailed)

Research Question 7: Is there a statistically significant relationship between CT skills and years of teaching?

Figure 4 was the scatterplot, which described the distribution of overall CT skill scores and years of teaching. The data points were spread far apart and lacked any linear consistency, which suggested a very low correlation. The Pearson product-moment correlation indicated that correlation coefficient r = -.18, n =102, p = .07 (See Table 16). There was a weak, negative correlation between overall CT scores and years of teaching, demonstrating that participants with more years of teaching generally had lower CT skill scores and vice versa. A statistically significant relationship did not exist between these two variables: overall CT scores and years of teaching ($p > .05$). The coefficient of determination r^2 = .03, showed that only 3% of the

variance in overall CT scores could be explained by years of teaching. The researcher of this study failed to reject the null hypothesis.

Figure 4

Relationship between Overall CT Scores and Years of Teaching

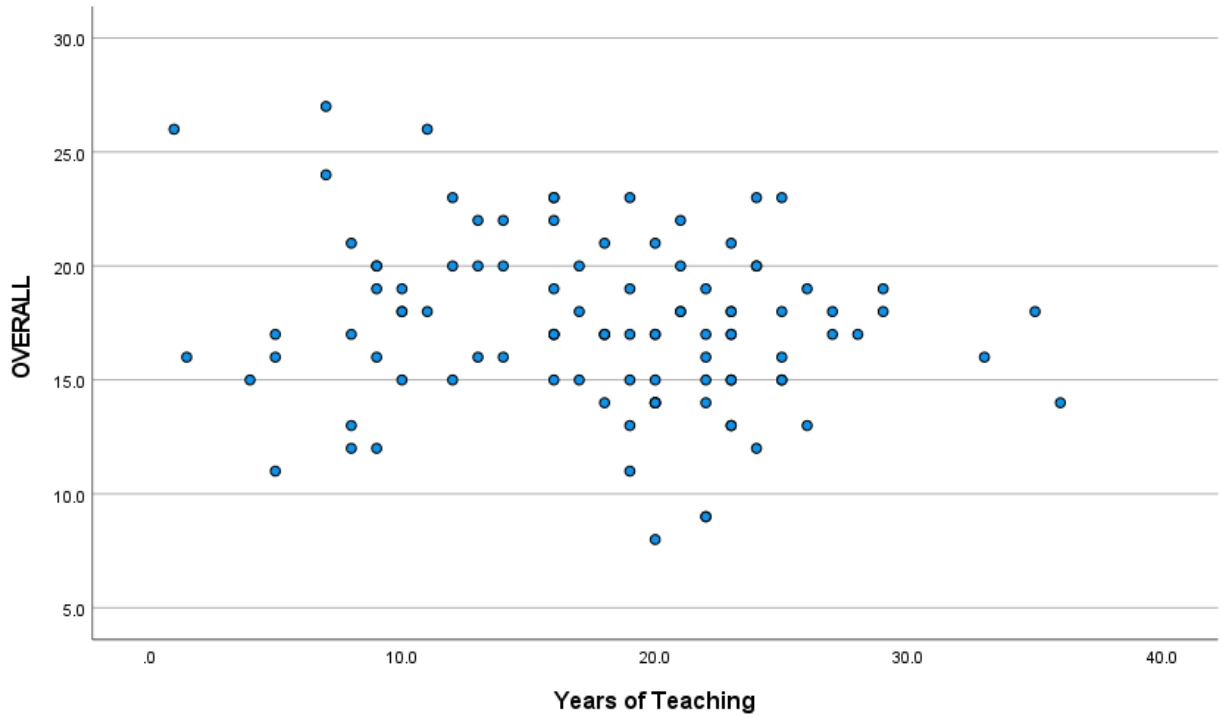


Table 16

Pearson Correlation between Overall CT Scores and Years of Teaching

		Overall CT Scores	Years of Teaching
Overall CT Scores	Pearson correlation	1	-.18
	Sig. (2-tailed)		.07
	n	102	102

Table 17 indicated there was a weak, negative correlation between *analysis* and years of teaching (r = -.10, p = .32), *inference* and years of teaching (r = -.14, p = .17), *evaluation* and years of teaching (r = -.12, p = .23), *induction* and years of teaching (r = -.20, p = .04), as well as *deduction* and years of teaching (r = -.10, p = .32). Although the relationship between *induction*

and years of teaching ($p < .05$) was statistically significant, the coefficient of determination $r^2 = .04$ indicated that only 4% of the variance in *induction* scores could be explained by years of teaching. There were no statistically significant relationships with regard to the remaining four correlations.

Table 17

Pearson Correlations between Five Core CT Skill Scores and Years of Teaching

		Analysis	Inference	Evaluation	Induction	Deduction
Years of Teaching	Pearson correlation	-.10	-.14	-.12	-.20	-.10
	Sig. (2-tailed)	.33	.17	.23	.04*	.32
	n	102	102	102	102	102

* Correlation is significant at the .05 level (two-tailed)

Research Question 8: Is there a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers based on the educational background?

A two-way ANOVA was conducted to explore the differences in CT skills between the ISEC teachers and non-ISEC teachers based on the educational background. Because there were only two teachers with bachelor's degrees and two teachers with post doctorate experiences, bachelors and post doctorates were not considered in the two-way ANOVA. Descriptive statistics indicated (see Table 18) the overall mean score of ISEC teachers (M = 17.33, SD = 3.47, n = 49) was slightly higher than that of non-ISEC teachers (M = 17.29, SD = 3.77, n = 49), with regards to the educational background. The mean score of ISEC teachers with master degrees (M = 17.81, SD = 3.46, n = 37) was higher than that of non-ISEC teachers with master degrees (M = 17.38, SD = 3.75, n = 40). The magnitude of the differences in the means (mean difference = 0.43) was very small (Cohen *d* = .1). The difference between these groups had little practical significance. The mean score of non-ISEC teachers with doctoral degrees (M = 16.89, SD = 4.11, n = 9) was higher than that of ISEC teachers with doctorates (M = 15.83, SD = 3.17,

n = 12). The magnitude of the differences in the means (mean difference = 1.06) was small (Cohen *d* = .3).

Table 19 showed the interaction effect between two educational background groups and two teacher groups was not statistically significant, $F_{(1, 94)} = .69$, p = .41, and the effect size was very small (partial eta squared = .007). There was no statistically significant main effect for two groups of teachers: ISEC and non-ISEC, $F_{(1, 94)} = .12$, p = .73. The effect size was very small (partial eta squared = .001). The main effect for the educational background, $F_{(1, 94)} = 1.89$, p = .17, did not reach statistical significance, either. The effect size was quite small (partial eta squared = .02), as well. The researcher of this study failed to reject the null hypothesis.

Table 18

Descriptive Statistics of Overall CT Scores for ISEC and Non-ISEC Teachers Based on the Educational Background

Educational Background	Group	Mean	SD	n
Total	ISEC	17.33	3.47	49
	Non-ISEC	17.29	3.77	49
	Total	17.31	3.61	98
Mater	ISEC	17.81	3.46	37
	Non-ISEC	17.38	3.75	40
	Total	17.58	3.60	77
Doctorate	ISEC	15.83	3.19	12
	Non-ISEC	16.89	4.11	9
	Total	16.29	3.55	21

Table 19

Two-Way ANOVA of Overall CT Scores for ISEC and Non-ISEC Teachers Based on the Educational Background

Source	F	Sig.	Partial Eta Squared
ISECNon-ISEC	.12	.73	.00
EducationalBack.	1.89	.17	.02
ISECNon-ISEC* EducationalBack	.69	.41	.01

*Indicates interaction effect

Research Question 9: Is there a statistically significant difference in CT skills between the ISEC teachers and non-ISEC teachers based on the discipline?

A two-way ANOVA was conducted to investigate the differences in CT skills between the ISEC teachers and non-ISEC teachers based on the discipline. Descriptive statistics showed (see Table 20) that there were 24 ISEC and 2 non-ISEC teachers who taught Economics, Management, Accounting, and Tourism. There were 16 ISEC and 31 non-ISEC teachers who taught Foreign Languages. Five ISEC and 3 non-ISEC teachers teaching Math, Physics and Chemistry. Four ISEC and one non-ISEC teachers teaching Education. Four ISEC and two non-ISEC teachers teaching Liberal arts (Chinese, Marxism, Literature). There were 6 non-ISEC teachers who taught Physical Education, but no ISEC teachers taught Physical Education. Given the uneven distribution of the disciplines among participants, disciplines were divided into two general groups: Science and Arts. Science consisted of Computer Science and Technology, Math, Physics and Chemistry, Economics, Management, Accounting and Tourism, and Physical education. Arts included Foreign Language, Education, and Liberal Arts.

Table 20 described descriptive statistics of the overall CT scores for ISEC and non-ISEC teachers based on two discipline groups. It indicated the overall CT scores of ISEC teachers (M = 17.39, SD = 3.63, n = 52) was slightly higher than that of non-ISEC teachers (M = 17.30, SD = 3.74, n = 50). The mean score of Arts ISEC teachers (M = 16.75, SD = 3.30, n = 24) was lower

than that of Arts non-ISEC teachers (M = 17.12, SD = 3.30, n = 34). The mean value of Science ISEC teachers (M = 17.93, SD = 3.86, n = 28) was higher than that of Science non-ISEC teachers (M = 17.69, SD = 4.63, n = 16). The difference between these groups had little practical significance (Pallant, 2010).

Table 21 showed that the interaction effect between two discipline groups and two teacher groups was not statistically significant, $F_{(1, 98)} = .16$, $p = .69$, and effect size was very small (partial eta squared = .002). There was no statistically significant main effect for two groups of teachers: ISEC and non-ISEC, $F_{(1, 98)} = .01$, $p = .93$. The effect size was very small (partial eta squared = .000). The main effect for the educational background, $F_{(1, 98)} = 1.33$, $p = .25$, did not reach statistical significance, either, and the effect size was quite small (partial eta squared = .01). The researcher of this study failed to reject the null hypothesis.

Table 20

Descriptive Statistics of Overall CT Scores for ISEC and Non-ISEC Teachers Based on the Discipline

Group	Discipline	Mean	SD	n
Total	Arts	16.97	3.28	58
	Science	17.84	4.10	44
	Total	17.34	3.66	102
ISEC	Arts	16.75	3.30	24
	Science	17.93	3.86	28
	Total	17.39	3.63	52
Non-ISEC	Arts	17.12	3.30	34
	Science	17.69	4.63	16
	Total	17.30	3.74	50

Table 21

Two-Way ANOVA of Overall CT Scores for ISEC and Non-ISEC Teachers Based on the Discipline

Source	F	Sig.	Partial Eta Squared
ISECNon-ISEC	.01	.93	.00
DisciplineGroup2.	1.33	.25	.01
ISECNon-ISEC* DsciplineGroup2	.16	.69	.00

*Indicates interaction effect

The Qualitative Data Analysis Procedures

For the qualitative research paradigm, the grounded theory model introduced by Strauss and Corbin (1998), was utilized to analyze data through interview protocol, field notes and verbatim transcriptions. The emerging themes were identified and the theory was generated by means of asking questions and making constant comparisons. This qualitative grounded theory model was appropriate because it concentrated on the process related to the topic. The overarching goal of data analysis in the grounded theory is to construct theory (Vollstedt & Rezat, 2019). The researcher of this study explored the process concerning the participants' actions in an effort to generate theory about the process (Leedy & Ormrod, 2013).

Participant Interviewees' Demographic Profiles

With the purpose of obtaining deep, detailed, and comprehensive information about the ISEC and non-ISEC teachers' perceptions, attitudes, and experiences of CT in institutions of higher education in the north of China, the researcher collected qualitative data through 12 interviews with 12 volunteer teachers, including six ISEC and six non-ISEC teachers. The participants were first purposefully selected and then volunteered. The interviews were carried out shortly after collecting the quantitative data, and they were conducted on Zoom. All

interviews were audio and video recorded via Zoom and transcribed by the researcher. Each interview addressing nine questions lasted from 40 to 60 minutes. To ensure all participants understood the questions completely and to gain more accurate information, all interviews were conducted in Mandarin. The recorded interviews were transcribed verbatim in Chinese by the researcher and member checked by each teacher participant for accuracy to guarantee the interview was congruent with what the teacher participants wanted to convey. Then, the researcher analyzed each interview transcript and the field notes using the grounded theory methodology. Table 22 described the demographic information of 12 participants.

Table 22

Demographic Information of Non-ISEC and ISEC Teacher Participants

Interview-ee	Age	Gender	Educational Background	Professional Rank	Years of Teaching	ISEC or Non-ISEC	CT Scores
P1	43	M	PhD	Associate prof.	1	Non-ISEC	26
P2	43	F	PhD	Associate prof	17	Non-ISEC	20
P3	46	F	MA	Associate prof	22	Non-ISEC	20
P4	45	F	MA	Associate prof	24	Non-ISEC	20
P5	46	F	MA	Associate prof	24	Non-ISEC	15
P6	49	M	PhD	Prof	25	Non-ISEC	15
P7	33	F	MA	Instructor	8	ISEC	24
P8	45	M	PhD	Prof.	23	ISEC	21
P9	34	F	MA	Associate prof	9	ISEC	20
P10	44	F	MA	Associate prof	21	ISEC	18
P11	44	M	PhD	Instructor	14	ISEC	16
P12	40	F	MA	Instructor	16	ISEC	15

Explanation of the Steps to Identify Categories

According to Strauss and Corbin (1998), there are three kinds of coding procedures that are vital to develop a theory grounded from data: open, axial, and selective coding. Following Strauss and Corbin's ideas, the researcher utilized open coding, axial coding, and selective coding to analyze all data, which "consisted of processing the data by naming and categorizing concepts and verifying concepts to each other until reaching the core category [or central category]" (Williams, 2017, p. 69). These procedures are not "precise procedures that are easily distinguishable. On the contrary, the procedures are neither clear-cut, nor do they easily define phrases that chronologically come one after the other" (Vollstedt & Rezat, 2019, p. 86). Following Strauss and Corbin, the three procedures or phases of data analysis were exemplified to demonstrate the most concrete guide to the grounded theory methodology.

Open Coding. Although coding phases do not occur in a strict sequence, open coding is always the first phase of data analysis. It means raw data are "broken down into discrete parts, closely examined, and compared for similarities and differences" (Strauss & Corbin, 1998, p. 102). This analytic process is the key to the grounded theory, because data need to be analyzed line by line for all transcriptions and field notes, as well as comparing the documents with the original audio or video recording for accuracy and authenticity. According to Strauss and Corbin (1998), open coding was aimed at conceptualizing and categorizing data via two analytic procedures: making constant comparisons and asking questions. The constant comparisons and questions were used to explore commonalities and differences. This process helped to identify initial categories, and eight emergent categories emerged in the stage of open coding: varied but fragmented understanding of CT; positive and supportive attitude towards CT; informal assessment on CT; application of teaching strategies; student-oriented challenges and difficulties

in the CT teaching process; Opinions of textbooks, their uses, and contents; Responses to school support on CT instruction; and usefulness of CT training. After the categories were "formed in the open coding, they are fleshed out in terms of their given properties and dimensions" (Brown et al., 2002, p. 176). Table 23 indicates the initial eight categories and corresponding codes.

Table 23

Initial Categories and Codes

Categories	Codes
Varied but fragmented understanding of CT	to be skeptical; to be critical; to be rational; to have independent and active thinking; to challenge and question; to have logical reasoning ability; to make a judgment or decision; to solve problems
Positive and supportive attitude towards CT	to be beneficial to students' learning, exams (especially with open questions), and future development; to offer a welcoming attitude to students' challenging and questioning; provide encouragement to students
Informal assessment on CT	medium or low level of students' CT; students' engagement in class; questions posed and answered; assignments completed; academic performance; medium level of teachers' CT; teachers' teaching practice; teachers' lived experiences
Application of CT teaching strategies	group discussion; open questions; presentation; writing assignment
Student-oriented challenges and difficulties	students' individual quality; students' non-cooperation; students' personality trait; students' attitude; lack of teacher- or –discipline-oriented factors
Opinions of textbooks, their uses, and contents	authoritative status of textbooks; usefulness of textbooks; textbooks as reference; conflicting responses to standardized answers; varied attitudes to the assignment of identifying flaws in the textbook

Responses to CT support from schools	no emphasis on CT instruction vs stress on CT instruction; obedience to authority
Usefulness of CT training	necessity of improving teachers' CT; feasible strategies; advocacy of effective CT integration into teaching practice

Axial Coding. As the second phase of data coding, axial coding "worked congruently with open coding" (Williams, 2017, p. 74). According to Strauss and Corbin (1998), the purpose of axial coding is to "begin the process of reassembling data that were fractured during open coding" (p. 124). They stressed that the focus of axial coding is to work out the relationships or links between concepts and categories emerged from the open coding process. Put another way, through the process of axial coding, data were continually compared, put together again by related concepts, and built connections linking categories in terms of properties and dimensions. Strauss and Corbin suggested that using a coding paradigm helps to build links between categories, which includes the following basic components: conditions (causal, intervening, and contextual), action/interaction, and consequences. According to Brown et al (2002), there are four analytical processes in axial coding: "(a) continually relating subcategories to a category, (b) comparing categories with the collected data, (c) expanding the density of the categories by detailing their properties and dimensions, and (d) exploring variations in the phenomena" (p. 177).

The researcher, in this phase, developed and refined the categories by merging or deleting some concepts after making possible relationships. New data continued to be coded, reexamined and compared until saturation was achieved, which meant no new or relevant data emerged from the extant data, the category was well developed, and the relationships among categories were well established. Three major categories or themes emerged through this stage: (a) barriers to CT

instruction, based on the subcategories of exam-oriented educational system, not enough support from schools, lack of professional knowledge of CT, student-oriented difficulties and challenges; (b) inadequacy of CT instruction, built upon the subcategories of adoption of informal CT assessment, limited application of CT teaching strategies, attitudes towards textbooks and standard answers; and (c) necessity of promoting CT, based upon the subcategories of the medium or low level of students' CT, the medium level of teachers' CT, imperative of CT training, and need of further school support. Table 24 described the refined categories and their subcategories.

Table 24

Refined Categories and Subcategories

Categories	Subcategories
Barriers to CT instruction	-exam-oriented educational system; -not enough support from schools; -lack of professional knowledge of CT; -student-oriented difficulties and challenges
Inadequacy of CT instruction	-adoption of informal CT assessment; -limited application of CT teaching strategies; -attitudes towards textbooks and standardized answers
Necessity of promoting CT	-the medium or low level of students' CT; -the medium level of teachers' CT; -imperative of CT training; -need of further school support

Selective Coding. As the final phase of data analysis, selective coding builds upon the first two phases of open coding and axial coding. According to Strauss and Corbin (1990), selective coding is "the process of selecting the central or core category, systematically relating it to other categories, validating those relationships, and filling in categories that need further

refinement and development" (p. 116). The core category is "the central phenomenon within which all the other categories are integrated" (Strauss & Corbin, 1990, p. 116). Strauss and Corbin (1998) elaborated that the central idea or core category has "the analytic power [to] pull the other categories together to form an explanatory whole, [and] should be able to account for considerable variation within categories" (p. 146). During the stage of selective coding, the core category was chosen systematically using Strauss and Corbin's (1998) procedures of data analysis. Put another way, this study conformed to the following six criteria presented by Strauss and Corbin (1998), in order to establish the central or core category.

1. It must be central; that is, all major categories can be related to it.

2. It must appear frequently in the data. This means that within all, or almost all, cases, there are indicators pointing to that concept.

3. The explanation that evolves by relating the categories is logical and consistent. There is no forcing of data.

4. The name and phrase used to describe the central category should be sufficiently abstract that it can be used to do research in other substantive areas leading to the development of a more general theory.

5. As the concept is refined analytically through integration with other concepts, the theory grows in depth and explanatory power.

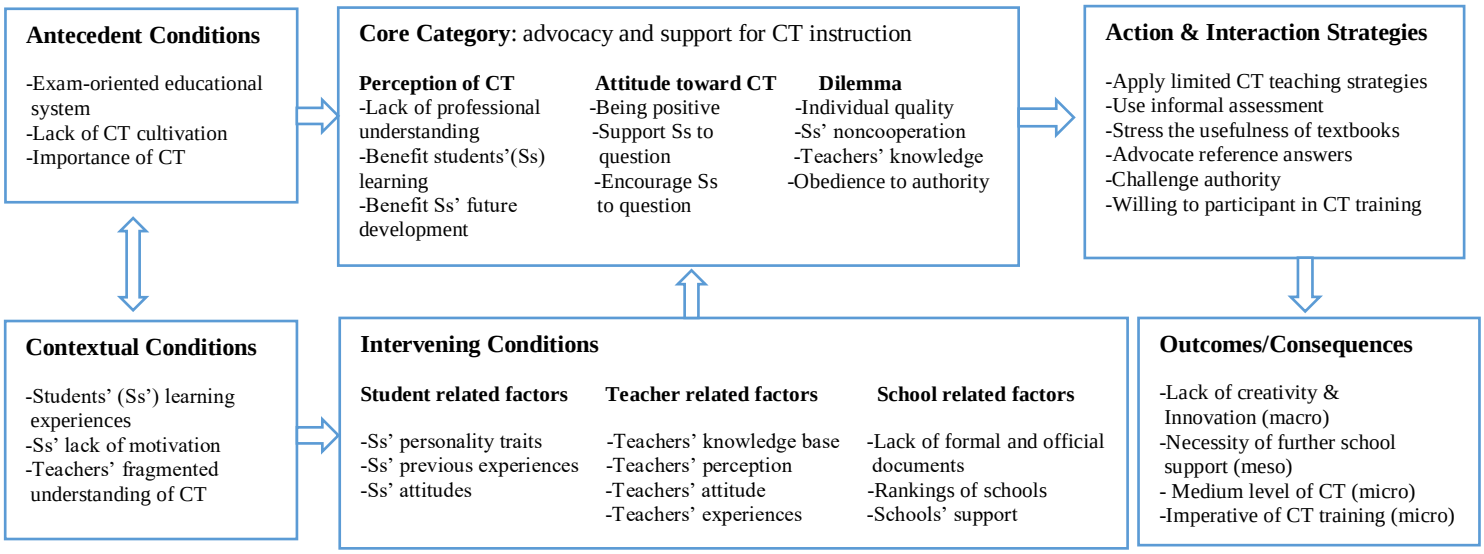
6. The concept is able to explain variation as well as the main point made by the data. (p. 147)

After following the aforementioned criteria, the core category grounded in the data was generated: *advocacy and support for CT instruction*. The conceptualization of the story is compelled through this core category, via the process of the open coding, the axial coding, and

selective coding. Teacher participants in institutions of higher education in the north of China advocate and support CT instruction, although they are faced with barriers at various levels, such as exam-oriented educational system (macro level); not enough support from schools (meso level); and lack of professional knowledge of CT (micro level). The core category "advocacy and support for CT instruction" systematically related to other categories, and those relationships among categories were validated by further refinement and development. The iterative characteristic of the data analysis required the researcher to repeatedly ask herself questions while examining the data. The researcher followed Strauss and Corbin (1998) to create a paradigm which indicated the relationships among categories. Figure 5 showed the summary of findings in line with Strass and Corbin's paradigm.

Figure 5

Coding Paradigm for the Research Findings

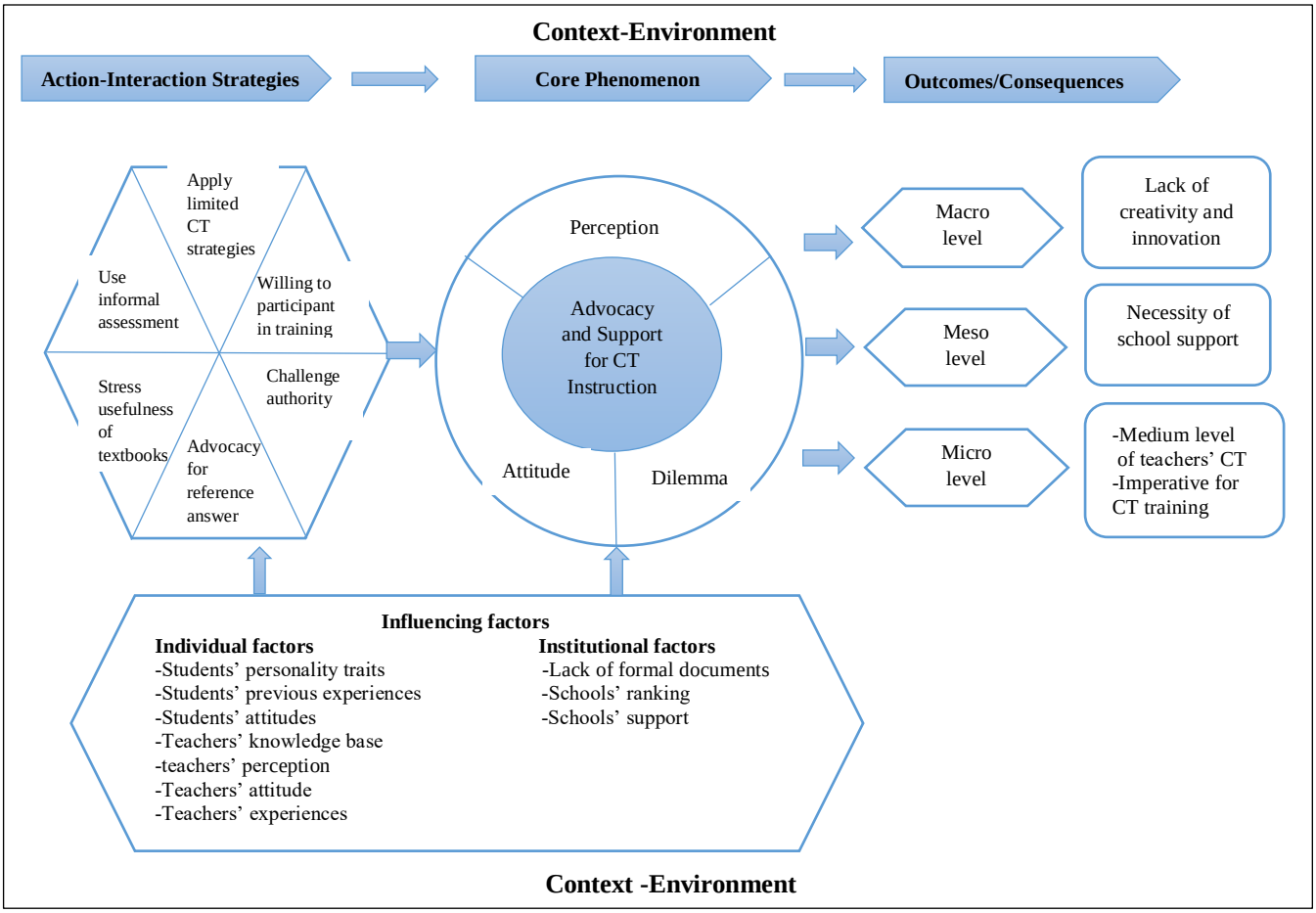


Then, the conditional/consequential matrix or model started to emerge "as soon as the diverse properties began to integrate" (Wanko Keutchafo, et al., 2022, p. 5). According to

Strauss and Corbin (1998), the conditional/consequential matrix or the matrix is an "analytic device", "conceptual guide", [composed of] a set of ideas" (pp. 190-193). Conditions/ consequences "represent the structural context in which action/interaction occurs" (Strauss & Corbin, 1998, p. 192). Structural context "has been broken down into different areas ranging from more macro to more micro" (Strauss & Corbin, 1998, p. 192). Figure 6 describes the conditional/consequential matrix or mental model of this study. With the help of the matrix, the researcher captured the interplay between "conditions, the responses of actors, and the consequences that result" (Strauss & Corbin, 1998, p. 193).

Figure 6

Conditional /Consequential Matrix or Mental Model



Interpretation of Conditional/Consequential Matrix

The data analysis procedures of Strauss and Corbin (1998) supported the emergence of a theory. The interpretive theory represented the process of Chinese teachers' perception of CT and experiences of CT instruction in higher education. The answer to the central question, "What is the perception, attitude, and practice regarding CT among the ISEC and non-ISEC teachers in institutions of higher education in the north of China?" was discovered in this study. Through the perspective of the participants, their experiences in the CT instruction offered clues for Chinese higher education. The meaning of the results is best understood from the explanation of three major categories: barriers to CT instruction; inadequacy of CT instruction; and necessity of promoting CT. Meanwhile, the core category of "advocacy and support for CT instruction" provided a descriptive narrative about the central phenomenon of the study. The following sections mainly deal with context and environment, action/interaction strategies, influencing factors, outcomes/consequences, relationship among categories, and the substantive theory.

Context and Environment

> According to Corbin and Strauss (2015), *context* is a term that consists of events, the set of circumstances or conditions that make up part of any situation, the meanings given to these (a problem, goal, etc.), the action and interaction persons take to manage or achieve desired outcomes, and the actual consequences that result from their action. (p. 155)

In this study, *context* referred to the conditional factors that influenced participants' understanding, behavior, and their teaching practice. The conditional factors included antecedent conditions, contextual conditions, and intervening conditions. Antecedent conditions stand for "a temporary state within dynamic natural and social systems that precedes and influences the onset

and magnitude of a hazard and its consequences. They are distinct from, but influenced by, what are commonly referred to as preconditions" (Crozier et al., 2013, p. 11). In this study, the exam-oriented educational system was one of the antecedent conditions. Contextual conditions refer to "the specific sets of conditions (patterns of conditions) that intersect dimensionally at this time and place to create the set of circumstances or problems to which persons respond through actions/interactions" (Strauss & Corbin, 1998, p. 132). For instance, students' lack of motivation and participants' fragmented understanding of CT belonged to these conditions. Intervening conditions are "those that mitigate or otherwise alter the impact of causal conditions on phenomena…" (Strauss & Corbin, 1998, p. 131), such as individual factors: students' previous experiences and participants' knowledge base, as well as institutional factors: lack of formal and official documents at schools. The *context* also involved participants' perception of CT, their attitude about CT, and the dilemma they met in their teaching process. The *context* encompassed participants' action and interaction strategies and the anticipated or actual consequences or outcomes, as well.

Within the Conditional /Consequential Matrix or model, the environment referred to the educational system, schools, as well as persons involved in the communicative encounter, that is, teachers and students. China implements a nine-year compulsory education from primary school (six years) to junior secondary education (three years). There are 3 years for senior secondary school, and four years in the standard university curriculum. China has improved the quality of education through a series of educational reforms. China's Education Modernization 2035 Plan has been launched to set the direction for the development of the education sector from "capacity" to "quality" (China Education Center, n.d., para. 2). However, the predominantly exam-oriented education system still exerts great influence on Chinese instruction and learning.

Teachers focus more on students' academic performance, and students obtain knowledge by rote learning and memorization. All of the above-mentioned created the context and environment for the core phenomenon in this study: advocacy and support for CT instruction.

Action/Interaction Strategies

According to Strauss and Corbin (1998), "strategic or routine tactics or the how by which persons handle situations, problems, and issues that they encounter are termed actions/ Interactions. These represent what people, organizations, social worlds, or nations do or say" (p. 133). In this study, participants self-reported some purposeful or deliberate acts that were taken to solve the problem: lack of CT cultivation, and "in so doing shape the phenomenon in some way" (Strauss & Corbin, 1998, p. 133). These action/interaction strategies included: apply limited CT teaching strategies; use informal assessment; stress the usefulness of textbooks; advocate for reference answers; challenge authority; and willingness to participate in CT training. These strategies are discussed respectively in the following sections.

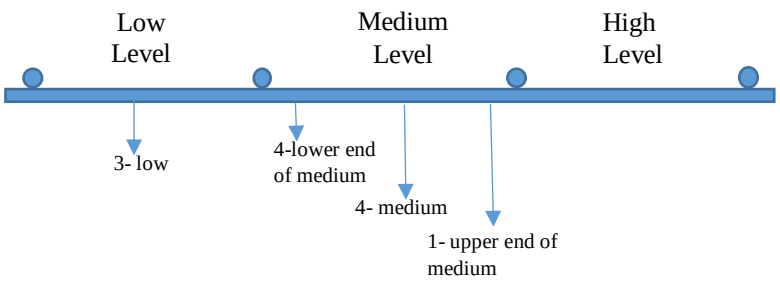
Apply limited CT teaching strategies. Effective CT teaching or instructional strategies play a vital role in cultivating and improving students' CT skills in the educational process. In this study, all ISEC and non-ISEC teachers in institutions of higher education in the north of China reported they applied some teaching strategies to cultivate students' CT in their teaching practice. The most frequent strategies they employed were group discussion, followed by open questions, presentations and writing assignments. Among the 12 teacher interviewees, only two ISEC teachers used the strategy of case study and one ISEC teacher employed a debate strategy in the teaching process. According to Behar-Horenstein and Niu (2011), there were specific CT instructional methods that had been investigated, and they were: concept mapping, scenario-based course exercises, active learning techniques, problem-based learning, inquiry-based

learning, question approach, guided practice, computer-based instruction, structured web-based bulletin boards, and online instruction. Compared to CT instructional strategies discussed by Behar-Horenstein and Niu, Chinese teachers utilized limited CT teaching or instructional methods or strategies**,** although they realized the importance and benefit of CT for students' learning, exams (especially with open questions), and future development.

Use informal assessment. "The most important tool available for effective teaching is assessment" (Rotenberg, 2010, p. 139). Assessment plays a vital role in both teaching and learning process. Participants reported their students' CT level was not high, belonging to medium or low level: three participants rated their students' CT as low level, four participants rated students' CT as medium level, four rated students' CT at the lower end of medium, and one rated students' CT at the upper end of medium (see Figure 7). The majority of these participants used informal ways of assessment to evaluate students' CT skills, such as students' engagement in class, questions that students posed and answered, assignments that student completed, as well as results from quizzes or exams. There was only one ISEC teacher (Participant #10) used a global CT assessment tool to assess students' CT. She explained this CT assessment tool was offered by her discipline leader, who got it when he participated in the ISEC program training. The assessment tool was a Chinese version and the format was Likert-scale.

Figure 7

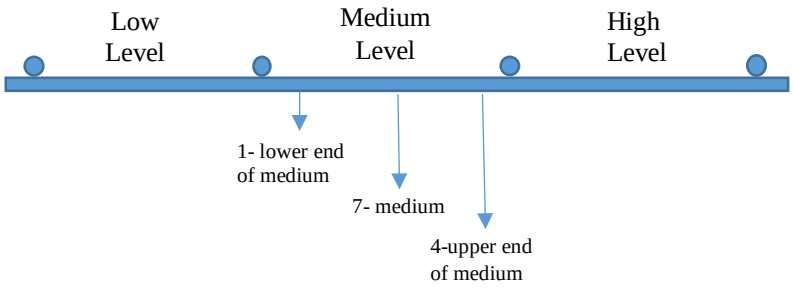
The Level of Students' CT Skills on a Continuum



When all participants were asked to rate their own CT skills, they reported their CT skills at the medium level. Seven participants rated their CT skills as medium, four participants rated them at the upper end of medium, and one rated herself at the lower end of medium (See Figure 8). Compared to the CT scores these participants received from the CCTST, most of them (eight teachers in total) underrated their CT skills, two rated themselves correctly, and one participant overrated his CT skills slightly, who got 16 points from the CCTST, and he rated himself at the upper end of medium. According to the criteria of CCTST User Manual (2021), his score belonged to medium. Among the eight participants who underrated themselves, two of them underrated themselves drastically. For instance, Participant #1, a non-ISEC teacher, got 26 points from the CCTST, belonging to the superior level based on the criteria of CCTST User Manual (2021), but he rated his CT skills at the upper end of medium. Participant #7, an ISEC teacher, got 24 from the CCTST, falling into the superior level as well, but she rated herself as the medium level. Most of them rated themselves based on their teaching practice and life experiences. Only one ISEC teacher (Participant #10), used the Chinese version of global CT assessment tool to evaluate herself. She also used the same tool to assess her students' CT skills. Except that, ISEC teachers stressed the importance of CT training program offered by the ISEC office. During the training, ISEC teachers were frequently asked to self-evaluate their CT skills. One non-ISEC teacher (Participant #1) pointed out that CT skills were closely related to the discipline: teachers of sciences must have higher CT skills than those of arts.

Figure 8

The Level of Participants' CT Skills on a Continuum



Stress the usefulness of textbooks. According to Jackson and Du (2022), textbooks are "central to teaching in Chinese schools and compulsory university courses. Not only must students memorize their contents, but most teachers depend heavily on textbooks in their teaching" (p. 198). In this study, most participants emphasized the usefulness of textbooks. However, they talked about the uses and contents of textbooks from different dimensions. Some mentioned the authoritative status of textbooks in the Chinese educational system. Some stressed the importance of textbooks in the teaching and learning process. Some claimed that textbooks should be regarded only as a reference.

Advocate for reference answers. With regards to standardized answers, some stated all standardized answers should be changed into reference answers. Some believed that objective questions need standardized answers, while subjective questions need suggested answers or answers for reference. Some claimed that whether standardized answers were needed or not depended on the different disciplines: disciplines of sciences, such as mathematics, need standardized answers. No matter what words they used, all participants advocated for reference answers.

Challenge authority. With respect to the assignment of identifying flaws in textbooks, there were varied opinions. Some viewed that picking flaws in textbooks was not suitable for college students, and college students need to do something more complicated to improve their

CT skills. Some claimed that identifying flaws in punctuation marks, spelling, and grammar was useless to train students' CT skills, while identifying flaws in the logic of the contents in the textbook seemed more effective to train students' CT. However, most participants stated that identifying flaws in textbooks was like asking students to challenge authority, so it was good to do such an assignment. There was an exception: one non-ISEC teacher, Participant #2 (associate professor, having a Ph.D and one-year experience as a visiting scholar in the U.S), held a unique opinion and claimed editors should avoid making mistakes in any textbooks before they were published, and it was not Chinese students' responsibility to pick out mistakes in the textbook. This type of challenging authority was more suitable for American students, which advocated independence, freedom, and democracy. She described her ideas by pointing out:

> In China, we should set up a type of authority, ask students to defer to authority, and lead students to learn and imitate… If there is no authority constructed, students are required to challenge, to do critical thinking [about everything], what about other issues, such as Chinese system, Chinse constitution and law? Do you still allow students to challenge and question those things? Therefore, there should be a standard for what issues should be regarded as authority, and what issues may be discussed and revised by the civilians and students.

When asked whether their schools emphasized and required the cultivation of students' CT, some participants responded by saying "yes." Some said "no." Some were not one hundred percent sure, for example, Participant #4, a non-ISEC teacher, described, "Based on the requirements of syllabus and the scheme of cultivation, our university should have required the cultivation of students' CT." One ISEC teacher, Participant #9, pointed out it was hard for the school to formulate a general requirement to stress CT cultivation. She added, "However, it was

likely that CT cultivation was mentioned in the scheme of cultivation in other programs or disciplines." For those participants who claimed there were no official or formal documents to advocate the cultivation of students' CT, when asked what recommendations or suggestions they would like to offer to their schools, they stated they were not willing to render any recommendations or suggestions to schools (it was an indirect way to challenge authority), because they were just ordinary faculty at schools who had no ability, expertise, and/or influence to provide suggestions to the school administration. They stressed their recommendations or suggestions would not be adopted even if they were given the opportunity. For them, their colleges or universities focused more on ideological and political education rather than CT education in the last few years.

Willing to participate in CT training. All participants regarded attending CT training as a useful way for them to improve their CT skills. Some participants stressed that teachers need first to learn about their own CT skills and then to improve students' CT. It was vital for teachers to understand what CT was and what connotations CT represented. In order to understand CT and its connotations, reading more books on CT was an effective way. Next, developing the awareness of CT was equally important. Some participants suggested building a scientific assessment system of teachers' CT to urge teachers to improve their CT actively, because of the fact that most of the teachers got access to CT via either the study-abroad experiences or the ISEC program. It meant most of them came into contact with CT relatively late, when they were almost or already adults and their thinking styles had been fossilized. Therefore, it was necessary for teachers to participate in the CT training program frequently.

With regard to how to improve teachers' CT, all participants rendered feasible strategies. These suggestions were: attending experts' presentations, participating in workshops, hosting

seminars, building interschool partnerships, sponsoring teacher abroad exchange programs, participating in teaching experiences sharing (teaching peer collaboration), attending teachers' demonstration class, and engaging in class observation. Some participants especially pointed out the most difficult and important issue in their teaching practice was how to effectively integrate CT into their disciplines. For some participants, teachers need to put different and effective strategies together and use them with flexibility, in order to solve this problem. Some participants summarized formulas for improving teachers' CT: "theory-practice-reflection" and "awareness-recognition-practice". It was worth noting that three participants, two ISEC teachers and one non-ISEC teacher, mentioned CT was not foreign to the Chinese people. Some ancient Chinese classics, such as *The Analects of Confucius*, *Zhuangzi*, and *The Tao Te Ching* (or *The Book of Tao and Teh*), did elaborate CT in their own ways. Therefore, CT was not solely the product of the West. The Chinese culture did have CT as well. Their ideas were congruent with Paton's (2005) argument: CT is not a special Western construct, and CT has been existed "in Chinese culture for at least the last thousand years" (p. 4).

Influencing Factors

For this model, influencing factors referred to the factors that impacted the participants' behaviors and their teaching practice. These influencing factors were summarized as *institutional factors* and *individual factors*. *Institutional factors* involved: schools' ranking, lack of formal or official documents on CT, and school support. For most participants, the school's ranking exerted a strong influence on their teaching results. For example, Participant #4, an English teacher, stated she was in a university of technology in Shenyang, Liaoning province, which was a provincial key university. Her university focused more on the electric appliance and mechanical programs, and the English major had been marginalized. The university recruited a

small number of English major students to form only one class during an academic year. Most participants claimed there was no formal or official document on the cultivation of students' CT, although their colleges and universities only offered some formats of support.

With regard to the problem of limited support from schools, some participants provided their ideas. For instance, Participant #5 believed the ways to solve this problem were closely related to what extent the school leaders attached importance to CT education. Participant #9 took it a bit further and elaborated,

> Under the Chinese system, if one policy needs to be implemented, the first thing is to alter the ideology of the administrations. If I had the chance to offer some suggestions, I would like to say that the importance of CT should be emphasized in the administrative training program for chairs, deans, provosts and presidents of colleges and universities.

Individual factors consisted of *student-related* and *teacher-related factors* respectively. *Student-related factors* included: students' personality traits, students' previous experiences, and students' attitudes. Most participants reported the difficulties and challenges in their teaching practice were student-oriented: introverted personality traits, medium or low English proficiency, inactive attitude, non-cooperation, etc. For example, Participant #3 described in her English class, some students were not willing to question or challenge their teachers because of their personality traits [introverted and shy] and restriction of their English proficiency [not high]. Participant #5 claimed her university was a local university, located in Changchun, Jilin Province. The quality and ability of the students recruited by her school were not high, and their CT skills were low. Participant #2 held a strong belief that students' CT skills were closely related to their previous experiences, such as, learning habits, individual ability, prior to attending college. She further explained that her students held inactive attitudes toward their

learning. They were short of motivation or drive to learn, and were not cooperative with their teachers. They demonstrated a state of "Lying flat" (a buzz word in China now, indicating more and more Chinese young people around 20 reject the rat race and choose to "lie flat" after they watch their friends work themselves to death).

Among these participants, only one, Participant #9, discussed teacher-oriented challenge or difficulty, that is, how to grade her students in an objective and fair way, based on rubrics. Only two participants pointed out that they did not have enough class hours to train students' CT in their class. The rest of them focused solely on students in terms of challenges and difficulties in their teaching practice.

Teacher-related factors were made up of teachers' knowledge base, teachers' perception, teachers' attitude, and teachers' experiences. Participants in this study pointed out when students posed deep academic, professional or specialized questions in class, teachers had to face the dilemma that they could not answer these questions. Under such circumstances, students' questioning and challenging demanded higher requirements for teachers with respect to their knowledge base and ability. For instance, Participant #6 stated CT was related to a person's knowledge and experiences. He took his mother as an example and described her as an illiterate old woman who lived in the countryside. She was easily influenced or even deluded by others' words and behaviors, especially during the COVID-19 pandemic. Participant #6 mentioned many Chinese rushed to queue up to buy a so-called effective medicine for curing COVID, without considering or questioning the truthfulness of the news. The consequence of this behavior was that the medication was sold out and people who lined up for a long time caught COVID. He stressed this was a typical example that people lacked CT.

Participant #11, an ISEC teacher, elaborated most ISEC teachers learned about CT through the ISEC program. Before participating in it, they did not have any awareness of or perception about CT. He claimed it was likely that a few teachers had access to CT, but they had a weak understanding about CT. He cited himself as an example.

> Before I taught the course of CT the first time, the provost office in our university invited me to teach a new course, named *mingbing xing siwei* (Critical Thinking). I told them I couldn't, because I didn't know what *mingbing xing siwei* was and I even never heard of this term before. Then they showed me the English name of the course---Critical Thinking. I suddenly realized I knew it quite well [because of experiences of studying abroad], but I translated CT into *pipan xing siwei* (Critical Thinking). We just didn't reach a consensus on the translation of this term…"

He then claimed it was vitally important for teachers to have an awareness of CT, understand its definition, recognize the importance of CT, and finally practice in their teaching process. Participant #9, an ISEC teacher, also emphasized the importance of CT awareness.

Most participants stated teachers' attitude also impact their CT instruction. For instance, Participant #7 believed holding an active attitude was vital. She suggested everything need to be considered and dealt with from different perspectives: listen to others' opinions, avoid stubbornness, and finally do more reflections. Participant #7 stressed teachers need to hold a positive attitude to cultivate and train students' CT skills actively.

Most participants claimed their experiences exerted influence on their CT and CT instruction. Participant #8, an ISEC teacher, reported his experiences and stated since 2015, when his university began to introduce the ISEC program, it was the first time for him to access

the ideology of CT. Then he became interested in CT and learned about it further. He benefited a lot from this process.

> Since then, I began to have my own CT skills, and I improved my CT to a great extent. As a teacher, you have to first equip yourself with CT skills and then to teach students to think critically. I think I have benefited from the ISEC program…

In addition to his ISEC experience, he also mentioned his other professional and academic experiences. He stressed his doctoral program, especially his dissertation writing and academic research were beneficial to improving his CT. He pointed out that during his academic research, he had to pose scientific questions, read various literature around the problem, and do a literature review. By means of scientific research methodology, he addressed his questions and finally drew a conclusion. He regarded the process of doing scientific research as the process of doing CT. Participant #7 shared her personal experiences and claimed educating and cultivating kids became very critical.

> Nowadays, you can't treat your kids in the way you did in the past, something like 'listen to your mum and do what your mum tells you to do'. Instead, mothers need to give more space to their kids, and allow their kids to play and think based on their own will.

However, Participant #4, an associate professor in her university, shared her personal and professional experiences with the researcher. Compared to other participants in this study, her experiences were unique. She confessed, except taking care of her kid and family, she put all her energy into teaching. She did not care too much about academic research. Most of the teachers in her university were similar to her and were in a state of "lying flat". The main reason for it was it was so hard to be promoted to full professor, because there was no position available there. She stated "it is hopeless. Teachers in our university lack deeper thinking about their teaching and

academic research. Although our main devotion is to our instruction, we do limited research and reflection on other aspects…" Participants' reports evidenced their personal, professional, and academic experiences impacted their CT and CT instruction.

Outcomes/Consequences

For this model, the outcomes/consequences were grouped into three levels: Macro, meso and micro. Macro level outcomes included *lack of creativity and innovation*. With the influence of a predominantly exam-oriented educational system, a lot of students obtain knowledge through rote learning and memorization, resulting in Chinese students' lack of creativity and innovation. Meso level outcomes consisted of *necessity of further school support*. Participants in this study claimed there was no formal or official document on the cultivation of students' CT, indicating it was necessary for colleges and universities to offer further support for teachers with more obvious goals or missions. Participant #5 mentioned even though her university offered some presentations on CT, they were on a shallow basis. Her university put forward the ideology of CT, but there was not much specific implementation of action.

Micro level outcomes covered *medium level of teachers' CT* and *imperative of CT training*. All participants in this study reported the level of their CT skills was medium, which was similar to the results from the CCTST. All participants believed participating in CT training was useful and necessary. Participant #5, a non-ISEC teacher, reported "I didn't participate in any systemic CT training, and I had little knowledge and theory about CT". She believed participating in CT training and other relevant activities, such as presentation, experience sharing, and class demonstration would be beneficial to improve teachers' CT skills. She added only after learning about more knowledge and theory of CT, could teachers integrate CT into their teaching practice consciously. Participant #3 expressed the similar idea and claimed

"through CT training, teachers may have a good understanding of CT connotation, its benefits, and ways to cultivate CT."

Participant #8, an ISEC teacher, stated teachers need to participate in more CT training programs, because teachers were at the age of 30 to 40 in colleges and universities, their CT skills could not retain longer. If they did not access CT training frequently, they would go back to the original state [lack of CT]. "If your CT was cultivated and trained from childhood, and CT was deeply entrenched in your mind, it is unnecessary for you to participate in the training program." However, he stressed most Chinese teachers got access to CT without a solid foundation or training, and it was imperative for teachers to participate in CT training on a frequent basis. Participant #10 elaborated people similar to her age (late 40s) were lazy to learn new things. She valued various CT training programs and regarded them as one of the effective ways to improve her CT. She stressed the ISEC program office hosted training for ISEC teachers annually, and this year ISEC office conducted Teaching Innovation activity. She taught the course of CT in her university. She and her colleagues formed their CT team and signed up for the Teaching Innovation activity, which started its first round on January 15, 2023, and would last half a year at least. During this period, they would receive systemic training.

Relationship among Categories

According to Branden (2012), one of the major steps in grounded theory is "to describe the relationships that exist between categories" (p. 137). All categories in this study emerged from the data and each category was interlinked with the core phenomenon. Put another way, all categories and subcategories in the conditional matrix or model were directly or indirectly connected. The core category or phenomenon in this model was "advocacy and support for CT instruction". This category consisted of teachers' perception of CT, teachers' attitude towards

CT, and the dilemma that teachers were in. All these categories were directly connected. The next category was action and interaction strategies that need to be performed to realize the core phenomenon: advocacy and support for CT instruction. This category was directly connected to the core category or phenomenon and intervening conditions. For example, although Chinese teachers lacked professional understanding of CT, they all believed CT benefited students' learning and future development. They held a positive attitude towards CT instruction, supported students to question or challenge in class, and created an environment that encouraged students to question. However, Chinese teachers were in a dilemma when they trained students' CT in class.

The dilemma included students' individual quality, students' non-cooperation, teachers' knowledge, and their obedience to authority. The intervening conditions, including individual-related and school-related factors, exerted great influence on teachers' perception, attitude and practice about CT in their teaching process. This indicated the core phenomenon or category, action and interaction strategies, and intervening conditions were closely related, all of which led to the outcomes or consequences. There were three levels of outcomes or consequences in this study: macro level, meso level, and micro level. For instance, because of student-related, teacher-related factors, as well as institution-related factors, teachers had varied and fragmented perception or understanding of CT, but they held a positive attitude towards CT. They applied limited teaching strategies to train or develop students' CT, used informal assessment to evaluate CT, stressed the usefulness of textbooks, and advocated reference answers, etc. All of these actions and interactions resulted in students' lack of creativity and innovation, necessity of further school support, and the imperative of CT training. The model also indicated teachers'

advocacy and support for CT instruction took place within a specific context, which was the institutions of higher education.

The Substantive Theory

Through development of the concepts, categories, and their relations to each other based on the data from 12 participants' interviews, a substantive theoretical model emerged and "continued to evolve throughout the GT [grounded theory] research process" (Branden, 2012, p. 138). The final product was the conditional/consequential matrix or model about Chinese teachers' CT, shown in Figure 6. According to Strauss and Corbin (1998), using the coding paradigm (see Figure 5) to answer "questions of who, when, where, why, how, and with what consequences" (p. 127), the researcher of this study moved into the conditional/ consequential matrix, a theoretical model was developed to add to the current knowledge of Chinese teachers' CT topic. This conditional/consequential matrix or model demonstrated Chinese teachers' advocacy and support for CT instruction as a process that other teachers could take as a reference in the context of institutions of higher education. This substantive theory grounded in data was Chinese teachers advocated and supported CT instruction in the north of China, which suggested that a combination of internal and environmental features of, and around the teachers, set forth the context in which an advocacy and support event took place. After discovering the context/environment and the teachers' decision to advocate and support CT instruction, teachers engaged in the advocacy and support process positively. Finally, teachers were willing to promote their CT instruction via various strategies, such as participating in CT training programs, obtaining more support from their colleges and universities, and thus altering the status quo of Chinese higher education: students' lack of creativity and innovation.

Summary

This study utilized a non-experimental causal-comparative methodology with an explanatory mixed methods research design. It was aimed to explore the status quo of CT skills among ISEC and Non-ISEC teachers in institutions of higher education in the north of China, as well as their perception, attitude and practice of CT in their teaching process. This chapter mainly focused on analyzing data obtained from both quantitative and qualitative research. For the quantitative paradigm, all data were collected via CCTST. The participants' demographic information was collected first. Then descriptive statistics were used to describe the characteristics of the data and help to address the first two research questions. Inferential statistics were employed to make predictions or inferences from the data set after descriptive statistics. Statistical tests, such as the independent *t*-test, two-way ANOVA, and Pearson correlation were employed to compare the differences and figure out the relationships among variables, and assist in addressing the remaining seven research questions.

For the qualitative paradigm, all data were collected via interviews, field notes, and verbatim transcripts. Twelve participant interviewees' demographic information was presented first. Then Strauss and Corbin's (1998) grounded theory approach was followed to analyze data: open coding, axial coding, and selective coding. After following the iterative process of asking questions and making constant comparisons, the core category emerged, and the substantive theory was explained and summarized. Chapter Five mainly concentrates on discussions, which includes research findings, implications of the findings, recommendations, contributions, and conclusion.

Chapter Five: Conclusions

This explanatory mixed methods research was designed to explore the status quo (current situation) of CT skills among the ISEC and non-ISEC teachers in institutions of higher education in the north of China, as well as their perception, attitude, and practice of CT in their teaching process. The quantitative methodology was dominant in this study, and was followed by the qualitative grounded theory methodology. In this chapter, a discussion of the findings is presented. The interpretation of the findings stresses the medium or moderate level of CT skills among ISEC and Non-ISEC teachers in institutions of higher education in the north of China. The core phenomenon or category emerged from the qualitative data: advocate and support CT instruction. The implications of the findings are put forward for the institutions of higher education in China. The recommendations for practitioners or others and for future study are offered. Contribution to the field is identified, as well, followed by the conclusion of this study.

Findings

Findings from the Quantitative Research

The purpose of this quantitative methodology was to investigate the status quo of CT skills among the ISEC and non-ISEC teachers in institutions of higher education in the north of China. Nine research questions were posed to examine whether there was any difference in CT skills between ISEC teachers and non-ISEC teachers, whether there was any difference in CT skills between ISEC and non-ISEC teachers based on their professional rank, education background, and discipline they taught, as well as whether there was any relationship between CT and age, and CT and years of teaching. In order to address these questions, 102 participants from 22 colleges and universities in the north of China were recruited to participate in this study. The standardized instrument, the CCTST, was utilized to measure the participants' overall CT

skills and their five CT subskills of *analysis*, *inference*, *evaluation*, *inductive reasoning*, and *deductive reasoning*. Both descriptive and inferential statistics were used to answer nine research questions. The following sections discuss the findings from the quantitative research.

The Status Quo (Current Situation) of Chinese Teachers' CT Skills (Question 1 & 2)

With respect to ***Question #1***, what is the level of CT skills of overall Chinese teachers in institutions of higher education in the north of China?, the descriptive statistics obtained from CCTST described the overall CT skills of Chinese teachers as medium or moderate, specifically falling into the upper range of the moderate level. Among the five core subskills, the *analysis* skill was the highest, standing in the lower range of the strong level. Three subskills of CT: *inference*, *inductive reasoning* and *deductive reasoning* fell into the moderate or medium level, and *evaluation* skill was in the lower range of the moderate level. For ***Question #2***, what is the level of CT skills of the ISEC and non-ISEC teachers in institutions of higher education in the north of China, respectively?, the overall CT skills of the ISEC teachers were slightly higher than that of non-ISEC teachers. With regards to the five core subskills, the mean scores of *analysis*, *inference*, and *deduction* among non-ISEC teachers were slightly higher than those of the ISEC teachers, and the mean scores of *evaluation* and *induction* among the ISEC teachers were slightly higher than those of non-ISEC teachers.

It was worth noting when the researcher of this study collected data from the end of 2022 to the beginning of 2023, the Chinese government just cancelled the zero-tolerance policy of COVID-19 pandemic. Since then, the majority of volunteer participants were suffering from COVID: some of them caught COVID and were quarantined at home, and some of them were recovering shortly after COVID. According to Cuffari (n.d.), COVID-19 has badly impacted human health and mental health. It not only affects the respiratory system, but also affects the

cardiovascular, gastrointestinal, and urinary systems (Cuffari, n.d.). Most importantly, COVID-19 also influences the neurological system, and

> these neurological effects are due to direct infection of the brain, a virus-induced hyperinflammatory response, hypercoagulation, and post-infectious immune-mediated processes. As a result, these neurological effects can lead to a wide range of psychological issues ranging from depression, anxiety, fatigue, and post-traumatic stress disorder (PTSD) (Cuffari, n.d., p.3).

Therefore, because most participants caught COVID and/or just recovered from it when they participated in this study, consequently their mental health may be impacted by COVID. This was an intervening factor that could not be disregarded when interpreting the status quo of the Chinese teachers' CT skills.

Although some studies stressed teachers play a crucial role in higher education to foster students' CT (Birjandi & Bagherkazemi, 2010; Boonjeam et al., 2017; Cave, 1993; Janssen, et al., 2019; Stedman & Adams, 2012), there is a paucity of studies concentrating on teachers' CT skills (Janssen, et al., 2019). After searching and screening the extant literature both in China and in the West, the researcher of this study could not find any literature focusing on CT skills of Chinese teachers, especially the ISEC and non-ISEC teachers. Only limited literature concentrated on teachers' perceptions of and/or attitudes toward CT in higher education (Choy & Cheah, 2009; Li, 2016; Stedman & Adams, 2012; Zhang, et al., 2020). Because there were no empirical studies available on the CT skills of Chinese teachers, the researcher of this study had to make a comparison with an aggregate sample of CCTST Four Year College Students, "the average percentile score of this group of participants is 62" (Insight Assessment Report, January 2023, p. 1). This means that roughly 61 people out of 100 score lower than this group of Chinese

teacher participants and 38 persons out of 100 score higher than this group of Chinese teachers in the national comparison group (CCTST User Manual, 2021).

Ng, et al (2022) did research on Chinese community college students' CT skills and CT dispositions. Their research results showed the mean score of Chinese community college students' CT skills was 17.82 ± 4.20, falling into the upper end of the moderate range, which was similar to the result of this research on Chinese teachers' CT skills (17.34 ± 3.66), also belonging to the upper end of the moderate level. It indicated Chinese community college students' CT skills were slightly higher than Chinese teachers', although the CT skills of Chinese teachers also fell into the upper end of the moderate range. Since teachers play a vital role in "students' CT-skill acquisition" (Janssen, et al., p. 311), teachers first need to "possess CT-skills themselves so that they can provide explicit instruction and integrate CT in their lessons" (Janssen, et al., p. 311). Therefore, it is imperative for Chinese teachers to improve their CT skills in order to be able to teach students' CT.

Significant Differences among Variables (Question 3, 4, 5, 8, & 9)

Variables here included the dependent variable of CT skill scores (ratio level of measurement), and independent variables: gender, professional rank, educational background, and discipline (nominal level of measurement). ***Question #3*** addressed whether there was a statistically significant difference in CT skills between the ISEC and non-ISEC teachers. Although the mean overall CT score of the ISEC teachers was higher than that of non-ISEC teachers, there was no statistically significant difference in the overall CT scores for the ISEC and non-ISEC teachers, after performing an independent t-test. Meanwhile, the scores of five core subskills for the ISEC and non-ISEC teachers were not statistically significantly different, either.

For ***Question #4***, a two-way ANOVA was conducted to examine the difference in CT skills between the ISEC and non-ISEC teachers, based on gender. The results of the ANOVA indicated the overall mean score of the female teachers was higher than that of the male teachers. The mean score of the female non-ISEC teachers was higher than that of the male non-ISEC teachers, too. However, the mean score of female ISEC teachers was lower than that of male ISEC teachers. The mean score of the female non-ISEC teachers ranked the highest, and followed by the male ISEC, female ISEC, and male non-ISEC. The interaction effect between gender and scores of CT skills of two groups was not statistically significant, and there was no significant main effect for two groups of teachers: ISEC and non-ISEC.

Some studies focused on students' CT skills, CT dispositions, and gender. The research results were equivocal or conflicting concerning CT and gender. Facione (1990c) reported there were no statistically significant gender differences in the pretest and control group data, indicating the CCTST was not gender-biased. However, significant difference between the genders emerged "after students complete[d] their college level CT course" (p. 3). Dow and Wood (2006) supported the idea that CT skills were perceived differently with regard to gender. They summarized females employed CT skills and problem-solving skills as much as males, but they used it in a less confrontational and direct style than males. Bers, et al. (1996) found statistically significant difference in mean scores of the CCTDI for community college students, and females' scores of CT disposition were much higher than those of males. The results of Leach's (2011) study on students' CT skills in higher education demonstrated the total scores of CT skills from CCTST was not significant between female and male students. It was consistent with the research results from this study: there was no significant difference in CT skill scores from CCTST between male and female teachers. The difference between these two studies was

the overall mean score for female students (M =15.9) was lower than that of male students (M = 16.3) in Leach's study, while in the present study, the overall mean score of female teachers (M = 17.48) was higher than that of male teachers (M = 17.11).

With respect to ***Question #5***, a two-way ANOVA was conducted to investigate the differences in CT skills for the ISEC and non-ISEC teachers based on professional rank. The ANOVA showed the overall mean score of instructors (M = 17.54) was the highest and followed by associate professors (M = 17.26) and professors (M = 16.30). The mean score of the ISEC instructors was the highest (M = 18.09), followed by non-ISEC associate professors (M = 18.07), ISEC professors (M = 18.00), non-ISEC instructors (M = 16.82), ISEC associate professors (M = 16.33), and non-ISEC professors (M = 15.17). There was no significant interaction effect between the two professional ranks: instructors and associate professors (ten professors were not further considered because of the small number) and two teacher groups (ISEC and non-ISEC). There was no significant main effect for the two teacher groups and the professional rank. Although there were no significant differences in CT skills for the ISEC and non-ISEC teachers based on the professional rank, there was practical significance between the ISEC instructors and ISEC associate professors (Cohen d = .5), as well as between the ISEC associate professors and non-ISEC associate professors (Cohen d = .6).

Zhang, et al. (2020) conducted multivariate analysis of variance of Chinese EFL teachers' perceptions of CT and CT teaching. They found the perceptions of integrating CT with language teaching "were significantly associated with academic [or educational] background ($F = 5.256$, $p < 0.05$), and that the interaction between gender and professional titles [or rank] was also significant ($F = 3.360$, $p < 0.05$)" (Zhang, et al., 2020, p. 488). It indicated teachers with higher

professional ranks had "significantly higher scores than lecturers and assistants among female teachers" (Zhang, et al., 2020, p. 488).

Question #8 was aimed to investigate the significant difference in CT skills between the ISEC teachers and non-ISEC teachers based on the educational background. A two-way ANOVA showed the overall mean score of the ISEC teachers (M = 17.33) was slightly higher than that of non-ISEC teachers (M = 17.29) based on the educational background. The mean scores of teachers with master's degrees (M = 17.58) were higher than those with doctoral degrees (M = 16.29). The mean scores of ISEC teachers with a master (M = 17.81) ranked the highest, followed by the non-ISEC teachers with a master (M = 17.38), non-ISEC teachers with a doctorate (M = 16.89), and ISEC teachers with a doctorate (M = 15.83). The interaction effect between two educational background groups (with a master's degree and with a doctorate) and two teacher groups was not statistically significant. There was no statistically significant main effect for two teacher groups and the main effect for the educational background did not reach statistical significance.

Zhang, et al. (2020) conducted research on Chinese teachers' perceptions of CT and CT teaching through a questionnaire and interviews. They performed a one-way ANOVA to compare teachers' education background and their perceptions of CT and its teaching. The results from the ANOVA indicated "significant differences ($F = 14.238$, $p < 0.001$)… Teachers with MA degrees and doctoral degrees had significantly higher scores than those with BA degrees" (p. 488). They summarized teachers with MA and doctorate were more positive about "integrating CT into their language teaching" (Zhang, et al., 2020, p. 488).

Q***uestion #9*** was designed to examine the difference in CT skills between the ISEC teachers and non-ISEC teachers based on discipline. After performing a two-way ANOVA, the

overall mean score of Science teachers (M = 17.84) was higher than that of Arts teachers (M = 16.97). The overall mean score of the ISEC teachers (M = 17.39) was slightly higher than that of non-ISEC teachers (M = 17.30) based on the discipline. The mean score of the ISEC teachers of Science ranked highest, followed by non-ISEC teachers of science, non-ISEC teachers of arts, and ISEC teachers of Arts. There was no statistically significant interaction effect between the two discipline groups and two teacher groups. No significant main effect for two groups of teachers was found and no significant main effect was found for the educational background, either.

According to Walsh and Hardy (1999), some research indicated science majors "were more likely to be more logical than nonscience majors and, therefore, were better critical thinkers. Other studies reported no differences" (p. 150). It demonstrated the research results were inconclusive with respect to associations between disciplines and CT (Walsh & Hardy, 1999). Facione (1990c) elaborated in his study, students were grouped into 6 clusters of academic majors, based upon "the epistemological and methodological similarities and differences hypothesized by this researcher [Facione] to obtain among the disciplines in each cluster" (p. 12). Six clusters of academic majors consisted of: (1) Letters, Languages, English, Liberal Studies, History, and Humanities; (2) Social Sciences, Psychology, Human Services, and Teaching; (3) Mathematics, Engineering, Statistics, and Computer Science; (4) Natural Sciences, Physical Sciences, Health Professions; (5) Business, Administration, Management, and Government; (6) Performance Studies, Drama, Art, Music, Physical Education (Facione, 1990c, p. 13). The ANOVA showed academic major was not statistically significant with respect to the CCTST pretest ($F = 1.47$; $p = .20$), but there were statistically significant differences between the academic major and the CCTST posttest ($F = 5.23$; $p = .00$).

Gadzella and Masten (2013) did research on CT and learning processes for students in two majors: one major was Psychology and Special Education (n = 23) from the College of Education, and the other was Sociology, Social Work, and Criminal Justice (n = 17) from the College of Liberal Arts and Science. The instrument they used was the Watson-Glaser Critical Thinking Appraisal (WGCTA). The results indicated "students in Psychology and Special Education scored significantly higher [than those in Sociology, Social Work, and Criminal Justice] on two subtests (inference and Evaluation of Argument)" (p. 256). Gadzella and Master's study made a comparison between students majoring in Psychology & Special Education and Sociology, Social Work, & Criminal Justice, and significant differences were found. The present study compared teachers of Science and Arts, and the significant difference did not exist.

Significant Relationships among Variables (Question 6 &7)

Variables here involved the criterion variable of CT skill scores (ratio level of measurement), and predictor variables of age and years of teaching, which were the ratio level of measurement. Q***uestion #6*** was designed to examine the relationship between CT skills and age. Pearson product-moment correlation (r = -.17) indicated there was a weak, negative correlation between overall CT scores and age, which meant the older people generally had lower CT skill scores and younger people had higher CT skill scores. This weak negative relationship between overall CT scores and age was not statistically significant. With regards to the five core CT skill scores and age, there were weak, negative correlations between the five core skills and age respectively: *analysis* (r = -.12), *inference* (r = -.10), *evaluation* (r = -.13), *induction* (r = -.24), and *deduction* (r = -.05). Only the correlation between i*nduction* and age was statistically significant, and the remaining correlations were not statistically significant. Friend and Zubek

(1958) did research on the effect of age on CT ability. They recruited a group of 484 participants ranging in age from 12 to 80 years old with diverse educational, professional, and economic backgrounds. They finally found that CT ability showed "a progressive increase from late childhood through to the mid-twenties, holds up to 35 years and then declines progressively through to the seventies" (pp. 412-413). The results of this study that the older people had lower CT skill scores and vice versa were consistent with the findings of Friend and Zubek.

Question #7 was designed to explore the relationship between CT skills and years of teaching. The Pearson product-moment correlation indicated there was a weak, negative correlation (r = -.18) between overall CT scores and years of teaching, but it was not statistically significant. There were weak, negative correlations between the five core skills and years of teaching: *analysis* (r = -.10), *inference* (r = -.14), *evaluation* (r = -.12), *induction* (r = -.20), and *deduction* (r = -.10). The results indicated participants with more years of teaching generally had lower CT skill scores and vice versa. A significant correlation was only found between *induction* and years of teaching, and the rest of correlations were not statistically significant. The researcher of this study failed to reject all null hypotheses.

Findings from the Qualitative Research

The purpose of this qualitative grounded theory was to generate a theory to elaborate the perception, attitude, and practice of CT among the ISEC and non-ISEC teachers in institutions of higher education in the north of China. In order to gain deep and comprehensive data about the ISEC and non-ISEC teachers' perception, attitude, and experience of CT, twelve volunteer participants (six ISEC and six non-ISEC teachers) were invited to take part in the interview via Zoom. The interview protocol consisted of nine open questions. All interviews were audio and video recorded through the recording function of Zoom. After collecting the data, the researcher

of this study followed Strauss and Corbin's (1998) grounded theory procedures to analyze the data. As the data were analyzed and coded, it was vitally important to guarantee trustworthiness in this process. Trustworthiness was established by ensuing accuracy and verification strategies. According to Creswell and Poth (2018), there were three lenses: the researcher's lens, participant's lens and reader's lens, covering nine strategies, for verification. In this study, some verification strategies were utilized, such as identifying and minimizing the researcher's biases, writing with thick and detailed descriptions, taking the narrative back to the participants in member checking.

After the iterative data analysis, the core phenomenon or emergent theory finally emerged from the data: "advocate and support for CT instruction", which was "central to the research problem" (Williams, 2017, p. 87). In this study, the conditional/consequential matrix or model (Figure 6) was constructed to offer an analytic device or conceptual guide for the researcher to capture the interaction between conditions, action/interaction strategies, and outcome/consequences that resulted, and thus better understand the central question: What is the perception, attitude, and practice regarding CT among the ISEC and non-ISEC teachers in institutions of higher education in the north of China? This central question was explored through three subquestions: (1) what is the perception of CT among the Chinese ISEC and non-ISEC teachers in institutions of higher education in the north of China? (2) What is the attitude toward CT among the Chinese ISEC and non-ISEC teachers in institutions of higher education in the north of China? (3) What is the experience of CT teaching among the Chinese ISEC and non-ISEC teachers in institutions of higher education in the north of China?

Participants' Perception of CT

Many conditional factors exerted great influence on the Chinese teachers' perception of CT. At the macro level, the exam-oriented educational system is still dominant in China, although a series of educational reforms have been implemented since the founding of the People's Republic of China (Tan, 2016). At the meso level, Chinese colleges and universities began to focus more on cultivation of creative and innovative talents, because the Chinese government advocate internationalization, modernization, and creative and innovative talents (Wei, 2003; Ryan, 2011; Zhu, 2019). Although colleges and universities have introduced CT into their campuses, they are confronted with many difficulties. At the micro level, Chinese teachers showed a great tendency for advocating and supporting CT instruction, under the aforementioned contexts and environments. However, the definitions of CT offered by Chinese teachers were varied and fragmented. Some of them regarded CT as being skeptical, critical and rational. Some took CT as independent and active thinking. Some viewed CT as challenging and questioning. Some believed CT was an ability of logical reasoning, and some stressed CT helped to make a judgement or decision and finally solve problems. This indicated Chinese teachers perceived CT in their own ways and they lacked professional understanding and perception of CT.

Li (2016) did a study on Chinese EFL teachers' conceptions, beliefs, and practice about integrating CT skills into their language instruction. Her research results showed Chinese EFL teachers "demonstrated varied but fragmented understandings of thinking skills… The findings suggest that teachers might find it hard to define and articulate thinking skills…" (p. 277). It indicated the results of the present study supported Li's findings.

Zhang, et al. (2020) conducted research on Chinese teachers' perceptions of CT and CT teaching through a questionnaire and interviews. Several central categories or themes emerged from their data: disposition of CT, reasonable and logical, to analyze, and work/life/problem-

solving (Zhang, et al., 2020). It indicated Chinese EFL teachers covered only limited aspects of CT when they defined it, which meant they had a narrow and fragmented understanding of CT. The finding of the present study was also consistent with the results of Zhang, et al.'s.

Choy and Cheah (2009) conducted research on teachers' perception of CT among students and its influence on higher education. Choy and Cheah elaborated most participants defined CT as the "intellectual stimuli which become the impetus to facilitate thinking among students in the classroom and enable students to enjoy the process of learning" (p. 200). It implied most participants in Choy and Cheah's (2009) study thought CT was beneficial for students' learning and "help[ed] them obtain better learning outcomes" (p. 201). Birjandi and Bagherkazemi (2010) took this idea further and claimed CT "is critical for students to perform well not only in educational systems, but also in future workplaces, and social and interpersonal contexts" (p. 135). In the present study, although Chinese teachers lacked holistic and professional understanding of CT, they all believed CT skills were beneficial to students' learning, their exams (especially with open questions), and their future development, which aligned with and corroborated Choy and Cheah's findings, as well as Birjandi and Bagherkazemi's ideas.

Participants' Attitude Toward CT

According to Asgharheidari and Tahriri (2015), "teachers' and students' attitude toward CT is a factor that can affect incorporating CT practice in the classroom" (p. 389). Therefore, holding a positive attitude towards CT is vitally important in teaching practice. In this study Chinese teachers believed CT played a vital role in the process of students' learning and personal development. They all held a positive and supportive attitude towards CT. All of them liked students to question and challenge them in class. They tried to create a welcoming environment

in class and encourage students to question and challenge via some teaching strategies, such as questioning, group discussion, and presentation. However, it was not an easy task to train students' CT and Chinese teachers were in a dilemma. For instance, students' individual qualities, such as learning habits, knowledge base, and attitude, exerted a great influence on CT instruction. Students held an inactive attitude and refused to cooperate or engage in class activities. Teachers sometimes worried about their own knowledge base, when students posed deep questions in class. Both teachers and students had been trained to obey authority since their childhood, which was closely related to the social norm and Chinese culture.

In Li's (2016) study, the majority of Chinese EFL teachers agreed or strongly agreed CT or higher order thinking skills were important for students' learning and their personal development. It demonstrated Chinese EFL teachers held positive attitudes towards CT. The finding of this study was consistent with Li's research results. Li also discovered the majority of Chinese EFL teachers' believed thinking skills or CT were vital for all subjects and they thought thinking skills improved "students' accuracy because they conceptualize thinking skills as a tool to enhance memory, learning strategies, metacognitive strategies, and intelligence" (Li, 2016, p. 281). The possible reason for this was learning a foreign language was closely related to obtaining linguistic knowledge via memorization in China (Li, 2016). In Li's study, Chinese EFL teachers also demonstrated positive attitudes towards "integrating thinking skills in language classroom" (p. 273).

Stapleton (2011) conducted research on 72 Hong Kong secondary school teachers' attitudes towards CT. He discovered the teachers had some conception of CT, but it was narrow and lacked a precise understanding of CT. Meantime, these participants expressed "strong support for the inclusion of CT in the curriculum" (Stapleton, 2011, p.14). They also conveyed

"a desire for training in how to teach it" (Stapleton, 2011, p.14). All of these indicated Hong Kong secondary school teachers held a positive attitude towards CT. The result of the present study corroborated Stapleton's findings.

Asgharheidari and Tahriri (2015) conducted a study on Iranian EFL teachers' attitude towards CT and CT instruction. The results indicated most of the EFL teachers in Iran had a clear idea about the concept of CT and believed "it is an important part of their job as a teacher to increase learners' critical thought.… Most of them conveyed a strong desire for more training in how to teach these skills" (p. 388). It showed Iranian EFL teachers held a positive attitude towards CT, with which the findings of the present study was congruent.

Participants' Practice of CT

Asgharheidari and Tahriri (2015) claimed although many teachers agreed with the importance of CT, they still utilized traditional ways to teach, and disregarded "incorporating CT skills in their classes and claim that they don't have enough time to focus on these skills" (p. 389). In this study, all participants claimed they applied CT teaching strategies in their practice. The most frequently used strategies were group discussion, followed by open questions, presentation, and writing assignments. Only two out of twelve participants used case study and debate in their teaching practice. It indicated Chinese teachers in this study employed limited CT teaching strategies in their teaching process. The majority of participants employed informal assessment to evaluate students' CT skills and their own CT skills. Only one ISEC teacher used a global CT assessment tool to measure herself and students' CT. All participants reported their students' CT skills were not high, falling into the medium or low level. The majority of participants rated their CT skills as medium or at the upper end of medium. Only one participant rated herself at the lower end of medium (in reality, her CT scores from the CCTST belonged to

medium). Compared to the CT skill scores obtained from the CCTST, most of the participants underrated their CT skills. With respect to standardized answers, participants held varied views. However, all participants advocated for referent answers. It is a good trend that Chinese teachers tend to not depend too much on so called standardized answers and are willing to design more open questions to develop students' thinking skills.

Chinese education has "a long tradition of valuing knowledge" (Mast, 2016, p. 43). Chinese people believe knowledge is a basis for understanding. Thus, Chinese students are encouraged to obtain as much knowledge as possible. The effective way for them to absorb knowledge in volume is through rote learning and memorization. Besides, Chinese language is based on characters, rather than pronunciations. In order to become literate, Chinese students have to memorize "4500 characters" (Mast, 2016, p. 43). In addition to this emphasis on knowledge, there is a strong belief among Chinese in the wisdom of elders. Students learn how to read Chinese through memorization of "classical literature throughout all compulsory years of education" (Mast, 2016, p. 44). Teachers play a key role in transmitting the message of the wise men or various knowledge. After teachers pose rhetorical questions, "like Confucius, they then give the answer in order to share the wisdom" (Mast, 2016, p. 44). This is why there is progress when Chinese teachers do not strongly value standardized answers and advocate for reference answers.

This study also showed Chinese teachers held various attitudes and views about textbooks. School textbooks, as an important resource in support of teaching and learning in China, play an important role in classrooms (Liu & Laohawiriyanon, 2013). Most participants in this study stressed the usefulness of textbooks. Some stated textbooks in China were regarded as authority. Some claimed textbooks should be taken as reference. Some also emphasized

textbooks played an important role in the educational process. There was a conflicting view among Chinese teachers concerning challenging authority. All Chinese teachers claimed they liked and encouraged their students to question and challenge them in class. Most of them also compared identifying flaws in the textbook to challenging authority. They thought it was good for students to challenge authority. However, for teachers themselves, when they were asked to offer suggestions to their schools concerning CT education, they refused to do so. For them, it was an indirect way to challenge authority, and it was unlikely that this bottom-up suggestion would be accepted by the top leaders.

The aforementioned implied that the Chinese education system is deeply rooted in Chinese culture (Mast, 2016). The traditional aspects of Chinese culture are entrenched in the pervasive influence of Confucius (Bush & Haiyan, 2000). Confucius was the first scholar to start private education and recruited disciples in the Spring and Autumn period in ancient China (Gu, 2006). Confucius emphasized education and hoped "through education, the ruler could become wise and the subjects could 'be civilized,' thus become good subjects" (Gu, 2006, p. 170). Put another way, education and a set of rites made people become obedient (Gu, 2006). Traditional Chinese culture is "reflected in continued respect for authority, collectivism and harmony in schools" (Bush & Haiyan, 2000, p. 58), indicating Confucianism still exerts strong influence on Chinese education. This explained why both Chinese teachers and students are reluctant to challenge authority in the educational practice.

This study also showed all Chinese teachers had a strong desire for CT training. They were willing and eager to participate in any CT training program. For them, if teachers know nothing about CT, how could they teach their students to think critically? One of the barriers for them to train students' CT skills was from the macro level: the exam-oriented education system

created difficulty in stimulating and cultivating students' creativity and innovation. It caused challenges for teachers to prioritize developing CT skills over exams. This finding corroborated the research results of Li (2016). Other studies, such as Craft (2015), Zohar (2008), and Zawojewski and McCarthy (2007), also claimed high-stakes exams hampered the development of students' thinking.

At the meso level, participants in this study reported there were no formal or official documents to guide them to integrate CT into their discipline. It indicated that for some colleges and universities, CT education was not carried out in depth and it was only superficial. When colleges and universities did not focus on CT education, teachers could not put their heart and soul into it without guidance and support. Therefore, colleges and universities need to elevate the slogan of "integrating CT into disciplines" to the policy level, and cover this slogan in the goals and missions of school development. It was also vitally important for colleges and universities to render enough support, encouragement, and help for teachers to teach CT in class.

At the micro level, teacher-related barriers or factors were: most Chinese teachers knew and accessed CT through either the ISEC program or the going abroad experiences. It meant they knew little about the definition and connotation of CT, as well as the ways to improve CT, before they reached adulthood. In their childhood, all of them were encouraged to obtain more knowledge, and cultivation of their thinking skills was neglected. When they became adults, their thinking mode had been focalized. They themselves lacked innovation and creativity. It meant Chinese teachers lacked knowledge and awareness of CT. Another barrier was when teachers reached the middle age, such as 40s or late 40s, laziness and being self-content get in the way from further developing and improving themselves for career advancement. The most challenging problem for most teachers was they did not have much idea about how to effectively

integrate CT into their discipline. It implied Chinese teachers were lacking in CT skills. Student-related factors or barriers included: students' previous learning habits in secondary schools hindered them from thinking critically. Students were taught to show respect for their teachers from their childhood. They became obedient and passive and were not willing to challenge authority (the teachers) in their learning process. This finding supported the research results of Zhang, et al. (2020). Students in Zhang, et al.'s study had no strong motivation to learn after they went to college. They exhibited an inactive attitude towards learning and did not cooperate with others in class. Many students were in a state of "lying flat".

To summarize, all participants in this study showed strong advocacy and support for CT instruction. They also demonstrated a strong desire to take part in CT training programs. They stressed the imperative of having CT knowledge and awareness, as well as the feasible strategies of teaching CT. The findings of this study aligned with several research results, such as Stapleton (2011), Asgharheidari and Tahriri (2015), Zhang, et al. (2020). All of these studies discovered their participants conveyed a strong desire for more CT training. With regard to how to improve teachers' CT, all participants in this study rendered some suggestions, such as attending more experts' presentations, participating in workshops and seminars, building interschool partnerships, taking part in teaching experiences sharing (teaching peer collaboration), attending teachers' demonstration class, and engage in class observation. Some participants suggested colleges and universities should sponsor some teacher exchange programs, in addition to CT training programs.

Implications of the Findings

Although teachers play a crucial role in the process of teaching CT skills, there is a paucity of literature concentrating on teachers' CT skills and dispositions (Janssen, et al., 2019).

Given the fact that there is a dearth of literature on teachers' CT, especially Chinese teachers' CT, an explanatory mixed methods research design was utilized in the current study to first explore the status quo of Chinese teachers' (including the ISEC and non-ISEC teachers) CT skills. The relationships and differences among variables were investigated between the ISEC and non-ISEC teachers, and these variables included scores of CT skills, age, gender, years of teaching, professional rank, educational background, discipline. Finally, the perception, attitude, and practice of Chinese teachers were examined, who came from the institutions of higher education in the north of China.

The results from the quantitative research indicated the Chinese teachers' CT skills were at the upper range of the moderate level. The CT skills of the ISEC teachers were slightly higher than those of non-ISEC teachers. However, there was no statistically significant difference in CT skills between the ISEC and non-ISEC teachers. There were no significant differences in CT skills for the ISEC and non-ISEC teachers based on gender, professional rank, educational background, and discipline they taught. There were no statistically significant relationships in CT skills for the ISEC and non-ISEC teachers based on age and years of teaching. All interview participants rated their CT skills at the medium or the upper end of the medium. It indicated the results from the qualitative research further supported the findings from the quantitative research.

In the meantime, the findings from the qualitative research further complemented and deepened results from the quantitative research. Qualitative research results showed Chinese teachers had varied and fragmented understandings or perceptions of CT. Although they all held a positive and supportive attitude toward CT education, they applied limited CT teaching strategies in class. The substantive theory grounded and emerged from the qualitative data was Chinese teachers advocate and support CT instruction. All interview participants displayed a

strong desire for CT training, because they believed if teachers did not understand CT, it was impossible for them to teach their students to do CT (Stedman & Adams, 2012). The findings from both quantitative and qualitative research offered a panoramic view about Chinese teachers' CT in institutions of higher education in the north of China.

The implications of the findings from this study included three levels or dimensions. At the micro level, although China has achieved remarkable progress in educational development, Chinese education is still faced with many problems and challenges (Zhu, 2019). For instance, the results of transforming from "exam-oriented" to "quality-oriented" education reform was not satisfying, because the current exam focus was still dominant within the long-term mainstream education (Zhou, 2013). Therefore, after one of the policy makers of Chinese education, the Ministry of Education (MOE) issued plans, policies, and guidelines, they need to render necessary guidance, interpretation, and supervision, ensuring the policies were thoroughly understood and applied by school leaders, and faculty and staff. It needs "further exploration to bridge the gap between policy, curriculum, and practice" (Li, 2016, p. 286).

At the meso level, some participants reported their schools did not render enough support for them to teach CT in class, and there were no formal documents to guide them to integrate CT into their disciplines. The first task for colleges and universities is to put their emphasis on CT education, in order to cultivate more students equipped with creativity and innovation in this globalized world. Colleges and universities need to provide guidelines or policies for teachers to effectively integrate CT into their disciplines. The guidelines or policies need to cover the definition of CT, connotation of CT, and the paradigm or parameters for CT. The positive influence of CT on both teachers and students "should drive the school to invest effort and financial resources to students [and teachers] develop their critical thinking abilities"

(Soeherman, 2010, p. 234). Finally, colleges and universities need to build a campus culture that advocates, encourages, and supports CT education. With the aim at constructing CT-based campus culture, colleges and universities need to offer more CT training programs to develop and improve teachers' CT skills. Only when teachers' CT skills are improved, will more likely students' CT skills be improved. Colleges and universities need to set up a systemic and scientific teachers' CT assessment system to push and encourage teachers to integrate CT into their teaching practice. The first step is to put students' CT cultivation into teachers' syllabus.

At the micro level, given the fact that the CT skills of Chinese teachers were not high, they had a fragmented perception of CT, and they applied limited CT teaching strategies into their practice, it is imperative for Chinese teachers to obtain more CT knowledge, develop CT awareness, and improve their CT skills effectively. According to Li (2016), teachers need to "understand what 'good' thinking skills are and be more open to various ideas of developing students' higher order thinking skill" (p. 285). Teachers also need to "develop pedagogical awareness and skills" (Li, 2016, p. 285), which they can effectively integrate CT into their educational process. Put another way, if teachers are lacking in knowledge, awareness, teaching methodologies, and skills of CT, it will be quite difficult for them to effectively implement teaching thinking skills in their educational practice (Mok, 2009).

Although all participants reported they applied CT teaching strategies in class, the number of these strategies were limited. Teachers should realize some of their teaching methodologies or strategies may have no or little impact on the improvement of students' CT skills. Teachers first need to construct a CT-based atmosphere in class, give students more open topics, such as current affairs and issues, to discuss. Then teachers need to offer students good examples of how to use CT to solve problems. Finally, teachers need to learn and utilize various

CT strategies in the teaching process, such as, case study, debate, concept mapping, scenario-based course exercises, active learning techniques, problem-based learning, inquiry-based learning, guided practice, computer-based instruction, structured web-based bulletin boards, and online instruction (Behar-Horenstein & Niu, 2011), in additional to four main strategies that participants self-reported: group discussion, open questions, presentation, and writing.

In order to improve knowledge, awareness, and skills of CT, teachers need to read more books concerning CT theories and CT instructions, and participate in various CT training programs "focusing on concrete CT pedagogy and assessment" (Zhang, et al., 2020, p. 491). Teachers also need to participate in CT workshops and seminars, build interschool partnerships at home and abroad, take part in teaching experiences sharing (teaching peer collaboration), attend CT experts' presentation and teachers' demonstration class, participate in teacher exchange programs, and engage in class observation. These strategies were presented by all participants in this study. Finally, some participants rendered two formulas for teachers to improve their CT and CT instruction: "theory-practice-reflection" and "awareness-recognition-practice", which can be taken as a reference or paradigm for others to try and test their effects in their own teaching practice.

Li (2016) stressed teachers need to consciously implement CT strategies in class, and "appropriate training/professional development opportunities are need[ed] to tailor teachers' need" (p. 285). She also recommended two CT teaching strategies: case study and video-based reflection. A case study "might be an effective way where teachers are able to see how thinking skills are promoted in similar classrooms using these strategies" (Li, 2016, p. 285). A video-based reflection is

> also useful in this case where teachers are able to take an instance of their classroom to critically review and assess whether there is a thinking moment and how task/activity could be realized to accommodate both pedagogical aim of developing language skills/knowledge and higher order thinking. (Li, 2016, p. 285)

Both strategies demand "a collaborative environment where teachers can learn and reflect in a community" (Li, 2016, p. 285).

Recommendations

This non-experimental causal-comparative study provided a broad view of teachers' CT in institutions of higher education in the north of China. The results of this study rendered some guidance for policy makers, school leaders, researchers, and educators to learn more about teachers' CT. Recommendations of this study are explored from two dimensions: recommendations for practitioners/others, and recommendations for future studies.

For Practitioners/Others

The recommendations for practitioners/others are covered from three perspectives: policy makers, school leaders, and educators. From the perspective of policy makers, they should formulate plans, policies, and guidelines based on the social and economic development of that culture. Take China as an example: since the founding of the People's Republic of China in 1949, China has focused on the integration of national development and personal development in education (Zhu, 2019). In 2010, the National Medium- and Long-Term Education Reform and Development Plan (2010–2020) was issued to further guide and advance the education reform in China. As China "has become the second largest economy in the world and achieved remarkable progress in educational development" (Zhu, 2019, p. 355), "China's Education Modernization 2035 Plan" was implemented in 2019 to cater to the educational development in the new era

(Zhu, 2019). The major policy makers of Chinese education, the Ministry of Education (MOE) should guide, advance, and supervise all schools to carry out the education modernization thoroughly, and avoid letting the implementation of education modernization remain at the policy level. Additionally, the ISEC program needs to continue advocating CT education in Chinese higher education. The ISEC office needs to continue hosting more CT training programs with additional focus on improving Chinese teachers' CT skills.

From the leaders' perspective, school leaders should first construct a CT-based campus culture, formulate policies or guidelines to advocate CT education, and encourage teachers to carry out critical thinking-based pedagogy in their teaching process. School leaders need to render "more CT-based resources to teachers, and design relevant teacher training programs focusing on concrete CT pedagogy and assessment" (Zhang, et al., 2020, p. 491). Finally, school leaders need to formulate a scientific CT assessment system to motivate teachers to integrate CT into the teaching practice.

From the perspective of educators or faculty members, teachers need to learn about the concept and theories of CT by various means, such as, do more reading on CT, attend more presentations, workshops, or seminars. According to Choy and Cheah (2009), "teachers' perceptions of CT among students influence their behaviors in the classroom" (p.198). If teachers have a fragmented perception or understanding of CT, they "may find it challenge to teach students critical thinking, as it is sometimes difficult to incorporate aspects of critical thinking into their lessons" (Choy & Cheah, 2009, p. 199). Teachers need to know they are not disseminators of knowledge, but mediators of students' learning (Choy & Cheah, 2009). In this regard, teachers should have a holistic perception of CT, act as a mediator of students' learning, and integrate CT into their teaching process. Despite utilizing limited CT strategies in class,

teachers held a positive attitude toward CT and CT instruction. It is imperative for teachers to take part in CT training programs to improve their own CT skills. To summarize, policy makers, school leaders, and educators can use the findings of this study to create appropriate strategies, policies, and procedures to increase the quality of their support to improve CT instruction, service, and learning.

For the Future Study

This study utilized an explanatory mixed methods research design to explore the status quo of Chinese teachers' CT skills, the significant differences and relationships among eight variables, as well as their perception, attitude and practice of CT in institutions of higher education in the north of China. Because this study delimited itself to the teachers' CT in institutions of higher education in the north of China, future studies can be conducted to center on the teachers' CT in the institutions of higher education in the south of China, in order to obtain an overview of teachers' CT all over China. Due to the big gap between the north and south of China with respect to climate, environment, economy, and politics, it is necessary to make a comparative study on CT between teachers in the north and those in the south. Thus, the research results will render a more holistic and comprehensive view of CT among the whole population of Chinese teachers in institutions of higher education.

This study employed a mixed methods research design. With regard to the quantitative research paradigm, the instrument of the Chinese version of CCTST was utilized to collect data. For the qualitative paradigm, a grounded theory approach was used to examine the perception, attitude and practice of CT in institutions of higher education in the north of China. All qualitative data were collected via interviews, verbatim transcripts, and field notes. In addition to the standardized questionnaire and interviews, future studies could include classroom

observations to offer "a clearer picture of how CT is actually taught [by teachers in their practice]" (Zhang, et al., 2020, p. 491). The findings were self-reported only from the perspective of teachers, "in-depth qualitative studies are needed to be conducted from students' perspective" (Zhang, et al., 2020, p. 491). Because this study used a small nonprobability sample (102 for quantitative), it is hard to generalize and/or transfer research results to a whole population of teachers in China. Future studies could replicate this study with a larger sample from a wider scope, not only in the north and/or south of China, but also in other countries.

Some research indicated there is "a significant relationship between culture and thinking" (Merrifield, 2018, p. 2). Nisbett, et al (2001) elaborated the cognitive processes caused by a specific cultural situation cannot be interpreted from their context at the time. Since cultures and contexts exert great influence on people's perception, attitude, and behavior about CT, future studies could explore the concept of "how thinking is situated in a given sociocultural context and to what degree thinking skills are universal and/or culturally specific" (Li, 2016, p. 287).

Contribution to the Field

Considering there is a dearth of literature on teachers' CT skills in the educational practice, the study bridged the gap in this field of literature. The results of the study added new knowledge to the literature on teachers' CT, especially Chinese teachers' CT in the educational process. The findings of the study help policy makers, school leaders, researchers, and educators learn more about the status quo of Chinese teachers' CT skills, their perception, attitude, and experience in the educational process. Furthermore, policy makers, school leaders, researchers, and educators can take the findings of the study as a reference to create appropriate and effective plans, policies, and guidelines to improve teachers' CT and the quality level of their instruction

and service, and thus improve students' CT and produce more talents with creativity and innovation.

Conclusion

This study utilized a non-experimental causal-comparative methodology with an explanatory mixed methods research design. It was aimed to explore the status quo of Chinese teachers CT skills, as well as the perception, attitude, and practice of CT among them in institutions of higher education in the north of China. The results from the quantitative research revealed the CT skills of Chinese teachers in institutions of higher education were not high, falling to the upper range of moderate level. There were no statistically significant differences in the CT skills for ISEC and non-ISEC teachers. There were no significant differences or relationships in CT skills for ISEC and non-ISEC teachers based on the variables: gender, professional rank, educational background, discipline they taught, age, and years of teaching, either. The results from the qualitative research indicated Chinese teachers advocated and supported CT instruction, and they had a varied and fragmented perception about CT. Although they held a positive attitude towards CT and CT instruction, they applied limited CT teaching strategies in their practice. All of them displayed a strong desire to participate in the CT training programs.

The finding from the qualitative paradigm was *advocacy and support for CT instruction*. It further supported, complemented, and deepened the findings from the quantitative paradigm, which offered a panoramic view of Chinese teachers' CT in institutions of higher education in the north of China. Since there is a paucity of literature focusing on teachers' CT, this explanatory mixed methods research design filled the gap in this field of literature. The results of this non-experimental causal-comparative study added new knowledge to the literature on

teachers' CT, especially Chinese teachers' CT in institutions of higher education. Future studies can be conducted to center on the teachers' CT in the institutions of higher education in the south of China, in order to obtain a holistic and comprehensive overview of the whole population of teachers' CT in China. Future studies could also include classroom observations to offer a more authentic picture of how teachers teach students CT, or replicate this study with a larger sample from a wider scope, not only in the north and/or south of China, but also in other countries, in order to generalize the findings to a larger population. Future studies could explore how sociocultural factors impact teachers' CT instructions, as well. To summarize, developing the CT skills of teachers and students is vitally important in this globalized world. However, if teachers know little about CT and CT instruction, it is unlikely that teachers can successfully integrate CT into their disciplines and thus improve students' CT. Therefore, in order to cultivate more students with CT, creativity, and innovation, it is a premise and must to develop and improve teachers' CT.

References

Adelman, C., Ewell, P., Gaston, P., & Schneider, C. G. (2014). The degree qualifications profile: A learning-centered framework for what college graduates should know and be able to do to earn the Associate, Bachelor's or Master's Degree. *Lumina Foundation for Education.*

Airasian, P. W. (1994). The impact of the taxonomy on testing and evaluation. In L. W. Anderson & L. A. Sosniak (Eds.), *Bloom's taxonomy: A forty-year retrospective: Ninety-third yearbook of the National Society for the Study of Education* (pp. 82–102). Chicago: University of Chicago Press.

Allen, R. M. (2017). A comparison of China's "Ivy League" to other peer groupings through global university rankings. *Journal of Studies in International Education*, 21(5), 395-411.

Asgharheidari, F., & Tahriri, A. (2015). A survey of EFL teachers' attitudes towards critical thinking instruction. *Journal of Language Teaching and Research*, 6(2), 388.

Atabaki, A. M. S., Keshtiaray, N., & Yarmohammadian, M. H. (2015). Scrutiny of critical thinking concept. *International Education Studies*, 8(3), 93-102.

Atkinson, D. (1997). A critical approach to critical thinking in TESOL. *TESOL Quarterly*, 31(1), 71–94.

Bailin, S., Case, R., Coombs, J. R., & Daniels, L. B. (1999). Conceptualizing critical thinking. *Journal of Curriculum Studies*, 31(3), 285–302.

Bailin, S., & Siegel, H. (2003). Critical thinking. In N. Blake, P. Smeyers, R. Smith, & P. Standish (Eds.), *The Blackwell guide to the philosophy of education* (pp. 181-193). Malden, MA: Blackwell.

Bali, M. (2013). *Critical thinking in context: practice at an American Liberal Arts University in Egypt* (Doctoral dissertation, University of Sheffield).

Bali, M. (2015). Critical thinking through a multicultural lens: cultural challenges of teaching critical thinking. In M. Davies & R. Barnett (Eds.), *The Palgrave Handbook of Critical Thinking in Higher Education* (pp.317-334). Palgrave Macmillan Publisher.

Ballard, B. (1996). Through language to learning: preparing overseas students for study in Western universities, in: H. Coleman (Ed.) *Society and the language classroom* Cambridge: Cambridge University Press.

Ballard, B. & Clanchy, C. (1991). *Teaching Students from Overseas*. Melbourne: Longman Cheshire.

Bartlett, D., & Cox, P. D. (2002). Measuring change in students' critical thinking ability: implications for health care education. *Journal of Allied Health,* 31(2), 64-69.

Behar-Horenstein, L. S., & Niu, L. (2011). Teaching critical thinking skills in higher education: A review of the literature. *Journal of College Teaching & Learning (TLC),* 8(2), 25-41.

Bers, T. H., McGowan, M., & Rubin, A. (1996). The disposition to think critically among community college students: The California critical thinking dispositions inventory. *The Journal of General Education*, 45(3), 197-223.

Birjandi, P., & Bagherkazemi, M. (2010). The Relationship between Iranian EFL teachers' critical thinking ability and their professional success. *English language teaching*, 3(2), 135-145.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives, The classification of educational goals, Handbook I: Cognitive domain.* New York: McKay.

Boonjeam, W., Tesaputa, K., & Sri-ampai, A. (2017). Program development for primary school teachers' critical thinking. *International education studies,* 10(2), 131-138.

Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language teaching*, 36(2), 81-109.

Borg, S. (2006). *Teacher education and language education: Research and practice*. London: Continuum.

Branden, P. S. (2012). *The nurse as advocate: A grounded theory perspective* [Doctoral dissertation, Villanova University]. ProQuest LLC.

Bransford, J. D., Sherwood, R. D., & Sturdevant, T. (1987). Teaching thinking and problem solving. In J. Baron & R. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (162-181). New York: W. H. Freeman Co.

Branford, J. D., & Stein, B. S. (1993). *The ideal problem solver: A guide for improving thinking, learning, and creativity*. San Francisco: Freeman.

Brown, S. C., Steven, R. A., Jr., Troiano, P. F., & Schneider, M. K. (2002). Exploring complex phenomena: Grounded theory in student affairs research. *Journal of College Student Development*, 43(2), 173.

Bush, T., & Haiyan, Q. (2000). Leadership and culture in Chinese education. *Asia Pacific Journal of Education*, 20(2), 58-67.

Butler, H. A., Dwyer, C. P., Hogan, M. J., Franco, A., Rivas, S. F., Saiz, C., & Almeida, L. S. (2012). The Halpern Critical Thinking Assessment and real-world outcomes: Cross-national applications. *Thinking Skills and Creativity*, 7(2), 112-121.

Canagarajah, A. S. (2002). *Critical academic writing and multilingual students*. Ann Arbor: University of Michigan Press.

Cave, L.M. (1993). *The relationship of teacher behavior and characteristics to critical thinking skills among middle-level student*. (Doctoral dissertation, Oregon State University)

Chen, L. (2017). Understanding critical thinking in Chinese sociocultural contexts: A case study in a Chinese college. *Thinking Skills and Creativity*, 24, 140-151.

China Education Center. (n.d.). Educational in China. https://www.chinaeducenter.com/en/cedu.php.

The State Council, People's Republic of China. (23 Feb, 2019). China's Education Modernization Plan towards 2035. www.gov.cn/xinwen/2019-02/23/content_5367987.htm

China Scholarship Council (2022). Introduction to the Organization. https://www.csc.edu.cn/about.

Chiu, C. (2020). Project 211, Project 985, C9 League, and Double First Class. https://www.china-admissions.com/blog/guide-to-chinese-universities-project-211-project-985-c9-league/

Choy, S. C., & Cheah, P. K. (2009). Teacher perceptions of critical thinking among students and its influence on higher education. *International Journal of teaching and learning in Higher Education*, 20(2), 198-206.

Clarke, A. M. (1986). Beyond IQ: A triarchic theory of human intelligence. *British Journal of Educational Studies*, 34(2), 205-207.

Clifford, J. S., Boufal, M. M., & Kurtz, J. E. (2004). Personality traits and critical thinking skills in college students: Empirical tests of a two-factor theory. *Assessment*, 11(2), 169-176.

Cohen, J.W. (1988). Statistical Power Analysis for the Behavioral Science (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Colucciello, M. L. (1997). Critical thinking skills and dispositions of baccalaureate nursing students—A conceptual model for evaluation. *Journal of Professional Nursing*, 13(4), 236-245.

Connerley, M. L., & Pedersen, P. B. (2005). *Leadership in a diverse and multicultural environment: Developing awareness, knowledge, and skills*. Sage Publications.

Corbin, J., & Strauss, A. (2015). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (4th ed.). Newbury Park: Sage.

Cortazzi, M., & Jin, L. (1997). Communication for learning across cultures. In D. McNamara & R. Harris (Eds.), *Overseas students in higher education: Issues in teaching and learning*. Routledge.

Craft, A. (2005). *Creativity in schools: Tensions and dilemmas*. London: Routledge.

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating Quantitative and Qualitative Research* (4th ed.). Pearson Education, Inc.

Creswell, J.W., & Poth, C. N. (2018). *Qualitative inquiry research design: Choosing among five approaches* (4th ed.). Sage Edge.

Crozier, M.J., Preston, N., Glade, T. (2013). Antecedent Conditions. In: Bobrowsky, P.T. (eds) *Encyclopedia of Natural Hazards*. *Encyclopedia of Earth Sciences Series*. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-4399-4_13

Cuffari, B. (n.d.). How has the COVID-19 pandemic impacted global health? *News Medical*. https://www.news-medical.net/health/How-has-the-COVID-19-Pandemic-Impacted-Global-Health.aspx

Dai, D. Y., Gerbino, K. A., & Daley, M. J. (2011). Inquiry-based learning in China: Do teachers practice what they preach, and why?. *Front. Educ China*, 6(1), 139–157.

Darling-Hammond, L., & Bransford, J. (Eds.). (2005). *Preparing teachers for a changing world: What teachers should learn and be able to do.* John Wiley & Sons.

Davies, M. (2013). Critical thinking and the disciplines reconsidered. *Higher Education Research & Development*, 32(4), 529-544.

Davies, M., & Barnett, R. (Eds.). (2015). *The Palgrave handbook of critical thinking in higher education*. Palgrave Macmillan Publisher.

De Bono, E. (1967). *The use of lateral thinking*. Penguin Books Ltd.

De Bono, E. (1969). *The mechanism of mind*. Penguin Books Ltd.

De Bono, E. (1985). The direct teaching of thinking as a skill. https://www.um.edu.mt/library/oar/bitstream/123456789/51782/1/Education2(1)A1.pdf

De Bono, E. (2002). New thinking. *Think*, 1(1), 39-48.

De Bono, E., & Zimbalist, E. (1970). *Lateral thinking*. London: Penguin.

Dewey, J. (1910). *How we think*. Courier Corporation. https://pure.mpg.de/rest/items/item_2316308/component/file_2316307/content

Dike, S. E., Kochan, F. K., Reed, C., & Ross, M. (2006). Exploring conceptions of critical thinking held by military educators in higher education settings. *International Journal of Leadership in Education*, 9(1), 45-60.

Dong, Y (2015). Critical thinking education with Chinese characteristics. In M. Davies & R. Barnett (Eds.), *The Palgrave handbook of critical thinking in higher education* (pp.351-368). Palgrave Macmillan Publisher.

Dow, B.J., & Wood, J.T. (2006). *The Sage handbook of gender and communication*. Thousand Oaks, CA: Sage.

Du, T. J., Yu, N., & Guo, S. Y. (2008). A study on the influence of PBL teaching model on students' critical thinking ability. *Journal of Higher Medical Education in China.*

Durkin, K. (2008a). The middle way: East Asian master's students' perceptions of critical argumentation in UK universities. *Journal of Studies in International Education*, 12(1), 38–55.

Durkin, K. (2008b). The adaptation of East Asian masters students to western norms of critical thinking and argumentation in the UK. *Intercultural Education*, 19(1), 15–27.

Egege, S., & Kutieleh, S. (2004). Critical Thinking: Teaching Foreign Notions to Foreign Students. *International Education Journal*, 4(4), 75-85.

Ennis, R. H. (1962). *A Concept of Critical thinking*. Harvard educational review.

Ennis, R.H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. B. Baron, & R.J. Sternberg (Eds), *Teaching critical thinking skills: Theory and practice* (pp. 9-26). New York: Freeman,

Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational researcher*, 18(3), 4-10.

Ennis, R. H. (1996). Critical thinking dispositions: Their nature and assessability. *Informal logic*, 18(2). 165-182. https://ojs.uwindsor.ca/index.php/informal_logic/article/view/2378

Ennis, R. H. (2011). Critical thinking: Reflection and perspective Part I. *Inquiry: Critical thinking across the Disciplines*, 26(1), 4-18.

Ennis, R. H. (2015). Critical thinking: A streamlined conception. In M. Davies & J. R. Barnett (Eds.), *The Palgrave Handbook of Critical Thinking in Higher Education* (pp. 31-47). Palgrave Macmillan Publisher.

Ennis, R. H., Millman, J., & Tomko, T. N. (2005). *Cornell Critical Thinking Tests Level X & Level Z Manual* (5th ed.). The Critical Thinking CO.

Ennis, R. H. & Weir, E. (1985). *Test Manual critical scoring sheet: An instrument for teaching and Testing*. Midwest Publications.

Facione, P. A. (1990a). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction (The Delphi Report).

Facione, P. A. (1990b). The California Critical Thinking Skills Test--College Level. Technical Report# 1. Experimental Validation and Content Validity.

Facione, P. A. (1990c). The California Critical Thinking Skills Test--College Level. Technical Report No. 3. Gender, Ethnicity, Major, CT Self-Esteem, and the CCTST.

Facione, P. A. (1991). Using the California Critical Thinking Skills Test in Research. Evaluation, and Assessment, 1-22.

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). *Test manual: The California critical thinking dispositions inventory*. The California Academic Press, Millbrae, CA.

Facione, N. C., & Facione, P. A. (1997). *Critical thinking assessment in nursing education programs: An aggregate data analysis*. The California Academic Press, Millbrae, CA, USA.

Facione, P. A., Sanchez, C. A., Facione, N. C., & Gainen, J. (1995). The disposition toward critical thinking. *The Journal of General Education*, 44(1), 1-25.

Farrell, T. S., & Kun, S. T. K. (2008). Language policy, language teachers' beliefs, and classroom practices. *Applied Linguistics*, 29(3), 381-403.

Fasko, Jr., D., & Fair, F. (Eds.). (2021). *Critical thinking and reasoning: Theory, development, instruction, and assessment*. Brill Sense.

Feucht, F. C., & Bendixen, L. D. (2010). Exploring similarities and differences in personal epistemologies of US and German elementary school teachers. *Cognition and Instruction*, 28(1), 39-69.

Fischer, S. C., & Spiker, V. A. (2000). A framework for critical thinking research and training. *Report Prepared for the US Army Research Institute*.

Flowerdew, L. (1998). A cultural perspective on group work. *ELT Journal*, 52(4), 323–328.

Forehand, M. (2010). Bloom's taxonomy. *Emerging perspectives on learning, teaching, and technology*, 41(4), 47-56.

French, J. N., & Rhoder, C. (1992). *Teaching thinking skills: Theory and practice*. Garland Publishing, Inc.

Friend, C. M., & Zubek, J. P. (1958). The effects of age on critical thinking ability. *Journal of Gerontology*, 13(4), 407-413.

Frisby, C. L. (1992). Construct validity and psychometric properties of the Cornell Critical Thinking Test (Level Z): a contrasted groups analysis. *Psychological Reports*, 71(1), 291.

Furst, E. J. (1994). Bloom's taxonomy: Philosophical and educational issues. In L. W. Anderson & L. A. Sosniak (Eds.), *Bloom's taxonomy: A forty-year retrospective: Ninety-third yearbook of the National Society for the Study of Education* (pp. 28–40). Chicago: University of Chicago Press.

Godzella, B. M., & Masten, W. G. (1998). Critical thinking and learning processes for students in two major fields. *Journal of Instructional Psychology*, 25(4), 256.

Gagné, R. M. (1965). *The conditions of learning*. New York: Holt, Rinehart and Winston.

Gagné, R. M. (1972). Domains of learning. *Interchange*, 3(1), 1-8.

Gagné, R. M. (1980). Learnable aspects of problem solving. *Educational Psychologist*, 15(2), 84-92.

Gagné, R. M. (1984). Learning outcomes and their effects: Useful categories of human performance. *American psychologist*, 39(4), 377-385.

Gao, H., Chen, J., Wang, B., Tan, S. C., Lee, C. M., Yao, X., Yan, H., & Shi, J. (2011). A study of air pollution of city clusters. *Atmospheric Environment*, 45(18), 3069-3077.

Gay, L. R., Mills, G. E., & Airasian, P. (2006). *Educational research: Competencies for analysis and applications* (8th ed.). Rearson Education, Inc.

Gelder, T. V. (2005). Teaching critical thinking: Some lessons from cognitive science. *College teaching*, 53(1), 41-48.

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.

Glen, S (n.d.). Snowball sampling: Definition, advantages and disadvantages. From StatisticsHowTo.com: Elementary Statistics for the rest of us! https://www.statisticshowto.com/probability-and-statistics/statistics-definitions/snowball-sampling/

Grimshaw, T. (2007). Problematizing the construct of 'the Chinese learner': Insights from ethnographic research. *Educational Studies*, 33(3), 299-311.

Grosser, M. M., & Lombard, B. J. J. (2008). The relationship between culture and the development of critical thinking abilities of prospective teachers. *Teaching and Teacher Education*, 24(5), 1364-1375.

Gu, Z. Z, & Liu, Z. H. (2015). *Critical thinking course*. Beijing: Peking University Press

Guan, Q., & Meng, W. (2007). China's new national curriculum reform: Innovation, challenges and strategies. *Frontiers of Education in China*, 2(4), 579-604.

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field methods*, 18(1), 59-82.

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring. *American psychologist*, 53(4), 449-455

Halpern, D. F. (2001). Assessing the effectiveness of critical thinking instruction. *The Journal of General Education*, 50(4), 270-286.

Halpern, D. F. (2003). Thought and knowledge: An introduction to critical thinking, (4th ed). Lawrence Erlbaum Associates, Inc.

Halpern, D. F. (2010). *Halpern Critical Thinking Assessment manual*. Vienna, Austria: Schuhfried GmbH.

Hassan, K. E., & Madhum, G. (2007). Validating the Watson Glaser critical thinking appraisal. *Higher Education*, 54(3), 361-383.

Hidayati, Y., & Sinaga, P. (2019). The profile of critical thinking skills students on science learning. *In Journal of Physics: Conference Series* (Vol. 1402, No. 4, p. 044075). IOP Publishing.

Hitchcock, D. (2018). History. Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/critical-thinking/history.html

Hu, G. (2002). Potential cultural resistance to pedagogical imports: The case of communicative language teaching in China. *Language, Culture and Curriculum*, 15(2), 93-105.

Huang, C. L. (2019). *A Study on College Critical Thinking Course Design from the perspective of general education curriculum*. [Doctoral dissertation, Shanghai Normal University].

Huang, C. Y. (2015). *Developing and validating college students' Critical Thinking Disposition Scale*. [Master's thesis, Southwest University].

Huitt, W. (1998). Critical thinking: An overview. *Educational psychology interactive*, 3(6), 34-50. http://www.edpsycinteractive.org/topics/cognition/critthnk.html

Insight Assessment. (2021). California Critical Thinking Skills Test: 2021 User Manual and Resource Guide.

Insight Assessment Report. (January, 2023). California Critical Thinking Skills Test (CCTST). Insight Assessment.

ISEC Office. (2022). Brief Introduction of ISEC Program in China.

Iwaoka, W. T., Li, Y., & Rhee, W. Y. (2010). Measuring gains in critical thinking in food science and human nutrition courses: The Cornell Critical Thinking Test, problem-based learning activities, and student journal entries. J*ournal of Food Science Education*, 9(3), 68-75.

Jackson, I., & Du, S. (2022). The impact of history textbooks on young Chinese people's understanding of the past: A social media analysis. *Journal of Current Chinese Affairs*, 51(2), 194-218.

Janssen, E. M., Mainhard, T., Buisman, R. S., Verkoeijen, P. P., Heijltjes, A. E., Van Peppen, L. M., & Van Gog, T. (2019). Training higher education teachers' critical thinking and attitudes towards teaching it. *Contemporary Educational Psychology*, 58, 310-322.

Johnston, M. (2000). New thinking for the new millennium, Edward de Bono. *Long Range Planning*, 33(1), 151-152.

Ju, X. (2017). *Research on ISEC teaching idea and teaching practice -- A case study of Baotou Teachers' College*. [Master's thesis, Inner Mongolia Normal University].

Karadağ, F., & Demirtaş, V. Y. (2018). The effectiveness of the Philosophy with children curriculum on critical thinking skills of pre-school children. *Education & Science/Egitim ve Bilim*, 43(195), 1-22.

Ku, K. Y. L. (2009). Assessing students' critical thinking performance: Urging for measurements using multi-response format. *Thinking Skills and Creativity*, 4 (1), 70-76. https://www.sciencedirect.com/science/article/pii/S1871187109000054

Lai, E. R. (2011). Critical thinking: A literature review. *Pearson's Research Reports*, 6(1), 1-50.

Leach, B. T. (2011). *Critical thinking skills as related to university students' gender and academic discipline* (Doctoral dissertation, East Tennessee State University).

Lee, T. Y. (1993). Comparisons of cognitive development, science process skills, and attitude toward science among Republic of China preservice teachers with different science backgrounds. *Science Education*, 77(6), 625-636.

Leedy, P. D., & Ormrod, J. E. (2013). *Practical research: Planning and design* (10th ed.). Upper Saddle River, NJ: Pearson.

Lewis, A., & Smith, D. (1993). Defining higher order thinking. *Theory into practice*, 32(3), 131-137.

Li, J. J. (2019). *Research on the Cultivation of critical thinking Ability of middle school students in Physics* [Doctoral dissertation, East China Normal University].

Li, L. (2016). Integrating thinking skills in foreign language learning: What can we learn from teachers' perspectives?. *Thinking Skills and Creativity*, 22, 273-288.

Li, R. (2017). *Critical thinking in Chinese higher education: A case study of knowledge transfer* [Doctoral dissertation, University of Nottingham].

Li, M., & Li, J. X. (2016). Critical thinking education at Tsinghua University. *Journal of Critical Thinking and Innovative Education Newsletter*.

Lincoln, Y.S., & Guba, E.G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.

Lipman, M. (1987). Critical thinking: What can it be?. *Analytic Teaching*, 8(1), 5-12.

Lipman, M. (1988). *Philosophy goes to school*. Temple University Press.

Lipman, M. (2003). *Thinking in education*. Cambridge university press.

Liu, D. (1998). Ethnocentrism in TESOL: Teacher education and the neglected needs of international TESOL students. *ELT Journal*, 52(1), 2–10.

Liu, L. (2015). Research on critical thinking of college students based on WGCTA. *Journal of Education and Teaching Forum*.

Liu, O. L., Frankel, L., & Roohr, K. C. (2014). Assessing critical thinking in higher education: current state and directions for next-generation assessment. *ETS Research Report Series*. https://onlinelibrary.wiley.com/doi/epdf/10.1002/ets2.12009

Liu, S., & Laohawiriyanon, C. (2013). Cultural content in EFL listening and speaking textbooks for Chinese university students. *International Journal of English Language Education*, 1(1), 82-93.

Lloyd, M., & Bahr, N. (2010). Thinking critically about critical thinking in higher education. *International journal for the Scholarship of Teaching and Learning*, 4(2), 1-19.

Luo, Q. X. (2002). *Critical Thinking Theory and Assessment Technology Research* [Doctoral dissertation, Nanjing Normal University].

Luo, Q. X., & Yang, X. H. (2001). A preliminary revision of the Chinese version of the California Critical Thinking Disposition Inventory. *Journal of Psychological Development and Education*.

Luo, Q. X., & Yang, X. H. (2002). A preliminary revision of the California Critical Thinking Skill Test. *Journal of Psychological Science*.

MacDonald, S. (2005). *A century of art and design education: From arts and crafts to conceptual art*. Cambrigge: Lutterworth Press.

Magnussen, L., Ishida, D., & Itano, J. (2000). The impact of the use of inquiry-based learning as a teaching methodology on the development of critical thinking. *Journal of Nursing Education*, 39(8), 360-364.

Mangena, A., & Chabeli, M. M. (2005). Strategies to overcome obstacles in the facilitation of critical thinking in nursing education. *Nurse Education Today*, 25(4), 291-298.

Marzano, R. J., & Kendall, J. S. (Eds.). (2007). *The new taxonomy of educational objectives*. Corwin Press.

Mast, R. (2016). How culture affects how Chinese students approach learning in western education environments. *The International Schools Journal*, 36(1), 40.

McBride, R.E., Xiang, P., Wittenburg, D., & Shen, J. (2002). An analysis of preservice teachers' dispositions toward critical thinking: A cross-cultural perspective. *Asia-Pacific Journal of Teacher Education*, 30(2), 131–140.

McPeck, J.E. (1981). *Critical Thinking and Education.* New York: St. Martin's Press in Ennis.

Merrifield, W. (2018). *Culture and critical thinking: exploring culturally informed reasoning processes in a Lebanese university using think-aloud protocols* [Doctoral dissertation, George Fox University]. https://digitalcommons.georgefox.edu/edd/112/

Ministry of Education. (MOE, 2001). Outline of curriculum reform for basic education (experimental) [in Chinese]. Retrieved October 1, 2022, from http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/moe_309/200412/4672. html

Ministry of Education. (MOE, 2007). *College English Curriculum requirements*. Beijing: Foreign Language Teaching and Research Press.

Mok, J. (2009). From policies to realities: Developing students' critical thinking in Hong Kong secondary school English writing class. *RELC Journal*, 40(3), 262–279.

Moore, T. (2004). The critical thinking debate: How general are general thinking skills?. *Higher Education Research & Development*, 23(1), 3-18.

Moore, T. J. (2011). Critical thinking and disciplinary thinking: A continuing debate. *Higher Education Research & Development*, 30(3), 261-274.

Naber, J., & Wyatt, T. H. (2014). The effect of reflective writing interventions on the critical thinking skills and dispositions of baccalaureate nursing students. *Nurse Education Today*, 34(1), 67-72.

Ng, S. Y., Cheung, K., & Cheng, H. L. (2022). Critical thinking cognitive skills and their associated factors in Chinese community college students in Hong Kong. *Sustainability*, 14(3), 1-15.

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review*, 108(2), 291–310. Retrieved from 10.1037/0033-295X.108.2.291

Norris, S. P., & Ennis, R. H. (1989). *Evaluating critical thinking*. Pacific Grove, CA: Critical Thinking Press & Software

Pacheco, C. S., & Herrera, C. I. (2021). A conceptual proposal and operational definitions of the cognitive processes of complex thinking. *Thinking skills and creativity*, 39, 100794, 1-10.

Pallant, J. (2010). *SPSS Survival Manual* (4th ed.). Mc Graw Hill.

Pang, Y. (2012). From examination-oriented education to quality education and the merging of higher institutions: China's educational reforms in the past two decades. *Making Connections: Interdisciplinary Approaches to Cultural Diversity*, 13(2), 40-45.

Paton, M. (2005). Is critical analysis foreign to Chinese students? In E. Manalo & G. Wong-Toi (Eds.), *Communication skills in university education: The international dimension* (pp. 1–11). Auckland: Pearson Education.

Paul, R. W. (1992). Critical thinking: What, why, and how? *New Directions for Community Colleges*, 1992(77), 3–24.

Paul, R. (2005). The state of critical thinking today. *New directions for community colleges*, 27-38.

Pennycook, A. (1996). Borrowing others' words: text, ownership, memory and plagiarism. *TESOL Quarterly*, 30(2), 201-230.

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). A study on the reliability and validity of the Critical Thinking Ability Scale. *Chinese Journal of Nursing*.

Possin, K. (2008). A field guide to critical-thinking assessment. *Teaching Philosophy*, 31(3), 201-228.

Privitera, G. K. (2015). *Statistics for the Behavioral Sciences* (2nd ed.). Sage Publisher.

Profetto-McGrath, J. (2003). The relationship of critical thinking skills and critical thinking dispositions of baccalaureate nursing students. *Journal of Advanced Nursing*, 43(6), 569-577.

Quitadamo, I. J., & Kurtz, M. J. (2007). Learning to improve: using writing to increase critical thinking performance in general education biology. *CBE—Life Sciences Education*, 6(2), 140-154.

Redding, D. A. (2001). Critical thinking disposition as it relates to academic achievement in baccalaureate nursing education. *Nurse Educator*, 26(3), 125-127.

Reed, J. H. (1998). *Effect of a model for critical thinking on student achievement in primary source document analysis and interpretation, argumentative reasoning, critical thinking dispositions, and history content in a community college history course* [Doctoral dissertation, University of South Florida].

Robinson, S. R. (2011). Teaching logic and teaching critical thinking: revisiting McPeck. *Higher Education Research & Development*, 30(3), 275-287.

Roohr, K., Olivera-Aguilar, M., Ling, G., & Rikoon, S. (2019). A multi-level modeling approach to investigating students' critical thinking at higher education institutions. *Assessment & Evaluation in Higher Education*, 44(6), 946-960.

Rotenberg, R. (2010). *The art and craft of college teaching*. Walnut Creek, CA: Left Coast.

Ryan, J. (2011). Introduction. In Ryan, J. (eds.), *China's higher education reform and internationalisation* (pp. 1-7). New York: Routledge.

Sainn, G., & Ugwuegbu, D. C. E. (1980). *Education psychology in a changing world*. London: UnwinHyman.

Salkind, N. J. (2000). *Exploring research* (4th ed.).New Jersey: Prentice Hall.

Scriven, M., & Paul, R. (1987, March). Critical thinking. In *The 8th Annual International Conference on Critical Thinking and Education Reform, CA* (Vol. 7, No. 9).

Scott, J. N., Markert, R. J., & Dunn, M. M. (1998). Critical thinking: change during medical school and relationship to performance in clinical clerkships. *Medical education*, 32(1), 14-18.

Siegel, H. (1988). *Educating reason: Rationality, critical thinking and education*. New York: Routledge.

Soeherman, S. (2010). *The relationships of critical thinking skills, critical thinking dispositions, and college experiences of theological students in Indonesia* [Doctoral dissertation, Biola University].

Song, C., Wu, L., Xie, Y., He, J., Chen, X., Wang, T., ... & Mao, H. (2017). Air pollution in China: status and spatiotemporal variations. *Environmental pollution*, 227, 334-347.

Stapleton, P. (2011). A survey of attitudes towards critical thinking among Hong Kong secondary school teachers: Implications for policy change. *Thinking Skills and Creativity*, 6(1), 14-23.

Statistics Times (2021a). United States vs China by population. Retrieved from https://statisticstimes.com/demographics/china-vs-us-population.php

Statistics Times (2021b). Comparing United States and China by economy. Retrieved from https://statisticstimes.com/economy/united-states-vs-china-economy.php

Stedman, N. L., & Adams, B. L. (2012). Identifying faculty's knowledge of critical thinking concepts and perceptions of critical thinking instruction in higher education. *NACTA Journal*, 56(2), 9-14.

Sternberg, R. J. (1983). Criteria for intellectual skills training. *Educational Researcher*, 12(2), 6-26.

Sternberg, R. J. (1985). Teaching critical thinking, Part 1: Are we making critical mistakes?. *Phi Delta Kappa,* 67(3), 194-198.

Sternberg, R. J. (1986). *Critical thinking: Its nature, measurement, and improvement*. https://eric.ed.gov/?id=ED272882

Sternberg, R. J., & Davidson, J. E. (1983). Insight in the gifted. *Educational Psychologist*, 18(1), 51-57.

Stewart, R. A. C., & Krivan, S. L. (2021). Edward de Bono: May 19, 1933–June 9, 2021. *Social Behavior and Personality: An international journal*, 49(8), e10946

Strauss, A., & Corbin, J. (1990*). Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage.

Tan, C. (2017). Constructivism and pedagogical reform in China: issues and challenges. *Globalization, Societies and Education*, 15(2), 238–247. http://dx.doi.org/10.1080/14767724.2015.1105737

Tan, C. (2020). Conceptions and practices of critical thinking in Chinese schools: An example from Shanghai. *Educational Studies*, 56(4), 331-346.

Tan, C., & Hairon, S. (2016). Education reform in China: Toward classroom communities. *Action in Teacher Education*, 38(4), 315-326.

Tang, D., Bu, T., Feng, Q., Liu, Y., & Dong, X. (2020). Differences in overweight and obesity between the North and South of China. *American Journal of Health Behavior*, 44(6), 780-793.

Tanner, L. N. (1988). The path not taken: Dewey's model of inquiry. *Curriculum Inquiry*, 18(4), 471-479.

Thayer-Bacon, B. J. (2000). *Transforming critical thinking: Thinking constructively*. Teachers College Press.

Thomas, K., & Lok, B. (2015). Teaching critical thinking: an operational framework. In Davies, M. & Barnett, R. (Eds.), *The Palgrave Handbook of Critical Thinking in Higher Education* (pp. 93-105). New York, NY: Palgrave Macmillan.

Thompson, C., & Rebeschi, L. M. (1999). Critical thinking skills of baccalaureate nursing students at program entry and exit. *Nursing and Health Care Perspectives*, 20(5), 248-248.

Tian, J. (2008). *Critical thinking of Chinese students.* York: Lap Lambert Academic Publishing.

Tian, J., &, Low, G. D. (2011). Critical thinking and Chinese university students: are view of the evidence. *Language Culture and Curriculum*, 24(1), 61-76.

Tiwari, A., Avery, A., & Lai, P. (2003). Critical thinking disposition of Hong Kong Chinese and Australian nursing students. *Journal of Advanced Nursing*, 44(3), 298–307.

Torff, B. (2005). Developmental changes in teachers' beliefs about critical thinking activities, *Journal of Educational Psychology*, 97(1), 13-22.

Tsui, L. (2002). Fostering critical thinking through effective pedagogy: Evidence from four institutional case studies. *The Journal of Higher Education*, 73(6), 740-763.

Vansieleghem, N., & Kennedy, D. (2011). What is philosophy for children, what is philosophy with children—After Matthew Lipman?. *Journal of Philosophy of Education*, 45(2), 171-182.

Vaughan-Wrobel, B. C., O'Sullivan, P., & Smith, L. (1997). Evaluating critical thinking skills of baccalaureate nursing students. *Journal of nursing education*, 36(10), 485-488.

Vollstedt, M., & Rezat, S. (2019). An introduction to grounded theory with a special focus on axial coding and the coding paradigm. In Kaiser, G., & Presmeg, N. (Eds.), *Compendium for early career researchers in mathematics education* (pp. 81- 100). Springer Open.

Walsh, C. M., & Hardy, R. C. (1999). Dispositional differences in critical thinking related to gender and academic major. *Journal of Nursing Education*, 38(4), 149-155.

Watson, G., & Glaser, E. M. (1980). *Watson-Glaser Critical Thinking Appraisal*. San Antonio, TX: Psychological Corporation.

Watson, G., & Glaser, E. M. (2008a).*Watson–Glaser Critical Thinking Appraisal, forms A and B manual*. Upper Saddle River, NJ: Pearson Education.

Watson, G., & Glaser, E. M. (2008b).*Watson–Glaser Critical Thinking Appraisal short form manual*. Pearson Education: Upper Saddle River, NJ.

Watson, G., & Glaser, E. M. (2010).*Watson–Glaser II Critical Thinking Appraisal: Technical manual and user's guide*. San Antonio, TX: NCS Pearson.

Wan, G. (1998). The educational reforms in the cultural revolution in China: a postmodern critique. Retrieved from http://eric.ed.gov/?id=ED427419

Wan, G. (1999). The learning experience of Chinese students in American universities: A cross-cultural perspective. Retrieved from https://eric.ed.gov/?id=ED439653

Wang, L. L. (2011). *The revise and testing of the self-designed questionnaire of college students' critical thinking*. [Master's thesis, Shantou University].

Wang, X. S. (2006). A study on critical thinking and skills. *Journal of Yangzhou University* (Higher Education Research Edition).

Wanko Keutchafo, E. L., Kerr, J., & Baloyi, O. B. (2022, October). A model for effective nonverbal communication between nurses and older patients: A grounded theory inquiry. In *Healthcare* (Vol. 10, No. 11, p. 2119). MDPI.

Wei, X. (2003). The strategic choice of Chinese higher education internationalization. *Journal of Taizhou Polytechnical Institute.*

Wen, Q. F., Wang, J. Q., Zhao, C. R., Liu, Y. P., & Wang, H. M. (2009). A theoretical framework for constructing Chinese foreign language college students' critical thinking scale. *Journal of Foreign Languages*.

Werner, P. H. (1991). The Ennis-Weir critical thinking essay test: An instrument for testing and teaching. *Journal of Adolescent & Adult Literacy*, 34(6), 494 - 495.

Wheeler, L. A., & Collins, S. K. (2003). The influence of concept mapping on critical thinking in baccalaureate nursing students. *Journal of professional Nursing*, 19(6), 339-346.

Williams, B. (2017). *Clinical psychologists' experiences with poly-victims of trauma: A grounded theory study.* [Doctoral dissertation, Capella University].

Wilson, K. (2016). Critical reading, critical thinking: Delicate scaffolding in English for Academic Purposes (EAP). *Thinking Skills and Creativity*, 22, 256-265.

Wu, H. Z. (2004). What is critical thinking?. Journal of Qinghai Teachers College.

Xiang, G. F. (2012). Review of foreign critical thinking research. Journal of Jixi University.

Xiao, X. (2005). Chinese talent-training strategy under the background of higher education internationalization. *Journal of Harbin Teachers College*.

Xiong, M. H. (2014). Theory and practice of critical thinking. *Journal of Huazhong University of Science and Technology.*

Yang, Y. T. C., & Chou, H. A. (2008). Beyond critical thinking skills: Investigating the relationship between critical thinking skills and dispositions through different online instructional strategies. *British Journal of Educational Technology*, 39(4), 666-684.

Yang, Y. T. C., Newby, T. J., & Bill, R. L. (2005). Using Socratic questioning to promote critical thinking skills through asynchronous discussion forums in distance learning environments. *The American Journal of Distance Education*, 19(3), 163-181.

Yuan, H., Kunaviktikul, W., Klunklin, A., & Williams, B. A. (2008). Improvement of nursing students' critical thinking skills through problem‐based learning in the People's Republic of China: A quasi‐experimental study. *Nursing & health sciences*, 10(1), 70-76.

Yue, W., & Zhu, J. (2009, November). A quantitative perspective on the leading role of top universities in China. *In Proceedings of 3rd International Conference on World-Class Universities: The Role of WCU in National Systems*, 109-119.

Zawojewski, J., & McCarthy, L. (2007). Numeracy in practice. *Principal Leadership*, 7(5), 32–38.

Zhai, J. Y. (2015). Critical thinking curriculum, why claim that Chinese students really need it?. *Chinese Education Daily*.

Zhang, H., Yuan, R., & He, X. (2020). Investigating university EFL teachers' perceptions of critical thinking and its teaching: Voices from China. *Asia-Pacific Edu Res*, 29(5), 483-493.

Zhang, L., & Kim, S. (2018). Critical thinking cultivation in Chinese college English classes. *English Language Teaching*, 11(8), 159-164.

Zhou, Q. (2013). Reflections of English teachers: The quality-oriented education reform in China's middle schools. Journal of Cambridge Studies, 8(1), 155-190.

Zhou, X., Cao, Z., Ma, Y., Wang, L., Wu, R., & Wang, W. (2016). Concentrations, correlations and chemical species of PM2. 5/PM10 based on published data in China: potential implications for the revised particulate standard. *Chemosphere*, 144, 518-526.

Zhu, R. (2017). A study on the relationship between critical thinking and creative thinking, [Doctoral dissertation, Minzu University of China].

Zhu, Y. (2019). New national initiatives of modernizing education in China. *ECNU Review of Education*, 2(3), 353-362.

Zhu, Z. X. (1985). A review of modern cognitive psychology. *Journal of Beijing Normal University.*

Zohar, A. (2008). Teaching thinking on a national scale: Israel's pedagogical horizons. *Thinking Skills and Creativity*, 3(1), 77–81

Appendix A

Online Survey Consent Form (English Version)

You are invited to participate in a research project about "A Comparative Study on Critical Thinking skills of ISEC teachers and Non-ISEC teachers in Institutions of Higher Education in the north of China". The main purpose of this study is to explore the critical thinking (CT) skills of ISEC and non-ISEC teachers in China. This online survey should take about 45-50 minutes to complete. Participation is voluntary, and responses will be kept confidential.

You have the option to not respond to any questions that you choose. Participation or nonparticipation will not impact your relationship or performance in your college or university. Submission of the survey will be interpreted as your informed consent to participate and you affirm that you are at least 18 years of age.

If you have any questions about the research, please contact the Principal Investigator, Liqin Tang, via email at liqin.tang@umontana.edu . If you have any questions regarding your rights as a research subject, contact the UM Institutional Review Board (IRB) at (406) 243-6672.

Please print or save a copy of this page for your records.

☐ I have read the above information and agree to participate in this research project.

☐ No, I am not interested in participating.

Appendix B

Online Survey Consent Form (Chinese Version)

参与网络调查同意书（汉语版）

诚挚邀请您参加“中国北方高校ISEC和非ISEC教师明辨性思维能力对比研究“的调研项目。此项研究的目的是探究中国ISEC和非ISEC教师的明辨性思维能力水平。完成此项网络调查大概需要45-50分钟。您自愿参与此项研究， 调研数据会保密，并且您的名字也不会在调研报告中提及。您可以随时退出本项研究，您的参与和退出不会对您造成任何影响。提交此项调研将意味着您的知情同意参与，并且您至少18岁。感谢您的积极参与。

如果您对此项研究有任何问题，请随时联系调研者唐利芹，她的邮箱是 liqin.tang@umontana.edu。 如果您对您的此次参与权有任何问题，请联系蒙大拿大学机构审查委员会（IRB），电话：(406) 243-6672.

_____我已经阅读了以上研究描述，我同意参加这次调研。

_____我对此项研究不感兴趣。

Appendix C

The Interview Protocol (English Version)

School name: _____ Age: _____ Gender: _____

Years of teaching: _____ Professional title: _____

Educational background: _____ College /Department: _____

Discipline: _____ Participation in CT training: ____Yes / No___

ISEC teacher: ___Yes / No___

Q_1 How do you understand critical thinking (CT)? What is your definition of CT?

Q_2 Do you think CT is helpful or useful for students' learning, exams, and their future development? Why or why not?

Q_3 Do you like those students who challenge or question you in your class? How do you deal with such situations in class?

Q_4 What do you think of your students' CT skills, high, medium, or low? How do you assess your students' CT skills?

Q_5 What do you think of your own CT skills, high, medium, or low? How do you know?

Q_{6a} Have you cultivated your students' CT skills in your course or classroom? If yes, what are your teaching strategies on CT instruction? What difficulties or challenges have you met?

Q_{6b} If no, what are the reasons for it?

Q_7 What do you think of textbooks and so-called standard answers? Recently, there is a popular assignment among senior high school students: Students are required to find the mistakes in their Chinese textbooks, including spellings, punctuation marks, wrong characters, or logical problems in the texts, or what aspects the textbook need to be corrected and improved, etc.. What do you think of this assignment? Are you willing to assign such homework to your

students? Do you think it is beneficial to students' learning or exams? Why or why not?

Q_{8a} Does your college or university emphasize and require the cultivation of students' CT? If yes, how does your school emphasize or require CT instruction or cultivation?

Q_{8b} If no, what recommendations would you like to present to your school?

Q_{9a} Do you have any ideas about how to improve teachers' CT skills?

Q_{9b} Do you think there is a need to participate in the CT training (program)? Why or why not?

Appendix D

The Interview Protocol (Chinese Version)

访谈提纲（汉语版）

所在大学/学院名称： _____ **年龄：** _____

性别： _____ **教学年限:**_____

职称： _____ **教育学历背景:**_____

所在院系： _____ **教授学科/科目：** _____

是否参加过明辨性思维培训： _____**是 / 否**__

是否是 ISEC 教师：____**是 / 否**__

1. 您如何理解明辨性思维？您如何定义明辨性思维？ · ·

2. 您认为明辨性思维对学生的学习有帮助么？对考试有帮助么？对学生今后的人生发展有帮助么？为什么？ · ·

3. 您喜欢学生在课堂上表现出很强的质疑能力吗？面对学生的质疑或者追问您会怎样处理？ · ·

4. 您认为您的学生的明辨性思维能力如何？属于高等水平，中等水平，还是低等水平？您如何评估学生的明辨性思维能力？

5. 您认为您自身的明辨性思维能力如何？属于高等水平，中等水平，还是低等水平？您如何评估自身的明辨性思维能力？

6. 您在课堂上是否有培养学生的明辨性思维？（1）如果有，通常您会采取哪些手段或者教授策略？您觉得还有哪些困难或者挑战？（2）如果没有，请谈一下原因。

7. 您怎样看待教材或者标准答案？最近有一项面对高中生的作业在中国很流行：学生被要求尽量找出语文课本上的种种错误，包括拼写，标点，错别字，或者是课文中存在的逻辑上的问题以及任何学生认为可以有待提高或更正的问题。您是如何看待这种作业的？您会愿意给学生留这种作业么？您认为这种作业对学生的学习或者考试有帮助么？为什么？

8. 您所在学校是否对明辨性思维的培养有所强调或是要求？（1）如果有，学校通过何种方式给予强调或要求？（2）如果没有，您会向学校提出哪些建议？·

9. （1）您认为应该如何提高教师的明辨性思维技能？（2）您认为参加明辨性思维培训项目有用吗？为什么？